PERGAMON

Social Science & Medicine 58 (2004) 57–74

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Area effects on health variation over the life-course: analysis of the longitudinal study sample in England using new data on area of residence in childhood

Sarah Curtis[a,*], H. Southall[b], P. Congdon[a], B. Dodgeon[c]

[a] Department of Geography, Queen Mary College, University of London, London E1 4NS, UK
[b] Department of Geography, University of Portsmouth, UK
[c] Centre for Longitudinal Studies, Institute of Education, University of London, UK

Abstract

There is a growing literature which demonstrates that (a) conditions throughout the life-course are important for health outcomes in older people and (b) 'contextual' conditions in the place of residence, as well as individual characteristics influence health variations. This paper contributes to this debate by presenting results of an analysis of data from the Office for National Statistics Longitudinal Study (LS) for England and Wales. The analysis makes use of a new set of variables, which have been added to the LS, describing the social and economic conditions in the 1930s in residential areas where members of the LS sample were registered as living in 1939. The analysis focuses on people who were aged 0–16 in 1939. The health outcomes considered are death between 1981 and 1991, and for those still living, whether long-term illness was reported in the 1991 census. Regression analysis is used to examine the effects of residence in 1981, and data on the registered place of residence in 1939. The analysis shows that individual characteristics and type of area of residence in 1981 were associated with health outcomes. In addition, some variables describing socio-economic conditions in the 1930s contribute independently to the regression models (notably measures of economic deprivation and unemployment). The results suggest that conditions in residential area in early life may help to explain relatively poor health experience of populations in some parts of Britain today.

Keywords: Mortality; Morbidity; Longitudinal effects; Place effects; Health inequality; England and Wales

Introduction

This research forms part of a project funded under the ESRC Health Variations Programme. The project has examined associations between health variation and area social and economic conditions during the period 1920–1970 (Congdon et al., 2001; Campos et al., 2003). This paper describes how data on socio-economic conditions in local government districts in the 1930s has been linked to individual data in the *Office for National Statistics Longitudinal Study* (LS) and analysed to examine the associations between health of individuals in late middle age and conditions in their residential area in 1939, when they were children (based on the Population Census 8 years before, in 1931). We have carried out analyses to show how these conditions related to their mortality and state of health in 1991. We present results showing how far area conditions in the location of residence during childhood are associated statistically with health in late middle age, controlling for individual characteristics and aspects of area of residence later in life.

Background

A large body of research has built up, especially in Britain, which demonstrates, both theoretically and empirically, that health variation is associated both with

*Corresponding author. Tel.: +44-20-7-975-5400; fax: +44-207-975-5500.
E-mail address: s.e.curtis@qmw.ac.uk (S. Curtis).

0277-9536/03/$ - see front matter
doi:10.1016/S0277-9536(03)00149-7

the socio-economic characteristics of *individual* people and also with the socio-economic characteristics of the wider *communities*, or *places*, in which they live.

The pattern of association with *individual* characteristics has been comprehensively reviewed in several summaries of the evidence for health inequality (Townsend, Davidson, & Whitehead, 1988; Dept of Health, 1998; Gordon, Shaw, Dorling, & Davey-Smith, 1999). The pattern could be crudely summarised as a gradient in health such that those who are most economically disadvantaged (according to measures such as occupational social class or relative advantage in terms of housing tenure) have the poorest health. The gradient is evident, though not identical, for men and women. Certain other social factors also affect this relationship; for example, being with a marital partner is often found to be protective, at least for men. A complex set of causal pathways are invoked to explain these relationships, relating to the interaction of socio-economic inequality, social determinants of health, and health outcomes (e.g. reviewed in Marmot & Wilkinson, 1999; Gordon et al., 1999).

The theoretical and empirical basis for 'place' effects in health inequalities in Britain, operating independent of individual characteristics has been discussed in detail elsewhere (for example, by Duncan, Jones, & Moon, 1993; Duncan & Jones, 1995; Macintyre, Maciver, & Soomans, 1993; Shouls, Congdon, & Curtis, 1996; Congdon, Shouls, & Curtis, 1997; Curtis & Jones, 1998; Popay, Williams, Thomas, & Gatrell, 1998; Sloggett & Joshi, 1998; Macintyre, 1999; Gatrell et al., 2001). This body of research derives mainly from studies which are cross sectional, relating to one point in time.It suggests that individuals with similar personal socio-economic characteristics vary in their health experience if they are living in different communities which vary in terms of their regional position and socio-economic profile. Whatever their individual socio-economic position, people living in poor and socially disadvantaged areas, tend to have worse health experience, than those in more affluent areas. Also, even after allowing for area socio-economic conditions, those living in northern England and Scotland have worse health than those in the South. There is also evidence from some studies that although, on average, populations in more affluent areas have better health, the health gradient between rich and poor individuals may be particularly strong in these more privileged areas (Shouls et al., 1996). Theoretical explanations for the association of health with characteristics of places invoke causal pathways resulting from variation in a range of geographical factors: the material landscape affecting living conditions, the ecological landscape and physical environmental hazards, local opportunities for health related consumption, distribution of wealth, social capital and cultural factors (Macintyre et al., 1993; Wilkinson, 1996; Curtis & Jones, 1998; Popay et al., 1998; Gatrell et al., 2001).

In empirical statistical studies, including a recent set of explicitly multi-level studies (e.g. Duncan et al., 1993; Duncan & Jones, 1995; Shouls et al., 1996; Congdon et al., 1997; Sloggett & Joshi, 1998; Macintyre, 1999), the associations with attributes of place are usually found to 'explain' a smaller part of health difference than individual characteristics and some authors have argued that place effects are insignificant, or are only the result of a failure to measure adequately the individual level variables which are important for health. However, findings from qualitative research, as well as a strong body of theory, indicate that an atomistic view of the processes involved in health inequality does not adequately reflect the ecological processes producing health inequality (Schwartz, 1994). These seem to involve an interaction between individual characteristics and factors in the wider social and economic environment.

The *Health Variations Programme*, funded by the Economic and Social Research Council, has contributed a significant body of new theory and evidence concerning the antecedents of heath inequalities in later life (Graham, 2001). It is clear from this, and from other research, that health disadvantage builds up cumulatively over the life-course and that it is important to consider the impact of conditions early in life (or even in previous generations) in order to explain the differences in health which are manifest later in life (Barker, 1992; Kuh & Ben-Shlomo, 1997; Davey-Smith, Hart, Blane, Gillis, & Hawthorne, 1997; Power & Matthews, 1997; Berney, Blane, Davey-Smith, & Holland, 2001; Benzeval, Dilnot, Judge, & Taylor, 2001).

Some studies in Britain are of particular relevance to the research described here because they have used samples of individuals to examine the association between aspects of place of residence in early life and health outcomes in later life. Strachan, Leon, and Dodgeon (1995) used an approach similar to the one described in this paper to analyse data on individuals from the LS (a national sample, described below). They found that region of residence in 1939, as well as in 1971, was associated with risk of death from cardiovascular causes between 1971 and 1988 among those born before 1939. Elford, Phillips, Thomson, and Shaper (1989) reported an analysis of individuals in the British Regional Heart Study which examined risk of ischaemic heart disease in a sample of middle aged men in 24 selected British towns. Aspects of their present region of residence were associated with risk of IHD events but, within these regions, the same risk of IHD was found, regardless of the region of birth.

Some ecological studies have also examined impact of place of birth and place of residence just before death for aggregated populations of older British born populations. For example, Ben Shlomo and Smith (1991) found that infant mortality in place of birth was not significantly associated with mortality after controlling for socio-economic circumstances in place of death. Socio-economic conditions were measured using an indicator proposed by Carstairs and Morris (1991) and social class composition. Barker (1992) suggested, on the contrary, that relatively low mortality among older women in London 1968–1978 might reflect the relatively good health status of a cohort of women who moved to London from healthier areas in their youth. Osmond, Slattery, and Barker (1990) also reported that place of birth, as well as place of death, was associated with geographical variation in mortality.

This interpretation of the literature on health inequalities leads to the view that (a) health inequalities can be explained in terms of attributes of individuals interacting with attributes of places and communities, and that (b) these processes operate throughout the life-course. However, there is relatively little empirical work which has explored how health inequalities in later life (especially morbidity differences) are related to conditions in the place of residence during childhood. This paper reports on a study that has explored longitudinal relationships, over the life-course, between conditions in the place of residence in childhood and health in later life.

Linkage of historical area data to the Longitudinal Study

The Office for National Statistics Longitudinal Study (LS) is a 1% representative sample of the population of England and Wales, drawn initially from the 1971 census. The LS data are anonymised and no individuals can be identified from the statistics provided. The sample has been followed up subsequently by linking information about the same individuals from the 1971, 1981 and 1991 censuses. Vital events (parenthood, death) since 1971 are also included (Fox & Goldblatt, 1982; Britton, 1990, Chapter 2). However, for those who were already adults by 1971, a much longer period needs to be studied to understand the life-course as a whole, and in particular the relationship between childhood environments and health in later life.

The LS does contain an item of information which, for some of the samples, dates much further back. The NHS was not established until 1948, but it made use of the National Registration numbers used during World War II to manage food rationing. These were originally issued through the National Registration carried out at the outbreak of war, on September 29th 1939. National Registration excluded the ca. 900,000 military personnel in England and Wales in 1939, overwhelmingly male. A much larger number of adults at the time, mostly male, lost their original number while in military service between 1939 and 1948.

With the major exception of these military personnel, the NHS numbers held within the LS identify the local government district (LGD) of residence of each LS member in September 1939, if they were alive in that year. Overall, 62% of those in the LS sample who were aged 0–16 in 1939 have retained this registration number (50% of men and 74% of women). The registration codes are more complete for those who were younger; as shown in Fig. 1, and there are valid codes for more than 70% of all those aged under 50 years in 1981. This information has already been used to study the relative importance of region of origin in 1939 and area of residence in 1971 for ischaemic heart disease and stroke (Strachan et al., 1995). Leon and Strachan (1993) made an analysis of the individuals for whom the 1939 registration data were not available, and who are excluded from their analysis (and from the present study). These men were more likely to be of the age for active service in 1939 and thus in the 60–80 age group by 1991, but, in terms of proportions who were owner occupiers, local authority tenants and car owners, they were similar to men who were included. These authors state that "Those included …do not seem to be appreciably different in terms of socio-economic composition from those excluded because of missing information in 1939 region. Thus, provided age is taken into account the conclusions may be reasonably generalised to the population of England and Wales from which the sample was drawn." (Leon & Strachan, 1993, p. 1449). The new research presented here has greatly extended that earlier study by adding systematic information on the characteristics of the 1930s districts to the information available in the LS.

Besides the exclusion of military personnel, one other limitation of the NHS numbers within the LS should be noted. Although the National Registration was carried out soon after the outbreak of war, a major evacuation of children had already taken place, from areas that were thought likely to be bombed. Many children brought up in urban industrial areas before 1939 are therefore listed as living in rural areas to which they were evacuated. The country was in fact classified into three types of area, as indicated in the 1939 report (see Table 1). Only 7.5% were classified as '*evacuation*' areas, but they included all London Boroughs and all five of the largest provincial cities, covering 31.2% of the 1939 civilian population. A further 18% of districts, including less industrial provincial towns and suburban areas such as the whole of outer Middlesex, were classified as '*neutral*', neither sending nor receiving evacuees, and contained 24.5% of the 1939 civilian population. *Reception* areas comprised 75% of the total number of districts, although they contained only 44.3% of the population. The actual distribution of LS members between these three categories is discussed below, but obviously children living in 'evacuation' or 'neutral'

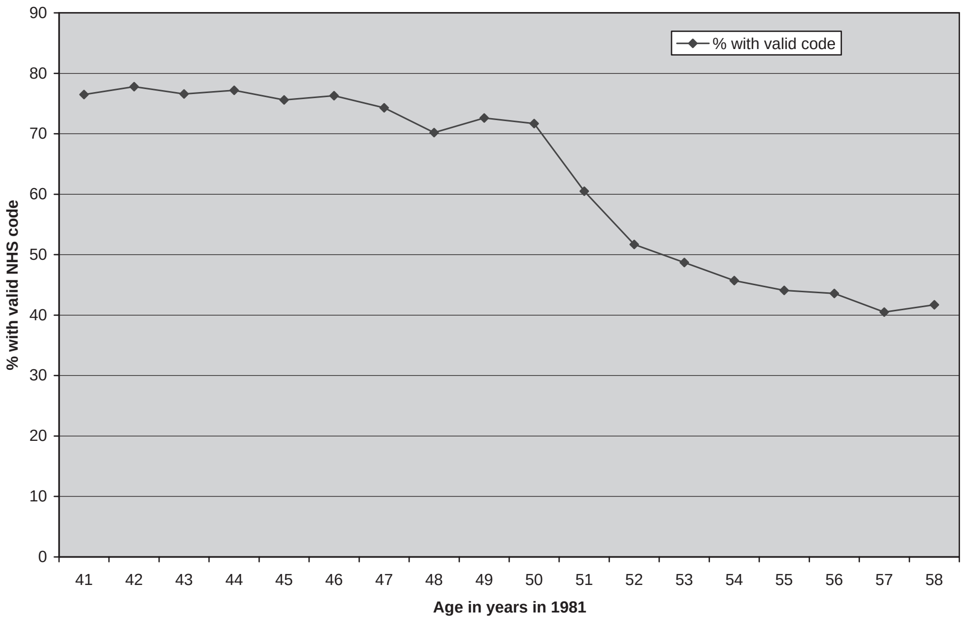

Fig. 1. Proportion of LS sample members aged 0–16 in 1939 retaining valid NHS code.

Table 1
Civilian and LS populations by type of area in 1939

	1939 Local govt. districts (number and % of total)	1939 Civilian population (numbers and % of total)	LS sample members in analysis
Reception area	1105	18,005,000	31,990
%	75	44	51
Evacuation area	110	12,681,000	14,979
%	8	31	24
Neutral area	257	9,966,000	15,750
%	18	25	25
Total	1472	40,652,000	62,719

Source: National Registration Report (1944).

areas had generally been brought up in them, while some of those in 'reception' areas had recently moved there. It seems likely that the actual impact of evacuation was mainly on London, not the previously depressed industrial cities of the north, and that most evacuees were in a small subset of the 'reception' areas. It was not feasible to concentrate in this study only on the populations who had lived in 'neutral' areas, since this would have involved excluding the majority of the sample.

The 1939 National Registration had been planned in parallel with the intended 1941 census, as an alternative should war break out, and used some of the same machinery (National Registration Report, 1944). However, the only personal characteristics recorded were gender and age. For information on socio-economic attributes of area of residence, we must turn to the 1931 census. This is problematic because of changes in the geographical boundaries used to define administrative areas during the 1930s. Although the basic architecture of local government remained constant between 1931 and 1939, consisting of County Boroughs, Municipal Boroughs (all urban units), Rural Districts and London Boroughs, the detailed geography of local government was greatly changed through a rolling programme of county reviews: the 1931 census reports on 1800 LGDs, but this had been reduced to 1472 in 1939. Further, many of the districts which were not abolished were altered through boundary changes: 289 (19.6%) of the 1939 LGDs were new creations or had been affected by boundary changes. A set of supplementary census reports was issued, re-tabulating 1931 data for the units that existed after the county reviews, but unfortunately these were again limited to gender and age.

Ensuring consistent geographic boundaries

Some method was therefore needed to re-cast 1931 data to 1939 units based only on the published information available to us. In studying mortality change over longer periods, we have used a Geographical Information System recording the changing boundaries of local government districts to redistrict data, on the assumption that the population transferred was proportionate to the area transferred (Gregory, Southall, & Dorling, 2000; Gregory, 2001). However, over the period 1931–1939, a more accurate method is possible: the 1931 census reports obviously list of the 1931 population of each LGD as it was defined in 1931, but the 1939 National Registration report also lists the 1931 population of each LGD as defined in 1939, and the lists of boundary changes in the Registrar General's annual *Statistical Reviews* give the 1931 populations of the areas transferred (Registrar General, 1938).

A geography conversion table (Simpson, forthcoming) was therefore constructed from the 1931 and 1939 reports plus the 1805 boundary changes listed for the intervening period, using 1931 populations rather than geographical areas. By linking this table to 1931 census data, we can cut data for the 1800 1931 districts up into 2916 fragments, and then reassemble them into the 1472 1939 units. The table was very carefully cross-checked by comparing the 1931 populations of 1939 units computed by applying the boundary change information to the 1931 census figures, with the 1931 populations listed in the 1939 report. This method avoids any problematic assumptions about population density. However, we are still assuming that the population of a district had uniform socio-economic characteristics; for example, there would have been an equal proportion of unemployed workers in the middle of a town and on the rural fringe. However, boundary changes by which part of an urban area was transferred to the surrounding rural district were very unlikely: most changes consisted of either part of a rural area being transferred into an urban unit, or very small urban units being abolished through merger with an adjacent urban area.

Data

This paper reports on an analysis of information for 62,719 individuals in the LS who were aged 0–16 years in 1939, who were still alive in 1981, and whose place of residence was recorded in the 1939 registration. This represents a reduction on the expected number of 66,397 of LS members in this age group who were recorded at civil registration in 1939. This is almost entirely due to eliminating LS members who were not known to have died between 1981 and 1991, but who nevertheless were not enumerated at the 1991 census, so that we have no information in order to classify their health status in 1991.[1]

Some of these individuals will certainly have recently moved to the area where they were registered in 1939 as a result of evacuation as discussed above, although we cannot identify them. About half of the individuals analysed were living in 'reception' areas, known to have received evacuees, though we have no way to identify which ones may have in fact been evacuees. The other individuals in this sample were either living in 'evacuation' areas, or in areas which are known to have received very few evacuees (these people are more likely to have been registered in areas which were their permanent home in 1939). Table 1 shows that the LS members aged 0–16 in 1939 were more likely to have been in 'reception' areas than the population as a whole, reflecting the fact that they were in the age groups most likely to be evacuated.

The data analysed include information on two health outcomes considered in this analysis: (1) whether or not the individual died between 1981 and 1991, and (2) for those still alive in 1991, the proportion who reported having long-term illness in the 1991 census. These give us measures of survival and of morbidity for the individuals studied in late middle age.

The analysis was intended to investigate associations between these two health outcomes in later life and a selection of variables describing the characteristics of the individuals and of their areas of residence at different points in the life-course. The choice of independent variables was informed by the literature reviewed above on factors associated with health variation in other studies. The selection was, to some extent, constrained by the range of variables available. Restrictions to protect confidentiality of the LS sample also limited the degree of detailed classification which was permissible for this analysis.

We included information in the analysis on four groups of variables. The first two groups of factors related to demographic and socio-economic characteristics of individuals, likely to be associated with illness or mortality. The other variables are indicators of characteristics of place of residence of people in the sample, in later life and in childhood. Place of residence in 1981

[1] Ninety-three were not traced at NHSCR, so they may have died without our knowledge. A small number of other individuals (153 cases) were also eliminated because their 1981 ward of residence had no data on the Carstairs index (explained below). In 16 cases the 3-digit NHS code had no corresponding LG District in our look-up table. Most of these unmatched codes had only 1 case, so they may be due to typographical errors. Others omitted from the analysis may have died abroad, or been abroad in 1991. Others may have failed to respond to the 1991 Census, or they gave a different date of birth from that recorded previously, so could not be identified as an LS member.

is characterised in terms of region and by an indicator of the socio-economic conditions in the electoral ward of residence. Characteristics of sample member's residential area when they were children are based on the place where individuals were registered in 1939. Details of the definitions of the variables are given below.

Group 1: Individual demographic characteristics

- age group (in 4 age bands);
- sex.

Group 2: Individual socio-economic characteristics, as at 1981

- social class based on occupational classifications (whether in professional/managerial, clerical or senior manual, semi/unskilled manual, or unclassified social classes);
- housing tenure (whether owner occupiers or not);
- marital status (whether married and living with spouse or not);
- whether unemployed in 1981: this variable indicated whether the individual was unemployed and seeking work in 1981.[2]

Group 3: Attributes of the area of residence in 1981

- Socio-economic conditions in the area of residence in 1981 were classified using an index developed by Carstairs (See Carstairs and Morris, 1991), which is based on small area statistics from the 1981 census.[3] Area of residence is based on the small area of residence (ward). The categories used here indicate whether the ward was classified in the most affluent quintile of areas, in the two intermediate quintiles, or in the two poorest quintiles, which have relatively high scores on the index.
- Broad regional location was indicated by three groupings of Standard Regions in 1981 (North, North West Yorkshire and Humberside; Wales; East and West Midlands; South West, East Anglia and the South East).

Group 4: Variables which typify the socio-economic characteristics of place of residence in 1939.

- An indicator was included to show whether the area was identified as a 'depressed area' in 1934 (sometimes also described as 'special' areas or 'distressed' areas in the policy documents at the time). We drew on a list of districts officially defined as 'depressed' (GB Parliament, 1934).[4]
- The level of population density in 1931 was classified in tertile categories of density (a measure of urbanisation).[5]
- The percentage of the area population in semi-skilled or unskilled manual work in 1931 was classed in tertile groups.[6]
- The proportion of the area population in overcrowded housing (over 1.5 persons per room) was classified in tertile groups.[7]
- Unemployment rate in the local population in 1931 was also categorised in tertile groups.[8]

[2] Most of those who were not in this category were in work (77%) or not economically active (21%). A small proportion (1.8%) of those who where not classified as unemployed at this time were permanently sick and unable to work.

[3] The Carstairs index is a composite indicator of deprivation based on 1981 census data for 1981 census wards (small electoral units of about 5000 people on average). The components of the index are measures of overcrowded accommodation, male unemployment, low social class and lack of household car.

[4] As this list is for a date part way through the "County Reviews", it had to be partially re-districted to 1939 units. In a very small number of cases, 1939 districts combined parts of 1934 districts which were both in the 'depressed areas' list and outside it; we included them here if more than half the 1931 population had been in a depressed area.

[5] The acreage of each district was listed in the 1939 National Registration report, permitting calculation of population density, while the grid reference of each district's centroid was extracted from our boundary GIS.

[6] Class was not tabulated by the census until 1951, but the Registrar General's Decennial Supplement for 1921–1930 does report social class from the 1931 census for the country as a whole and also for crude regions (see Decennial Supplement, 1931, Part IIa Occupational Mortality, Table 1, "Aggregate mortality in each Occupation Unit distinguished at the census and in Social Classes …', and Table 11 'Areal Distribution of Social Class Mortality Data …'). We were therefore able to precisely compute numbers of males in each social class for large towns by assigning each of the 591 detailed categories to the class given in the Decennial Supplement. For other districts, very precise estimates were computed by dividing up each mutually exclusive group on the assumption that the proportion in each social class was the same as that for the relevant county as a whole. This was greatly assisted by the particular detailed categories of occupational group in the reports selected for tabulation, which included, for example, all the major groups comprising social class 5. This work is reported on in more detail elsewhere.

[7] Information on housing 'amenities', meaning hot water, baths and so on, was not gathered until 1951. However, the 1931 census counted the numbers of rooms, excluding sculleries, bathrooms and so on, and exhaustively cross-tabulated household size against numbers of rooms. We simplified this to numbers of people living at different densities: over 3 persons per room, between 2 and 3 persons, and so on. In the 1931 census the definition of overcrowding was over 1.5 persons per room.

[8] The census reports for the period provide separate information on numbers 'out of work', on numbers of 'operatives' and on numbers 'retired and unoccupied', all broken down by gender. Being 'out of work' rather than 'unoccupied' was self-defined; 1931 claimant count data are available from the Ministry of Labour's Local Unemployment Index, but not for the geographical units used here.

Differences in mortality and illness

As a preliminary to the regression models, Table 2 shows mortality and illness rates by the available socio-demographic and area categories. Table 2 shows that overall 7% of the individuals analysed had died by 1991. The proportion who had died varies in association with all of the 'independent' variables included in this analysis. The chi-squared values for the tabulations of these differences all have probabilities of less than 1%, suggesting that these associations are statistically significant. Older people and males were more likely to have died. Those in work in 1981 were least likely to have died by 1991 and, as would be expected, those who were permanently sick in 1981 had a particularly high rate of mortality over the following 10 years. Those unemployed or economically inactive for other reasons in 1981 also had worse 10-year mortality rates than the employed. The expected gradient of mortality with social class is evident, with the highest rate of mortality among those in class IV or V. Those who were unmarried and those who were not owner occupiers of their homes in 1981 were also more likely to have died by 1991.

The categories of area of residence in 1981 also show associations with survival. Area disadvantage (as measured by categories of the Carstairs index) also showed a graded relationship with survival in that people from poorest areas were most likely to have died. Those who in 1981 were living in regions in the North and West of the country had worse mortality rates, especially when compared with the South East.

This much was expected from earlier work on health inequalities. Table 2 also shows new information on the relationship between mortality and aspects of area of residence in 1939 (classified according to 1931 area typologies). Survival rates were worse for people who in 1939 had been living in areas classed as urban in 1931, areas with larger proportions of workers in semi- and unskilled jobs, areas with higher levels of overcrowding, higher unemployment, or in areas which in 1934 had been identified as depressed.

Those reporting long-term illness in 1991 comprised 21% of the sample. Table 2 also shows that the relationships between the illness outcome and the other characteristics were similar to those found for mortality. Reporting of long-term illness was, not surprisingly, even more strongly associated than mortality with being permanently sick and unable to work in 1981. Again, the results of particular interest in terms of their originality are those which show that the likelihood of reporting illness in 1991 was worse for people who in 1939 had been living in urban areas, areas with larger proportions of workers in semi- and unskilled jobs, areas with higher levels of overcrowding, higher unemployment, or in areas which in 1934 had been identified as depressed.

The statistical analysis

We now describe the regression analyses which were intended to demonstrate whether characteristics of area of residence in 1939 were associated with individual health outcomes after controlling for (a) individual demographic and social factors, and (b) for aspects of area of residence later in life, namely in 1981.

We first used a series of non-hierarchical, logistic regression to examine variation in each of the two health outcomes. We present separate analyses for the genders. Other work (e.g., Townsend et al., 1988; Dept of Health, 1998) has suggested that male mortality and illness are more strongly related to socio-economic factors than the comparable female outcomes and we wanted to compare the models for the two sexes. A series of non-hierarchical models were considered. Thus model 1 considers only the impact of demographic attributes. Model 2 includes in addition the individual's socio-economic characteristics in 1981. Model 3 introduces information on the area of residence in 1981. By comparing the results of model 3 with model 2 it is possible to examine whether these area variables have any power to predict the health outcome once individual factors are controlled. Model 4 includes variables on area of residence in 1939. When compared with model 3, model 4 gives an impression of the extent of association between the health outcome and type of residential area in childhood allowing for the other factors tested. Each of the variables describing area of residence in 1939 was tested individually and the final versions of the models, shown here, only include those area variables which resulted in significant reduction in scaled deviance when they were included.

Controlling for other variables tested, the results show whether the health outcome was significantly different for this category of individuals, compared with the reference category for the variable. The reference category was usually that expected to have the lowest level of morbidity or mortality. Only those variables showing some significant association with the health outcome in the multivariate model were retained in the final models shown here. Fit is assessed by a criterion which corrects the deviance for the number of parameters, and so penalises models which might improve fit only slightly but involve many extra parameters. Specifically, when the number of parameters p is known the Akaike Information Criterion (Akaike, 1973) involves adding twice that number to the deviance D to give the AIC as $(D + 2p)$. A better model has a lower AIC. The results reported here focus especially on the difference in risk of morbidity in 1991, or of death between 1981 and 1991, associated with differences in the 1939 area of residence variables, after controlling for the other factors in the models.

Table 2
Proportions of different groups in sample who had died by 1991 and illness among those still living

Attribute of individual or area		% Alive in 1991	% Dead by 1991	Number of subjects	% Alive no illness	% Alive with illness	Number of subjects
Individual attributes							
Age group 1981	40–44	97	3	15,651	86	14	15,168
	45–49	95	5	20,936	80	20	19,883
	50–54	92	8	16,193	75	25	14,861
	55–59	87	13	9939	71	29	8694
Sex	Male	92	8	25,226	78	22	23,311
	Female	94	6	37,493	80	20	35,295
Individual EC status 1981	Not unemployed	94	6	60,461	74	19	56,581
	Unemployed (seeking work)	90	10	2258	65	35	2025
Individual social class 1981	I or II	95	5	13,373	83	17	38,431
	IIINM or IIIM	94	6	22,128	71	29	20,175
	IV or V	93	7	12,961	81	19	48,289
	Unclassed	92	8	14,257	71	29	10,317
Individual marital status 1981	Married, lived with spouse	94	6	51,326	85	15	12,666
	Not married, divd. sepd.	91	9	11,393	81	19	20,851
Individual tenure 1981	Owner-occupier	95	5	40,528	76	24	12,023
	Other	91	9	22,191	73	27	13,066
1981 area attributes							
Area carstairs group 1981	Most affluent 20%	95	5	11,473	86	14	10,924
	Middle 2 quintiles	94	6	22,612	82	18	21,332
	Poorest 40%	92	8	28,634	74	26	26,350
Region 1981	N, NW, Y&Hum	92	8	19,932	74	26	18,406
	Wales	93	7	4036	70	30	3740
	Wmids & Emids	93	7	11,762	80	20	10,960
	SW,EA,SE	94	6	26,989	83	17	25,500
1939 area attributes							
Special area 1934	Depressed area in 1934	92	8	4677	69	31	4284
	Not depressed	94	6	58,042	80	20	54,322
Area POP density 1930s	Low	94	6	13,634	82	18	12,820
	Medium	93	7	11,918	78	22	11,126
	High	93	7	37,167	78	22	34,660
Area PCT classes IV&V 1930s	Low % IV&V	94	6	21,202	82	18	19,950
	Middle tertile	93	7	21,387	77	23	19,942
	High % IV&V	93	7	20,130	77	23	18,714
Area housing density 1930s	Low % crowded	95	5	14,327	83	17	13,541
	Middle tertile	94	6	17,043	81	19	15,987
	High % crowded	93	7	31,349	76	24	29,078
Area unemployment 1930s	Low	94	6	15,677	83	17	14,800
	Medium	94	6	16,984	81	19	15,992
	High	93	7	30,058	76	24	27,814
Total		93	7	62,719	79	21	58,606

Groupings for individual characteristics are based on data for 1981; types of area of residence relate to 1981 or 1939.

We next extended the analysis using a modelling strategy which is more suited to analysis of independent variables which vary 'hierarchically' (at individual or household level and at area level). Hierarchical modelling with these data is complex because we are dealing with geographical information which relates mainly to *type* of area, rather than area *location*. Also the areas in 1939 and in 1981 have different and overlapping boundaries, so that this dataset does not represent a classic 'nested' hierarchical multi-level problem.

The strategy here is therefore to represent levels in the data in terms of *categories* of area in 1981 and 1939, which are cross-classifications of the area categories available at the two time points. We set up a variable for 1981 area of residence which has 12 categories, obtained by cross-classifying the Carstairs category (3 categories) with region in England and Wales (4 categories). The 1939 area variables have been simplified to reduce the number of categories for analysis. For variables relating to population density, percent classes IV and V, housing density, the areas with the highest values are contrasted with the rest, reducing a trichotomy to a dichotomy. Thus 1939 area type is defined by cross-classifying four binary area characteristics: high vs. lower population density, high vs. lower percent classes IV and V, high vs. lower housing density, and whether or not the district had 'special area' (1934) designation. This produces 16 categories and when crossed with 1981 area type, the 1939–1981 composite area variable has 192 categories.

To analyse the effects of 1939 area type, in conjunction with those of individual variables and 1981 area type, we used a Bayesian modelling approach with the program WINBUGS. This approach provides a way to assess model fit, in the presence of random effects, when the true number of model parameters is not known, and measures of fit are required which penalise the deviance to account for 'extra model' complexity (Spiegelhalter et al., 2002). The Bayesian analogue provides an estimate of the number of parameters (p_e) and a criterion known as the 'deviance information criterion' (DIC), lower values of which indicate better fit.

Controlling for individual characteristics, the model examines whether the association between health outcomes and area conditions in 1981, varied according to attributes of place of residence in 1939. To illustrate effects of 1939 area type, we consider the random effects (i.e. varying intercepts) for cross-classified types of area containing over 50 people in the LS sample. To compare health outcomes between groups we take exponentials of the 1939–1981 area random effects to generate 'smoothed SMR' measures, with pooling of strength (i.e. smoothing) over the 1939–1981 area types. These 'smoothed SMRs' are standardised morbidity (or mortality) ratios which control for all the 'individual level' variables in the model. 'Smoothed SMR' values of 100 for a sub-group in the sample would represent rates of illness (or death) which are typical of the whole sample population, while values above 100 represent health which was relatively poor and values below 100 indicate relatively good health for the group. 'Smoothed SMR' values were compared for different categories in the 1939–1981 classification, to demonstrate the impact of the combined effect of 1981 and 1939 area. The greater frequency of illness, compared with death, as an outcome means that the area variables are more significantly associated with morbidity than with the mortality in these data. We have therefore focused mainly on morbidity here.

Controlling for individual characteristics and for 1981 area type, we have also examined the contrasts between 2 larger groups of people in the sample, constructed by grouping together some of the 192 cross-classified area types. The first group ('affluent areas') comprises those from areas where conditions in the 1930s were relatively affluent (having lower population density, and lower overcrowding and lower proportions in class IV and V and being outside the special areas). These are compared with the rest of the sample ('other areas'). Population weighted averages of 'smoothed SMRs' were calculated for these aggregates. From these we obtained a 'relative risk' ratio of illness (or of death) for people originating from relatively affluent 1930s areas, compared with the rest of the sample.

Results

Table 3a and b shows the results of the multivariate analysis for mortality with males and females modelled separately. Model 1 incorporates only demographic data. Examination of the change in the AIC shows that power to predict survival over the period 1981–1991 improves as individual socio-economic status and marital status are included (model 2), and also as information about place of residence in 1981 is added (model 3). Thus, after controlling for individual level variation in social factors, some additional association with area conditions is still significant. With some qualifications (discussed in the conclusions below) this supports the argument that area factors, as well as individual characteristics in later life, had a bearing on mortality outcome for the sample.

Model 4 in Table 3a and b shows that, for men and for women, some further improvement in the predictive power of the model results from incorporating information on area unemployment in the place of residence in the 1930s. This applies after controlling for individual characteristics, region of residence in 1981 and local area socio-economic conditions in 1981. Men and women who had lived in areas with highest unemployment showed significantly higher probability of death than those who were in areas of lowest unemployment. This is

Table 3
Regression analysis of mortality for Men[a] and Women[b]

Predictor variables	Model 1: 1981 demographic data only					Model 2: demographic and individual attributes					Model 3: individual data and area data for 1981					Model 4: individual data and area data for 1981, and 1939				
	B	SE	Odds ratioX 100	Odds ratio confidence interval		*B*	SE	Odds ratioX 100	Odds ratio confidence interval		*B*	SE	Odds ratioX 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval	
				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)
(a) Men																				
Constant individual attributes	−3.309	0.061				−3.704	0.075				−3.911	0.099				−3.911	0.099			
Age (Reference 40–44)																				
45–49	0.571	0.074	177	153	205	0.563	0.075	176	152	203	0.562	0.075	175	151	203	0.564	0.075	176	152	204
50–54	1.248	0.075	348	301	403	1.216	0.075	338	291	391	1.219	0.075	338	292	392	1.224	0.075	340	293	394
55–59	1.895	0.081	665	567	780	1.804	0.082	607	517	713	1.794	0.082	601	512	706	1.795	0.082	602	512	707
Unemployed						0.222	0.089	125	105	149	0.174	0.090	119	100	142	0.163	0.090	118	99	140
Not owner occupier						0.450	0.053	157	141	174	0.413	0.054	151	136	168	0.416	0.054	152	136	168
Not married						0.426	0.058	153	137	172	0.422	0.058	153	136	171	0.426	0.058	153	137	172
Social class[c] (Classes I and II reference)																				
III-NM or III-M						0.114	0.062	112	99	127	0.057	0.063	106	94	120	0.056	0.063	106	93	120
IV or V						0.234	0.073	126	110	146	0.173	0.074	119	103	137	0.171	0.074	119	103	137
Unclassed						0.602	0.114	183	146	229	0.571	0.115	177	141	222	0.571	0.115	177	141	222
Area of residence 1981																				
Carstairs (reference group: affluent quintile)																				
Middle 2 Quintiles											0.046	0.077	105	90	122	0.046	0.077	105	90	122
Poorest 40%											0.177	0.078	119	102	139	0.153	0.079	117	100	136
Region (reference: SE and SW)																				
N, NW, Yorks & Humb											0.296	0.060	135	120	151	0.233	0.064	126	111	143
Wales											0.223	0.099	125	103	152	0.160	0.102	117	96	143
W. Midlands & E. Midlands											0.160	0.071	117	102	135	0.127	0.072	114	99	131
***Area of residence* 1939**																				
Unemployment 1931 (reference: low)																				
Medium																−0.046	0.073	96	83	110
High																0.129	0.069	114	99	130
AIC	12878					12622					12588					12583				

Predictor variables	Model 1: 1981 demographic data only					Model 2: demographic and individual attributes					Model 3: individual data and area data for 1981					Model 4: individual data and area data for 1981, and 1939				
	B	SE	Odds ratioX 100	Odds ratio confidence interval		*B*	SE	Odds ratioX 100	Odds ratio confidence interval		*B*	SE	Odds ratioX 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval	
				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)
(b) Women																				
Constant	−3.607	0.071				−3.961	0.096				−4.191	0.111				−3.911	0.099			
individual attributes																				
Age (reference 40–44)																				
45–49	0.458	0.085	158	134	187	0.450	0.085	157	133	185	0.449	0.085	157	133	185	0.453	0.085	157	133	186
50–54	0.973	0.080	264	226	310	0.920	0.081	251	214	294	0.919	0.081	251	214	294	0.924	0.081	252	215	295
55–59	1.462	0.080	432	369	505	1.341	0.080	382	327	448	1.336	0.080	380	325	445	1.334	0.080	380	324	445
unemployed						0.140	0.141	115	87	152	0.104	0.142	111	84	146	0.101	0.142	111	84	146
Not Owner Occupier						0.402	0.046	149	137	164	0.327	0.048	139	126	152	0.329	0.048	139	127	153
Not Married						0.327	0.053	139	125	154	0.319	0.053	138	124	153	0.322	0.053	138	124	153
Social class[c] (Classes I and II reference)																				
III-NM or III-M																				
IV or V						−0.102	0.083	90	77	106	−0.118	0.084	89	75	105	−0.119	0.084	89	75	105
Unclassed						0.053	0.086	105	89	125	−0.003	0.087	100	84	118	−0.005	0.087	100	84	118
Area of residence 1981						0.429	0.077	154	132	179	0.403	0.077	150	129	174	0.403	0.077	150	129	174
Carstairs (reference group: affluent quintile)																				
Middle 2 quintiles																				
Poorest 40%											0.094	0.074	110	95	127	0.091	0.074	110	95	127
Region (reference: SE and SW)											0.330	0.073	139	121	161	0.307	0.074	136	118	157
N, NW, Yorks & Humb																				
Wales											0.168	0.054	118	106	132	0.105	0.059	111	99	125
W. Midlands & E. Midlands											0.091	0.095	109	91	132	0.024	0.098	102	84	124
Area of residence 1939											0.155	0.062	117	103	132	0.113	0.064	112	99	127
Unemployment 1931 (reference: low)																				
Medium																0.006	0.066	101	88	115
High																0.143	0.063	115	102	131
AIC	16,231					16,004					15,957					15,954				

[a] Outcome: whether individual died during 1981–1991, $n = 25,226$.

[b] Outcome: whether individual died during 1981–1991, $n = 37,493$.

[c] Classes I and II professionals, managers, entrepreneurs, Classes III-NM (routine non-manual) and III-M (skilled manual), Classes IV and V (semi- and unskilled manual).

Table 4
Regression analysis of illness for Men[a] and women[b]

Predictor variables	Model 1: 1981 demographic data only					Model 2: demographic and individual attributes					Model 3: individual data and area data for 1981					Model 4: individual data and area data for 1981, and 1939				
	B	SE	Odds ratio X 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval	
				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)
(a) Men																				
Constant	−1.785	0.033				−2.484	0.046				−2.907	0.060				−3.122	0.075			
Individual attributes																				
Age (reference 40-44)																				
45–49	0.511	0.041	167	154	181	0.532	0.042	170	157	185	0.539	0.043	171	158	186	0.545	0.043	172	158	188
50–54	0.947	0.045	258	236	282	0.969	0.047	263	240	289	0.996	0.047	271	247	297	1.006	0.047	273	249	300
55-59	1.244	0.059	347	309	389	1.194	0.061	330	293	372	1.184	0.061	327	290	369	1.188	0.062	328	291	370
Unemployed						0.396	0.064	149	131	168	0.301	0.065	135	119	153	0.295	0.065	134	118	153
Not owner occupier						0.456	0.036	158	147	169	0.390	0.036	148	137	159	0.388	0.037	147	137	158
Not married						0.298	0.042	135	124	146	0.295	0.043	134	124	146	0.302	0.043	135	124	147
Social class[c] (Classes I and II reference)																				
III-NM or III-M						0.494	0.042	164	151	178	0.390	0.043	148	136	161	0.390	0.043	148	136	161
IV or V						0.694	0.050	200	182	221	0.567	0.051	176	160	195	0.570	0.051	177	160	195
Unclassed						1.650	0.087	521	439	618	1.613	0.089	502	422	597	1.612	0.089	501	421	597
Area of residence 1981																				
Carstairs (reference group: affluent quintile)																				
Middle 2 Quintiles											0.148	0.053	116	105	129	0.144	0.053	115	104	128
Poorest 40%											0.415	0.053	151	136	168	0.365	0.054	144	130	160
Region (reference: SE and SW)																				
N, NW, Yorks & Humb											0.505	0.040	166	153	179	0.415	0.043	151	139	165
Wales											0.754	0.064	213	188	241	0.625	0.069	187	163	214
W. Midlands & E. Midlands											0.185	0.048	120	110	132	0.162	0.049	118	107	129
Area of residence 1939																				
population density 1931 (reference: low)																				
Medium																0.187	0.056	121	108	135
High																0.199	0.048	122	111	134
% Classes IV and V (reference: low)																				
Middle tertile																0.156	0.044	117	107	127
Upper tertile (high %IV&V)																0.147	0.048	116	105	127
Depressed area in 1934																0.260	0.060	130	115	146
AIC	24,080					22,956					22,556					22,514				

Predictor variables	Model 1: 1981 demographic data only					Model 2: demographic and individual attributes					Model 3: individual data and area data for 1981					Model 4: individual data and area data for 1981, and 1939				
	B	SE	Odds ratio X 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval		*B*	SE	Odds ratio X 100	Odds ratio confidence interval	
				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)				Lower limit (2.5%)	Upper limit (97.5%)
(b) Women																				
Constant	−1.781	0.033				−2.260	0.051				−2.666	0.061				−4.191	0.111			
individual attributes																				
Age (Reference 40-44)																				
45–49	0.245	0.041	128	118	139	0.240	0.042	127	117	138	0.245	0.042	128	118	139	0.252	0.042	129	118	140
50–54	0.536	0.040	171	158	185	0.484	0.041	162	150	176	0.489	0.041	163	150	177	0.497	0.041	164	152	178
55-59	0.775	0.043	217	200	236	0.651	0.043	192	176	209	0.654	0.044	192	177	210	0.655	0.044	192	177	210
unemployed						0.300	0.086	135	114	160	0.257	0.087	129	109	153	0.253	0.087	129	109	153
Not Owner Occupier						0.503	0.028	165	157	175	0.414	0.029	151	143	160	0.418	0.029	152	143	161
Not Married						0.434	0.033	154	145	165	0.427	0.033	153	144	164	0.431	0.034	154	144	164
Social Class[c] *(Classes I and II reference)*																				
III-NM or III-M						−0.054	0.049	95	86	104	−0.061	0.049	94	85	104	−0.062	0.049	94	85	103
IV or V						0.174	0.051	119	108	131	0.105	0.051	111	101	123	0.108	0.051	111	101	123
Unclassed						0.522	0.046	169	154	184	0.499	0.046	165	150	180	0.501	0.046	165	151	181
Area of residence 1981																				
Carstairs (reference group: affluent quintile)																				
Middle 2 quintiles											0.184	0.045	120	110	131	0.182	0.045	120	110	131
Poorest 40%											0.430	0.044	154	141	168	0.392	0.045	148	136	162
Region (reference: SE and SW)																				
N, NW, Yorks & Humb											0.365	0.033	144	135	154	0.282	0.036	133	123	142
Wales											0.506	0.055	166	149	185	0.397	0.059	149	133	167
W. Midlands & E. Midlands											0.175	0.039	119	110	129	0.134	0.040	114	106	124
Area of residence 1939																				
Unemployment 1931 (reference: low)																				
Medium																0.002	0.041	100	93	109
High																0.140	0.043	115	106	125
Population density 1931(reference: low)																				
Medium																0.111	0.045	112	102	122
High																0.090	0.040	109	101	118
Depressed area in 1934																0.125	0.053	113	102	126
AIC	35,009					34,086					33,713					33,676				

[a] Outcome: whether individual died during 1981–1991, $n = 23,311$.

[b] Outcome: whether individual died during 1981–1991, $n = 35,295$.

[c] Classes I and II professionals, managers, entrepreneurs, Classes III-NM (routine non-manual) and III-M (skilled manual), Classes IV and V (semi and unskilled manual).

after controlling for the other variables in the model. When this group was compared with people who had lived in areas with low unemployment in the 1930s, the relative risk of dying between 1981 and 1991 was 14–15% higher. The other variables for 1939 place of residence did not contribute significantly in the model and are excluded here.

Table 4a and b shows results of logistic regression of the morbidity outcome for males and females. The change in deviance indicated significant improvements in the predictive power of the model as individual socio-economic characteristics were introduced and also when the 1981 Carstairs classification of area of residence and region of residence are included. Again, we can see independent associations with morbidity for variables relating to individual characteristics and area of residence variables.

For both men and women, Model 4 provides a further improvement in the predictive power when variables on area of residence in 1939 (as represented by the Census, 1931) are included. The probability of reporting illness in 1991 was greater for men who in 1939 had lived in: areas which in 1931 were of medium or high population density (more urban areas) as compared with rural areas; in areas with medium, or high proportions of workers in class IV or V, as compared with areas with fewer semi/unskilled workers; and in areas which were classified as depressed areas in 1934 compared with those that were not. Unemployment in the 1930s did not show a significant association when these other variables were included (though unemployment will have been particularly high in depressed areas).

Results for women, shown in Table 4b, are similar for the individual variables and for variables relating to place of residence in 1981. The differences in relative risks are less extreme, however. Women were more likely to have reported illness in 1991 if their area of residence in the 1930s had high unemployment rates, was a depressed area or had medium or high population density.

The hierarchical model predicting chronic illness for males showed a very similar pattern of individual level impacts on illness as for the simple logistic regression analysis in Table 4a above. With the intercept varying randomly over the 1939–81 area composite with 192 levels, the DIC is reduced to 22536.5 (i.e. the model 'fit' has improved) as compared to models in Table 4a. The 'smoothed' standard illness ratios vary across the 192 area categories from just over 50 to nearly 250. Thus using a hierarchical model, and including cross-classified data on 1939 and 1981 conditions, we can more effectively differentiate groups in the sample with varying levels of morbidity.

Fig. 2 illustrates 'smoothed' SMRs for a selection of the 192 'cross-classified' area categories showing particularly interesting patterns. Table 5 shows, for these categories, the residential area conditions in 1981 and in 1939. There are five categories of men who were all living in the poorest wards of northern regions in 1981 (labelled PN1–PN5). Three groups have relatively high SMRs (labelled PN1, PN2, PN3 in Fig. 2). For two other categories of men living in the poorest northern regions in 1981, the SMRs are lower (labelled PN4 and PN5 in Fig. 2.). The differences in SMR are associated with area conditions in 1939. Categories PN1, PN2, PN3

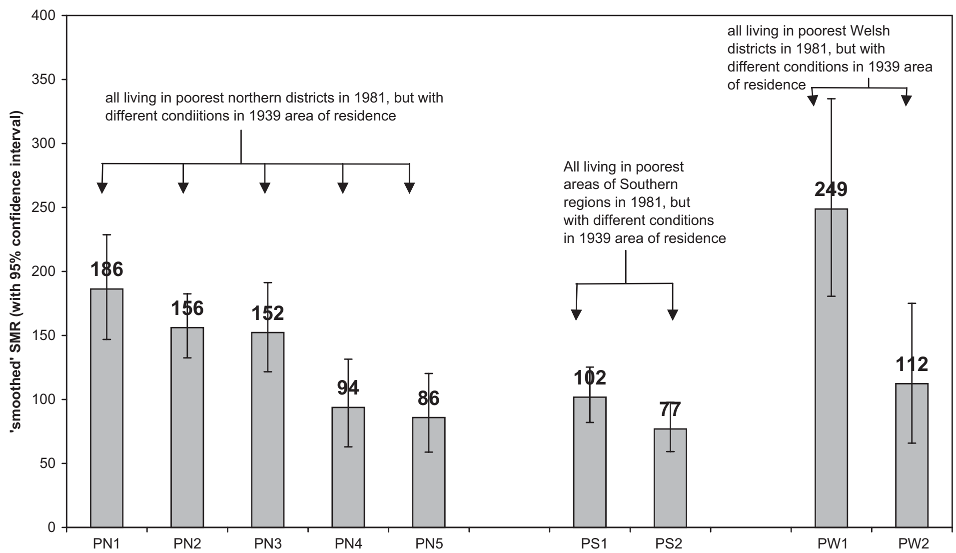

Fig. 2. Smoothed standardised Morbidity Ratios, controlling for individual factors, showing association with cross classification of 1981 and 1939 area conditions, for males from selected area types.

all comprise men who lived in areas in 1939 with high concentrations of populations in lower social classes and with high housing densities. Two of these areas were also more urban areas with high population density and one was in a special area. Categories PN4 and PN5 comprise men living in relatively rural areas (low population density) and neither group was living in special areas in the 1930s.

Two further categories of men were living in the poorest districts of Wales in 1981 (PW1 and PW2). The numbers in these groups is relatively small, so the confidence interval for the 'smoothed SMR' is wide. However, category PW1 has higher 'smoothed SMRs' than category PW2. Group PW1 comprises men who lived in special areas in 1939, while group PW2 did not. Categories PS1 and PS1 are both groups of men living in the poorest Southern districts in 1981. In comparison with the categories in the North and in Wales levels of illness are lower, but this is clearer for group PS2 than for PS1. In 1939, men in PS1 were living in more urban areas, with greater levels of overcrowding, than group PS2.

Table 6 illustrates the impact of relatively affluent area conditions in the 1930s residence. For 4151 men originating in 'affluent areas', the average 'smoothed SMR' for illness is shown (82) and this is less than for 19160 men originating from 'other' areas (104). This comparison is regardless of 1981 area type and involves comparing the aggregate of 12 standard illness ratios (over the 1981 area of residence types) with the remaining 180 illness ratios. The 'relative risk' ratio is 0.79 (with 95% interval 0.73–0.86) showing a long-term 'protective' effect of early life residential environment (i.e. 1930s area type) on later illness chances. Similarly, women originating from 'affluent areas' had an average 'smoothed SMR' for illness of 83. Compared with other women, their relative illness risk is 0.79 (with 95% interval 0.73–0.84), again showing a long-term 'protective' effect of early residential environment on later illness chances. Table 6 also shows, for men and women combined, that the average 'smoothed SMR' for mortality is less for people originating in the 'affluent areas' in 1939 and their relative risk of death compared with the others is 0.93. This disparity in risk for

Table 5
Male illness; illustrations of results from multi-level models to analyse 1931–1981 area composite

1981 Category	No of persons in category	Popn density, 1930s	% Classes IV and V, 1930s	Housing density, 1930s	Special Area, 1930s	'Smoothed SMR'	Lower confidence interval 2.5%	Higher confidence interval 97.5%	Label in Fig. 2
1 Poorest (North)	424	High	High	High	Yes	186	147	229	PN1
2 Poorest (North)	1056	High	High	High	No	156	133	183	PN2
3 Poorest (North)	372	Low	High	High	No	152	122	191	PN3
4 Poorest (North)	149	Low	Low	High	No	94	63	132	PN4
5 Poorest (North)	128	Low	High	Low	No	86	59	120	PN5
6 Poorest (South)	630	High	Low	High	No	102	82	125	PS1
7 Poorest (South)	366	Low	High	Low	No	77	59	98	PS2
8 Poorest (Wales)	150	Low	Low	High	Yes	249	181	335	PW1
9 Poorest (Wales)	54	Low	Low	High	No	112	66	175	PW2

Categories illustrated in Fig. 2.

Table 6
Comparison of illness and mortality for individuals originating from 'affluent areas' in 1939, compared with 'other' areas

Area conditions in 1939	Illness for males			Illness for females			Mortality for males and females		
	Number in group	Average of smoothed SMR[a]	Relative risk (affluent/ others)	Number in group	Average of smoothed SMR[a]	Relative risk	Number in group	Average of smoothed SMR[a]	Relative risk
Relatively affluent[a]	4151	82	0.79	6028	83	0.79	10825	95.5	0.93
other areas[a]	19160	104		29267	106		51867	102.5	

[a] See the text for explanation of these terms.

mortality is less extreme than for illness. This may result from real differences in the strength of the association, or because death is a less frequent outcome than illness, so the statistical relationships are based on smaller number of cases.

Conclusions

Conclusions from these results need to be qualified, partly because of the limitations imposed on the analysis by the data available. One difficulty is the omission of those who were not covered by civil registration in 1939, or who subsequently lost their registration numbers. As we have seen, this means that the analysis only relates to 62% of the whole sample of people in the relevant age group and the data are less complete for men than for women. Also, because of evacuation in the early phases of World War II, place of residence in 1939 may not have been the area where much of childhood had been spent. Furthermore, we have no data on individual characteristics earlier in life, although (as indicated in the review above) other cohort studies suggest such early circumstances might have been important for health later in life. It is possible that variables describing area of residence in the 1930s may be acting as surrogate indicators of these *individual* level differences in early life, rather than reflecting important impacts of *places* in the 1930s.

Confidentiality restrictions have also meant that we were not permitted to analyse more detailed classifications on the variables of interest. Thus the categories used to describe differences between groups in the sample are rather broad ones, and may not be sufficiently complex to distinguish key aspects of differences associated with the health outcomes. This could mean that area variables which appear significant in the models are substituting for insufficiently specified associations with individual characteristics.

However, allowing for these caveats, the results suggest that both individual and area characteristics measured here contributed independently to variation in health outcomes in late middle age (especially the chance of limiting illness). In particular, we have reported on analysis of new data which showed that people who had lived as children in disadvantaged areas, with high levels of unemployment had higher relative risk of illness and death in later life. This finding holds after allowing for more recent circumstances.

Thus conditions in the area of residence during childhood appear to have had a measurable association with health outcomes later in life. For example, results in Table 3a and b show that those that had lived in areas classified in 1934 as 'depressed areas' had a relative risk of mortality, or of illness reporting, which was 14–15% higher than those who had not been registered as living in such areas in the 1939. These 'depressed areas' were mainly in the north of England (and in Wales) and they had particularly high levels of unemployment during the 1930s. They include mining areas in regions such as the north east of England, which also had high levels of unemployment (especially for men) in the 1990s and have been shown to have particularly high prevalence of long-term illness reported in the 1991 census. Some authors (e.g., Haynes, Bentham, Lovett, & Eimermann, 1997) have suggested that this may have been affected by local labour market conditions in the late 1980s and early 1990s and that in areas of high unemployment, people were more likely to declare themselves to have long-term illnesses preventing them from working. The results shown here certainly support the view that men who were unemployed in 1981 had relatively poor health outcomes by 1991 (high relative risks of death or illness for men seeking work in Table 3a and b). The effects of unemployment on health outcomes appear to have been less striking for women (indicated by less extreme relative risks for those seeking work in Tables 4b and 5). It therefore might be argued that the data on 'depressed area' status is acting as a marker for areas of special health disadvantage in the 1980s, rather than the 1930s. However, data on individual employment status in 1981, local unemployment levels in 1981 and broad region of residence have been included in the models described here, which still show an independent effect of area deprivation in the 1930s. Furthermore the relative risks reported here for poor health outcomes linked to 'depressed area' conditions in the 1930s are similar for men and women. This suggests a general, early influence of community level deprivation, distinct from later effects of individual unemployment and labour market difficulties in the 1980s. Therefore another possible explanation for poor health today, in areas like the northeast of England and Wales, may be that this is a legacy of deprived environments experienced in childhood.

If the associations reported here are indicative of causal links between contextual conditions during childhood and later health outcomes, and if similar processes apply today, then current levels of socio-equality between geographical areas may have long-term implications for health inequalities in England and Wales. Fifty years from now, children who are now living in areas of high unemployment would then be expected to show poorer health as ageing adults than those who now live in more affluent areas, and social and geographical inequalities of health will have been perpetuated for another generation.

Acknowledgements

This research was supported by a grant from the Economic and Social Research Council, Health

Variations Programme L128251051. We gratefully acknowledge advice and assistance from Professor David Strachan, St. Georges Hospital Medical School, in the work undertaken to link data from the 1930s to the LS. The views expressed here are entirely those of the authors.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csaki (Eds.), *Proceedings of the second international symposium in information theory*. Budapest: Akademiai Kiado.

Barker, D. (1992). *Fetal and infant origins of adult disease*. London: British Medical Journal Publishing Group.

Ben Shlomo, Y., & Smith, G. (1991). Deprivation in infancy or in adult life—which is more important for mortality risk. *Lancet*, *337*, 530–534.

Benzeval, M., Dilnot, A., Judge, K., & Taylor, J. (2001). Income and health over the life course: Evidence and policy implications. In: Graham, H. (Ed.), *Understanding health inequalities* (pp. 96–112). Milton Keynes: Open University.

Berney, L., Blane, D., Davey-Smith, G., & Holland, P. (2001). Life course influences on health in early old age. In: Graham, H. (Ed.), *Understanding health inequalities* (pp. 79–95). Milton Keynes: Open University.

Britton, M. (1990). Sources of Data and Limitations. In M. Britton (Ed.), *Mortality and geography: A review in the mid-1980s England and Wales OPCS series DS no. 9* (pp. 6–17). London: HMSO.

Campos, R.M., et al. (2003). Locality level mortality and socio-economic change in Britain since 1920: First steps towards analysis of infant mortality variation. In: P. Boyle, et al. (Eds.), *The geography of health inequalities in the developed world: Views from Britain and North America*. Aldershot: Ashgate (in press).

Carstairs, V., & Morris, R. (1991). *Deprivation and health in Scotland*. Aberdeen: Aberdeen University Press.

Census, 1931: *Classification of occupations*. London: HMSO, 1934.

Congdon, P., Shouls, S., & Curtis, S. (1997). A multi-level perspective on small area health and mortality: A case study of England and Wales. *International Journal of Population Geography*, *3*, 243–263.

Congdon, P., Campos, R. M., Curtis, S. E., Southall, H. R., Gregory, I. N., & Jones, I. R. (2001). Quantifying and explaining changes in geographical inequality of infant mortality in England and Wales since the 1890s. *International Journal of Population Geography*, *7*, 35–51.

Curtis, S., & Jones, I. (1998). Is there a place for geography in the analysis of health inequality. *Sociology of Health and Illness*, *20*(5), 645–672.

Davey-Smith, G., Hart, G., Blane, D., Gillis, C., & Hawthorne, V. (1997). Lifetime socioeconomic position and mortality: Prospective observational study. *British Medical Journal*, *314*, 547–552.

Dept of Health (1998). *Independent inquiry into inequalities in health: Report*. London: The Stationary Office.

Duncan, C., & Jones, K. (1995). Individuals and their ecologies: Analysing the geography of chronic illness within a multi-level modeling framework. *Journal of Health and Place*, *1*, 27–40.

Duncan, C., Jones, K., & Moon, G. (1993). Do places matter? A multi-level analysis of regional variation health related behaviour in Britain. *Social Science & Medicine*, *37*, 725–733.

Elford, J., Phillips, A., Thomson, A., & Shaper, A. (1989). Migration and geographic variations in ishemic heart disease in Great Britain. *Lancet*, *1*, 343–346.

Fox, A. J., & Goldblatt, P. (1982). *Longitudinal study: Socio-economic mortality differentials, 1971–75, Series LS no1*. London: HMSO.

Gatrell, A., Thomas, C., Bennett, S., Bostock, L., Popay, J., Williams, G., & Shahtahmasebi, S. (2001). Understanding health inequalities: Locating people in geographical and social spaces. In: Graham, H. (Ed.), *Understanding health inequalities* (pp. 156–169). Milton Keynes: Open University.

GB Parliament (1934). *Official reports 5th series parliamentary debates*: Commons 1934–5 Vol 295 5th December 1934.

Gordon, D., Shaw, M., Dorling, D., & Davey-Smith, G. (1999). *Inequalities in health: The evidence presented to the independent inequiry into inequalities in health, Chaired by Sir Donald Acheson*. Bristol: Policy Press.

Graham, H. (2001). *Understanding health inequalities*. Milton Keynes: Open University.

Gregory, I. N. (2001). The accuracy of areal interpolation techniques: Standardising 19th and 20th century census data to allow long-term comparisons *Computers Environment and Urban Systems* (in press).

Gregory, I. N., Southall, H. R., & Dorling, D. (2000). A century of poverty in England & Wales, 1898–1998: A geographical analysis. In J. Bradshaw, & R. Sainsbury (Eds.), *Getting the measure of poverty. The early legacy of Seebohm Rowntree* (pp. 130–159). Bath: Policy Press.

Haynes, R., Bentham, G., Lovett, A., & Eimermann, J. (1997). Effect of labour market conditions on reporting of limiting long term illness and permanent sickness in England and Wales. *Journal of Epidemiology and Community Health*, *51*(3), 282–288.

Kuh, D., & Ben-Shlomo, Y. (Eds.) (1997). *A lifecourse approach to chronic disease epidemiology*. Oxford: Oxford University Press.

Leon, D., & Strachan, D. (1993). Socioeconomic characteristics of interregional migrants in England and Wales 1939–1971. *Environment and Planning A*, *25*(10), 1441–1451.

Macintyre, S. (1999). Geographical inequaities in mortality, morbidity and health related behaviour in England. In D. Gordon, M. Shaw, D. Dorling, & G. Davey-Smith (Eds.), *Inequalities in health: The evidence presented to the independent inquiry into inequalities in health, Chaired by Sir Donald Acheson* (pp. 148–154). Bristol: Policy Press.

Macintyre, S., Maciver, S., & Soomans, A. (1993). Area, class and health: Should we befocusing in places or people. *Journal of Social Policy*, *22*(2), 213–234.

Marmot, M., & Wilkinson, R. (1999). *Social determinants of health*. Oxford: Oxford University Press.

National Registration Report (1944). *National Register: United Kingdom and Isle of Man: Statistics of Population on 29th*

September, 1939 By Sex, Age and Marital Condition: Report and Tables. London: HMSO.

Osmond, C., Slattery, J., & Barker, D. (1990). Mortality by place of birth. In M. Britton (Ed.), *Mortality and geography: A review in the mid-1980s. England and Wales. Office of Population Censuses and Surveys*. London: HMSO.

Popay, J., Williams, G., Thomas, C., & Gatrell, A. (1998). Theorising inequalities in health: The place of lay knowledge. In M. Bartley, D. Blane, & G. Davey-Smith (Eds.), *The sociology of health inequalities* (pp. 59–84). Oxford: Blackwells.

Power, C., & Matthews, S. (1997). Origins of health inequalities in a national population sample. *Lancet*, *350*, 1584–1589.

Registrar General (1938). *The Registrar General's decennial supplement: England and Wales, 1931, Part IIa. Occupational mortality*. London: HMSO.

Schwartz, S. (1994). The fallacy of the ecological fallacy: The potential misuse of a concept and the consequences. *American Journal of Public Health*, *84*, 819–824.

Shouls, S., Congdon, P., & Curtis, S. (1996). Modelling inequality in reported long term illness in the UK: Combining individual and area characteristics. *Journal of Epidemiology and Community Health*, *50*(3), 366–376.

Simpson, L. (forthcoming). Geography conversion tables. *International Journal of Population Geography*, submitted for publication.

Sloggett, A., & Joshi, H. (1998). Deprivation indicators as predictors of life events 1981–1992 based on the UK ONS longitudinal study. *Journal of Epidemiology and Community Health*, *52*(4), 228–233.

Spiegelhalter, D., Best, N., Carlin, B., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B*, *64*, 583–639.

Strachan, D. P., Leon, D. A., & Dodgeon, B. (1995). Mortality from cardiovascular disease among interregional migrants in England and Wales. *British Medical Journal*, *310*, 423–427.

Townsend, P., Davidson, N., & Whitehead, M. (1988). *Inequalities in health*. London: Penguin Books.

Wilkinson, R. (1996). *Unhealthy societies: The afflictions of inequality*. London: Routledge.