

# Modelling inequality in reported long term illness in the UK: combining individual and area characteristics

Susanna Shouls, Peter Congdon, Sarah Curtis

## Abstract

**Study objective** – To assess the nature of the relation between health and social factors at both the aggregated scale of geographical areas and the individual scale.

**Design and setting** – The individual data are derived from the sample of anonymised records (SAR) from the census of 1991 in Great Britain, and are combined with area data from this census. The ecological setting (context) was defined using multivariate methods to classify the 278 districts of residence identifiable in the SAR. The outcome health variable is the 1991 census long-term limiting illness question. Health variations were analysed by multilevel logistic regression to examine the compositional variation (at the level of the individual) and the contextual variation (variability operating at the level of districts) in reported illness.

**Participants** – 10 per cent randomised subsample of the SAR who are aged 16+ and are resident in households.

**Main results** – The multi-level modelling revealed that area factors have a significant association with individual health outcome but their effect is smaller than that of individual attributes. The results show evidence for both compositional and contextual effects in the pattern of variation in propensity to report illness.

**Conclusions** – The results suggest generally higher levels of ill health for individuals who are older, not married, in a semi/unskilled manual social class, and socioeconomically deprived (as measured by a composite deprivation score). All individuals living in areas with high levels of illness (which tend to be more deprived areas) show greater morbidity, even after allowing for their individual characteristics. However, within affluent areas, where morbidity was generally lower, the health inequality (health gradient) between rich and poor individuals was particularly strong. We consider the implications of these findings for health and resource allocation policy.

(*J Epidemiol Community Health* 1996;50:366-376)

This paper examines some aspects of health inequalities and their associations with socio-economic conditions in the British population. While most measures of population health

demonstrate inequalities in health status, the patterns observed depend on the type of indicator used and the aspect of health it measures.

The measure of health considered here comes from the 1991 census question on long-term limiting illness, which is a self assessed measure of health status. It may be more affected by subjectivity and imprecision than other health measures such as mortality and physiological measures. However, premature mortality can be the result of chronic ill health, and perceived health status is a good predictor of mortality.<sup>1</sup> Moreover, chronic illness encompasses disabling conditions not usually associated with mortality, and provides a more comprehensive health status measure. Self reported health is also associated with physiological health<sup>2</sup> and general practitioner and hospital utilisation.<sup>3</sup> The 1991 census measure has been shown to be ecologically associated with mortality rates.<sup>4,5</sup> The census question therefore provides an indicator which is likely to reflect varying patterns of health and health care need. It is of particular interest because it has been collected for all individuals enumerated in the 1991 census, and therefore provides considerable scope for analysis and potential for applications in health planning. This study examines evidence of health inequalities revealed by this measure.

## EVIDENCE FOR COMPOSITIONAL AND CONTEXTUAL EFFECTS IN HEALTH DIFFERENCE

Socioeconomic differentials in health are associated with a number of characteristics, apparent both at the level of individuals and over broader geographic areas. Inequalities in health have been shown to be associated with sex, age, marital status, and ethnicity, as well as with indicators of socioeconomic position, such as employment, class, education, housing tenure and quality, income, and wealth.<sup>6</sup> Typical geographical dimensions of variation include a broad north/south divide,<sup>7</sup> disparities between urban and rural areas,<sup>8</sup> and differences between rich and poor areas.<sup>9,10</sup>

The source of spatial contrasts may not be simply an aggregation of individual level characteristics within areas. Several recent studies such as Duncan *et al*,<sup>11</sup> Duncan and Jones,<sup>12</sup> and Macintyre *et al*<sup>13</sup> warn against confusing *compositional* effects with *contextual* effects. Compositional effects operate simply because of the varying distribution of types of people whose individual characteristics influence their

Department of  
Geography, Queen  
Mary and Westfield  
College, Mile End  
Road, London E1 4NS  
S Shouls  
P Congdon  
S Curtis

Correspondence to:  
Dr P Congdon.

Accepted for publication  
November 1995

health. That is, similar types of people will have similar health experience no matter where they live. Contextual effects operate where the health experience of a particular type of individual depends not only on his or her own characteristics but also on the area where he or she lives, so that similar types of people have different health status from one part of the country to another. For example, it is well known that, within the UK, there is a north-south mortality gradient within any given social class, such that northern areas have consistently worse mortality.<sup>7,14</sup>

The evidence of social and geographical differences in the health of resident populations in different parts of Britain comes from a very large literature – see reviews by Britton<sup>15</sup> and Townsend, Davidson and Whitehead.<sup>6</sup> Contextual effects are of particular interest to geographers and epidemiologists, since they suggest that place may be germane in some way to the processes which affect health. Broadly, there are two types of approach to the analysis of such effects. One uses tabulations, separate regression analyses or regression analyses with area level variables to make comparisons of associations between health and social conditions in different areas. Another strategy uses multilevel modelling or other techniques to test for the simultaneous effects of both compositional and contextual influences.

Using the tabulation method on data from the longitudinal study, Fox *et al*<sup>16</sup> showed that males in similar categories of class, tenure, and employment had different mortality ratios according to the type of electoral ward (small area) in which they lived. Blaxter<sup>17</sup> used data from the health and lifestyle survey (HALS) to show that residents in industrial areas typically had poorer health than those in rural/resort areas. Health differences between manual and non-manual groups seemed to be more pronounced in industrial areas. In addition, the north/south divide in health depended upon the type of ward in which individuals lived.

Sloggett and Joshi<sup>18</sup> have also examined individual and area effects on mortality by introducing area variables representing type of area of residence into a regression analysis of individuals. Their results suggest that there is little residual variation in mortality associated with area variables once individual differences are accounted for, although some residual variation seemed to be linked to the north/south divide.

Other studies have examined whether regional setting influences the micro-scale ecological relationships between small area population health and small area social profiles. Phillimore and Reading<sup>19</sup> examined ward level indicators of aspects of health such as premature mortality and birthweight. Inequalities in health between more and less deprived wards were more pronounced in urban areas than in rural areas. Eames *et al*<sup>20</sup> reported that the association between ward level mortality and ward level deprivation depended on which regional health authority the ward was situated in.

These studies provide some evidence to suggest that, for groups of individuals with a similar

social position or for small areas with a similar social profile, health may vary according to the region and the type of area in which they are living or located. This supports the idea of contextual effects, although in some cases these effects are quite weak. However, some authors<sup>11,21</sup> suggest that the “tabulation” and separate regression methods of comparing social group differences for different areas are not very efficient strategies to test for contextual effects, and are also possibly biased in neglecting autocorrelation within districts. Hence multi-level modelling, or some alternative, would be more appropriate so that the full “nested” structure of the data can be represented.

Congdon<sup>22</sup> reports an application of multi-level analysis in an ecological study of ward level health differences (in East Anglia and Greater London) nested within local authorities and family health service authorities. Ward level associations between three health indicators and deprivation were tested. This analysis suggested that ward level differences in health were more pronounced in metropolitan suburbs and the inner city than in rural areas.

Other studies have used multilevel modelling with data for individual people at the first level of the model. For example, Humphreys and Carr-Hill<sup>21</sup> and Duncan and Jones<sup>12</sup> used the multilevel modelling approach with the HALS data to examine differences relating to four health measures. Most of the differences in health were found to relate to individual variables, although some residual variation appeared to be associated with area level variables.

Gould and Jones<sup>23</sup> used multilevel modelling to examine the long term limiting illness data collected in the 1991 census. The individuals (level 1) were nested within the sample of anonymised records (SAR) districts (level 2). Again, their health measure was explained mainly by the level one variation but a significant, albeit smaller, area effect remained at level two. The research described here provides a more extended and detailed analysis of the SAR data.

Since health is the outcome of a range of individual behaviours, attributes, and life experience, it is not surprising that most of the studies considered above suggest that individual characteristics are at least as important as geographic context in determining health differences. Indeed, it is probably true to say that individual characteristics “explain” more of the differences in health between people than the characteristics of the areas where they live. On the other hand, most of these studies have shown some evidence that the geographic context has some independent effect on the individual level associations between socioeconomic conditions and health. These contextual effects seem to operate with respect to a variety of measures of socioeconomic position and of health status. Typical contextual effects suggested by the studies reviewed above are as follows:

- There is evidence of a north/south divide in the way that social conditions and health

are associated. Some studies suggest that socioeconomic health inequalities are most striking in the north and west of Britain.

- The relationship between health and socioeconomic conditions may also depend on whether the context is rural or urban.<sup>10 19 24</sup> The effects of urbanisation may operate in outer urban areas as well as in the inner cities. Some results suggest that the dimensions of social position which are salient for health may differ in urban and rural areas. For example poorer access to health care facilities in remote areas may be detrimental, making car ownership especially important in such areas.<sup>8</sup>
- The area typologies of Webber<sup>25</sup> and Craig<sup>26</sup> may also reflect contexts in which there are differences in the relationship between individual social status and health. The strongest differences seem to be between clusters which are clearly rural or semirural and those which are urban/industrial. However, it seems likely that these clusters are picking up differences between rich and poor areas as well as between urban and rural areas.

This suggests that it might be fruitful to examine three particular dimensions of geographic context for possible effects on health differences; these are urban/rural, rich/poor, and north/south disparities.

### Method

A linked data set was established including information on individuals from the SARs from the 1991 census and information on the districts where the individuals lived from data in the 1991 census area files. Multi-level modelling was applied to these data to examine compositional variation (at the level of individuals) and contextual variation (at the level of districts).

### DATA ON INDIVIDUAL CHARACTERISTICS

Two subsample sets were compiled, for males and females separately, using a 10% pseudo-random sampling procedure on the SARs. This was followed by a filtering procedure which removed records of residents aged under 16 years, visitors, and residents living in communal establishments. The two subsamples of the SARs were compared with the original data sets and found to be sufficiently similar in terms of the variables being used.

The dependent variable is binary; it equals 1 if long term limiting illness is present, and zero otherwise. The independent variables are:

- Age;
- Age squared (to represent the possibly non-linear age gradient);
- Social class IV/V (binary);
- Non-white ethnicity (binary);
- Married (binary);
- Deprivation indicator.

Social class is defined by the family head of the household, and if this is not applicable it is described by the individual's social class.

The individual variables were chosen on the basis of preliminary logistic regression analysis of associations between individual characteristics and long standing illness. The deprivation indicator is intended as a measure of material deprivation and is formed by the summation of the following five binary outcomes:

- Living in a household with density greater than 1 person per room;
- Living in a non-owner occupied household;
- Living in a household with no car;
- No access to separate bathroom;
- Unemployed.

Thus, the possible scores range from zero to five. We explored various formulations of the measures of deprivation and found that the combination of variables used here was the most clearly related to the health variable.

### THE TYPOLOGY OF GEOGRAPHICAL CONTEXT

For this study, information on geographical context can be provided at the scale of local government districts or combinations of these. This relatively crude geographical scale is determined by the constraints of the data available in the SAR, which for reasons of confidentiality, only indicate the local government district of residence.<sup>27</sup> Where districts had less than 120 000 population in the mid-1989 estimates they were amalgamated with other districts. The result is a total of 278 districts in Britain with the following population statistics which can be used to identify the place of residence of individuals in the SAR.

- Minimum SAR district population 113 725
- Maximum SAR district population 961 041
- Average SAR district population 197 442

In order to summarise the complexity of area contexts, multivariate techniques were used to (a) combine information on those dimensions of socioeconomic structure which are likely to be most important for health and (b) develop a typology of districts based on these dimensions which can be used as a basis for assessing contextual variation.<sup>28</sup> The analysis was based on a selection of area variables to reflect two of the three key dimensions which were identified above as significant for health inequality and contextual variability: urban/rural and rich/poor. The third, north/south position, is assessed independently because it is a locational attribute with different characteristics from socioeconomic variables.

The details of the classification procedure are discussed elsewhere.<sup>28</sup> To summarise, 25 area variables were selected to represent levels of deprivation/affluence and urbanisation/rurality. The choice of variables was influenced by other work which also aimed to distinguish districts by their socioeconomic structure within these dimensions (for example, by Craig).<sup>26</sup> As it was anticipated that many of the variables would be interrelated, the original variables were reduced to eight significant factors through an oblique rotation factor analysis. The advantage of using oblique rotation is that

Table 1 Descriptions of factors and clusters

(A) Description of factors (with absolute loading above 0.6)	Factor loading
Factor 1 – “inner cities with large ethnic minorities”:	
Ethnic group (persons describing themselves as “white”)	-0.94
Population density	0.89
Overcrowding	0.87
Married persons	-0.84
Single persons	0.81
Flats	0.76
Detached housing	-0.62
Car ownership	-0.61
Unemployment	0.61
Factor 2 – “affluence and service sector”:	
Educational level	-0.92
Social class IIIM	-0.95
Service industry	0.91
Social class I/II	0.85
Social class IV/V	-0.79
Manufacturing industry	-0.75
Single, non-OAP households	0.61
Factor 3 – “Familiism”:	
Dependent children in households	0.93
Lone pensioner households	-0.78
Children as a percentage of population	0.91
Factor 4 – “low proportions of clerical”:	
Social class IIIN	-0.85
Factor 5 – “dual career affluence”:	
Female economic activity	0.96
Unemployment	-0.69
Owner occupancy	0.62
Factor 6 – “deprivation”:	
Owner occupied	-0.82
Housing association, council	0.93
Lone parent family	0.70
Car ownership	-0.80
Unemployment	0.70
Overcrowding	0.68
Flats	0.67
Factor 7 – “commuting areas”:	
Commuting across district	0.93
Social class IV/V	-0.62
Factor 8 – “rurality”:	
Agricultural work	0.90
Detached housing	0.88
Car ownership	0.69
Unemployment	-0.62
Population density	-0.62
Lone parent households	-0.68
Married residents	0.65

(B) Description of clusters	Cluster	Summary of main features
A	Manufacturing, north and central	Manufacturing bias in workforce, high manual and unskilled social classes
B	Socially polarised, London	High social class I/II; high council tenancy, lone parent households, and unemployment
C	Affluent semirural	Low population density; high detached housing and car ownership
D	Affluent suburbia	High commuting, located around London and other metropolitan hinterlands (eg Bristol)
E	Rural	Very low population density and agricultural bias
F	Deprived, northern	High levels of unemployment and council tenancy
G	Affluent seaside resorts	Large population over retirement age; located primarily around coastal regions
H	Middle sized towns	Self contained middle sized towns located around Britain, especially in northern and central areas
I	Socially mixed, London	Social class distribution similar to whole of Britain, high council tenancy, lone parent households and unemployment

the factors are allowed to be correlated with one another, reflecting the real world association between the dimensions. The eight factors accounted for 93% of the original variability in the data. Table 1(A) shows the description of the eight factors in terms of the variables which most strongly influenced the definition of factors; that is, those with absolute variable loadings (eigenvalues) over 0.6.

The factor scores provided summary indices of the SAR area socioeconomic structure by which to derive a cluster analysis (using the iterative relocation method). Nine clusters (labelled A to I) were chosen as a “natural” grouping on the basis of first and second differences in the residual sum of squares. The differences in socioeconomic structure between these are described briefly in table 1(B). The appendix shows the complete list of variables used in the factor analysis and the average

values of each of these variables for districts in each cluster.

The geographical (ecological) associations between population health and the typology of districts developed for this study are examined in more detail in a separate paper.<sup>28</sup> That analysis confirmed that the typologies are associated with area differences in the reported prevalence of long term illness. Area standardised illness ratios (SIRs) were calculated from the census illness data, using indirect standardisation to measure area health differences. Table 2 shows correlation coefficients between the factor scores and SIRs of the SAR districts. Factors 5, 6, and 8 display the strongest linear correlations with SIRs. Factor 5, the “dual career affluence” and factor 8, “rurality” are both negatively correlated with SIR. Factor 6 is associated with deprivation and has a positive association with SIR.

Two factors showed non-linear relationships with SIR: factor 8 and factor 1. Factor 8, the rurality factor, has a “U” shaped association with health since semirural areas have lower SIRs than either urban or deeply rural areas. This is in accordance with Bentham’s<sup>24</sup> finding using mortality rates. The other non-linear relationship exists between factor and SIRs. This factor, which defines inner cities (especially inner London) with large ethnic minorities, has a complex bifurcated relationship with SIR.

The average SIR over districts in each cluster is shown in table 3. Higher than average standard illness ratios occurred in clusters A, B, F, and I for males and females of both age groupings. Cluster A (“manufacturing, north and central”) and cluster F (“deprived northern”) revealed the highest average SIR. These clusters may be reflecting the north/south divide in Britain. Clusters B and I, the two London clusters, have higher average SIR for females compared with males, which suggests a different contextual effect may be occurring between sexes.

Clusters C (“affluent semirural”) and D (“affluent suburbia”) show the lowest average SIRs ratios for both males and females. Clusters E (“rural”) and G (“affluent seaside resorts”) show a lower than average SIR. These clusters appear to reflect the better health enjoyed in more rural and more affluent areas.

We also tested the relevance of the area clusters to reported illness using separate logistic regression analyses on groups of individuals in our sample from the SAR, selected according to the type of area where they were living. Table 4 illustrates some of the results derived from these analyses. In this table, the odds ratios for long term illness are compared with the base case of an individual who is not in social class IV/V, who describes themselves as of white ethnicity and has a deprivation score of zero. (For the method of calculation see, for example, Hosmer and Lemshow).<sup>29</sup> The table includes information from separate regressions for people living in three of the clusters of districts described above, shown here to illustrate some of the differences between area types in the strength of association between health and other characteristics. The table also

Table 2 Correlation between factors and standardised illness ratios

Factors	Correlation coefficients			
	Males 16+	Females 16+	Males 45-59	Females 45-59
1 Inner city with large minority ethnic groups	0.20	0.30	0.22	0.38
2 Affluence and service sector	-0.40	-0.37	-0.36	-0.30
3 Familism	0.13	0.22	0.10	0.17
4 Low proportions of clerical	0.21	0.22	0.20	0.22
5 Dual career affluence	-0.70	-0.62	-0.72	-0.63
6 Deprivation	0.56	0.58	0.58	0.63
7 Commuting areas	-0.34	-0.25	-0.35	-0.29
8 Rurality	-0.46	-0.55	-0.44	-0.55

Table 3 Average standardised illness ratios of adult residents in clusters

Clusters	Average standard illness ratios			
	Males 16+	Females 16+	Males 45-59	Females 45-59
A Manufacturing, north and central	128.51	123.67	138.85	134.61
B Socially polarised, London	110.32	113.62	116.32	130.02
C Affluent, semirural	82.57	87.42	71.01	74.89
D Affluent, suburbia	83.47	88.74	73.03	77.81
E Rural	92.17	89.42	89.63	82.76
F Deprived, northern	132.28	125.93	149.63	143.23
G Affluent, seaside resorts	92.39	90.36	89.24	85.19
H Middle sized towns	99.58	99.42	97.00	96.81
I Socially mixed, London	119.58	123.23	131.71	145.40

Table 4 Relative risk of assessed morbidity by area types and standard regions using logistic regression

(The base case describes an individual who is not in social class IV/V, is white, married, with a deprivation score of zero)

	GB	Area types			Aggregated standard regions		
		Cluster A (manufacturing, north and central)	Cluster C (affluent, semirural)	Cluster F (deprived, northern)	North	Central	South
<b>Males</b>							
Base	1.00	1.00	1.00	1.00	1.00	1.00	
Social class IV/V	1.27	1.27	1.26	1.34	1.37	1.24	
Ethnic group ("non-white")	1.20	1.24	1.82	1.37	1.50	1.48	
Not married	1.33	1.39	1.28	1.22	1.45	1.15	
Deprivation indicator over 2	1.89	1.73	2.00	1.44	1.64	2.01	
<b>Females</b>							
Base	1.00	1.00	1.00	1.00	1.00	1.00	
Social class IV/V	1.29	1.23	1.24	1.15	1.35	1.21	
Ethnic group ("non-white")	1.08	0.85	1.12	0.85	0.84	1.36	
Not married	1.07	0.91	1.01	1.03	0.97	1.02	
Deprivation indicator over 2	1.98	1.93	2.32	2.13	1.88	2.06	

shows results of separate regression analyses for subjects in aggregated standard regions, which show disparities in the relative risks.

The factor scores and clusters used to typify areas of the country therefore seemed appropriate to distinguish types of area with differing patterns of reported illness. We proceeded to test the evidence for contextual effects associated with this area typology using multilevel modelling, which allows further exploration of health outcomes in individuals, taking account of both individual attributes, and the attributes of the area of residence.<sup>30</sup>

#### STRATEGY FOR MULTILEVEL MODELLING

Males and females were modelled separately to allow for any gender differentiation of relationships. The multilevel analysis described below used information only on the selected males and females aged 16 to 59. This age range was chosen because health inequalities tend to be more marked for these age groups than for people at older ages.

A two level structure of individuals (level 1) nested within SAR districts (level 2) was adopted. The individual data at level 1 were as described above. The logarithmic form of the

deprivation score was used after adding one. The logarithmic form of the deprivation score is used because it yielded greater gain in deviance in the multilevel analysis. The factor scores and the clusters described above were used as alternative ways to characterise the district level (level 2) effect on the illness outcome of residence in different SAR districts. A logistic regression formulation was appropriate, and was applied within a multilevel context; allowance was made for departures the binomial assumption (which turns out to be minimal). In total seven multilevel models are described in the following section.

#### Results

Tables 5A, 6 and 7 show the results of the multilevel models for men and women. These tables include information on models which demonstrate different aspects of the findings from this analysis. The model fit is measured in terms of deviance, or the log likelihood ratio: as the goodness of fit improves, the deviance falls. Changes in fit are thus indicated by reductions in deviance as compared to the simplest model (model 1, described below). The greater the reduction in deviance, the better

Table 5 Multilevel models: incorporating multivariate sample of anonymised records analysis

	Model 1 Variance components		Model 2(A) Deprivation added as fixed effect		Model 2(B) Deprivation varies at level 2		Model 3 More fixed effects added	
	Parameter	(SE)	Parameter	(SE)	Parameter	(SE)	Parameter	(SE)
Males (16–59)								
Level 2 variances:								
Intercept variance	0.049	(0.018)	0.031	(0.016)	0.073	(0.026)	0.076	(0.027)
Intercept-slope covariation	—		—		–0.041	(0.018)	–0.042	(0.019)
Deprivation slope variance	—		—		—		—	
Level 1 variance:								
Extra-binomial variance	1.003	(0.008)	1.000	(0.008)	0.996	(0.008)	0.999	(0.008)
Fixed parameters:								
Constant	–2.730	(0.028)	–2.992	(0.035)	–2.993	(0.037)	–5.304	(0.100)
Age	—		—		—		0.064	(0.002)
Non-white	—		—		—		0.188	(0.109)
Social class IV/V	—		—		—		0.308	(0.057)
Married	—		—		—		–0.457	(0.058)
Log of deprivation score	—		0.720	(0.049)	0.735	(0.048)	0.732	(0.054)
Reduction in deviance from model 1	0		1501.000		1498.204		7824.100	
Females (16–59)								
Level 2 variance:								
Intercept variance	0.095	(0.023)	0.080	(0.022)	0.148	(0.038)	0.152	(0.039)
Intercept-slope covariation	—		—		–0.099	(0.046)	–0.106	(0.048)
Deprivation slope variance	—		—		0.094	(0.071)	0.108	(0.076)
Level 1 variance:								
Extra-binomial variance	1.003	(0.008)	1.000	(0.008)	0.997	(0.008)	0.992	(0.008)
Fixed parameters:								
Constant	–2.792	(0.032)	–3.004	(0.038)	–3.003	(0.041)	–4.122	(0.074)
Age*Age	—		—		—		0.001	(0.000)
Non-white	—		—		—		0.096	(0.116)
Social Class IV/V	—		—		—		0.321	(0.059)
Married	—		—		—		–0.248	(0.059)
Log of deprivation score	—		0.615	(0.055)	0.620	(0.058)	0.643	(0.064)
Reduction in deviance from model 1	0		999.420		1006.000		6312.240	

the fit to the data. The importance of reduction in the deviance from one model to another needs to be interpreted in terms of the statistical significance of the reduction and the information

criterion (goodness of fit in terms of the numbers of predictor variables in the model).

Table 6 Multilevel models: incorporating multivariate sample of anonymised records analysis

	Variable slopes with log of deprivation. Fixed variables included at the district level			
	Males		Females	
	Parameter	(SE)	Parameter	(SE)
Model 4				
Level 2 variances:				
Intercept variance	0.012	(0.016)	0.071	(0.032)
Covariance: intercept and deprivation	–0.028	(0.009)	–0.061	(0.041)
Deprivation variance	—		0.073	(0.070)
Level 1 variances:				
Extra-binomial variance	0.994	(0.008)	0.992	(0.008)
Fixed parameters:				
Constant	–5.320	(0.099)	–4.142	(0.072)
Age*age	0.064	(0.002)	0.001	(0.000)
Non-white	0.267	(0.113)	0.205	(0.122)
Social class IV/V	0.274	(0.057)	0.295	(0.059)
Married	–0.467	(0.057)	–0.263	(0.059)
Deprivation score	0.695	(0.055)	0.623	(0.064)
Factor 1	–0.044	(0.028)	–0.118	(0.038)
Factor 2	–0.105	(0.028)	–0.103	(0.035)
Factor 3	0.005	(0.029)	–0.027	(0.033)
Factor 4	0.040	(0.026)	0.038	(0.030)
Factor 5	–0.103	(0.026)	–0.074	(0.033)
Factor 6	–0.022	(0.025)	0.038	(0.035)
Factor 7	–0.058	(0.025)	–0.004	(0.032)
Factor 8	–0.119	(0.029)	–0.141	(0.036)
Reduction in deviance from model 1	8459.4		7008.6	
Model 5				
Level 2 variances:				
Intercept variance	0.040	(0.019)	0.088	(0.033)
Covariance: intercept and deprivation	–0.043	(0.009)	–0.072	(0.042)
Deprivation variance	—		0.077	(0.070)
Level 1 variances:				
Extra-binomial variance	0.994	(0.008)	0.993	(0.008)
Fixed parameters:				
Constant	–5.424	(0.114)	–4.272	(0.096)
Age*age	0.064	(0.002)	0.001	(0.000)
Non-white	0.229	(0.109)	0.147	(0.119)
Social class IV/V	0.285	(0.057)	0.304	(0.059)
Married	–0.468	(0.057)	–0.257	(0.059)
Deprivation score	0.700	(0.053)	0.625	(0.064)
Cluster A	0.390	(0.079)	0.424	(0.094)
Cluster B	–0.090	(0.154)	–0.065	(0.184)
Cluster D	–0.164	(0.098)	–0.009	(0.106)
Cluster E	0.023	(0.105)	–0.059	(0.122)
Cluster F	0.377	(0.079)	0.371	(0.108)
Cluster G	0.045	(0.116)	0.234	(0.123)
Cluster H	0.197	(0.091)	0.129	(0.109)
Cluster I	0.105	(0.146)	–0.141	(0.187)
Reduction in deviance from model 1	8267.5		6826.8	

MODEL 1

Model 1, the base case (variance components), has the data composed into two levels without additional information on individuals or districts. The intercept variance parameter is positive and significant, reflecting the geographical clustering of higher or lower individual propensity to report morbidity. Thus, before allowing for the impact of individual characteristics, there is evidence of significant autocorrelation in illness levels within districts. The proportion of total variation which is attributable to the individual level is, however, greater than the district level variation. The “district effect” is equivalent to around 5% and 10% of the total variation for males and females respectively.

MODEL 2

Model 2(A) includes the measure of individual multiple deprivation. There is a major reduction in deviance as compared to model 1, with the statistically significant and positive coefficient associated with individual deprivation showing that more deprived individuals are significantly more likely to report illness than less deprived people. The intercept variance parameter remains significant, indicating that in certain districts the population shows a greater propensity to report illness, after allowing for their individual deprivation characteristics (that is, controlling for an important compositional effect). However, the introduction of information on the composition of the population has reduced the intercept variance slightly, suggesting that some of the clustering apparent in model 1 was due to the compositional differences in deprivation.

Table 7 Cross level interaction models. Model 6 – multilevel models with cross level interactions

	Males		Females	
	Parameters	(SE)	Parameters	(SE)
Level 2 variances:				
Intercept variance	0.000	(0.000)	0.068	(0.031)
Covariance	—	—	-0.054	(0.041)
Deprivation variance	—	—	0.063	(0.069)
Level 1 variance:				
Extra-binomial variance	1.003	(0.008)	0.993	(0.008)
Fixed parameters:				
Constant	-5.332	(0.099)	-4.155	(0.072)
Age/Age*Age	0.064	(0.002)	0.001	(0.000)
Non-white	0.207	(0.111)	0.190	(0.122)
Social class IV/V	0.271	(0.058)	0.297	(0.059)
Married	-0.465	(0.058)	-0.264	(0.059)
Deprivation score	0.698	(0.059)	0.662	(0.065)
Factor 2	-0.124	(0.038)	-0.099	(0.053)
Factor 5	-0.174	(0.037)	-0.171	(0.044)
Factor 7	-0.007	(0.036)	-0.106	(0.040)
Factor 8	-0.087	(0.036)	-0.139	(0.047)
Fa1*dep	—	—	-0.027	(0.071)
Fa2*dep	-0.016	(0.053)	0.143	(0.065)
Fa5*dep	0.151	(0.052)	0.035	(0.060)
Fa7*dep	-0.071	(0.054)	—	—
Fa8*dep	-0.006	(0.055)	0.001	(0.073)
Reduction in deviance from model 1	8458.694		8715.844	

Model 2(B) introduces contextuality in the deprivation effect by allowing the association between health and deprivation (the health “gradient” in relation to deprivation) to vary at the level of districts. The size and significance of the district level intercept variance increases as we pass from model 2(A) to model 2(B), and there is a small but significant gain in goodness of fit: the deviance gain of model 2(B) over model 2(A) is 6.9 for females (for the loss of 2 degrees of freedom) and 6.6 gain in deviance for males (for the loss of 1 degree of freedom). Therefore, by allowing the deprivation gradient in health to vary between districts, we can clarify the clustering effect (intercept variance) and improve model fit. However, the form of contextual variation differs in relation to gender: only for females is there significant variation in slopes *per se*, while for both males and females there is a negative intercept-slope covariance term. This implies that in areas with higher levels of long term illness (usually less affluent areas), the effect of deprivation on illness is less marked than in more affluent areas. Thus in areas where levels of illness are lower overall, the illness “gradient” associated with deprivation is particularly steep at the individual level. This suggests that district context may influence the individual impact of deprivation on health.

### MODEL 3

Model 3 expands model 2(B) by including further information on individual attributes. The large change in deviance reflects the importance of individual variables in explaining differences in reported illness. However, there is no reduction in the intercept variance at the level of districts in going from model 2(B) to model 3, and if anything there is an increase for both males and females, suggesting that areal differences are greater when the compositional effect of their populations is allowed for (see Jones and Bullen,<sup>31</sup> p 1420). The coefficients associated with the individual characteristics show that propensity to report illness

increases with age and is significantly higher for people in classes IV and V. This class effect is evident after controlling for material deprivation since the deprivation score is also in the model. Married people report illness significantly less. The association with ethnic minority status in this analysis appears to be rather weak. This seems to suggest that deprivation is more strongly associated with illness reporting than ethnicity *per se*. However, we would not make this an inqualified assertion. It might be argued that the one tailed significance test is more appropriate for this variable, and on this basis, there would be evidence here for a significant association at the 5% level of probability for males, but not for females. It should also be noted that the pattern of variation of health between ethnic minorities may be variable between different ethnic groups. Since in many parts of the country, the numbers of individuals in the sample belonging to specific groups was very small, we have used a rather crude indicator of ethnicity, distinguishing between “white” and “non-white” categories only, and this may mask the detail of ethnic differences.

### MODELS 4 AND 5

Models 4 and 5 include information about the SAR districts, using the information on factors and clusters to describe the district context. The inclusion of factor scores in model 4 has some effect on the individual level fixed parameters when compared to model 3. The effect of ethnicity increases in both size and significance (moving from non-significance to significance for females) but there is a slight reduction in the social class IV/V parameter. The district level fixed parameters (the factor scores) have additional explanatory power with respect to long standing illness, though with some differentiation between males and females. For males, four factors are of importance: factors 2, 5, 7, and 8. The first three of these describe different aspects of affluence and have a negative impact on the probability of illness. Interestingly, factor 5 previously showed strong negative correlation with male SIRs in SAR districts (see table 2). Factor 8 (“rurality”) was also negatively correlated with male SIRs.

For females, factors 2, 5, and 8 have similar significant effects but factor 1 (“inner cities with large ethnic populations”) is further associated with lower chances of illness. As would be expected, the inclusion of ecological fixed part variables reduces the intercept variance at the district level considerably, and more so for males than females. Therefore, it appears that morbidity is better explained by the socio-economic typology of areas in the case of males. Other studies of health differences have also shown a stronger ecological relationship with socioeconomic variables for males compared with females.<sup>6</sup>

In model 5, clusters replace factors to represent the type of area. Cluster C (“affluent semirural”) was chosen as the reference cluster as it previously revealed the lowest average SIRs for males aged over 16 years (table 3).

Clusters also show clear differentiation in illness rates between areas within the multilevel regression. Compared with cluster C, there is an enhanced risk of illness in cluster A, "manufacturing, north and central", and cluster F, "deprived, northern". These areas also showed the highest average SIR for males over 16 years (table 3). Both clusters are mainly located in the north of Britain and may reflect aspects of the north/south divide.

Residence in two area clusters, D and H, also play a role in explaining male long term illness. These are the "affluent suburbia" and "middle sized towns" clusters which have negative and positive effects on illness respectively. The significant difference shown between cluster C and D was unexpected as both had similar average SIR (see table 3). Clusters C and D both represent affluent regions but cluster D represents more affluent commuting regions around London and other metropolitan hinterlands (see table 1). In this case there appears to be a contextual advantage associated with cluster D. The positive association with cluster H might indicate some urban-rural effect, but it may also reflect the fact that many districts in this cluster were in northern or central parts of the country. For females the likelihood of long term illness is higher in cluster G (retirement areas), in addition to clusters A and F, which is possibly a residual effect of age structure even though age is explicitly included as a predictor.

To examine whether the type of district has an effect on the relationship between individual deprivation and illness it is necessary to include cross level interactions between district attributes (that is, factor scores or cluster group) and individual multiple deprivation scores.

#### MODEL 6

Model 6 (table 7) examines cross level interactions of the individual multiple deprivation score with area characteristics to test whether the deprivation relationship is dependent upon the type of area an individual lives in. This is the "contrast" approach to assessing contextual effects but within a multilevel setting (see Robertson).<sup>32</sup> In these models we retained only the significant factor effects shown in model 4. For males, the only cross level interaction of significance occurs for factor five, the "dual career affluence" factor. It was noted above that there was a significant negative covariation in model 4 between slope and intercept at the district level, interpretable as the more marked effect of individual deprivation on the chance of illness in affluent areas. Introduction of the cross-level interaction involving district affluence scores eliminated both this covariance effect, and the intercept variance. This suggests that the source of the covariation at the district level has been effectively modelled by explicit representation of area structure as a fixed effect, and in particular, its interaction with individual deprivation. The change in scaled deviance for males between model 4 and 6 is insignificant in statistical terms (that is, goodness of fit remains the same). However, it should also be

noted that model 6 is more parsimonious (has fewer predictor variables) than model 4, so some improvement of fit occurred.

For females the cross level interaction term for factor 2 is significant. This is also an affluence factor, though with a slightly different interpretation to factor 5 (see table 1). There is a significant improvement in goodness of fit between model 4 and model 6 for women. The contextual effects represented by the intercept and covariance coefficients at the district level are also reduced in model 6 compared with model 4.

The introduction of the cross level interactions had some effect on the fixed level two parameters; the impact of scores on the "affluence factors" two and five became larger in absolute terms, for both males and females. Thus including information on the interaction between individual and area characteristics has made the "district" effect on health difference clearer. In particular the sources of the negative covariation between slope and intercept at the district level (in models 2(B) to 5 for males) can be more explicitly quantified: the main effect of male individual deprivation increases the probability of long term illness by around three and a half times (comparing scores 5 and 0) but this effect is enhanced or diminished according to the score of the district of residence on the affluence factor 5; the greater the affluence of the district the more marked the individual deprivation effect, and the greater the deprivation of the district the more attenuated the individual deprivation effect. For females a similar effect follows from the interaction between factor 2 and the individual deprivation score.

#### Discussion

This paper has investigated the variables associated with reported illness for a sample of individuals resident in Britain taken from the 1991 census. We have particularly focused on the association between health and individual material deprivation. While we concur with others (for example, Sloggett and Joshi<sup>18</sup>) who suggest that individual variation is of prime importance in explaining health inequalities, we report here results suggesting some contextual effects are also significant.

Geographic context has been summarised here using factor analysis and clustering techniques to classify the 278 areas of residence identified in the SAR. A review of other studies suggested that population health is likely to vary according to area deprivation, urbanisation and rurality, and, in the UK, the north/south divide. Our classification of areas distinguishes areas along these dimensions.

A multilevel analysis has shown that there are areas where high levels of illness tend to be clustered and that this is not fully explained by the individual characteristics of the people of whom the district population is composed. Especially in the case of males, much of this clustering was associated with area deprivation (in more deprived areas individuals of all types had higher overall levels of illness). There is



less evidence in favour of a contextual effect associated with rural as against urban areas, although the factor indicating rurality did show some negative association with propensity to report illness. There is also a contextual effect emerging from the present study which shows that the health inequality between more and less deprived individuals is especially marked in more affluent areas, rather than in more deprived areas. This emerges in two ways: as a negative covariation between slopes and intercepts at the district level in models without cross-level interactions, and in the form of significant interactions between individual deprivation and district affluence scores in models allowing such interdependence.

Alternative methods of assessing variation gave similar results, but with unclear statistical associations. Inadequacies in using tabular and separate regression analyses can occur depending upon how the data is structured. Tabular and separate regression models also assume that no autocorrelation exists within districts and this is also true for regression methods which include area level variables at the individual level. Humphreys and Carr-Hill<sup>21</sup> describe this as an intra-area clustering of outcomes – individuals living within an area resemble individuals living in the same area more than individuals living in different areas. By assuming autocorrelation does exist and modelling it by allowing variation to occur at the higher level it is possible to explore area differences more explicitly. Our study showed that much of the district level variation seemed to be accounted for by area specific variables or by cross-level interactions between individual and district. In this type of study, it is possible that differences which appear to operate at an area level may arise because the compositional effects have been incompletely modelled. However, in our study, we noted an increase in area variation after individual level variables were included, suggesting that differences between areas are greater when the compositional effects of their population is allowed for.

Nevertheless, it needs to be borne in mind that the incomplete specification of individual fixed effects may have important effects on the interpretation of multilevel analysis. Some aspects of individual variation which are not measurable in census data (for example, income) might explain what appear in this analysis as area variations.

We report here analysis of over 68 000 individuals, being a 10% pseudorandom subsample from the SAR, designed to make the analysis manageable in the available version of the multilevel modelling software. An analysis using the whole SAR would be more statistically powerful, although the SAR is itself a sample of the total population in every respect. The significance levels associated with our results show those patterns which are unlikely to have arisen from chance variation between samples. It is, however, possible that other subsamples drawn from the SAR would have produced different patterns of association and this would be interesting to test in further analysis.

The finding of an apparently stronger ecological effect operating for males than for females, parallels findings from other studies which have shown area inequalities for middle-aged males to be particularly striking. In this analysis it is possible that this is because the specification of individual characteristics used here is less powerful for males than for females. An alternative interpretation is that the effects of area of residence have a particularly strong influence on men.

It is interesting to speculate beyond the empirical evidence here to consider reasons why these dimensions of geographic setting might be important. Our study has shown that, regardless of their individual characteristics, people in more socio-economically disadvantaged areas seem to report more illness. Macintyre *et al*<sup>13</sup> suggest that there are five broad ways in which socio-environmental factors might influence health for all people living in disadvantaged areas. They suggest that physical features of the environment such as air and water quality and climate may be important. The effect of the domestic and working environment in the local area may also have an impact. They also suggest that one should consider the provision of services including health and social care, and educational quality, as well as socio-cultural features of neighbourhoods and the reputation of an area which may affect psychological health and morale. The operation of such factors are illustrated in a comparison of two areas of Glasgow. In a similar vein, Phillimore and Morris<sup>34</sup> discuss a range of factors which might account for the differences between mortality in wards in Middlesbrough and Sunderland. They suggest that environmental pollution is the most likely differentiating factor between the two areas.

Various possible reasons could also be postulated for the relatively strong differentials in health between more and less deprived people living in relatively affluent areas. The effect may be a statistical one, resulting from more extreme differences in wealth between individuals in more affluent areas, so that stronger socio-economic inequality is reflected in clearer differences in health. One possibility is that individuals with low deprivation scores in more affluent areas are in fact extremely affluent and that the deprivation score is not adequately representing the difference between these very rich individuals and deprived people (a "floor effect" in the representation of affluence by a deprivation score). This might suggest that a measure of material conditions explicitly measuring individual affluence might have an additional role to individual deprivation in explaining spatial context effects in individual health inequalities. Another effect may be a socioeconomic one, in that the impact of a given level of poverty is greater for an individual living in a generally more affluent district: perhaps because a sense of relative deprivation is more acute and has a more severe impact on psychological state, because community support for deprived minorities is less in affluent areas, or perhaps because the cost of achieving a reasonable standard of living is greater so that

the poor suffer particularly marked material hardship in predominantly wealthy areas.

The results from this study suggest that both individual socioeconomic characteristics and the socioeconomic attributes of place of residence may be significantly associated with health differences. This finding is pertinent to community health strategies which aim to reduce health inequalities by improving the health of the poorest members of society. It seems to suggest that in areas where illness levels are generally high, it may be necessary to take action at the scale of whole communities, as well as for deprived individuals within those communities, and that such action might have benefits for the whole population. The finding of a particularly steep health gradient in more affluent areas underlines the need to be able to identify pockets of deprivation within generally advantaged areas, perhaps with a view to targeting resources effectively towards these groups of individuals. The results here also remind us that the health effects of deprivation do not always follow the same pattern in every place and that it is important to work out local strategies for reduction of health inequalities which are sensitive to geographical differences.

We would like to thank ESRC for its support in funding this research work (Grant number H507 25 5121). The work is based on the SARs and 1991 census area statistics provided through the Census Microdata Unit of the University of Manchester with the support of ESRC/JISC/DENI. All original data is used with the permission of the Controller of Her Majesty's Stationery Offices and is subject to Crown Copyright.

**Appendix**

For the factor analysis of SAR district characteristics, 25 variables were used from the 1991 census data. These variables were collated for the SAR districts. The variables were as follows:

- % of working population engaged in agriculture;
- % of working population engaged in manufacturing;
- % of working population engaged in services;
- % of economically active (ea) population (males and females) who were unemployed;
- % of female population of working age who were economically active;
- % of households living in flats;
- % of households living in detached houses;
- % of households living in owner occupied housing;
- % of households living in council rented accommodation;
- % of ea population in social class I (professionals) or II (managers);
- % of ea population in class III non-manual (junior clerical occupations);
- % of ea population in class III manual (skilled/supervisory manual workers);
- % of ea population in class IV/V (semi- or unskilled manual workers);
- % of households including dependent children;
- % of households including pensioners;
- % of households headed by lone parents;
- % single person households;
- % households including married couples;
- % population classed as "white";
- % households without cars;
- % population aged 30-44 with low educational attainment;
- % population in overcrowded housing; population density per hectare;
- % population who are under 16 years;
- % ea population who commute to work.

The average values of these variables for the different clusters derived from the analysis are shown in the table.

Cluster description – average percentages by SAR district (see table 5 for the names of the clusters)

Cluster	Work type			Economic activity		Housing type/tenure				Social class			
	Agriculture	Manufacturing	Service	Unemployed	Female	Flats	Detached	Owner occupied	Council etc.	III	IIIN	IIIM	IV/V
A	0.9	25.0	32.9	11.1	64.9	11.7	14.9	44.3	27.3	23.8	19.6	24.6	25.9
B	0.2	7.6	59.1	14.5	68.4	70.8	0.9	41.4	37.6	43.4	19.8	12.5	15.8
C	2.2	19.7	40.1	6.1	69.0	11.1	31.1	52.1	17.8	36.2	21.6	19.5	19.0
D	0.9	14.8	44.2	7.5	68.5	19.3	18.4	53.6	17.6	34.9	26.4	18.6	16.4
E	7.8	15.5	36.4	7.3	65.7	9.6	37.8	40.8	18.7	30.2	18.8	21.4	23.8
F	0.8	16.9	40.6	14.0	63.3	27.9	9.4	38.1	40.1	24.7	21.4	21.7	24.2
G	3.1	13.2	41.4	8.1	66.8	17.6	32.1	45.6	12.8	32.3	23.0	19.8	20.7
H	1.5	16.9	41.3	8.9	68.5	20.1	17.0	47.8	23.4	30.4	22.7	20.4	21.6
I	0.1	11.5	50.2	17.0	64.7	55.0	1.7	39.5	42.1	29.8	22.5	17.3	20.2
GB	2.2	17.6	40.2	9.2	66.9	18.5	21.9	46.5	23.1	31.0	21.9	20.6	21.3

  

Cluster	Family					General							
	Dependent children	Pensioner h/h's	Lone parent	Single h/h's	Married	Ethnicity "white"	Cars	Education attain	Over-crowding	Population density	Children	Commuting	Number districts
A	32.3	24.2	4.6	10.2	59.5	94.6	62.4	18.8	2.2	13.3	20.1	30.6	49
B	22.4	19.7	6.4	26.5	39.5	78.7	46.9	51.9	4.9	85.9	15.9	57.4	8
C	31.8	22.8	2.9	9.8	62.1	97.7	77.9	30.1	1.4	4.2	18.9	42.3	55
D	30.7	23.3	3.4	10.8	59.5	93.5	74.7	26.7	2.0	16.4	18.8	49.3	41
E	29.4	27.4	3.0	9.4	62.1	99.3	75.1	24.1	1.5	1.7	18.1	21.9	32
F	31.3	24.3	5.9	12.3	56.6	97.4	53.2	21.7	3.0	16.3	19.8	32.3	27
G	25.8	32.8	3.2	10.1	59.6	99.0	72.2	23.4	1.4	7.8	16.7	28.0	27
H	29.3	24.8	4.2	12.6	57.6	96.4	66.1	26.4	2.0	12.3	18.8	21.9	31
I	29.5	20.4	7.1	18.7	47.0	67.4	49.4	34.8	6.5	65.0	20.7	57.3	8
GB	30.2	24.8	4.0	11.3	58.9	95.2	68.4	25.8	2.1	13.8	18.8	35.2	278

h/h's = households, see notes explaining variables.

- 1 Östlin P. Occupational history, self reported chronic illness, and mortality: a follow up of 25 586 Swedish men and women. *J Epidemiol Community Health* 1990;44(1):12-16.
- 2 Wannamethee G, Shaper AG. Self-assessment of health status and mortality in middle-aged British men. *Int J Epidemiol* 1991;20(1):239-45.
- 3 Judge K, Benzeval M. *Limiting long-standing illness and the 1991 census. Implications for health planning in the 1990s*. London: The King's Fund Institute, 1990.
- 4 Charlton J, Wallace M, White I. Long-term illness: results from the 1991 census. *Population Trends* 1994;75:18-24.
- 5 Congdon P. Socio-economic structure and health in London, Dept of Geography, QMWC. *Urban Studies* 1995;32(3):523-49.
- 6 Townsend P, Davidson N, Whitehead M. *Inequalities in health*. London: Penguin, 1988.
- 7 Britton M, Fox A, Goldblatt P, Jones D, Rosato M. The influences of socio-economic and environmental factors on geographic variation in mortality. In: Britton M. ed. *Mortality and geography: A review in the mid-1980s England and Wales*. London: HMSO, 1990;57-78.
- 8 Watt IS, Franks AJ, Sheldon TA. Health and health care of rural populations in the UK: is it better or worse? *J Epidemiol Community Health* 1994;48:16-21.
- 9 Townsend P, Phillimore P, Beattie A. *Health and deprivation: inequality and the north*. London: Croom Helm, 1988.
- 10 Reading R, Openshaw S, Jarvis S. Are multidimensional social classifications of areas useful in UK health service research? *J Epidemiol Community Health* 1994;48:192-200.
- 11 Duncan C, Jones K, Moon G. Do places matter? A multi-level analysis of regional variations in health-related behaviour in Britain. *Soc Sci Med* 1993;37:725-33.
- 12 Duncan C, Jones K. Individuals and their ecologies: analysing the geography of chronic illness within a multilevel modelling framework. *Journal of Health and Place* 1995;1:27-40.
- 13 Macintyre S, Maciver S, Sooman A. Area, class and health: should we be focusing on places or people? *Journal of Social Policy* 1993;22(2):213-34.
- 14 Green A. The north/south divide in Great Britain: an examination of the evidence. *Transactions of the Institute of British Geographers* 1988;13:179-88.
- 15 Britton M. Geographic variation in mortality since 1920 for selected causes. In: Britton M. *Mortality and geography: A review in the mid-1980s England and Wales*. London: HMSO, 1990;27-45.
- 16 Fox A, Jones D, Goldblatt P. Approaches to studying the effect of socio-economic circumstances on geographic differences in mortality in England and Wales. *Br Med Bull* 1984;40(4):309-14.
- 17 Blaxter M. *Health and lifestyles*. London: Tavistock/Routledge, 1990.
- 18 Sloggett A, Joshi H. Higher mortality in deprived areas: community or personal disadvantage? *BMJ* 1994;309:1470-4.
- 19 Phillimore P, Reading R. A rural advantage? Urban-rural health differences in northern England. *J Public Health Med* 1992;14(3):290-9.
- 20 Eames M, Ben-Shlomo Y, Marmot M. Social deprivation and premature mortality: regional comparison across England. *BMJ* 1993;307:1097-102.
- 21 Humphreys K, Carr-Hill R. Area variations in health outcomes: Artefact or ecology. *Int J Epidemiol* 1991;20(1):251-8.
- 22 Congdon P. The impact of area context on long term illness and premature mortality: an illustration of multi-level analysis. *Regional Studies* 1995;29:327-44.
- 23 Gould MI, Jones K. Analysing perceived limiting long-term illness using UK census microdata. *Proceedings of the Sixth International Medical Geography Symposium Vancouver*. British Columbia, Canada, July 12th-16th, 1994.
- 24 Bentham CG. Mortality rates in the more rural areas of England and Wales. *Area* 1984;16:219-26.
- 25 Webber R. *The classification of residential neighbourhoods: an introduction to the classification of wards and parishes*. PRAG Technical Report TP23. London: Centre for Environmental Studies, 1977.
- 26 Craig J. Local authority urban-rural indicators compared. *Population Trends* 1988;51:30-8.
- 27 Marsh C. Samples of anonymised records from the 1991 census. *Population Trends* 1992;69:17-26.
- 28 Shouls S, Congdon P, Curtis P, Green M. Geographic variation in health: the development of a relevant area typology for SAR districts. *Health and Place* (In press).
- 29 Hosmer DW, Lemshow S. *Applied logistic regression*. New York: John Wiley & Sons, 1989.
- 30 Paterson L, Goldstein H. New statistical methods for analysis social structures: an introduction to multilevel models. *Multilevel Modelling Newsletter* 1991;3(1):4.
- 31 Jones K, Bullen N. A multi-level analysis of variations in domestic property prices: Southern England, 1980-87. *Urban Studies* 1993;30:1409-26.
- 32 Robertson C. Participation in post-compulsory education in Scotland. *J R Stat Soc* 1993;156:423-42.
- 33 Jones K. Specifying and estimating multi-level models for geographical research. *Transactions of the Institute of British Geographers* 1991;16(2):148-60.
- 34 Phillimore P, Morris D. Discrepant legacies: premature mortality in two industrial towns. *Soc Sci Med* 1991;33(2):139-52.