Are multidimensional social classifications of areas useful in UK health service research?

Richard Reading, Stan Openshaw, Stephen Jarvis

Abstract

Objectives – To show the advantages and disadvantages of a multi-dimensional small area classification in the analysis of child health data in order to measure social inequalities in health and to identify the types of area that have greater health needs.

Design – Health data on children from the district child health information system and a survey of primary school children's height were classified by the census enumeration district of residence using the Super profiles neighbourhood classification.

Setting – County of Northumberland, United Kingdom.

Subjects – One cohort comprised 21 702 preschool children age 0–5 years resident in Northumberland, and another cohort 9930 school children aged 5–8.5 years.

Main outcome measures – Variations between types of area in the proportions of babies with birthweight less than 2.8 kg; births to mothers aged less than 20 years; pertussis immunisation uptake; child health screening uptake; and mean height of school children.

Results – Areas with the poorest child health measures were those which were most socially disadvantaged. The most affluent areas tended to have the best measures of health, although rural areas also had good measures. Problems in analysis included examples of the “ecological fallacy”, misleading area descriptions, and the identification of the specific factors associated with poor health measures. Advantages included a wider view of social circumstances than simply “deprivation” and the ability to identify characteristic types of areas with increased child health needs.

Conclusions – There is a limited place for multidimensional small area classifications in the analysis of health data for both research and health needs assessment provided the inherent drawbacks of these data are understood in interpreting the results.

(J Epidemiol Community Health 1994;48:192-200)

There is a burgeoning interest in small area analysis of health data which is driven by three factors – firstly, the increasing power and accessibility of computers that can handle large health databases and carry out complex geographical and statistical analyses; secondly, the resurgence of academic research into health inequalities; and thirdly, the increased demand by health service purchasers for information on the health needs of communities. Although caution has been urged against the indiscriminate use of these methods, they do have potential value provided there is guidance on how best to interpret and apply their results.

Much of the work so far has concentrated on the use (and misuse) of various deprivation indices. These are unidimensional scores which are computed from small area census data by adding together individual standardised (and sometimes weighted) census variables which reflect different aspects of “deprivation”. Notwithstanding the debate about the various merits and demerits of these scores, they have the advantage of simplicity in allowing areas to be ranked against each other for statistical comparisons.

But therein also lies their main disadvantage. It has been argued that conceptualising deprivation as a unidimensional variable disregards the multifaceted nature of the problem both for individuals and between areas. For example, in comparing deprivation in urban and rural areas, Bradley states “It is not the place of policy makers (or academic commentators, for that matter) to decide whether infrastructural decay and environmental pollution are worse settings for poverty than service deprivation and inaccessibility”. When investigating the relationship between social factors and health, a unidimensional view may hide more than it reveals, given the complexity of interactions between socioeconomic, environmental, and behavioural influences on health.

In order to reflect the social diversity of areas in the investigation of geographical health variation, the use of multidimensional area classifications has been urged by some. These are produced by cluster analysis of selected census variables. This results in a descriptive classification in which areas sharing similar distinctive patterns of socioeconomic, demographic, housing, and environmental features are grouped together.

The theoretical advantages of these classifications are as follows: (1) they provide a multidimensional view of social circumstances; (2) areas which have a combination of socioeconomic and environmental factors with a particularly potent influence on health may be identified; (3) they may shed new ideas on structural and lifestyle influences on health; (4) arbitrary and possibly inaccurate decisions about one area being more or less “deprived”
Utility of multidimensional social classification of areas

than another are avoided; and (5) identification and mapping of localities likely to have increased health needs is made more reliable.

There are, however, corresponding disadvantages: (1) there is no rank order to the classification so trends cannot be measured; (2) there is a danger of drawing post hoc conclusions with no prior hypothesis being tested; (3) interpretation of the results depends on the classification reflecting the characteristics it purports to reflect; (4) important causes of health variations are virtually impossible to tease out from associated or incidental factors reflected in the classification; and (5) the hierarchical structure of some of these classifications may be difficult to understand.

Multidimensional classifications of areas have been produced specifically for health services research but more commonly they have been developed primarily for commercial use, for example ACORN and Super profiles. While it may seem preferable to use classifications designed for health research to measure health variations between areas, these are less readily available so much of the research has used commercially available multidimensional classifications.

This paper is an attempt to go beyond the theoretical debate over different methods of area classification by presenting the results of a study which used a multidimensional classification to investigate small area variations in child health in Northumberland, a county in the north of England. We describe not only the variations we found, but how they stimulated further avenues of investigation, and how the results could be applied in planning local health services. The results also illustrate certain inherent disadvantages of multidimensional small area classifications which limit their value in health service research.

Methods

The collection of the health data has been described elsewhere. Briefly, one source of this data was the district child health information system, from which we obtained birthweight, maternal age, immunisation, and child health screening coverage on a cohort of all children resident in Northumberland at the time the data were drawn for analysis (October 1990) born between January 1985 and September 1990. The full cohort included 21 702 births. Low birthweight was defined for this study as births less than 2·8 kg as a proportion of all live single births; teenage motherhood was defined as the proportion of births to mothers aged less than 20 years; immunisation coverage was defined as the proportion of children aged at least 15 months who were not fully immunised against pertussis (the variations in coverage of the other routine childhood immunisations were similar to that of pertussis); six week screening coverage was measured by the proportion of inborn babies born after June 1987 (when recording of screening coverage began) who were not screened at six weeks; and 18 month screening was measured by the proportion of inborn children, born after June 1987, who were not screened at 18 months.

The second source of data was a survey of heights of children in three school years in all schools in Northumberland, which gave a cohort of 9930 children aged between 5·0 and 8·5 decimal years. Heights were standardised to standard deviation (SD) scores to allow for the effects of age and sex – 0·2 SD scores being roughly equivalent to lcm around the age of 6 years. Details of the reproducibility of measurements and the standardisation procedure are given in the fuller description of the data collection. For both sources, the data on each child included the postcode of their address, and this was assigned to a census enumeration district by geographically matching the grid references of the postcode to the grid reference of the nearest enumeration district centroid. In about 7% of cases the system was unable to match a postcode to an enumeration district.

The classification of enumeration districts used was Super profiles. This is a multidimensional "socio-demographic" classification based on cluster analysis of 55 selected census variables. It was developed for academic and commercial use jointly at the Centre for Urban and Regional Development Studies, University of Newcastle upon Tyne, and at the Centre for Urban Studies, University of Liverpool. The classification allocates virtually every enumeration district in Britain into one of 150 "clusters", each of which share a typical social and demographic composition. These may be aggregated to 36 Target markets, which still share many common characteristics, although they are not so finely discriminated as the 150 clusters. Further aggregation of the Target markets is possible, into 22 Groups and into 10 Lifestyles. As the level of aggregation increases, the description of the areas becomes broader, but the categories are still intended to distinguish between different types of residential neighbourhood. An additional feature of the Super profiles classification is that the different types of area are also ranked according to estimated levels of affluence.

In this study we have chosen to use the highest level of aggregation, into the 10 Lifestyles, for the initial demonstration of health variations, and have used the level of the 36 Target markets for more detailed analysis. The distribution of children in our study population between the 10 Lifestyles is shown in table 1, which also shows the distribution of the population between Lifestyles for the whole of the country for comparison. Brief descriptions of each of the Lifestyles are given in the appendix.

In the tables the totals vary in each column and none add up to the total cohort number. This is partly because full health information was not available on all cases, some postcodes could not be matched to enumeration districts, and some enumeration districts with very small populations were unclassified, but more because the denominators were different in each health measure. For example screening
Table 1  Distribution of the study population between Super profiles “Lifestyles”, in comparison to the distribution of the United Kingdom population

<table>
<thead>
<tr>
<th>Lifestyle (in descending order of affluence)</th>
<th>Description</th>
<th>Proportion of the study population (%)</th>
<th>Proportion of the UK population (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Affluent minority</td>
<td>6.3</td>
<td>8.9</td>
</tr>
<tr>
<td>B</td>
<td>Metro singles</td>
<td>1.0</td>
<td>3.1</td>
</tr>
<tr>
<td>C</td>
<td>Young married suburbia</td>
<td>1.9</td>
<td>9.8</td>
</tr>
<tr>
<td>D</td>
<td>Rural Britain</td>
<td>14.5</td>
<td>8.6</td>
</tr>
<tr>
<td>E</td>
<td>Older suburbia</td>
<td>4.6</td>
<td>8.2</td>
</tr>
<tr>
<td>F</td>
<td>Lower middle class</td>
<td>10.0</td>
<td>16.0</td>
</tr>
<tr>
<td>G</td>
<td>Multi-ethnic areas</td>
<td>0.8</td>
<td>7.2</td>
</tr>
<tr>
<td>H</td>
<td>Fading industrial</td>
<td>1.0</td>
<td>10.4</td>
</tr>
<tr>
<td>I</td>
<td>Council tenants</td>
<td>21.7</td>
<td>15.8</td>
</tr>
<tr>
<td>J</td>
<td>Underprivileged Britain</td>
<td>8.7</td>
<td>10.8</td>
</tr>
</tbody>
</table>

Notes: The figures in the column describing the distribution of the study population are percentages of children in the study population, the figures in the column describing the distribution of the UK population are for the total population, thus the columns are not strictly comparable.

The distribution of the population of the United Kingdom between the different “Super profiles Lifestyles” is taken from reference 16.

The totals in each column do not quite add up to 100%, the remainder of the population live in unclassifiable enumeration districts.

coverage was restricted to births after June 1987, and 15 month immunisation coverage and 18 month screening coverage could only be measured on children in the cohort who had already attained those ages.

Cross tabulations were performed using the SPSSX package and confidence intervals were calculated using the CIA programme. Ethical permission for the study was given by the Northumberland Health Authority Ethical Committee, and the study was also approved by the Education Department of Northumber-

Results

MEASURES OF CHILD HEALTH ACCORDING TO SUPER PROFILES’ LIFESTYLES

Table 2 shows the variation between the different Super profiles’ Lifestyles for all six measures of child health. In general, the “Underprivileged Britain” and “Multi-ethnic areas” Lifestyles show the poorest measures. Numbers are small in the “Multi-ethnic areas” Lifestyle, but even so, the confidence intervals of the estimates suggest that these are reliable findings, particularly as all the health measures are poor in these types of area.

There is less consistency in the types of area where health measures were favourable. The “Affluent minority” Lifestyle has the lowest proportion of low birthweight babies and the tallest mean height, closely followed by the “Rural Britain” Lifestyle. The “Affluent minority” Lifestyle also has the lowest proportion of births to teenage mothers and the highest uptake of 18 month screening (bar the “Metro singles” Lifestyle which contains too few numbers to allow valid conclusions to be drawn). However, uptakes of pertussis immunisation and six week screening were highest in the “Young married suburbia” Lifestyle, the “Affluent minority” Lifestyle having rather poor levels of uptake.

It is also worth noting that many of the other Lifestyles have good records in one or other of the child health measures shown. For example, “Older suburbia” has a low proportion of low birthweight babies and a high uptake of pertussis immunisation, while perhaps more strikingly, “Fading industrial” has high levels of uptake of immunisation and screening at

Table 2  Child health measures according to Super profiles “Lifestyle”

<table>
<thead>
<tr>
<th>Lifestyle</th>
<th>Births &lt; 2.8 kg</th>
<th>Mothers’ age &lt; 20 y</th>
<th>Pertussis immunisation</th>
<th>Six week screening</th>
<th>18 month screening</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% rate</td>
<td>95% CI (no/total)</td>
<td>% not</td>
<td>95% CI (no/total)</td>
<td>% not</td>
<td>95% CI (no/total)</td>
</tr>
<tr>
<td>Affluent minority</td>
<td>9.9</td>
<td>8.1 to 12.1</td>
<td>1.1</td>
<td>0.5 to 2.1</td>
<td>22.2</td>
<td>19.3 to 25.1</td>
</tr>
<tr>
<td>Metro singles</td>
<td>15.7</td>
<td>10.3 to 21.1</td>
<td>3.7</td>
<td>1.4 to 7.9</td>
<td>20.0</td>
<td>13.3 to 26.7</td>
</tr>
<tr>
<td>Young married suburbia</td>
<td>11.1</td>
<td>9.9 to 12.3</td>
<td>2.1</td>
<td>1.6 to 2.8</td>
<td>17.2</td>
<td>15.6 to 18.8</td>
</tr>
<tr>
<td>Rural Britain</td>
<td>10.0</td>
<td>8.9 to 11.2</td>
<td>4.8</td>
<td>4.0 to 5.7</td>
<td>18.5</td>
<td>16.6 to 20.1</td>
</tr>
<tr>
<td>Older suburbia</td>
<td>10.8</td>
<td>8.8 to 12.9</td>
<td>5.8</td>
<td>4.3 to 7.7</td>
<td>17.9</td>
<td>15.1 to 20.8</td>
</tr>
<tr>
<td>Lower middle class</td>
<td>10.8</td>
<td>9.0 to 12.2</td>
<td>7.2</td>
<td>6.0 to 8.5</td>
<td>20.2</td>
<td>18.2 to 21.1</td>
</tr>
<tr>
<td>Industrial areas</td>
<td>19.5</td>
<td>13.9 to 25.0</td>
<td>25.5</td>
<td>18.7 to 32.3</td>
<td>33.3</td>
<td>15.6 to 40.7</td>
</tr>
<tr>
<td>Fading industrial</td>
<td>13.5</td>
<td>12.4 to 14.7</td>
<td>11.2</td>
<td>10.1 to 12.3</td>
<td>19.8</td>
<td>18.3 to 21.4</td>
</tr>
<tr>
<td>Council tenants</td>
<td>15.2</td>
<td>14.2 to 16.3</td>
<td>13.6</td>
<td>12.5 to 14.6</td>
<td>24.0</td>
<td>22.6 to 25.4</td>
</tr>
<tr>
<td>Underprivileged Britain</td>
<td>17.5</td>
<td>15.8 to 19.2</td>
<td>16.2</td>
<td>14.0 to 17.9</td>
<td>28.0</td>
<td>25.7 to 30.0</td>
</tr>
<tr>
<td>Total</td>
<td>13.0</td>
<td>24.969 to 19.007</td>
<td>9.1</td>
<td>15.8 (17.286)</td>
<td>21.2</td>
<td>32.9 (15202)</td>
</tr>
</tbody>
</table>
MEASURES OF CHILD HEALTH ACCORDING TO SUPER PROFILES "TARGET MARKETS"

The analysis of health measures between different Lifestyle, the highest level of aggregation of the classification, has raised a number of questions. What are the factors related to the poor health in the “Underprivileged” and “Multi-ethnic” Lifestyles, is this simply a matter of socioeconomic disadvantage and if so, are there subgroups of the other disadvantaged Lifestyle, “Council tenants”, which also suffer these levels of poor health? At the other end of the spectrum, how do the health measures vary between the more affluent areas, why are some aspects of child health less favourable in certain types of affluent area, in contrast to the deprived areas where most of the aspects of health were consistently poor? Finally, why do areas in the “Rural Britain” Lifestyle have such favourable measures of low birthweight and height, particularly as the social heterogeneity in these areas might suggest that the measures of health would tend towards the overall average?

Pointers to the answers to some of these questions may be found by examining the distribution of health measures at a more disaggregated level of the classification. The 10 Lifestyles are constituted from 36 Target markets. In using the “Target markets” level of the classification we rely both on their affluence ranking, between 1 and 36, and on their description, when interpreting differences in health between them. The brief descriptions given here are sketches and inevitably incomplete, fuller descriptions of the individual characteristics of each Target market are given elsewhere.

Table 3 shows the child health measures for the Target markets in the “Multi-ethnic areas” Lifestyle compared with those from the two poorer Lifestyles, “Underprivileged Britain” and “Council tenants”. Only a single Target market is represented in the “Multi-ethnic areas” Lifestyle, and its description, “Unskilled ethnic families in council flats”, gives a clue to the possible determinants of ill health. A more detailed description of this Target market shows it to consist of areas with high unemployment, low car ownership, and a high proportion of unskilled manual workers. It may be that the reasons for the poor health are more to do with socioeconomic disadvantage than with ethnic differences. This is investigated further later.

The table also shows that within the “Underprivileged” Lifestyle, Target markets 33 and 35 have very similar health measures. There is greater heterogeneity between Target markets in the “Council tenants” Lifestyle, but no Target markets within this Lifestyle have consistently poor or good health measures. Target markets 27 and 28, described respectively as “Middle aged and older couples in council flats” and “Blue collar workers in council houses”, seem to have the most favourable child health measures, while Target market 31, “Matute blue collar workers in mining areas”, has the worst. The health measures follow the affluence ranking of the Target markets reasonably well, although there are exceptions, for instance Target market 34.

Table 4 looks at the other end of the social spectrum. The inconsistency in health measures shown in the analysis by Lifestyle is just as evident when Target markets are considered. Target market 7, which is in the “Metro singles” Lifestyle is an obvious anomaly. (There were some other Target markets in this Lifestyle but they are not shown because they consisted of such small numbers.) This Target market has particularly poor measures of height and low birthweight. It is possible that families with young children in these areas do not share the socioeconomic characteristics of the majority of the inhabitants, most of whom, according to the description of this Target market, do not have any children. These areas are also those most likely to have undergone change since the 1981 census so that area labelling may no longer be as appropriate as it once was.

The Target markets in the “Affluent minority” Lifestyle all show generally good measures of low birthweight, proportion of teenage mothers, and mean height, but poor immi-
<table>
<thead>
<tr>
<th>Lifestyle</th>
<th>Target market</th>
<th>% low birthweight</th>
<th>% teenage mothers</th>
<th>% not immunised for pertussis</th>
<th>Mean height in SD scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affluent minority</td>
<td>1</td>
<td>10.4 (56/541)</td>
<td>0.5 (2/403)</td>
<td>23.6 (110/447)</td>
<td>0.001 (348)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.4 (2/83)</td>
<td>0.0 (0/65)</td>
<td>28.6 (18/63)</td>
<td>0.197 (60)</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>11.5 (30/260)</td>
<td>1.3 (3/233)</td>
<td>18.4 (41/233)</td>
<td>0.170 (76)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>18.2 (25/137)</td>
<td>3.8 (5/132)</td>
<td>16.5 (18/109)</td>
<td>-0.250 (58)</td>
</tr>
<tr>
<td>Metro singles</td>
<td>5</td>
<td>10.5 (41/389)</td>
<td>3.0 (10/328)</td>
<td>14.5 (48/331)</td>
<td>0.196 (219)</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>12.1 (92/820)</td>
<td>1.8 (14/764)</td>
<td>17.9 (118/661)</td>
<td>0.070 (430)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>11.2 (154/130)</td>
<td>2.0 (25/1222)</td>
<td>17.6 (197/1120)</td>
<td>0.006 (718)</td>
</tr>
</tbody>
</table>

Note: values are % or SD scores, figures in parentheses are the numbers in each group. Figures for child health screening have been omitted because of small numbers.

Target market descriptions:
1. Lifestyle - Brief description of characteristics of the Target market;
2. Affluent minority - Middle aged families in exclusive suburbs;
3. Affluent minority - Older families in select suburban properties;
4. Affluent minority - Middle aged families in upmarket semi-detached houses;
5. Metro singles - Young single white collars in rented property;
6. Young married suburbia - Younger families in suburban detached houses;
7. Young married suburbia - Younger families in larger semi-detached houses;
8. Young married suburbia - Young well to do married in high turnover semi-detached houses.

<table>
<thead>
<tr>
<th>Lifestyle</th>
<th>Target market</th>
<th>% low birthweight</th>
<th>% teenage mothers</th>
<th>% not immunised for pertussis</th>
<th>Mean height in SD scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural Britain</td>
<td>14</td>
<td>9.7 (91/939)</td>
<td>4.4 (36/819)</td>
<td>20.1 (149/740)</td>
<td>0.103 (485)</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>10.1 (123/1239)</td>
<td>5.7 (62/1097)</td>
<td>19.3 (100/984)</td>
<td>0.153 (651)</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>10.6 (57/540)</td>
<td>3.3 (15/451)</td>
<td>13.8 (60/435)</td>
<td>0.176 (227)</td>
</tr>
<tr>
<td>Lower middle class</td>
<td>12</td>
<td>10.4 (41/393)</td>
<td>5.7 (20/350)</td>
<td>17.3 (54/312)</td>
<td>0.032 (154)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>11.3 (48/423)</td>
<td>6.2 (24/300)</td>
<td>17.7 (58/327)</td>
<td>0.000 (216)</td>
</tr>
</tbody>
</table>

Note: values are % or SD scores, figures in parentheses are the numbers in each group. Figures for child health screening have been omitted because of small numbers.

Target market descriptions:
1. Lifestyle - Brief description of characteristics of the Target market;
2. Rural Britain - Affluent farming communities;
3. Rural Britain - Older and retired better off rural dwellers;
4. Rural Britain - Less affluent rural workers and pensioners;
5. Rural Britain - Young married suburbia - Middle aged white collar couples;
6. Rural Britain - Young married middle class - Lower middle class metropolitan semi-detached houses;
7. Rural Britain - Young married middle class - Lower middle class provincial semi-detached houses;
8. Rural Britain - Young married middle class - Upper working class in council housing;
9. Rural Britain - Young married middle class - Upper working class in terraced and semi-detached houses.

The overall conclusion from this table is that, by and large, the measures of health reflect the levels of affluence in these types of areas, except for immunisation uptake. Either the characteristics of areas classed into the “Affluent minority” Lifestyle include something which predisposes to particularly high levels of uptake, or certain characteristics of the “Affluent minority” Lifestyle result in uptake being little higher than that in the poorest areas in the “Underprivileged” Lifestyle.

The final question concerned child health measures in rural areas. Table 5 shows the health measures in the Target markets comprising the “Rural Britain” Lifestyle, along with Target markets of roughly equivalent affluence levels in more urban settings. The proportions of low birthweight babies and of teenagers mothers in the rural Target markets are very similar to those in the other Target markets, none of the differences being outside the 95% confidence intervals, and although immunisation uptake varies quite markedly within the rural Target markets, it corresponds with

Note: values are % or SD scores, figures in parentheses are the numbers in each group. Figures for child health screening have been omitted because of small numbers.

Target market descriptions:
1. Lifestyle - Brief description of characteristics of the Target market;
2. Rural Britain - Affluent farming communities;
3. Rural Britain - Older and retired better off rural dwellers;
4. Rural Britain - Less affluent rural workers and pensioners;
5. Rural Britain - Middle aged white collar couples;
6. Rural Britain - Young married suburbia - Lower middle class metropolitan semi-detached houses;
7. Rural Britain - Young married middle class - Lower middle class provincial semi-detached houses;
8. Rural Britain - Young married middle class - Upper working class in council housing;
9. Rural Britain - Young married middle class - Upper working class in terraced and semi-detached houses.

Meanwhile, low birthweight, the proportion of teenage mothers, and mean height in the Target markets in the “Young married suburbia” Lifestyle are comparable with the levels found in the “Affluent minority” Lifestyle. It is also of some interest that within the “Young married suburbia” Lifestyle mean height in Target market 5 is significantly greater than in either Target markets 9 or 10.

The overall conclusion from this table is that, by and large, the measures of health reflect the levels of affluence in these types of areas, except for immunisation uptake. Either the characteristics of areas classed into the “Affluent minority” Lifestyle include something which predisposes to particularly high levels of uptake, or certain characteristics of the “Affluent minority” Lifestyle result in uptake being little higher than that in the poorest areas in the “Underprivileged” Lifestyle.

The final question concerned child health measures in rural areas. Table 5 shows the health measures in the Target markets comprising the “Rural Britain” Lifestyle, along with Target markets of roughly equivalent affluence levels in more urban settings. The proportions of low birthweight babies and of teenage mothers in the rural Target markets are very similar to those in the other Target markets, none of the differences being outside the 95% confidence intervals, and although immunisation uptake varies quite markedly within the rural Target markets, it corresponds with
Utility of multidimensional social classification of areas

the range of uptake in the other Target markets. What is notable is the difference in mean heights with values in all the rural Target markets being greater than in virtually any other Target market and significantly so at the 5% level when compared to Target markets 21 and 22.

\[
\begin{align*}
(14-21) &= 0.201, \quad 95\% \text{ CI } = 0.059 \text{ to } 0.343, \quad p < 0.05 \\
(17-21) &= 0.232, \quad 95\% \text{ CI } = 0.097 \text{ to } 0.367, \quad p < 0.05 \\
(19-21) &= 0.274, \quad 95\% \text{ CI } = 0.103 \text{ to } 0.445, \quad p < 0.05 \\
(14-22) &= 0.213, \quad 95\% \text{ CI } = 0.058 \text{ to } 0.368, \quad p < 0.05 \\
(17-22) &= 0.244, \quad 95\% \text{ CI } = 0.095 \text{ to } 0.393, \quad p < 0.05 \\
(19-22) &= 0.286, \quad 95\% \text{ CI } = 0.104 \text{ to } 0.468, \quad p < 0.05
\end{align*}
\]

FURTHER INVESTIGATIONS

The demonstration of health variations between different types of area is not an end in itself. The classification may suggest explanations for these variations or point to further necessary investigations. As an example we describe one such investigation.

One of the unresolved questions from the earlier analyses was the reason for the poor child health measures in the “Multi-ethnic areas” Lifestyle. Although the descriptions of the Target markets in this Lifestyle suggested that economic factors may have accounted for the poor health, the intriguing feature of the finding was that areas with concentrations of ethnic minorities are not known to exist in the county. The proportion of non-white families in Northumberland is low and they are dispersed widely.

Therefore, the enumeration districts classed into this Lifestyle, of which there were only three, were identified. One of the enumeration districts was small, situated in a small rural market town and had unremarkable health measures. The other two enumeration districts were larger, situated in council housing estates in larger towns in the more urban south east of the county and both had extremely poor health measures. The poor health of the Lifestyle as a whole was accounted for by the health measures in these latter two enumeration districts.

Discussions were then held with local health visitors about the health and social circumstances of children in these areas. According to the health visitors none of the areas had a high proportion of families of non-white ethnic origin. The first enumeration district described above, with average health measures, was in an ordinary residential housing estate with no excess of deprived families. Why it was so classed is uncertain and it probably represents a classification error. On the other hand, the latter two enumeration districts with the poor health measures were both in council estates with deprived, mobile populations, high proportions of young and single parents, high rates of unemployment and marital disruption, and very poor housing conditions. One health visitor’s comment was that the families in one of these enumeration districts generated enough problems for her to be able to fill all her working time.

Thus, the multidimensional classification has identified a group of enumeration districts with exceptionally poor health which, apart from one anomaly, share common social characteristics. However, the description given by the classification of these characteristics was misleading. By investigating this finding in conjunction with local health professionals, we have identified areas with excess health needs in an objective way. This type of analysis may help in planning the local distribution of resources.

Discussion

This example of the analysis of health data using a multidimensional social classification of areas illustrates some of the problems associated with their use and indicates possible applications of them in health service research. The results also raise more fundamental questions about the utility of census data in small area health research.

PROBLEMS IN THE USE OF MULTIDIMENSIONAL AREA CLASSIFICATIONS

These relate mainly to the interpretation of the variations shown. Demonstration and testing of correlations and trends, for example using regression techniques, may be complex because the categorical nature of multidimensional classifications results in equations that are difficult to interpret. The classification we have used, Super profiles, includes an approximate affluence ranking which allows a compromise solution to this problem, although not without making assumptions which we have shown in the example of the “Multi-ethnic areas” Lifestyle, to be open to inaccuracy.

Next, are the descriptions of areas reliable or relevant to the questions we are asking? A potential weakness with multidimensional census classifications is that some areas can have very distinctive characteristics based on a minority of households. We found the description of the “Multi-ethnic areas” Lifestyle to be misleading, and similar problems arise in the interpretation of results from the “Older suburbia” Lifestyle – in which the description stresses the preponderance of older people – and from the “metro singles” Lifestyle – in which the description is exclusively of single people, either young and unmarried or elderly.

What the socioeconomic characteristics of families with young children are in these areas is not made clear. More detailed descriptions of the areas\textsuperscript{17,20,21} do not shed much more light on this question either.

This is a striking example of the “ecological fallacy”,\textsuperscript{22} which describes how the subjects of interest in an ecological group (in this case a type of case) may not share the general characteristics of the rest of the group. However, there is more to the ecological fallacy than this. Sometimes the discordance between the features of the subgroup of interest and the majority of the group has epidemiological relevance of itself\textsuperscript{23} because segregation of the atypical subgroup may lead to the adverse consequences. To be living in conditions of
poverty in an otherwise affluent area, for example, may give rise to a greater sense of deprivation and social isolation, and families with young children living in an area consisting mainly of single childless people may have similar experiences. Another feature of the ecological fallacy is the likelihood of confounding. Certain characteristics of areas may spuriously seem to be associated with poor health only because they are associated incidentally with real causal factors. Thus, features of the “Multi-ethnic areas” Lifestyle were strongly associated with poor measures of health, yet these were not related to the ethnic composition of the areas.

These problems are compounded by the descriptions of areas being couched in marketing jargon. The Super profiles labels are intended to provide thumbnail sketches that make some sense in marketing circles, hence the flowery terminology. We gave our reasons for using a commercial classification in the introduction, but we are not convinced that using a classification designed specifically for health service use would have altered matters very much. Control over the make up of these classifications is restricted to the choice of variables entered into the cluster analysis and specification of the number of clusters to be produced. Even changing these may not make much difference to the outcome of the cluster analysis. Thus, when one of us (SO) set out with colleagues to produce the Super profiles classification, had it been intended primarily for health service research the final clusters would probably have been similar, although the description of them would certainly have been worded differently.

The final problem in interpretation is the teasing out of factors which are associated with poor health. For example, are the favourable measures of height and low birthweight in the “Rural Britain” Lifestyle related to rurality, socioeconomic factors, environmental differences, differences in individual’s lifestyle, or demographic factors? The answer is not clear yet the classification purports to measure differences in all these aspects between areas.

POSSIBLE APPLICATIONS FOR MULTIDIMENSIONAL CLASSIFICATIONS
Multidimensional classifications may be used in the measurement of health inequalities. There is no theoretical reason why small area inequalities in health should not be measured using a multidimensional classification of areas rather than a unidimensional deprivation score. In an earlier study, admittedly using data from many fewer subjects, we compared a deprivation score with a multidimensional area classification as a means of measuring health inequalities and found little to choose between the two methods. Using the child health data described in this paper, we have also measured small area variations with a unidimensional deprivation score and comparison of those results with this study show that the width of variations in health measures between deprived and affluent areas is remarkably similar; if anything, the multidimensional classification shows slightly wider inequalities in health.

The second application in which multidimensional classifications are useful is in the identification of areas with increased health needs. Our experience mirrors that of Thunhurst in showing that the knowledge and experience of local health professionals is complementary to the analytical capabilities of these classifications. If harnessed together, they provide a very effective means of identifying areas where socioeconomic, environmental and demographic circumstances are associated with increased health needs. We have examined child health only, it may be that different types of area are associated with increased needs of other groups, for instance the elderly.

A third application is in the early stages of research into causes of health variations. When using multidimensional area classifications in epidemiological research it is tempting to bypass the crucial step of testing a specific hypothesis. Making inferences about causes of health differences on the basis of area variations has been cautioned against; however, the choice of these classifications in the analysis of health data may contribute to the earlier stage of hypothesis formation. An example of a hypothesis arising out of this study might be “at equivalent levels of affluence or material deprivation, children are taller in rural than in urban areas”. Such a conclusion is not warranted on the evidence produced here alone but it is supported by some previous research and, on the face of it, seems a hypothesis worth further investigation.

A final research application is in the measurement of variations in health related behaviour, particularly in studies examining the effect of interventions. Small area variations are of interest because interventions designed to change such behaviour are more likely to be directed at geographical areas than at specifically identified individuals. Multidimensional classifications of areas are the obvious choice for such research because they are intended to reflect different patterns of lifestyle and experience of their use in marketing suggests that they do reflect consumer habits accurately. An example of a finding worth pursuing which has arisen from this study is the different levels of uptake of immunisation between the two affluent Lifestyles, “Affluent minority” and “Young married suburbia”.

WHY USE CENSUS BASED SMALL AREA CLASSIFICATIONS?
This study has shown how a multidimensional census based classification of small areas can provide valuable information for both research and local planning. We have shown that it is possible to identify areas which have excess health needs reliably. However, we can not be certain of why these areas have different experiences of health, even if we think we know what we are measuring, for instance “depriva-
tion”. Why bother, therefore, with census data when classifying areas? Two alternative approaches are either to classify neighborhoods descriptively using the subjective views of local people, as has been carried out in Newcastle upon Tyne by Wright, or to produce a more objective classification of areas using a combination of health and demographic data, on the grounds that it is better to measure health needs directly rather than social characteristics which tend to be associated with poor health. Scott-Samuel argued for this approach some time ago. There is probably scope for further developments in all of the offered methods as each offers different insights into particular problems. The advantage of census data is that it provides such a rich source of comparative data relatively cheaply.

In conclusion, the answer to the question “are multidimensional social classifications of areas useful in health service research?” is a qualified “yes”. We have described possible applications of these, although we have not been able to show any unequivocal advantage over unidimensional deprivation scores. Our findings illustrate the potential pitfalls of uncritical acceptance of the results of analyses using multidimensional classification and the difficulty of drawing valid epidemiological conclusions. As a result of our study we can not fully endorse the comment made by Thunhurst and Macfarlane that “multivariate analyses, particularly those centring on the use of cluster analyses, are not as intuituive as they are often presented”. In our view there is a limited place for multidimensional classifications in health service research but they need to be used with caution.

We would like to thank Mrs Pat Waugh and Miss Gwen Chlopton who carried out the measurements of height on the Northumberland school children. Northumberland Health Authority generously supported the project financially and allowed the use of their child health data.


Appendix

Brief descriptions of the 10 Super profiles Life- styles (extracted from reference 21 with the authors’ permission).

(A) AFFLUENT MINORITY
This most affluent of the Lifestyles is characterised by large, detached, owner occupied housing which accommodates highly qualified, multi-car owning, professional worker households with few children, in low density, surburban and semi-rural areas from which the majority of workers commute by car and train to office jobs.

(B) METRO SINGLES
Typified by young single, well qualified, professional and other white collar workers with some single elderly living in small furnished and unfurnished rented flats, often lacking in basic amenities, in areas of ethnically mixed population with a high residential turnover well served by rail and tube, resulting in an unusually low level of car ownership and use for work travel.

(C) YOUNG MARRIED SUBURBIA
Young married couples with younger children, with very few older and elderly people, in households of well qualified, professional and other white collar workers, owning their own detached and semi-detached houses in leafy suburban estates, virtually all owning one or more cars, one of which is likely to be used for the journey to work.
(D) RURAL BRITAIN
Rural areas which are mainly associated with farming but also popular locations, some urban or semi-urban, for retirement and holiday homes, providing spacious out of town accommodation for multi-car owning and car commuting managerial households and tied properties for agricultural workers.

(E) OLDER SUBURBIA
Stable suburban areas populated largely by smaller households of retired and elderly in their own semidetached housing together with some managerial and other white collar workers, some working part time, with average car ownership, low unemployment and close to average performance with respect to many other indicators.

(F) LOWER MIDDLE CLASS
A Lifestyle distinguished from “Older suburbia” (Lifestyle (E)) by its much younger age structure and the presence of families with young children living in owner occupied semi-detached property, including families of those in the armed forces, more working wives and skilled manual workers, many using a car for the work journey from the car owning majority of households.

(C) MULTI-ETHNIC AREAS
Areas in which large ethnic families with young children are most evident, living in overcrowded rented property, in generally poor condition, some occupied by students and single workers, high unemployment and most households without a car, but those in employment working in the manufacturing and transport sectors in semi skilled and unskilled jobs to which many commute by rail and tube.

(H) FADING INDUSTRIAL
Generally typical national age structure with rather more older people, living in poorest quality unimproved terraced property accommodating a largely semi and unskilled blue collar labour force in older industrial and mining areas, with a level of male unemployment a little above average, and many of those working travelling to work on foot.

(I) COUNCIL TENANTS
In common with Lifestyle (J), a very high proportion of council rented property, mostly in the form of flats in a good state of repair, but having a much older age structure than (J) with more single female OAPs, no overcrowding and a rather lower but still relatively high level of unemployment amongst its blue collar workforce, which includes many part-timers, and most of whom rely upon foot and bus travel to reach their manufacturing jobs.

(J) UNDERPRIVILEGED BRITAIN
This Lifestyle accounts for those areas in which the worst conditions of social stress and deprivation are concentrated. It is characterised by very large families, including young children, living in cramped and overcrowded conditions in council flats that are in generally good condition, but with the highest levels of unemployment amongst a very poorly qualified, unskilled labour force, with very low car ownership, those in employment thus reliant upon the bus for the work journey.
Are multidimensional social classifications of areas useful in UK health service research?

R Reading, S Openshaw and S Jarvis

*J Epidemiol Community Health* 1994 48: 192-200
doi: 10.1136/jech.48.2.192

Updated information and services can be found at:
http://jech.bmj.com/content/48/2/192

**Email alerting service**

These include:

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/