

Residential area characteristics and health of the population: a census based analysis

Said Shahtahmasebi

School of Mathematics and Statistics &
Centre for Applied Research & Support (C.A.R.S)
CPIT
Christchurch
Email: said@cpit.ac.nz

**Presented at SIRC 2001 – The 13th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
December 2nd-5th 2001**

ABSTRACT

Past work suggest an increasing disparity in mortality and morbidity outcomes, health care service provision and uptake between the various population groups. To aid developing and targeting health care service provision it may be more useful to the provider to have an understanding of the population groups' area of residence. This paper is based on a public health report for the Yorkshire Region in the UK which combines routine health data together with the census, GIS and other data sources to construct and map an epidemiological profile of the Region by area of residence.

Acknowledgment: This paper is based on a Public Health Annual Report (Bensley et al 1994) which was produced by the Analysis Unit at the former Yorkshire Regional Health Authority, Harrogate, U.K.

1. INTRODUCTION

Quality information is required for two main purposes; i) to understand the process in order to design appropriate policies that may help shift health outcomes towards a desired target, and ii) to monitor and measure the magnitude and direction of the shifts in health outcomes. That is, as well as the medical diagnosis of a condition, it may be important to understand the extent to which other processes, such as social, environmental and individual characteristics influence mental health. These influences may either have a direct effect on the process outcome, i.e. a mental illness condition, or an indirect effect through affecting the quality of life.

One of the major difficulties in epidemiological studies of health in a social context is the paucity of social data. Even where improvements and modifications to data gathering and intelligence have been a high priority, data is often incomplete with a high rate of missing information (e.g. see Middle and Macfarlane 1995). Such limitations restrict the use of existing National Health Service (NHS) data. Most social epidemiology of ill health relies on additional survey data (e.g. see Meltzer, Gill et al 1996).

Lack of a unified data base means that service prevalence rates may well be under-estimated. The fact that current record keeping systems are not linked will have an affect on diagnosis, intervention and prevention. There is, for instance, ample evidence to suggest that some mental illness conditions are under-diagnosed and under-treated e.g. depression in the elderly population living in the community (Wattis 1995), post natal depression (Oates 1995, Warner et al 1996) which if untreated will increase the risk of more complex mental illness with social and domestic consequences (Godfroid and Charlot 1996, Cooper and Murray1995). Such cases increase the cost of care provision by an unknown factor, as the health care system may need to react to "crisis" situation which may also include the social network of the subjects (e.g. Caldock 1990).

However, in the absence of a unified linked data base it is possible to use a number of data sources in combination to form an epidemiological profile of the service users e.g. Regional mortality/morbidity records together with data from the 1991 census. The idea is that by gaining additional insight into characteristics of its service users, the provider will be able to develop flexible policies for the short term needs while being responsive to the long term requirements.

Furthermore, such insight should allow the inclusion of the public and its health needs, quality of life, development and targeting of services into the process of policy formulation.

2. AIM AND OBJECTIVES

This paper is intended to demonstrate the utilisation of routinely collected health and other data sources and is based on a public health annual report (Bensley et al 1994). The aims of the annual report were:

- To respond to the health targets set by the then UK government taking into account socio-economic factors.
- To better understand health beyond that explained by the boundaries of the District Health Authorities.
- Provision of an aid to service development and provision.

3. METHODOLOGY

Health service data are held anonymously and contain minimum information on individuals. Apart from hospital activity details (e.g details of hospital spells, episodes), causes of death, residential address (post code), age and sex very little else is collected. There is provision to collect data on social class and ethnicity, however, an analysis of these variables suggested that they had been treated as optional and often had not been filled. There is ample evidence about geographical variation in health outcomes (e.g see Townsend & Davison 1981). The health records do provide an opportunity to study health outcomes beyond the better understood age and sex effect through the postal address.

To link the health outcome of the population to their place of residence we require a method of linking health data to the census. Traditionally, a small number of census variables such as house and car ownership, employment status have been used to derive an index of deprivation e.g the Jarman index (Jarman 1983) or Townsend index (Townsend 1986) to be used as part of a formula to target health services. These indices have been argued to be single dimension sliding scale being inappropriate to the multidimensional health and deprivation link. Others, tried utilising other indices based on a large number of census variables such as Acorn (CACI 1991) which were produced for commercial purposes e.g marketing. Unlike the deprivation indices the commercial indices are based on a principal component analysis of a large number of census variables to generate relatively homogeneous clusters of the population or areas of residence. However, Openshaw (1983) argued that such indices are inappropriate and too general to be used when addressing different health and social issues and proposed a similar method to generate what he termed “super profiles” based on small area statistics. These super profiles may be tailor made relevant to the issues being studied. Unlike other deprivation indices, which are based on a handful of census variables, the super profiles are derived from a large number of census variables. Using the smallest enumeration areas referred to as enumeration districts (EDs) as building blocks to characterise residential areas within given boundaries such as health authorities. For all intents and purposes, EDs are assumed to be homogeneous. The super profiles based on Openshaw’s methodology became commercially available in the early 90s (see Brown and Batey 1994a &b) in which other data sources on economic activities were also used to further enhance and validate the clustering. It was decided to use the super profile for a number of reasons:

- Super profiles is one of the major commercially available census based index which exclude health variables.
- Through a national deal with the suppliers, super profiles became available at a third of its cost.
- Convenience!

A summary of the super profile classifications is provided in the appendix. Briefly, super profiles provide clustering groups at three levels of details: at the highest level of detail there are 160 clusters characterising geographical areas, there are 40 and 10 at the lower levels. Clearly, it is more convenient and much more manageable to work at the lower level with 10 groups than 160 or even 40 groups. The 10 groups still retain their detailed definition, yet, although they classify areas in terms of affluence/non-affluence, super profiles is not a sliding scale of deprivation.

Super profiles are then matched with the health records and used as a proxy for socio-economic background in a number of ways. One of the advantages of linking the health data to the census was that we could map mortality/morbidity by their geographical boundaries, now we can represent additional information about the area’s socio-economic background on the same map. In addition, comparison of health outcomes could be made between and within areas using super profiles. Furthermore, super profiles could be used as a proxy variable for socio-economic at individual level for modelling purposes. For this paper only the first two analysis were carried out

although two subsequent Public Health Reports (Shahtahmasebi 1997 & 1999) report all three analyses. The maps are, however, complemented with an acetate overlay of towns and roads map to aid interpretation of choropleth map allowing identification of specific areas of the maps, and are based on AA data.

Source of population and boundary data.

The population data used in this review consist of 'Small Area Statistics' (SAS) data from the 1991 Census. SAS data were used at the basic (and smallest) geographical unit employed for collection of Census data, namely the enumeration district (ED). Analysis of Census data is not permitted at an **'individual'** level, except for the anonymised records from the OPCS Longitudinal Study, because of Census confidentiality. The size of an ED is such that it provides a reasonable workload for a single Census enumerator and normally contains between 175 and 225 households, although there can be fewer in rural areas.

Sources of mortality data

The mortality data used in this report are based on the years 1987 to 1991 from the OPCS mortality data tapes. The Regional age sex specific death rates are based on unamended underlying cause. These were then standardised by age and sex. England and Wales populations and mortality were used for the same period as the reference group. The SMRs and confidence intervals by super profiles are represented with a bar. It should be noted that the dotted horizontal line at 100 does not represent the Super Profile specific SMR but represents the SMR for all of England and Wales, for Super Profile groups. This allows all SMRs to be presented on the same scale.

4. RESULTS

The map "Yorkshire Regional Health Authority" shows the distribution of the super profiles for the whole of Yorkshire. It can readily be seen that the more deprived areas appear to be associated with the inner cities and more densely populated areas. On the other hand, as shown in Table 1, super profiles with highest average ED population are "nest builders" and "producers". However these EDs tend to be larger in area as opposed to the inner cities EDs. It is also noticeable that the "country life" stands out. These EDs are mostly farmland and country locations which consists of larger than average EDs and lower than average ED population.

The map of Leeds is included to demonstrate the distribution of super profiles over one Health Authority. At this level, these maps provide a visualisation of the socio-economic make up of the health authority areas. One way of improving this visualisation in terms of the spread of a health outcome (morbidity or mortality) is to create a density overlay map. One such application is demonstrated for mental health service uptake within the Leeds Health Authority.

The fact that super profiles provide additional information about the population and its characteristics it should help to distinguish differing issues for the same health problem. For example areas with higher concentration of ethnic minorities demand attention to develop more relevant services. However, the differentials in health needs and outcome of ethnic minorities should also be noted. For example, it has been suggested that the incidence and prevalence of breast cancer is lower in women of ethnic minorities than in their white counterpart. This may explain, at least in part, the pattern shown in figure 4.9 as the super profile "urban venturers" contains a high proportion of women from ethnic minorities and has the lowest SMR for breast cancer.

The remainder of the paper is concerned with the reporting of the variations in mortality rate between super profiles. is reported. Mortality was based on the main causes identified by the government as priorities for which targets had been set (Department of Health 1992). These were

Coronary Heart Disease (ICD 410-414)

- i) All persons aged under 65
- ii) All persons aged 65-74

Stroke (ICD 430-438)

- i) All persons aged under 65
- ii) All persons aged 65-74

Cancer of the Female Breast (ICD 174)

- i) All females aged 50-64

Cancer of the Cervix (ICD 180)

- i) All females of all ages

Cancer of the Lung (ICD 162)

- i) All females aged under 75
- ii) All males aged under 75

Smoking Attributable Causes

This 'cause' is defined as a composite of seven separate diseases, using a method described by Peto et al(4). This method attributes different proportions of male and female deaths, due to these seven diseases, to the effects of smoking tobacco (Table 4.1).

Suicides and Undetermined Injury (ICD E950-E959, E980-E989)

- i) All persons

Accidents (ICD E800-E949)

- i) All persons aged under 15
- ii) All persons aged 15-24
- iii) All persons aged 65 and over

The SMRs along with their 95% confidence intervals are plotted for all the super profiles for each of the above causes (figures 4.1 – 4.23). It can be seen that in general mortality increases with deprivation as described by the affluence ranking of the super profiles. There are a couple of exception to this trend.

- 1- The SMR for breast cancer, figure 4.9, for the whole of Yorkshire appears to be below the national average and no statistically significant differences between the groups are apparent. This counter intuitive result prompted a second analysis of cancer registrations by super profiles. These analysis showed an even more interesting patterns: the standardised registration rate appear to decrease with deprivation which is a reversed trend. The question that is raised is that what is special about the affluent groups the reverse a high incidence rate to a low mortality rate and why such a reversal had not occurred for their counterparts in the more deprived areas? One explanation could be that affluence and education may make health services more accessible to those in higher affluence groups than their counterparts in other groups leading to the early detection and treatment of cancer.
- 2- The pattern for accidents mortality, figures 4.21 & 4.22, is slightly different in that it appears more erratic. It is possible that there are similarities between “affluent achievers”, “settled suburbans”, “nest builders”, “urban venturers” as the low risk groups and the rest as the higher risk groups. Similarly, it is also possible that the observed pattern may well be due to the compound effects of other variables not accounted for such as type and nature of accidents.

However, it is also worth noting that although the standardised prevalence of long standing illness (OPCS 1993) on the whole appear to increase with deprivation, figure 4.23, the super profiles “senior citizens” and “hard-pressed families” defy the affluence ranking and have rates below that of the national average. This proved somewhat a puzzle as check and rechecking of data and calculations suggested no error – it therefore require further investigations.

CONCLUDING COMMENTS

The aim of this paper was to demonstrate the utilisation of routinely collected data beyond the traditional distribution of health outcomes over demographic variables. All the data used in this paper are routinely collected and are readily available. The only add-on costs were the cost of super profiles which worked out about 1000 pounds (GB) paid for by the Region. As part of the deal the super profiles became freely available to all the health districts and all health service staff. The other cost were related to the purchasing of the AA digital road maps, additional digitised boundaries (ED-Line) and printing of the final report.

The implications of such a study and approach are in its flexibility of use. At the simple mapping level it provides an idea of socio-economic make up of the residential area under study. Initially, such an aid would enable the service provider to assess or compare policy with practice. Subsequently, the service provider should be able to develop, reassess and re-evaluate strategy/policy and targets.

At a slightly more complex level the approach can be used to monitor health outcomes whilst accounting for area of residence and socio-economic background. Over a long periods e.g. between census it is possible to compare changes in super profiles of areas. This comparison may provide insight into the effects of health, social, environmental and housing policies that had been implemented 10-20 years earlier on areas of residence and health outcomes.

At a more traditional research level, it raises a number of questions about the dynamics of the health process such as the conflicting trends observed about cancer outcomes. Furthermore, it can be used to assist with the design of studies to investigate health service research issues. The approach has already been used in assisting with sampling, surveys, case-control studies in teenage pregnancy, childbirth and analysis of health service uptake.

It is also possible to utilise Super Profiles in an action research type health care delivery programme where appropriate services are developed and targeted taking into account the area ethnic mix, socio-economic while monitoring service uptake and health outcomes.

Like any other approach there are limitations associated with the super profiles too. Post codes are needed to extract an associated super profile which give rise to problems associated with using post codes. There are two problems. The first is that the post code directory has a mis-match rate of around 2%. The second is the incompleteness of post code data when using the hospital records and mortality data. However, the latter is an issue of data collection and recording. Another problem is that the population estimates at ED level are only available at the time of the census. Furthermore, because special EDs (non-classified and non-residential areas e.g shipping) are without an assignable super profile such population and deaths are excluded from the analysis. Although, a comparison of age-specific death rates using Regional and ED data suggested a very small difference in rates and SMRs (see Bensley et al 1994).

Application of Super Profiles to New Zealand: the author's information at the time of the writing of this manuscript, suggested that no super profiles or such like indices were in use in New Zealand. There are however, indices similar to Jarman and Townsend such as NZDep96 are in common use. Due to the differences in social and geographical structure between New Zealand and the UK, there may be a problems associated with the use of similar census variables to construct an index of deprivation. For example house and car ownership may not necessarily imply higher affluence. One of the advantages of using super profiles will be to disentangle and establish a theoretical framework for the use of the census variables.

One of the attractions working with the New Zealand census is that it is more frequent. As opposed to the UK where census are every 10 years, in New Zealand they are collected every five years which may make it easier to track and observe change.

However, one of the areas of intense activity is likely to be the matching of a census based super profile to Area Units, post codes and other data sets such as the health service records in New Zealand. Like any other study of this kind there will be limitations. This area of study may well lead to the identification of some of them.

REFERENCES

AA Automap Data. Adept Scientific Micro System Ltd. 6 Business Centre West, Avenue 1, Letchworth, Herts.

- Bensley D, Shahtahmasebi S, Merrick D et al (1994). A census based view of the population and its health - statistical review. Harrogate, Yorkshire Regional Health Authority.
- Bhopal R (1997) "Is research into ethnicity and health racist, unsound, or important science?" *BMJ* 314: 1751-6.
- Brown B and Batey P (1994). "Design and construction of a geodemographic targeting system: Super Profiles." Liverpool: Credit and Data Marketing Services.
- Brown, P. & Batey, P. (1994). "The development of the 1991 based Super Profiles geodemographic classification system". URPERRL Working Paper 40. Department of Civic Design, University of Liverpool.
- CACI 1991. *Acorn Classification*. London, CACI.
- Caldock, K. (1990). *Domiciliary Services and dependency - A Meaningful relationship?*. CSPRD, University of Wales, Bangor.
- Cooper, P. J., & Murray, L. (1995). Course and recurrence of postnatal depression - Evidence for the specificity of the diagnosis concept. *British Journal of Psychiatry*, 166(2), 191-5.
- Department of Health (1992) *The Health of the Nation: a strategy for health in England*. London, HMSO.
- ED-LINE digitised boundaries*. M R Data Graphics, Greenford House, 309 Ruislip Road East, Greenford, UB6 9BL.
- Godfroid, I.O., & Charlot, A. (1996). Postpartum psychiatry. *Rev. Med. Brux.*, 17(1), 22-23, 6.
- Jarman, B. (1983) "Identification of underprivileged areas". *BMJ*, 286, pp1705-9.
- Meltzer H, Gill B, Petticrew M, et al. (1996). The prevalence of psychiatric morbidity among adults living in institutions, Office of National Statistics.
- Middle, C. and Macfarlane, A. (1995). "Recorded Delivery". *Health Service Journal* 31 August: 27.
- Oates, M. (1995). Risk and Childbirth in Psychiatry. *Advances in Psychiatric treatment*, 1, 146-153.
- OPCS (1992). *1991 Census Definitions Gt. Britain*. London HMSO.
- OPCS (1993). *Population Trends*. Issue 73. Autumn 1993. London: HMSO.
- OPCS (1993). *1991 Census. Limiting Long-term Illness*. Great Britain. London: HMSO.
- Openshaw, S. (1983). *Multivariate analysis of census data. the classification of areas*. In Rhind, D. (ed). A Census User's Handbook. Methuen, London.
- Peto R et al (1992). Mortality from tobacco in developed countries; indirect estimation from national vital statistics. *Lancet*; 339:1268-78.
- Shahtahmasebi, S. (1997) "An epidemiological profile of mental health service users in Leeds." *Internal Report*, Leeds Community and Mental Health Services Trust, Meanwood Park Hospital, the Mansion, Tongue Lane, Leeds.
- Shahtahmasebi, S. (1999) "An Epidemiological Profile of Service Uptake in Salford." *Internal Report*, on behalf of the PHRRC, University of Salford for the Salford health authority, Salford, UK.
- Townsend, P. & Dvidson, N. (1981) Inequalities in health- the Black Report. The Chaucer Press, Suffolk, UK.

Townsend, P., Philimore, P. & Beatie, A. (1988) Health and Deprivation. Inequality and the North. London: Croom Helm.

Wattis J P (1995). Difficulties in diagnosis of depression in the elderly. Spectrum International, XXXV, pp6-7

Werner, R., Appleby, L., Whitton, A. et al. (1996). Demographic and obstetric risks factors for postnatal psychiatric morbidity. *British Journal of Psychiatry*, 168(5), 607-11.

APPENDIX

Super Profile Groups

The Super Profiles can easily be utilised with the individuals' Post Codes. The names attached to the different Super Profiles are an attempt to capture the wider characteristics of the groups in a name that can be easily referred to.

The descriptions of the Super Profile groups that follow are based on a comparison of Census variables between particular profile groups and Great Britain as a whole. These have to be treated with some degree of caution since the Great Britain figures include diverse areas where characteristics may have different connotations. For example, a low proportion of car ownership in Greater London may not imply the same level of deprivation as a low proportion of car ownership in other parts of the country, because of the extremely well developed public transport that exists in London. In Scotland there is a greater proportion of flat dwellers and the particular rural characteristics associated with the Highlands may not apply in other areas of Great Britain.

Group I - 'Affluent Achievers'

Very high income professionals in exclusive areas
Mature families with large detached properties in 'stockbroker belts'
Mature families in select suburban properties

'Affluent Achievers' are principally middle aged couples with older children, many of whom are at university or college. They generally live in large, well appointed, detached homes, which they mostly own themselves, are well qualified and have managerial or professional occupations.

Group II - 'Country Life'

Prosperous and farming communities
Small holders and rural workers (mainly Scotland)

Mostly young and middle aged adults, these people generally live in detached residences, many working where they live. This points to a high percentage of farmers. There is also high second home ownership in this group. The predominantly male workforce is mostly made up of manual workers, many being self employed.

Group III - 'Thriving Greys'

High income households in genteel neighbourhoods
Affluent ageing couples, many in purchase property
Older professionals in retirement area
Comfortably well off older owner occupiers
Affluent ageing couples in rural areas

'Thriving Greys' are predominantly older middle aged residents, whose children have left home. They mostly own their own houses, which tend to be semi-detached or detached. Being well qualified, they generally have professional or managerial jobs.

Group IV - 'Settled Suburbans'

White collar families in owner occupied suburban semis
Mature white collar couples in established suburban semis
White collar couples in mixed suburban housing

In this group there is a high proportion of middle aged married couples with children. Most own their own semi-detached houses, many having white collar occupations. People within this group maintain a fairly high standard of living. There are many housewives, who also work part time.

Group V - 'Nest Builders'

Mortgaged commuting professionals with children in detached properties
Double income young families in select properties
Military families
Young white collar families in small semis and terraces
Young white collar families in smaller semis
Young blue and white collar families in semis and terraces
Young families in terraces, mainly council

This group largely consists of young families, many of whom have mortgages on their semis and terraces. Mostly they are employed in white collar or skilled manual occupations.

Group VI - 'Producers'

Older blue collar owner occupiers in semis
Older workers established in semis and terraces
Older and retired blue collar workers in small council properties

These are mainly older residents living in council owned semi-detached or terrace houses. They chiefly have relatively few qualifications and so are largely employed in blue collar professions.

Group VII - 'Senior Citizens'

Retired white collar workers in owner occupied flats
Older residents and young transient singles many in seaside towns
Old and young buying terraces and flats
Retired blue collar workers in council flats, mainly Scotland

The 'Senior Citizens' are primarily older, retired couples and single old ladies living alone. The majority live in flats and terraced houses. Relatively few own cars, mostly getting around on foot.

Group VIII - 'Urban Venturers'

High income young professionals mainly renting (mainly Greater London)
Young white collar workers in multi-racial areas (mainly London)
Young professionals buying property
Young families buying terraces in multi-racial areas
Young families renting basic accommodation
Young white collar singles sharing city centre accommodation

This group contains a very high relative proportion of ethnic minorities, mostly young singles, though also with some young families. They principally rent their fairly small accommodation. Mainly white collar workers in the service distribution trade make up this group, though there is also a significantly high proportion of members of the armed forces. Many of these people do not own cars, tending to use the train to travel to work.

Group IX - 'Hard-Pressed Families'

Blue collar families in council properties
Young blue collar families in council terraces
Manufacturing workers in terraced housing

This group has high numbers of young people, which also contains a fairly high proportion of single parent families. A large number of them live in medium sized council terraces. This group has a high rate of unemployment, those with jobs working in blue collar skilled and semi-skilled professions. With a relatively low level of car ownership, they often use buses as their main means of transport.

Group X - 'Have-Nots'

Families in council flats in multi-racial areas with high unemployment

Blue collar young families in council properties with high unemployment
 Young families, many single parent, with high unemployment
 Young singles and pensioners in council flats with high unemployment

This, the most deprived group, includes many single parent families and a significantly low number of married couples. They largely live in crowded accommodation, which, with a very low level of owner occupancy, is in the main made up of council of housing association flats and terraces. In this group, where most workers are unskilled, there is a very high level of unemployment. Again a relatively low level of car ownership means that a lot of people in this group rely on public transport.

Table 1. EDs and population in SuperProfiles.
 Yorkshire Region

Super Profiles	Number of Eds		Population		ED Average Population
I:Affluent Achievers	607	7.9%	285722	8.0%	471
II:Country Life	495	6.4%	97345	2.7%	197
III:Thriving Greys	937	12.1%	387845	10.9%	414
IV:Settled Suburban	923	11.9%	492836	13.8%	534
V:Nest Builders	1117	14.4%	591613	16.6%	530
VI:Producers	1205	15.6%	569661	16.0%	473
VII:Senior Citizens	595	7.7%	234648	6.6%	394
VIII:Urban Venturers	567	7.3%	252232	7.1%	445
IX:Hard-Pressed Families	599	7.7%	307504	8.6%	513
X:Have-Nots	687	8.9%	341026	9.6%	496
Total	7732	100%	3560432	100%	460