Providing census data for general practice. 2. Usefulness

CHRISTOPHER FOY, MSc
ALLEN HUTCHINSON, FRCGP
JANE SMYTH, MA

SUMMARY. Computerized census data are described in relation to a general practice population. The previously published methods for scoring deprivation — underprivileged areas score and material deprivation score — are applied to the data. Wards and enumeration districts within a single practice area are ranked by both methods and examples show the wide variation in deprivation scores for enumeration districts within single wards. The value of these data to a general practice is discussed with particular reference to developing a profile of the practice and to planning prevention and anticipatory care.

Introduction

IN our feasibility study we demonstrated that it is possible to provide census data which relate to a general practice population.[1] We now address the question of how the data can be used.

Census data can be displayed in the form of one or more variables obtained directly from the data files held on computer using the SASPAC package.[1] As such it is difficult to read and use. Alternatively, the data can be combined to provide an overall score which allows for a more easily understandable comparison between two or more census populations. Two such scoring methods have been described in the UK, the first by Jarman (underprivileged areas score),[2] the second by Townsend and colleagues (material deprivation score).[3]

Jarman[2] developed the underprivileged areas score by asking a large sample of 1802 general practitioners to identify factors they felt contributed to their workload. From these were derived a weighting for each of 10 census variables — percentage elderly alone, percentage aged under five years, percentage one parent families, percentage unskilled (social class 5), percentage unemployed, percentage living in overcrowded conditions, percentage changed address within one year, percentage ethnic minorities, percentage non-married couple families, percentage living in households lacking basic amenities. These were then combined into a score based on census data at ward level. More recently, Townsend and his colleagues[3] used ward level data in describing the relationship between a measure of deprivation and a measure of health. Again, four census variables — percentage of population unemployed, percentage of households with no car, percentage of households not owner occupied and percentage of households overcrowded — were combined into a score, in this instance without an externally derived weighting measure such as that used by Jarman.

Both the underprivileged areas and material deprivation scores were originally described for ward data. Townsend and colleagues discussed the problem of heterogeneous groups within wards but were constrained from using data at the enumeration district level by the need for matching health data which are only available at ward level. Jarman also suggested that it would be possible to use the underprivileged areas score for health planning for enumeration districts. A more flexible approach is offered by the possible linkage of postcodes to census data, a technique considered in some detail by Carstairs and Lowe[4] and to which we refer again in the discussion.

In this second paper we create a census data profile for a general practice in Ashington, including each enumeration district where more than 11 practice patients were domiciled, to show how enumeration district data reflect considerable differences within wards and to review the value of such data to the practice. We have chosen not to compare the relative merits of the two scoring methods but to show that both can be used for practice populations.

Since this is a preliminary study we have accepted that there may be inaccuracies in the attribution of patients to individual enumeration districts, particularly at the boundaries. However, the relatively large number of people in each enumeration district (approximately 500) and proposals by the Office of Population Censuses and Surveys (OPCS) to offer postcoded census data and to improve geographical resolution of postcodes suggest that any positive value from the data in this study will be increased by future changes. Indeed, the proposed OPCS changes in themselves make it timely to study the potential for providing census data at enumeration district level in general practice.

Method

Underprivileged areas scores and deprivation scores were computed for each of 98 enumeration districts, using the definitions of Irving[5] and Townsend and colleagues[3] respectively. Irving's 10 variables differ slightly from Jarman's: she includes percentage aged over 65 years, and excludes percentage of non-married couple families. Further analysis consisted of cross-tabulations of the proportion of patients in the study practice in each enumeration district with the underprivileged areas and deprivation scores, and a construction of ward scores to allow for comparison with Townsend's data.

Briefly, there are three steps in the construction of the underprivileged areas score:

1. To reduce skewness, each of the 10 census variables, expressed as a proportion, is normalized by the arc sine square root transformation $\sin^{-1}\sqrt{V}$.
2. Each of the transformed variables is standardized by subtracting the mean for England and Wales, and dividing by the standard deviation.
3. The underprivileged areas score is a weighted sum of the standardized variables; the weights are those derived from Jarman's survey of general practitioners.

The material deprivation score is also calculated in three steps:

1. The four variables are expressed as percentages. Two of these (overcrowding and unemployment) are normalized by a logarithmic transformation $\log_e(1+V)$. The other two are untransformed.

Christopher Foy, Lecturer in Medical Statistics, Health Care Research Unit, University of Newcastle upon Tyne; Allen Hutchinson, General Practitioner, Ashington, and Senior Lecturer in Health Care Research; and Jane Smyth, Research Associate, Health Care Research Unit, University of Newcastle upon Tyne.

© *Journal of the Royal College of General Practitioners,* 1987, 37, 451-454.

2. The resulting variables are standardized, as for the underprivileged areas score.
3. The material deprivation score is the (unweighted) sum of the standardized variables.

Results

On census day 1981 a population of 48 113 lived within the 98 enumeration districts in the study; of these 14 196 were registered with the study practice.

Scores for the seven wards containing the majority (85%) of the practice population were created and are ranked by both underprivileged areas and material deprivation scores in Table 1. Only the highest and lowest ranked wards are ranked equally by both scoring methods.

In Table 2 is shown the ranges of scores for the enumeration districts in the highest and lowest ranked of the seven wards on Table 1 and the range of scores across all 98 enumeration districts. Both underprivileged areas and deprivation scores are shown, negative values indicating the best (or least deprived) score. The variation in the census variables making up these scores for three sample enumeration districts in three different wards is shown on Table 3. Within each of these sample enumeration districts the proportion of the population registered with the practice is greater than 50%.

As expected, both underprivileged areas and material deprivation scores show a normal distribution curve across the 98 enumeration districts (Figures 1 and 2). Cross-tabulation of the proportion of the population in each enumeration district who are in the practice with the proportion of enumeration districts more or less underprivileged than average results in different distributions when underprivileged areas scores and deprivation scores are compared (Table 4). While the underprivileged areas score shows that the proportion of underprivileged enumeration districts is virtually the same in the districts where less than 50% of residents are practice patients and in those where at least 50% are practice patients, the material deprivation score shows some association between above average deprivation and a higher proportion of practice patients. Although this may indicate a trend, the difference is not statistically significant at conventional levels (chi-square test with continuity correction).

Table 1. The seven wards containing the majority of the practice population, ranked by underprivileged areas (UPA) and material deprivation scores.

Ward	UPA score	Ward	Material deprivation score
Hirst	13.8	Hirst	4.0
College	9.7	Park	2.8
Bothal	8.3	Central	1.8
Haydon	8.1	College	0.8
Central	3.6	Haydon	0.03
Park	−3.6	Bothal	0.0
Seaton	−7.5	Seaton	−0.1

Table 2. Variation in underprivileged areas (UPA) and material deprivation scores for enumeration districts within lowest and highest ranked of the seven wards on Table 1 and across all 98 districts.

	Range of scores	
	UPA scores	Material deprivation scores
Enumeration districts within:		
Lowest ranked ward (Seaton)	−64.1 to 16.1	−7.8 to 1.9
Highest ranked ward (Hirst)	−33.2 to 65.6	−1.2 to 7.4
All 98 enumeration districts	−76.9 to 65.6	−8.6 to 7.6

Table 3. Variation in underprivileged areas (UPA) and material deprivation scores and the census variables which make up the scores between three sample enumeration districts in three different wards.

	Location of enumeration district		
	Seaton ward	Hirst ward	Bothal ward
Variables used to calculate UPA score[a]			
% over 65 yrs	*16*	*5*	*36*
% lone OAP	*8*	*4*	*23*
% under 5 yrs	*9*	*6*	*3*
% single parent	*0*	*1*	*3*
% unskilled	*0*	*7*	*0*
% unemployed[b]	*5*	*21*	*8*
% living in households lacking amenities	*0*	*7*	*0*
% living in overcrowded households[b]	*2*	*20*	*1*
% changed address	*12*	*9*	*7*
% ethnic minority	*0*	*0*	*0*
UPA score	−12.5	−0.6	37.1
Variables used to calculate material deprivation score			
% unemployed[b]	*5*	*22*	*7*
% households lacking car	*38*	*71*	*80*
% households not owner occupied	*28*	*59*	*98*
% households overcrowded[b]	*1*	*10*	*1*
Material deprivation score	−3.6	5.7	1.8

[a] Irving's variables. [b] Data apply to the same three enumeration districts — differences between the two scoring systems in the proportions unemployed and overcrowded arise because different criteria are used to obtain the variables.

Discussion

Applying both the underprivileged areas score and the material deprivation score to electoral wards has been demonstrated to give useful data concerning large populations such as local medical committee areas[2] or local authority areas.[3] However, many wards have large numbers of residents and the aggregated sociodemographic data available from the census may hide large variations between the smaller, more homogeneous groups living within the sort of communities identified by enumeration districts of about 200 people. In this study we have shown how populations from enumeration districts differ very considerably within and between wards by using both single variables and multiple variables combined as scores. Unless they work in a very rural setting, most family doctors look after patients in more than one ward — perhaps many wards in an urban environment. Given that large variations between enumeration district populations within wards do exist, therefore, small area statistics relating to enumeration districts are much more relevant to general practice than data aggregated in wards or postcode sectors.

Having shown that these data are available to general practitioners (albeit not easily) and that the differences between the small enumeration district populations do exist, we must now

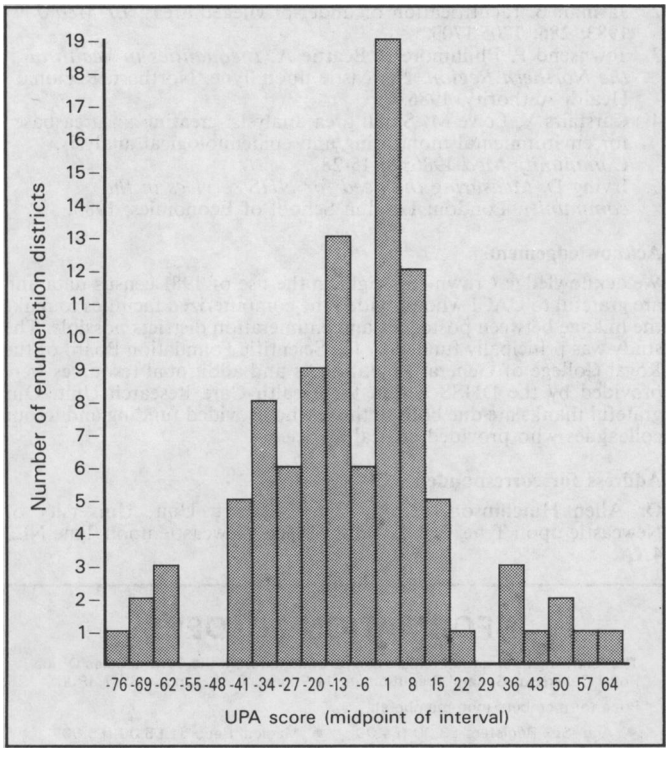

Figure 1. *Distribution of underprivileged areas (UPA) score across 98 enumeration districts.*

address the question of the value of the information. Since it currently takes a lot of effort and money to obtain the data, does it actually benefit a general practice to know the distribution and sociodemographic profile of its patients? Of course, many health professionals will feel that they can already differentiate between sub-populations in a practice. Unless survey data are used to quantify those differences, however, descriptions of small populations will always remain anecdotal.

We believe census data to be of particular benefit in two spheres: first, in allowing for comparison of the practice population with the general population of the area, and, secondly, in providing information which is valuable in planning services within the practice.

By comparing the relative proportion of practice patients in each enumeration district with the enumeration district scores it is possible to demonstrate whether a practice population scores differently from the general population. In addition to speculating why such differences might exist, a high score for a particular enumeration district might stimulate further analysis of individual variables, such as overcrowding or unemployment. By these means a detailed sociodemographic picture of the practice and its constituent enumeration districts is developed. In any negotiation for new resources, such as extra attached staff, information that the practice population is concentrated in the more disadvantaged enumeration districts could be crucial.

A decision on the balance within a primary care team might be similarly swayed. Since each general practice is a unique organization representing a mix of both the professionals and the population it serves, key decisions on how to arrange care might be made, or altered, in the knowledge that (say) 70% of people in an area did not own a car and were thus dependent on public transport when visiting the surgery. Indeed, Townsend and his colleagues[3] showed that not possessing a car very accurately reflected a high degree of both material deprivation and ill health.

Although the data in this paper were taken from the 1981 census, this study looks to the future. It is likely that changes within general practice and in information technology will enhance the value of data from the next census. Planning primary care at district level is now becoming a reality — indeed, in the area of the study practice, district-wide data collection and planning ventures have already been undertaken by the local medical committee. Census based demographic data will be a valuable addition to denominator and process data already available from family practitioner committees and, increasingly, general practitioners.

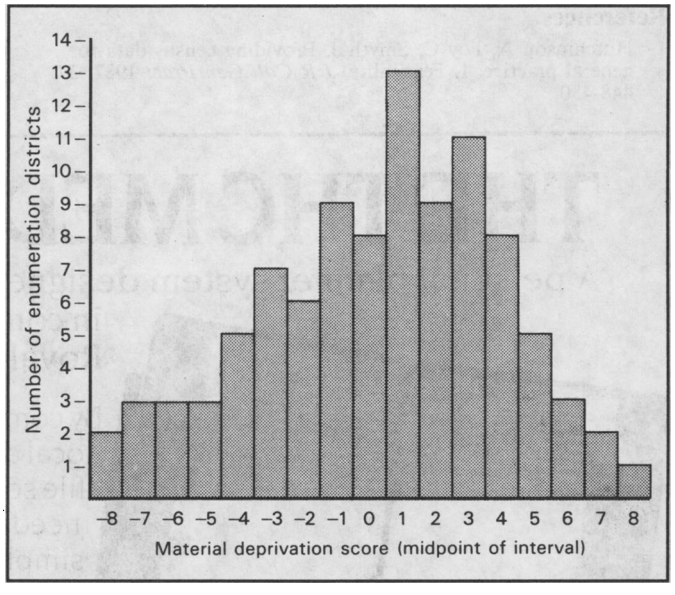

Figure 2. *Distribution of material deprivation score across 98 enumeration districts.*

Table 4. Comparison of underprivileged areas and material deprivation scores with practice population in enumeration district.

Measure of deprivation	Proportion of enumeration district population registered with practice	No. (%) of enumeration districts					
		Less underprivileged than average		More underprivileged than average		Total	
Underprivileged areas score	50% or more	12	*(57)*	9	*(43)*	21	*(100)*
	Less than 50%	46	*(60)*	31	*(40)*	77	*(100)*
	Total	58	*(59)*	40	*(41)*	98	*(100)*
Material deprivation score	50% or more	5	*(24)*	16	*(76)*	21	*(100)*
	Less than 50%	32	*(42)*	45	*(58)*	77	*(100)*
	Total	37	*(38)*	61	*(62)*	98	*(100)*

Development in information technology hold the key to this information explosion. The resources we expended in our feasibility study[1] to access census data make it an unrealistic proposition for an individual practice without further improvements in data handling. Such is the pressure for access to these data, however, that new software for microcomputers is already available commercially and there will undoubtedly be a range from which to choose by 1991. Of all the prospects we regard the proposals by the OPCS to link census data and postcodes as the most promising. Enumeration districts can suffer from the same inflexibility as wards, cutting across natural population groupings such as housing estates. By using groups of postcodes — aggregation will be required to preserve the confidential status of census data — an almost infinitely flexible access to small area statistics would be available. Artificial boundaries created by wards and enumeration districts would disappear and census data with real meaning to practitioners — relating to sections of their community which they can recognize — could be accessed for the first time. We hope the OPCS will proceed with the proposal.

As the outreach principle of viewing a general practice population as a clearly defined community has become established, so has the need for more specific and accurate data about that community. Computerization of age–sex registers, postcoding of patient addresses and improvements in information technology hold out the prospect of valuable demographic information previously unavailable to most general practitioners.

References

1. Hutchinson A, Foy C, Smyth J. Providing census data for general practice. 1. Feasibility. *J R Coll Gen Pract* 1987; **37:** 448-450.
2. Jarman B. Identification of under-privileged areas. *Br Med J* 1983; **286:** 1705-1709.
3. Townsend P, Phillimore P, Beattie A. *Inequalities in health in the Northern Region.* Newcastle upon Tyne: Northern Regional Health Authority, 1986.
4. Carstairs V, Lowe M. Small area analysis: creating an area base for environmental monitoring and epidemiological analysis. *Community Med* 1986; **8:** 15-28.
5. Irving D. *Measuring the need for NHS services in the community.* London: London School of Economics, 1983.

Acknowledgements

We acknowledge Crown copyright in the use of 1981 census data and are grateful to CACI who provided the computerized facilities to make the linkage between postcodes and enumeration districts possible. The study was principally funded by the Scientific Foundation Board of the Royal College of General Practitioners and additional resources were provided by the DHSS within the Health Care Research Unit. Our grateful thanks are due both to those who provided funding and to our colleagues who provided critical advice.

Address for correspondence

Dr Allen Hutchinson, Health Care Research Unit, University of Newcastle upon Tyne, 21 Claremont Place, Newcastle upon Tyne NE2 4AA.