# Where not to live: a geo-demographic classification of mortality for England and Wales, 1981-2000

Nicola Shelton[1], Mark Birkin[2] and Danny Dorling[3]

[1] Research Fellow, Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, London, WC1E 6BT.

[2] Director of Informatics Institute, School of Computing / School of Geography, University of Leeds, Leeds, LS2 9JT.

[3] Professor of Human Geography, Department of Geography, University of Sheffield, Winter Street, Sheffield, S10 2TN.

Correspondence to: Dr Nicola Shelton Tel: 020 7679 5648, Fax: 020 7813 0280, email: n.shelton@ucl.ac.uk

## Abstract

The aim of this paper is to pilot a method for geo-demographic classification for mortality patterns in Britain. Age and sex directly standardised mortality ratios (DSMRs) for 100 grouped International Classification of Disease series 9 causes of death (ICD-9) were calculated. The 84 European Parliamentary (EP) constituencies as defined in 1999 were used as the spatial basis of this study to allow regional comparisons to be made while comparing units of similar population sizes. Scotland was excluded from the final analysis, leaving 76 regions. This paper is a preliminary investigation of the patterns of the causes of death over time and space in England and Wales using cluster analysis to summarise some of the structure in the data. Seven major and three minor cluster profiles were developed.

Keywords: Geo-demographics; Health geography; DSMRs; ICD-9; Resource allocation

## Introduction

Geo-demographic classification systems, such as ACORN, Mosaic and SuperProfiles, are commonly used within both commercial and public sector organisations as a basis for target marketing and resource allocation (CACI, 2004; Birkin et al., 2002; Harris et al., 2004). Traditionally, these systems classify areas on the basis of census data profiles such as ethnicity and unemployment. It is increasingly common to incorporate attitudinal and behavioural data relating to things like newspaper readership and internet usage, which also allows the classifications to be presented at ever more disaggregate geographical scales (Birkin, 1995). Applications are widespread in both the UK and North America (Claritas, 2004), but less common in Europe and elsewhere despite the existence of a variety of international classification systems (Experian, 2004; EuroDirect, 2004).

The data inputs to the major classification systems have been reviewed by Vickers et al., (2003), who report that most systems ignore health data completely. A minor exception are the most recent ONS classifications, which include self-reported long-term limiting illness (Office for National Statistics (ONS), 2004). Nevertheless geo-demographic profiles have been used to try to explain health and health care utilisation (Openshaw and Blake, 1995; Hedges et al., 1997; Tickle et al., 2000). Whilst area-based geographic inequalities of mortality have been widely analysed and reported (Shaw et al., 2000, 2002; Davey Smith et al., 2001; Mitchell et al., 2000; Gordon et al., 1999) area-based classifications of mortality profiles have not been derived.

Geo-demographic classifications have been developed at a variety of spatial scales. At the most extreme level of disaggregation, individual households can now be classified using systems like PRiZM Household (Geobusiness, 2004). At the other extreme, area-based classifications remain of interest to both researchers and policy-makers. For example, ONS have consistently made available classifications of UK districts (Webber and Craig, 1978) and a similar concept of 'statistical neighbours' is widely used by education authorities for performance benchmarking and target-setting (e.g. Southwark County Council, 2003).

This paper uses a rich time series database of mortality in order to derive such an area-based profile based on cause of death data. The study was co-funded by the Economic and Social Science Research Council and Department of Trade and Industry (Grant No. RES-149-25-0001) as a small part of a larger project 'Health Care Planning with Data Driven Resource Allocation (HYDRA)'. The purpose of the whole project was to demonstrate ways in which the analysis of spatially or temporally rich databases and visualisation of patterns could lead to more effective resource allocation, in this case by investigating which areas had similar, different or changing mortality profiles. Previously, analysis of the factors affecting access to and uptake of health care has been explored by using geographic information systems. The effects of sociodemographic factors and/or distance to health care setting (Gatrell et al., 1998; Hyndman et al., 2000) are often considered the key factors in these analyses, although other factors such as the nature of the setting, e.g. presence of a female GP (Gatrell et al., 1998), could be important. Health data are used less often and background mortality profile is not usually considered as the data are not readily available. This paper seeks to address this omission.

## Method

Geographically referenced cause of death data were obtained for 1981–2000 from the Office of National Statistics and General Registrar Office (Scotland). The data for the early years were made available in two blocks of 5-year and 4-year time periods; to be consistent four periods were used here: 1981–1985, 1986–1990, 1991–1995, 1996–2000. The data provided, for everyone who had died in Britain, their date of death, cause of death, age and sex, and place of residence at death (place of residence was recorded as at an institution, e.g. hospital, if the deceased had lived there for 6 months or more). Date of death was year of registration until 1993 and then year of occurrence. This change had little impact on the analysis as the 1991–1995 period was considered in aggregate.

The International Classification of Diseases (ICD) was designed to promote international comparability in the collection, processing, classification and presentation of mortality statistics. The principal tabulated cause is called the underlying cause of death; this was used here. The other reported causes are the non-underlying causes of death. Deaths reported here were classified using ICD-9, the ninth revision to the classification and in use from 1979 to 2000/1 until it was replaced by ICD-10 (Office for National Statistics (ONS), 2002). There was one change to ICD-9 during the period: the introduction of classifications for HIV infection in 1996. Cause of death was supplied by the statistics Offices coded as ICD-9 except for deaths in Scotland in the year 2000 which were supplied as ICD-10. The ICD-10 data were back coded to ICD9 and all the data were summarised into 100 causes of death by George Davey-Smith (2003);(Dorling, 2005).

Directly standardized mortality ratios (DSMRs) are the directly age/sex standardized ratio of the observed number of deaths in an area to the number of deaths that would be expected if national age-and sex-specific death rates were applied to each area by age and sex. The standardisation means that each cause of death has a mean of 0 and a standard deviation of 1. These were calculated for each cause of death for each of the four time periods and for each of the 84 European parliamentary districts (EPDs) in England, Wales and Scotland (see Table 1 and Fig. 1 for the geography used). The choice of EPDs was for several reasons. They provided a stable geography over the time period. They were sufficient in number to provide comparative analysis for this preliminary feasibility study.

Table 1. European parliamentary district number and name

| **No.** | **Name** | **No.** | **Name** |
|---|---|---|---|
| 1 | London Central | 43 | Essex South |
| 2 | London East | 44 | Essex West & Hertfordshire East |
| 3 | London North | 45 | Gloucestershire |
| 4 | London North East | 46 | Itchen, Test & Avon |
| 5 | London North West | 47 | South Downs West |
| 6 | London South & Surrey East | 48 | Wight & Hampshire South |
| 7 | London South East | 49 | Hertfordshire |
| 8 | London South Inner | 50 | Kent East |
| 9 | London South West | 51 | Kent West |
| 10 | London West | 52 | Lancashire Central |
| 11 | Greater Manchester Central | 53 | Lancashire South |
| 12 | Greater Manchester East | 54 | Leicester |
| 13 | Greater Manchester West | 55 | Lincolnshire |
| 14 | Merseyside East & Wigan | 56 | Norfolk |
| 15 | Merseyside West | 57 | Northamptonshire & Blaby |
| 16 | Sheffield | 58 | Cleveland & Richmond |
| 17 | Yorkshire South | 59 | North Yorkshire |
| 18 | Northumbria | 60 | Nottingham & Leicestershire North West |
| 19 | Tyne & Wear | 61 | Nottinghamshire North & Chesterfield |
| 20 | Birmingham East | 62 | Hampshire North & Oxford |
| 21 | Birmingham West | 63 | Herefordshire & Shropshire |
| 22 | Coventry & North Warwickshire | 64 | Somerset & North Devon |
| 23 | Midlands West | 65 | Staffordshire East & Derby |
| 24 | Leeds | 66 | Staffordshire West & Congleton |
| 25 | Yorkshire South West | 67 | Suffolk & South West Norfolk |
| 26 | Yorkshire West | 68 | Surrey |
| 27 | Bedfordshire & Milton Keynes | 69 | Worcestershire & South Warwickshire |
| 28 | Thames Valley | 70 | Sussex West |
| 29 | Bristol | 71 | Wiltshire North & Bath |
| 30 | Buckinghamshire & Oxfordshire East | 72 | Mid & West Wales |
| 31 | Cambridgeshire | 73 | North Wales |
| 32 | Cheshire East | 74 | South Wales Central |
| 33 | Cheshire West & Wirral | 75 | South Wales East |
| 34 | Cornwall & West Plymouth | 76 | South Wales West |
| 35 | Cumbria & Lancashire North | 77 | Central Scotland |
| 36 | Peak District | 78 | Glasgow |
| 37 | Devon & East Plymouth | 79 | Highlands & Islands |
| 38 | Dorset & East Devon | 80 | Lothian |
| 39 | Durham | 81 | Mid Scotland & Fife |
| 40 | East Yorkshire & North Lincolnshire | 82 | North East Scotland |
| 41 | East Sussex & Kent South | 83 | South of Scotland |
| 42 | Essex North & Suffolk South | 84 | West of Scotland |

They were aggregates of wards, which meant that the data could be compared to other data from parliamentary districts such as wards, using the "Linking Censuses Through Time" LCT software (http://cdu.mimas.ac.uk/lct/ ) in the future. The districts had similar populations to each other for comparison and lastly their names offer a meaningful geography. One of the constraints in choosing any geography for health research was that whereas socio-demographic data were available at parliamentary level, health access information would be available at health areas and the two geographies were not contiguous at any level of dis-aggregation. Additionally patients can choose their GPs and to some extent the hospital by which they are treated within a health area, which further distorts the geography. The analysis here provides a health care access requirement (i.e. mortality inequalities) at a basic, large, parliamentary geography.

The DSMRs were then analysed using k-means cluster analysis in SPSS 11.1. Cluster analysis aims to group together cases given information on variables on those cases where no a priori classification is known. After preliminary cluster analysis the data for Scotland were removed as the clustering pattern for Scotland (and particularly Glasgow) was unlike that of England and Wales (results not shown). The analysis comparing districts in England and Wales with those in Scotland was not meaningful, presumably due to the higher overall rates of mortality in that country, and the undue influence of Glasgow's significantly higher mortality. (The life expectancy at birth in Scotland in 2000 was 2.2 and 1.5 years below that of the UK average for men and women, respectively (Office for National Statistics (ONS), 2003).)

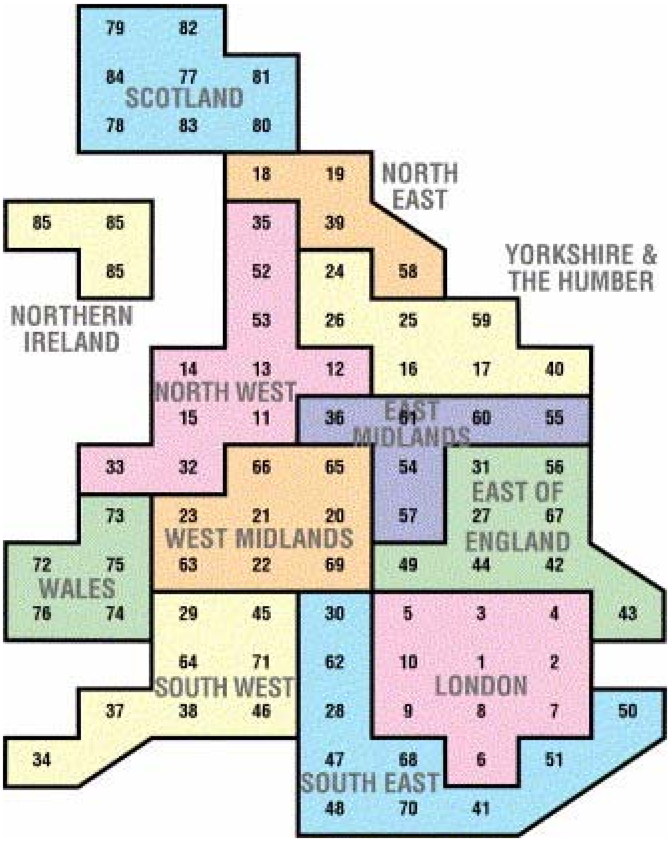

Fig.1.The location of European Parliamentary Districts.

The data for England and Wales then comprised a matrix of 76 districts by four time periods by 100 causes (a total of 3040 time-date-cause points). These results were appended to a single file of 304 temporo-spatial series for 100 causes. Data reduction was then applied to these, using k-means cluster analysis specifying an output with 10 clusters. This procedure attempts to identify relatively homogeneous groups of cases based on selected characteristics, using an algorithm that can handle large numbers of cases. The algorithm requires specification by the user of the number of clusters and so was repeated with 8, 9, 11 and 12 clusters. The most commonly used distance (simple Euclidean - the geometric distance between the objects) was used;

The cluster composition results were then mapped using schematic cartograms. Cartograms are maps that show projections of topographic areas which have been converted into statistical proportions (Dorling, 1996). In this case the projection was based on the population of each area rather than the physical size. The consequence of this is that rural areas seem smaller and urban areas which may be physically small in area, and otherwise invisible on a standard Cartesian chloropleth map, are made visible by being scaled in proportion to their population size.

## Results

Ten clusters were achieved from the 304 time-area points. Specification of a smaller number of clusters led to fragmentation of the smallest clusters and larger numbers led to amalgamation of these, but little change in the largest cluster. Of these ten final clusters, three contained districts from one or two time-area points only. The remaining seven clusters contained between six and 143 time-place points out of a total of 304. Table 3 summarises the broad geography. Four clusters - 1, 5, 6 and 8 - comprised districts in London. Cluster 3 comprised some coastal areas. Clusters 4 and 9 were comprised of former mining areas; cluster 9 also included areas of high deprivation. Cluster 7 comprised a large number of areas in the south east of England and a few northern districts. Cluster 10 comprised urban areas throughout England, but mainly outside of London.

## Significant causes

Table 2 shows the national ranking for each cause of death for 1981 - 2000 for the 20 causes with the highest aggregate mortality over the 20 years. Each cluster corresponds to a set of 100 DSMRs. These were ranked by their DSMRs, and the five causes with highest and lowest DSMRs are shown in Table 3 for each cluster. The ranked (out of 100 causes) contribution of each cause to national rates over the period is also shown.

Table 2. Total number of deaths from the 20 major causes for England and Wales, 1981-2000

| **Cause of death** | **Number of deaths** |
|---|---:|
| Heart attack | 1 654 563 |
| Cerebrovascular disease | 1 273 334 |
| Chronic heart disease | 1 152 483 |
| Pneumonia | 845 870 |
| Lung cancer | 606 481 |
| Chronic lower respiratory diseases | 491 614 |
| Unspecified cancers | 267 887 |
| Other heart disease | 253 713 |
| Breast cancer | 229 124 |
| Other digestive disorders | 214 657 |
| Colon cancer | 197 364 |
| Other circulatory disorders | 172 583 |
| Stomach cancer | 155 924 |

| | |
|---|---|
| Aortic aneurysm | 147 615 |
| Senile and presenile organic psychotic conditions | 143 757 |
| Prostrate cancer | 141 604 |
| Other cancers | 129 212 |
| Diabetes mellitus | 119 234 |
| Pancreatic cancer | 105 744 |
| Lymphatic cancer | 103 453 |

The rarer the cause of death, the more scope there is for extreme spatial variations in its rates, even once they are standardised. Cluster formations seemed to be primarily influenced by relatively minor causes of death. The key thing to note is that although they are not major causes of death, the particular combination of causes varies between clusters. The causes of death that are labelled suicide/accident are where the registrar has classed them as having been undetermined whether accidentally or purposely inflicted. The majority of these, however, are generally regarded to be suicide, whereas an accident is a death from unintentional injury.

The London clusters (1, 5, 6 and 8) had higher than average mortality levels of exogenous causes including accidents, suicides, assaults, drug-and alcohol-related deaths, road traffic incidents and some infectious diseases. They have lower than average mortality for bronchitis and several chronic conditions. The southeastern areas (clusters 7 and 2) also had higher levels of some exogenous causes; however, these differed and included (the very rare) air accidents and excluded infectious diseases and most assaults which were lower than average. These clusters also had below average rates of death from unknown causes. The coastal cluster (3) had unsurprisingly higher than average mortality from water transport accidents (and drowning and exposure - data not shown), but also lower levels of infectious diseases compared to the other clusters. The former mining clusters (4 and 9) had higher than average mortality from industrial lung diseases, and also chronic diseases such as respiratory diseases, and other mortality possibly attributable to occupation such as by machinery, with lowest levels of infectious diseases and some accidents/suicides. The urban (non-London) cluster (10) had high levels of some exogenous causes such as deaths from assaults by cutting (stabbing) relative to the other clusters, but relatively low levels of deaths from skin cancer and air accidents, with higher than average levels of deaths resulting from chronic respiratory diseases. This cause of death had the high contribution to overall mortality in the five causes with most variation in any of the clusters.

### Stability and change over time

The distribution of the clusters is shown in Figs. 2–5 for each of the four time periods respectively; there was little cluster change over time. Just less than 82% of the areas did not change cluster membership during the 20 years covered by this analysis. The relative stability of the individual clusters is illustrated in Fig. 6. The figure illustrates the stability of each cluster, as represented by the numbered columns. Each cell on the chart represents a transition from one 5-year period to the next. When a constituency stays in the same cluster from one time period to the next, then a cell is coloured light grey. When a constituency started in the cluster and ended in another, then a cell is coloured black in the column for the origin cluster. When a constituency started in another cluster, then a cell is coloured dark grey.

The number of transitions is shown to the right of each column. For example, there were five transitions from cluster 3 into another cluster; three transitions involving no change in cluster 3; and seven transitions into cluster 3 from another cluster. On this analysis, clusters 10 and 6 appear to be the most stable; clusters 7 and 9 less so; while clusters 3 and 4 are rather more dynamic. Thus it appears that both the former mining areas (clusters 4 and 9) and the coastal areas (cluster 3) exhibit greater mobility in their mortality profiles than London (cluster 7) and the urban areas (cluster 9).

Table 3 Mortality profile of each cluster for five highest and five lowest DSMRs in each cluster, England and Wales, with ranking out of 100 for overall contribution to national mortality 1981–2000 Cluster 1 - London (1); Cluster 2 - Southeast (1); Cluster 3 - Coastal; Cluster 4 - Former mining areas (1); Cluster 5 - London (2); Cluster 6 - London (3); Cluster 7 - Southeast (2); Cluster 8 - London (4); Cluster 9 - Former mining areas (2); Cluster 10 – Urban England.

| **Lowest DSMRs in cluster** | **Rank** | **Highest DSMRs in cluster** | **Rank** |
|---|---|---|---|
| 1. London (1 time place points) | | | |
| Chronic heart disease | 3 | Assault by firearms | 100 |
| Influenza | 72 | Due to drugs | 69 |
| Other motor vehicle accidents | 40 | HIV disease infections | 80 |
| Senile and presenile organic psychotic conditions | 15 | Hunger, thirst, exposure, neglect | 96 |
| Congenital malformations of nervous system | 62 | Suicide/accident by jumping | 83 |
| Cluster 2. South East (2 time place points) | | | |
| Ill-defined and unknown causes | 74 | Assault by firearms | 100 |
| Signs and symptoms | 92 | Accident - electric current | 95 |
| Epilepsy | 64 | Air accidents | 99 |
| Hepatitis | 89 | Pregnancy and childbirth | 98 |
| Pedestrian and motor vehicle accidents | 54 | Atherosclerosis | 26 |
| 3. Coastal (13 time place points) | | | |
| Chronic lower respiratory diseases | 6 | Water transport accidents | 97 |
| Tuberculosis infections | 72 | Suicide/accident by firearms | 85 |
| Fire | 73 | Caused by machinery | 93 |
| HIV disease infections | 80 | Assault by firearms | 100 |
| Ill-defined and unknown causes | 74 | Signs and symptoms | 92 |
| 4. Former mining (16 time place points) | | | |
| Assault by cutting | 90 | Hunger, thirst, exposure, neglect | 96 |
| Senility without mention of psychosis | 15 | Industrial lung diseases | 66 |
| Other assaults | 86 | Bronchitis | 70 |
| Other intestinal infections | 88 | Falls | 30 |
| Other suicide/accidents | 57 | Caused by machinery | 93 |
| 5 London (2 time place points) | | | |
| Caused by machinery | 93 | HIV disease infections | 80 |
| Parkinson's disease | 37 | Ill-defined and unknown causes | 74 |
| Multiple sclerosis | 65 | Hunger, thirst, exposure, neglect | 96 |
| Suicide/accident by firearms | 85 | Due to drugs | 69 |
| Congenital malformations of nervous system | 62 | Assault by cutting | 90 |
| 6. London (24 time place points) | | | |
| Signs and symptoms | 92 | HIV disease infections | 80 |
| Falls | 30 | Hepatitis | 89 |
| Influenza | 72 | Suicide/accident by jumping | 83 |
| Suicide/accident by gases | 60 | Assault by firearms | 100 |
| Caused by machinery | 93 | Assault by cutting | 90 |
| 7. South-east (143 time place points) | | | |
| Assault by cutting | 90 | Air accidents | 95 |
| Fire | 73 | Suicide/accident by firearms | 85 |
| Other suicide/accidents | 57 | Pedal cycle accidents | 84 |
| Tuberculosis infections | 72 | Other motor vehicle accidents | 40 |
| Ill-defined and unknown causes | 74 | Suicide/accident by gases | 60 |
| 8. London (6 time place points) | | | |
| Congenital malformations of nervous system | 62 | Assault by cutting | 90 |

| | | | |
|---|---|---|---|
| Senile and presenile organic psychotic conditions | 15 | HIV disease infections | 80 |
| Industrial lung diseases | 66 | Ill-defined and unknown causes | 74 |
| Influenza | 71 | Assault by firearms | 100 |
| Bronchitis | 70 | Other suicide/accidents | 5 |
| | | | |
| 9. Former mining (32 time place points) | | | |
| Multiple sclerosis | 65 | Industrial lung diseases | 66 |
| Hepatitis | 89 | Caused by machinery | 93 |
| Other motor vehicle accidents | 40 | Other accidental poisoning | 48 |
| Suicide/accident by gases | 60 | Rheumatic heart disease | 42 |
| Water transport accidents | 97 | Railway accidents | 94 |
| | | | |
| 10. Urban England (65 time place points) | | | |
| Multiple sclerosis | | Pedestrian and motor vehicle accidents | 54 |
| Accident - electric current | | Laryngeal cancer | 63 |
| Air accidents | | Assault by cutting | 90 |
| Suicide/accident by gases | | Hunger, thirst, exposure, neglect | 96 |
| Skin cancer | | Chronic lower respiratory diseases | 6 |

Cluster 1 - London (1); Cluster 2 - Southeast (1); Cluster 3 - Coastal; Cluster 4 - Former mining areas (1); Cluster 5 - London (2); Cluster 6 - London (3); Cluster 7 - Southeast (2); Cluster 8 - London (4); Cluster 9 - Former mining areas (2); Cluster 10 – Urban England.

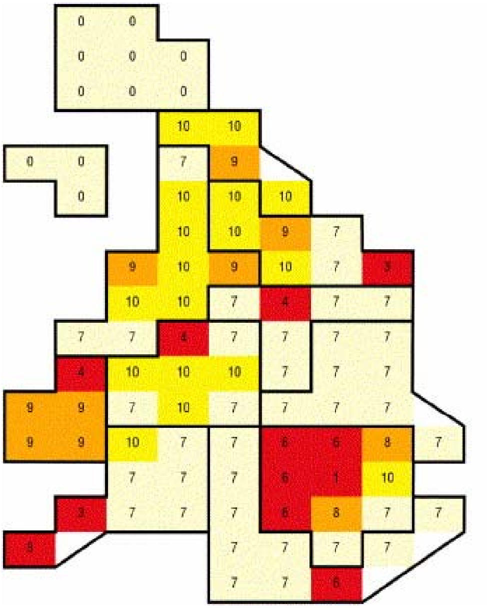

Fig. 2. Cluster location 1981–1985.

The geography of stability is assessed further in Fig. 7. In this cartogram, we have simply differentiated those clusters which ended in the same cluster they began from those which changed clusters between 1981 and 2000. There appear to be some clear spatial patterns in change over time. In particular, the relative stability of the constituencies in the southeast (outside London) is noticeable, whereas there is a relatively distinct band of constituencies in the ‘north midlands’ which have changed cluster, including both Staffordshire East and Staffordshire West, Sheffield, South Yorkshire, East Yorkshire, Cleveland, North Lancashire and Manchester East.

The dynamic transitions between clusters are shown explicitly in Table 4. The cells in the table represent the number of transitions from one cluster to another between quinquennial periods, i.e. between time period 1 (1981–1985) and time period 2 (1986–1990); from time period 2 to time period 3 (1991–1995); and from time period 3 to time period 4 (1996–2000).

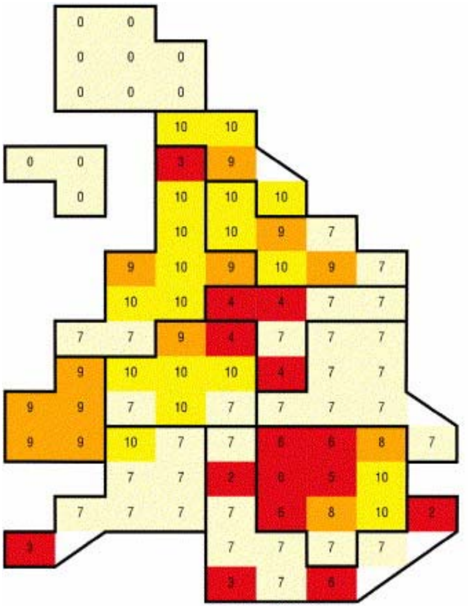

Fig. 3. Cluster location 1986–1990.

Although we have seen that many of the clusters are relatively stable, this does not preclude the possibility of persistent long-term trends which might see the continued expansion of some clusters and the diminution or elimination of others. The stability of the clusters can be explored using a test for the symmetry of the transition matrix (Upton and Fingleton, 1989). The test assumes that the probability of a transition from cluster i to cluster j is identical to the probability of a transition in the other direction. For example, Table 4 shows that there are only two transitions from cluster 10 to cluster 7, and four transitions from cluster 7 to cluster 10. If the clusters are in steady state, then we would expect three transitions in either direction. Once a set of expected probabilities has been derived in this way for each pair of clusters, then a chi-squared statistic can be computed to compare the actual and expected transitions.[1] The data in Table 4 yielded a chi-squared value of 7.76; with 45 degrees of freedom, this implies little or no support for the hypothesis that there is a significant pattern to the transitions. This suggests that the pattern of mortality clusters in Fig. 5 would be reasonably stable in the foreseeable future.[2]

---

[1] The formal specification for this test is therefore eij = eji = (fij + fji ) /2, where fij is the actual number of transitions from cluster i to cluster j, and eij is the expected number of transitions from cluster i to cluster j. Chi-squared is calculated in the usual way: $\chi^2 = \sum_{ij} (f_{ij} - e_{ij})^2 / e_{ij}$

[2] The test statistic can also be calculated for the cluster transition matrices between each pair of time periods. The associated chi-squared values are shown below. In each case, stability of the transition matrix over time is indicated.

| First time period | Second time period | Chi-squared |
|---|---|---|
| 1981–1985 | 1986–1990 | 10.0 |
| 1981–1985 | 1991–1995 | 14.0 |
| 1981–1985 | 1996–2000 | 8.0 |
| 1986–1990 | 1991–1995 | 11.3 |
| 1986–1990 | 1996–2000 | 10.3 |
| 1991–1995 | 1996–2000 | 15.3 |

An alternative perspective on this question would be to ask what might happen to the distribution of clusters if we assume that the pattern of transitions from 1980 to 2000 is in fact not random but systematic. This is explored in Table 5, which shows how we expect clusters to change at 15-year intervals over the next century. A transition matrix was calculated from the first time period (1980–1985) to the last time period (1995–2000). There is a 15-year spread between these periods.

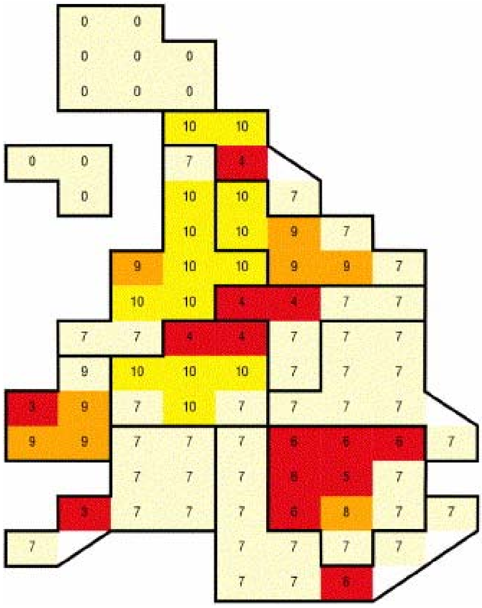

Fig. 4. Cluster location 1991–1995.

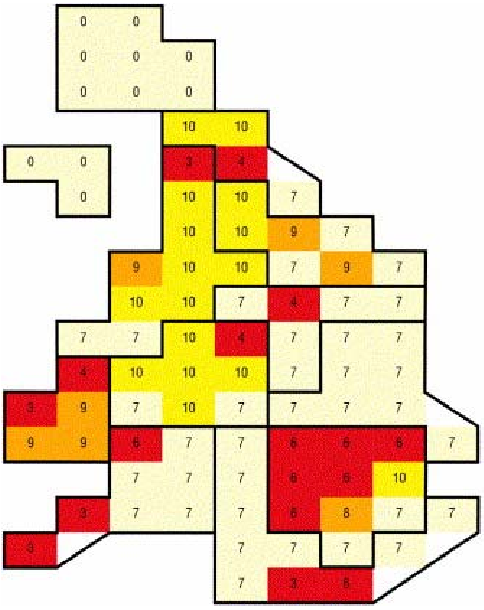

Fig. 5. Cluster location 1995–2000.

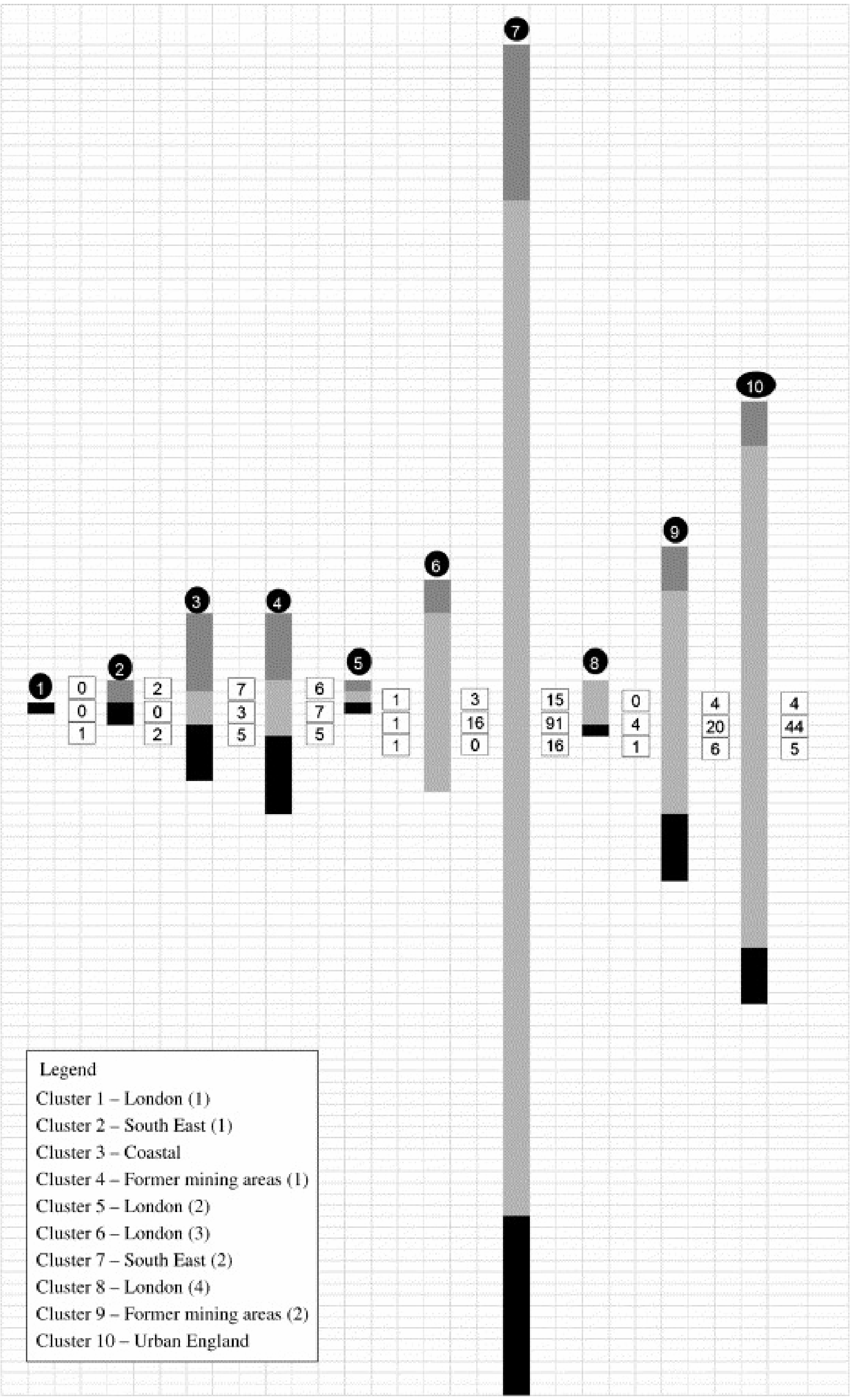

Fig. 6. Cluster stability.

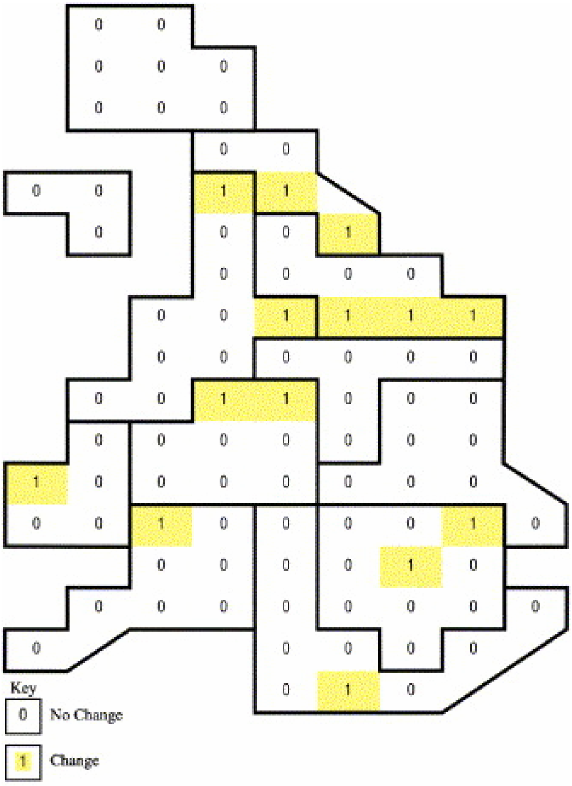

Fig. 7. Summary Cluster Change, 1981–2000.

Table 4 Transitions between clusters

| From | To 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 |
| 3 | 0 | 0 | 3 | 0 | 0 | 0 | 6 | 0 | 1 | 0 |
| 4 | 0 | 0 | 0 | 7 | 0 | 0 | 4 | 0 | 2 | 0 |
| 5 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 1 | 16 | 1 | 1 | 0 | 0 |
| 7 | 0 | 2 | 5 | 2 | 0 | 0 | 91 | 0 | 2 | 4 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 |
| 9 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 20 | 1 |
| 10 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 1 | 44 |

These transitions were then extrapolated into the future at 15-year intervals. Clearly the table shows a rather slow change, which is consistent with our findings regarding the stability of the clusters. However, there is also a distinct trend away from clusters 9 and 10, and towards cluster 6.

## Discussion

This paper has shown that it is possible to derive an area-based mortality profile using cause of death data. Area profiles tended to have high stability over time. Using DSMRs to characterise areas led to exogenous causes influencing mortality profiles especially in urban areas. This was due to the relatively small number of cases involved (still tens of thousands), but also the influence of lifestyle and circumstance on several causes such as drug-related deaths and skin cancer. Occupational mortality was still important in former mining areas.

Roughly half the population of England and Wales lived in areas with very similar mortality profiles, when defined using these methods. These are the cluster 7 areas on the maps of the four periods shown. This group lived generally in the south of England excluding London in the more affluent rural areas of Britain. In these areas people were more likely on average to die in certain types of suicides or accidents than people in other areas. The accident types were air, pedal cycle and other motor vehicle. These accidents may be a result of affluence through holiday/recreational activities and partly location or circumstance (most road deaths in Britain occur outside built-up areas on single carriageway roads; however, many higher risk roads are in the north and west as well as the southeast (EuroRAP, 2004)). Deaths classed as suicide or accidents are usually included in suicide statistics. Suicide/accident by gases and suicide by use of firearms were more common in cluster 7 than in other areas. Suicide by gases is often from car exhaust fumes from a cold start and hence requires access to a car and usually a garage (and access to a car and garage is still the privilege of the more affluent). As much as one third to one half of the deaths classed as suicide/accidents from firearms are from shot guns (Dept. of Treasury, 1996). Shotgun licenses are available to anyone of any age without a criminal record. Those involved in shooting sports and farmers would presumably have more access to shotguns. Handguns are now generally outlawed in Britain and the distribution of shotguns - as seen through the geography of murders committed by them - largely reflects the more affluent areas of the country (Shaw et al., 2004).

One of the questions raised by the approach taken and results of the analysis is why do variations in deaths from causes with low contribution to overall national mortality rates matter? One argument is that they matter because many of these deaths could be avoidable; the Government has set targets for the reduction of accidental deaths. The white paper Our Healthier Nation contains a target ‘to reduce death rates from accidents by at least one fifth by 2010’ (DoH, 1999). Additionally the geography of accidental death could be useful in resource allocation for the location and nature (e.g. burns units) of accident and emergency facilities. The geography of some individual causes was extreme, but remains important. The results of our work show that had we classified causes together before clustering, as was suggested to us, the assumptions we would have made doing this could have been erroneous. For example, air and water and all other accidents might have been grouped together. This would have lost variations that are essentially nothing to do with epidemiology and everything to do with geography!

Table 5
Predicted future cluster transitions

| **Cluster** | **1980–85** | **1995–2000** | **2010–2015** | **2025–2030** | **2040–2045** | **2055–2060** | **2070–2075** | **2085–2090** |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 3 | 5 | 6 | 7 | 7 | 7 | 7 | 7 |
| 4 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 5 | 8 | 9 | 11 | 12 | 12 | 14 | 15 |
| 7 | 37 | 36 | 36 | 36 | 36 | 36 | 36 | 36 |
| 8 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 9 | 8 | 6 | 5 | 4 | 3 | 3 | 3 | 3 |
| 10 | 17 | 16 | 15 | 15 | 14 | 13 |  |  |

The other noticeable feature was the change in composition of the former mining clusters. Planning of future health service allocation will need to take into account occupational causes of death which are currently higher in former mining clusters and may be replaced by ones that are currently lower than average there, or possibly higher than average in other similar areas that are non-mining. The postindustrial economy in former mining districts will presumably influence whether they end up in the urban cluster or the coastal and affluent ones. Most geographies of health tend to describe and explain inequalities between more deprived areas and more affluent areas. This paper highlights the homogeneity in the health geography of these affluent areas and the more heterogeneous nature of the mortality patterns of the remainder.

The transitions observed indicated that the mortality patterns of northern (urban) areas were likely to increasingly resemble those of London in the future. The long-term steady state is for the whole country to adopt the profile of cluster 6, because there are transitions into cluster 6, and none away from it. Since this pattern would take several hundred years to establish itself, it is probably safe to assume that further changes in the transition pattern will actually occur before then. Prediction of future health geographies is inevitably fraught with speculation, and deaths and survival rates from certain causes will inevitably change over time. Whether the mechanisms that created the underlying geographical inequalities shown here change in the same way remains to be seen.

Future research using these data could involve aggregating groups of causes to reduce the effects of potentially random variation in causes with small numbers of deaths. To avoid loss of variation between, for example, water, air and road traffic accidents being lost, any aggregation would need to be done taking into account both the cause of death and its geography. More work is required to understand why the mortality profiles of Scotland are so different to those seen in England and Wales. It would also be an interesting extension to repeat the analysis using smaller areas such as wards in a future study. The size of districts used here ensured that we were dealing with larger populations at risk; when considering rare causes of death, large areas were the most robust to use. It would have been interesting to repeat this analysis using health areas, to investigate the role of health care access and uptake, but this was beyond the scope of this pilot study.

Although we have demonstrated that the pattern of mortality groupings has not yet reached an equilibrium in its spatial distribution, the relative consistency of these patterns over time is nevertheless striking. There is an underlying geography to the variations of cause of death in Britain which has been little studied since Melyvn Howe produced his seminal atlas in the 1960s (updated at the end of that decade; Howe, 1970), despite the production of an atlas for the Victorian era by Woods and Shelton in 1997. Given that 20 years of data of late twentieth century mortality can now be modelled on a PC, rather than through the laborious processes that Howe was forced to employ almost four decades ago, it is perhaps time we health geographers took another look at the medical geography of our countries.

## Acknowledgements

The authors wish to thank Professor George Davey-Smith for kindly supplying us with the suggestions for the cause of death aggregations used and Graham Allsopp for the cartograms.

## References

Birkin, M., 1995. Customer targeting, geodemographics and lifestyles approaches. In: Longley, P., Clarke, G. (Eds), GIS for Business and Service Planning. Longman, London, Chapter 6

Birkin, M., Clarke, G., Clarke, M., 2002. Retail Geography and Intelligent Network Planning. Wiley, Chichester.

CACI (2004) Smart solutions for intelligent marketing and information systems, CACI, London. (http://www.caci.co.uk)

Claritas (2004) Adding intelligence to information, Claritas, San Diego. (http://www.claritas.com)

Davey Smith, G., 2003. Per. comm.

Davey Smith, G., Dorling, D., Shaw, M., (eds) 2001. Poverty, inequality and health: 1800-2000 - a reader. The Policy Press, Bristol.

DoH, 1999. Saving Lives: Our Healthier Nation Chapter 7 Saving lives: accidents The Stationery Office, London.

Dorling, D (1996) Area Cartograms: Their Use and Creation, Concepts and Techniques in Modern Geography (CATMOG), 59

Eurodirect (2004) Cameo: Painting a clear picture, Eurodirect, Bradford.

EuroRAP, 2004. European Road Assessment Programme Road Ratings (http://www.eurorap.org/roadratings/risk_rate_maps.htm).

Experian (2004) Mosaic UK, Experian Ltd, Nottingham. (http://www.experian.co.uk/business/products/data/113.html)

Gatrell, A., Garnett, S., Rigby, J., Maddocks, A., Kirwan, M., 1998. Uptake of screening for breast cancer in South Lancashire. Public Health, 112, 297-301.

Geobusiness (2004) PRiZM Household, Geobusiness Solutions, Wendover. (http://www.geobusiness.co.uk/Products/Data/Lifestyle&Geodem )

Gordon, D., Davey Smith, G., Dorling, D., Shaw, M., 1999. Inequalities in Health: the evidence. The Policy Press, Bristol.

Harris, R., Sleight, P., Webber, R. (2004) Geodemographics, GIS and Neighbourhood Targeting. Mastering GIS: Technology, Applications & Management, Wiley, Harlow.

Hedges, B., di Salvo, P., Purdon, S., 1997. Health variations by 'ACORN' area classifications Chapter 9 in Prescott-Clarke, P., Primatesta, P. The Health Survey for England 1996. The Stationery Office, London.

Howe, G. M., 1970. National atlas of disease mortality in the United Kingdom. Revised and enlarged edition. Nelson, London.

Hyndman, J., Holman, C., Dawes, V., 2000. Effect of distance and social disadvantage on the response to invitations to attend mammography screenings. Journal of Medical Screening, 7,141-145.

Mitchell, R., Dorling, D., Shaw, M., 2000. Inequalities in life and death: what if Britain were more equal? The Policy Press, Bristol.

Treasury Department, 1996. Written Answers to Questions Suicides (http://www.parliament.the-stationery-office.co.uk/pa/cm199596/cmhansrd/vo960509/text/60509w03.htm)

Office for National Statistics, 2002. ICD-10 for mortality (http://www.statistics.gov.uk/about/classifications/icd10/default.asp)

Office for National Statistics, 2003. Key demographic and health indicators, 1976 onwards: Health Statistics Quarterly 20. (http://www.statistics.gov.uk/statbase/Product.asp?vlnk=6725)

Office for National Statistics (2004) National Statistics 2001 Area Classification (http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=4804)

Openshaw, S., Blake, M., 1995. Geodemographic segmentation systems for screening health data, Journal of Epidemiology and Community Health, 49, S34-S38.

Shaw, M., Dorling, D., Gordon, D., Davey Smith, G. 2000. The Widening Gap: health inequalities and policy in Britain, 2nd edition. The Policy Press, Bristol.

Shaw, M., Dorling, D., Mitchell, R., 2002. Health, Place and Society. Pearson Education, Harlow.

Tickle, M., Blinkhorn, A., Brown, P. 2000. A geodemographic analysis of the Denplan patient population in the North West Region. British Dental Journal 189, 9, 494-499.

Vickers, D., Rees, P., Birkin, M. (2003) A new classification of UK local authorities using 2001 Census Key Statistics, Working Paper 03/03, School of Geography, University of Leeds.

Webber, R., Craig, J. (1978) A socio-economic classification of local authorities in Great Britain, London, HMSO.

Woods, R., Shelton, N., 1997. An Atlas of Victorian Mortality. Liverpool University Press, Liverpool.