Purpose

This is the fifth in a series of technical briefings produced by the Association of Public Health Observatories (APHO), designed to support public health practitioners and analysts and to promote the use of public health intelligence in decision making.

In this briefing we take a summary overview of segmentation and then focus on one approach to segmentation using geodemographic methods. We then compare the most commonly available geodemographic segmentation tools, highlighting the possible applications within the health sector and discussing some of the factors that should be considered when looking to invest in a system.

Further materials including tools to support our technical briefing series will be made available through our website at http://www.apho.org.uk

For more information about social marketing, including a case study database and other resources, go to the National Social Marketing Centre website at: http://www.nsmcentre.org.uk

Contents

Introduction 2
Geodemographic Classifications 3
Geodemographic Applications in Population Health 7
Summary 11

Authors

Jake Abbas
Helen Carlin
Anne Cunningham
Dan Dedman
Dominic McVey (National Social Marketing Centre)

Contributors

Mark Dancox
Paul Fryers
Clare Humphries
Simon Orange
Stuart Simms
Karen Tocque
Introduction

Geodemographics

Geodemographics has been defined as the ‘analysis of people by where they live’. The term has come into common use to describe the classification of small areas and the use of geography to help us draw general conclusions about the characteristics and behaviours of the people who live in them. The underlying premise is that similar people live in similar places, do similar things and have similar lifestyles – in other words, that ‘birds of a feather flock together’.

Already in widespread commercial use, the geodemographic approach is gaining currency in the public sector as a means of examining spatial patterns of crime, health and other social issues, and designing services to address them. The complex interplay between deprivation, the housing market, environmental issues and access to services ensures that ‘place’ remains an important factor in public health and health inequalities.

More subtle aspects of a neighbourhood, such as its socio-cultural features and reputation, may also have an impact on health. In health, as in other arenas, the identification of vulnerable neighbourhoods may pave the way towards area-based services and interventions.

Social Marketing and Insight

Social marketing is an adaptable approach which is increasingly being used to achieve and sustain behavioural goals on a range of social issues. The Government endorsed this approach in its White Paper Choosing Health: making healthier choices easier and set up the National Social Marketing Centre to lead on work in the field. The National Social Marketing Centre describes social marketing as “the systematic application of marketing alongside other concepts and techniques to achieve specific behavioural goals for a social good”.

Any social marketing intervention is predicated on gaining a deep insight into the citizen's life. Sir David Varney, in his Service Transformation review, defined insight as: “a deep truth about the citizen based on their behaviour, experience, beliefs, needs or desires, that is relevant to the task or issue and rings bells with targeted people”. One of his key conclusions for the public services was that “we need to exploit customer insight as a strategic asset”.

Social marketing interventions are built on a comprehensive understanding of the consumer and the drivers of behaviour change at the individual and societal level. Social marketing approaches not only include traditional public education mass media advertising campaigns, such as anti-smoking TV advertisements, but can also include interventions aimed at influencing policy and legislative change. For example, the ban on smoking advertising and the ban on smoking in public places were the result of concerted lobbying and campaigning, often applying social marketing techniques, to influence the policy and legislative agenda.

Social marketing techniques can also be applied to create a more supportive environment for individual behaviour change, e.g. designing workplace anti-smoking programmes, smoking cessation services and helpline support. Moreover, they have been used successfully to redesign services around customer need.

Multi-faceted social marketing interventions are developed against a set of criteria which define the process. The National Social Marketing Centre website at www.nsmcentre.org.uk contains many case study examples, along with a fuller description of each of the criteria. One of the key criteria which defines and underpins effective social marketing is the process of audience segmentation.

The Concept of Segmentation

Segmentation is a process of looking at the audience or ‘market’ and seeking to identify distinct sub-groups (segments) that may have similar needs, attitudes or behaviours. One of the central tenets of social marketing, it can be a powerful tool in helping to understand diverse sub-groups and focus resources where they are most needed.

We all regularly segment populations into groups. We talk about adults who are working and adults who are unemployed, single mothers who smoke and those who do not; and we subdivide these further by social class, ethnicity, level of income, use of public services, neighbourhood type, and by attitudes and motivation to change. The aim of any segmentation should be to define a small number of groups so that:

- all members of a particular group are as similar to each other as possible, and
- they are as different from the other groups as possible.

Traditionally, segmentation in the health arena has focused on the use of individual and household attributes such as age, gender, household composition, income, social class and physical status. However, adding in “attitudinal” and “psychographic” factors (personality traits, values, beliefs, preferences, habits and behaviours) results in a much more rounded understanding of subgroups of the population. This in turn will result in more appropriately targeted and tailored interventions.

An example of such segmentation underpins the Department of Health’s Ambitions for Health programme. The ‘Healthy Foundations’ model includes variables on claimed behaviour, environmental factors, lifestyle, personal motivation and health beliefs. The work will provide a detailed insight within each life-stage of how personal motivation and beliefs interact with environmental factors to encourage behaviour change. The completed segmentation model will be available in the autumn of 2009.

As the title suggests, this briefing will concentrate on geodemographic segmentation – the classification of
populations according to where they live. Multi-level modelling has confirmed that neighbourhood classifications have an explanatory power over and above that of the socio-demographic characteristics of the individual.\textsuperscript{8} In the world of public health, where we often struggle to obtain health and lifestyle information at the individual level, there is a growing appreciation of the potential of geodemographic segmentation tools to help us make inferences based on residence.

Geodemographic Classifications

History

An early example of geodemographic classification is Charles Booth’s 19th century ‘Poverty Map of London’, which allocated streets to one of seven classes ranging from ‘Lowest class, vicious, semi-criminal’ to ‘Upper-middle and upper classes, wealthy’.\textsuperscript{9} The geodemographic classification industry as we know it today can be traced back to the work of Webber in the 1970s, when he was commissioned by OPCS (the forerunner of ONS) to produce a classification based on the 1971 Census.\textsuperscript{1,9,10} As recognition grew of the commercial applications of geodemographic classifications, the following two decades saw the emergence of now familiar names such as ACORN and Mosaic, increasing competition in the marketplace and successive refinement of the product range with each new set of Census results. Since 2001, the commercial classifications have been joined by the freely-available Census-based Output Area Classification (see Figure 1).

A typical segmentation tool

As the market in geodemographic classifications has matured, we have come to expect the typical product to have certain characteristics:

- The classification will allocate neighbourhoods to categories with evocative names such as ‘Tower Block Living’ or ‘Affluent Blue Collar’.
- It will be available at various different levels of detail. Six or seven categories is considered a good number for mapping and data visualisation, around 20 for a conceptual understanding of a customer base, and anything up to 50 (probably without names) for more specialised purposes.\textsuperscript{11}
- It will provide narrative (and often pictorial) ‘profiles’ or ‘pen portraits’ of the typical environs and inhabitants of each neighbourhood type, describing everything from their socio-demographic characteristics to their choice of newspaper. These will often make use of data over and above those which were used to produce the classification itself.
- Commercial classifications nowadays all incorporate non-Census data at the derivation and/or description stages. Typical variables include house prices, unemployment, share ownership, TGI data and Council Tax band (TGI = Target Group Index, a large market research survey).\textsuperscript{10}

---

Figure 1 – The Output Area Classification (OAC) in and around the Wirral peninsula.
Creating a Geodemographic Segmentation Tool

Until recently, the methods employed in the creation of geodemographic classifications have largely remained a mystery, mainly due to commercial sensitivity between rival vendors. They all involve Cluster Analysis – statistical methods of grouping similar units into groups that are as different as possible from each other – but each vendor will add their own unique twist, such as different clustering algorithms or different sets of variables.

In Figure 2 we summarise the general process, based on the description by Vickers et al.\textsuperscript{11} Anybody trying to devise their own classification would require statistical expertise.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Creating a geodemographic classification.}
\end{figure}

- Variable Selection
  - Assemble variables which capture as much information as possible in the domains of interest.
  - Reject any which do not satisfy criteria such as timeliness, coverage, accuracy, validity.

- Data Reduction
  - The aim is to explain as much variability as possible with as few variables as possible.
  - If two variables are highly correlated (i.e. tell us much the same thing), discard one of them.
  - Statistical data reduction techniques can also be used to refine the list.

- Standardisation and Weighting
  - If the variables were used in their raw state, those involving big numbers would automatically dominate the classification. It is therefore usual to ‘standardise’ or scale them all to occupy a similar range. If there is a desire to give some variables more influence than others, differential weights can then be applied according to judgement.

- Cluster Generation
  - Clustering aims to allocate the areas to groups, in such a way that the areas within a group are as similar as possible, but the groups are as different as possible from each other.
  - Sometimes the number of clusters is pre-determined, but in practice, finding the ideal number of clusters is likely to be a matter of trial and error. There comes a point where nothing much is gained by adding more.

- Choose Number of Clusters
  - Each cluster is described in terms of the variables used to construct it and/or other variables.
  - Emphasis is often placed on those characteristics that make a cluster different from the average.
  - A name is chosen to encapsulate or stereotype the cluster in terms of its characteristics or typical residents.
Choosing a Geodemographic Segmentation Tool

With all the products that are now on the market (or available free), how is the user to make an informed choice between them? Some of the main characteristics of each are summarised in the table below:

<table>
<thead>
<tr>
<th>Supplier</th>
<th>CACI</th>
<th>CACI</th>
<th>Experian</th>
<th>ONS</th>
<th>ONS</th>
<th>Beacon Dodsworth</th>
<th>Acxiom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool</td>
<td>ACORN</td>
<td>Health ACORN</td>
<td>Mosaic Public Sector</td>
<td>Output Area† Classification (OAC)</td>
<td>2001 Area Classification</td>
<td>People &amp; Places P²</td>
<td>Geo/Household</td>
</tr>
<tr>
<td>Categories</td>
<td>Categorisation &amp; Nomenclature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Category (5)</td>
<td>Group (4)</td>
<td>Group (11)</td>
<td>Super Group (7)</td>
<td>Super Group (7)</td>
<td>“Tree” (14)</td>
<td>Behavioural Category (4 life stage/ 5 affluence)</td>
</tr>
<tr>
<td></td>
<td>Group (17)</td>
<td>Type (25)</td>
<td>Type (61)</td>
<td>Group (21)</td>
<td>Group (20)</td>
<td>“Branch” (41)</td>
<td>Group (20)</td>
</tr>
<tr>
<td></td>
<td>Type (56)</td>
<td>Sub Type (59*)</td>
<td>Sub Type (243)</td>
<td>Sub Group (52*)</td>
<td>Sub Group (53*)</td>
<td>“Leaf” (157*)</td>
<td>Behavioural Cluster (60/52)</td>
</tr>
<tr>
<td>Some areas unclassified?**</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Top Tier categories reflect:</td>
<td>Affluence</td>
<td>Health Outcome</td>
<td>Deprivation, Affluence &amp; Lifestyle</td>
<td>-</td>
<td>-</td>
<td>Affluence</td>
<td>Behaviour, Lifestyle and Affluence</td>
</tr>
<tr>
<td>Data Used</td>
<td>2001 Census data, proprietary survey, and other data</td>
<td>2001 Census data, proprietary survey, and other data</td>
<td>2001 Census data, proprietary survey, and other data</td>
<td>2001 Census data only</td>
<td>2001 Census data only</td>
<td>2001 Census data, proprietary survey, and other data</td>
<td>2001 Census data, proprietary survey, and other data</td>
</tr>
<tr>
<td>Variables</td>
<td>All companies utilise UK 2001 Census data, and may also use other sources such as electoral roll, consumer credit activity, Post Office address files, land registry house prices and council tax information, and ONS local area statistics. Some companies may also use proprietary survey data, further details of which can be accessed using the links at the bottom of the table.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>Annual licence, free for teaching purposes only (via Essex Data Archives)</td>
<td>Annual licence, free for teaching purposes only (via Essex Data Archives)</td>
<td>Household &amp; Postcode: annual licence, LSOA using aggregated data: free for academic use</td>
<td>Free</td>
<td>Free</td>
<td>OA: annual licence, LSOA using a rebuild of data: free for NHS</td>
<td>Annual licence &amp; free trial</td>
</tr>
<tr>
<td>Smallest Geographical Level</td>
<td>Postcode</td>
<td>Output Area</td>
<td>Household &amp; Postcode</td>
<td>Output Area</td>
<td>Super Output Area**</td>
<td>Postcode (built around Output Area)</td>
<td>Household &amp; Postcode</td>
</tr>
</tbody>
</table>

†Output Areas are the smallest areas for which Census data are published. *Categories left unnamed. **Some tools leave areas ‘unclassified’ if they do not fit easily into a category. Examples might include areas with an unstable population, or communal establishments such as halls of residence, prisons and army barracks. ***The ONS 2001 Area Classification is a blanket term for classifications ranging from LA level down to OA level. The OA Classification is considered in a separate column. Classifications for areas larger than an SOA are less useful for geodemographic purposes because of the wide variety of neighbourhoods contained within them.
It is important that the user understands some of the potential limitations of the products and is able to ask appropriate questions to make an informed decision. Some of these will relate to the specifics of how the segmentation tool works, and some will be about the more strategic aspects of embedding it in the organisation:

*Figure 3 – Investing in geodemographic tools – asking the right questions.*

Agonising unduly over the choice of segmentation tool could however be counterproductive. One authority on the subject gives the useful advice that the differences between the products do not actually matter all that much – ‘they will all do the job’. The important thing is to start reaping the benefits which geodemographics can provide. If in doubt, one route would be to start with the free products available from ONS.
Geodemographic Applications in Population Health

1. Population Health Profiling

Geodemographic classifications can help us towards a better understanding of the problems and needs of our population, and which groups within it are particularly in need of help. They may also be available for areas smaller than an LSOA, permitting the detection of inequalities that might otherwise be missed and finer targeting of interventions to overcome them.

Case Study (i) – Access to Healthy Foods in Great Yarmouth & Waveney

Great Yarmouth and Waveney PCT investigated the low consumption of fruit and vegetables in parts of their area. They made use of the fact that the Health Survey for England, which had asked about fruit and vegetable consumption, was coded to the Mosaic classification (see Box 1). People consuming not even one portion of fruit or vegetables per day were concentrated in six Mosaic groups nationally, which were then mapped locally (see Figure 4).

Case Study (ii) – Alcoholic Liver Disease

This study sought to establish whether geodemographic segmentation could effectively pinpoint heavy episodic drinkers, in order to target them with a social marketing programme. Mosaic codes were added to Hospital Episode Statistics (HES) data on admissions for alcoholic liver disease. The admissions proved to be concentrated in a few geodemographic groups, typified by low levels of income, social class, education and social cohesion (see Figure 5). Interventions could be targeted at these groups right down to postcode level.

Figure 4 – Lowestoft: Great Yarmouth & Waveney PCT’s analysis of location of Mosaic groups most likely to eat no fruit or veg (red and orange dots).

Figure 5 – Some of the Mosaic groups which are under- or over-represented among people hospitalised for alcoholic liver disease.

Box 1. Predicting characteristics based on geodemographic type

If a survey collects respondents’ postcodes along with data on, say, their smoking habits, then an estimated smoking rate can be worked out for each geodemographic group. These estimates may then be generalised to all neighbourhoods in the same group, wherever they occur. In Case Study (i), the Mosaic classification attached to the Health Survey for England enables its findings about healthy eating to be applied to small neighbourhoods in a locality which may not even have been sampled. The British Crime Survey was one of the first to have Mosaic classifications attached to it, and the ONS is now committed to attaching the OAC classification to its Expenditure and Food Survey, English Housing Survey, and Family Resource Survey, so that their results can be applied locally in the same way.

Even if our data are complete, we may still want to generalise from them if we suspect that they represent the ‘tip of an iceberg’. In Case Study (ii), HES admissions for alcoholic liver disease are matched to Mosaic categories to find out which ones are most over-represented (Figure 5). Focusing efforts to tackle alcohol problems upon these types of neighbourhood may help to reduce the need for such hospitalisations in future.

These generalisation techniques can be valuable when local data simply do not exist, or would be unreliable due to small numbers. Linked data may also be less influenced by local peculiarities in data recording or service provision, and thus be more reflective of true need.

Source: Winters & Barnes. Copyright Experian Ltd.

Source: Powell, Tapp & Sparks. Copyright Experian Ltd.
2. Service provision and utilisation

Geodemographics can help a service to understand its customer base, and/or their level of usage. By comparing service usage with need, we can identify any inequity of provision.

Case Study (iii) – GP practice profiles in Yorkshire & Humber

At the simplest level, radar plots prepared for each GP practice in Yorkshire & Humber show how the patients on their list are distributed between Health Acorn groups. In this example, the largest single group is category 2.5 – “disadvantaged multi-ethnic younger adults, with high levels of smoking” (see Figure 6).

Case Study (iv) – take-up of cervical screening

A cervical screening study in Yorkshire & Humber compared the geodemographic profile of unscreened women with that of all eligible women aged 25-34. Figure 7 presents the results as index numbers, indicating that ACORN Group N (‘Struggling Families’) had the greatest over-representation of unscreened women. When charts are drawn for each individual PCT, a similar pattern emerges. This means that further qualitative research focusing on Group N need only be undertaken once, in the knowledge that the findings will help to inform the design of appropriate interventions in every PCT.

Case Study (v) – Bradford A&E attendance

In an ongoing exercise aimed at optimising access to emergency care by those who need it, YHPHO has produced geodemographic profiles of those attending A&E departments in West Yorkshire. It may be that some users could more appropriately consult their GP or pharmacy instead.

Figure 8 shows the index of individuals attending A&E in Bradford by OAC Group (see Box 2). This may help us to generate hypotheses and focus our research effort. Clusters with both high indices and high populations – e.g. cluster 7a (‘Asian Communities’) – would be particularly worth following up, perhaps via a focus group or other qualitative approach, to try and understand their heavy use of A&E.

Case Study (vi) – Admissions for mental health conditions

A study in the North West made use of the fact that the ‘People & Places’ (P) categories can be ranked in order of affluence. This lets us not only see the inequalities between categories, but identify any categories which ‘buck the trend’. In Figure 9, it can be seen that the ‘New Starters’ category (circled) has an admission rate well above what might be expected given its level of deprivation. It would be informative to try and work out why that might be, as a first step to designing suitable services to meet the needs of this group.

Figure 6 – Radar plot showing how practice patients are split between Health Acorn groups.

Figure 7 – Women aged 25 to 34 in Yorkshire and Humber - uptake of Cervical Cancer Screening.

Box 2. Index numbers

In Case Study (iv), Group N accounted for 25% of unscreened women, but only 18% of eligible women overall. This has been summarised in Figure 7 by dividing one percentage by the other, and multiplying by 100, to obtain an index. Group N has an index of (25/18)*100 = 139. Hence 100 is the average percentage of women unscreened in the whole population, and an index of 139 indicates that group N has 39% higher percentage unscreened than the average.

These index numbers are commonly used in geodemographic work, but should be treated with caution: when comparing proportions or percentages like this, odds ratios should be used rather than these prevalence ratios, but the two are similar if the percentages are small (i.e. much closer to 0% than 100%). Such index numbers (or odds ratios) do not tell us the rate or propensity in absolute terms, but can help us identify which groups have more of a characteristic than others.
Box 3. Age standardisation

When comparing rates for different geodemographic groups, it may not always be obvious whether to age-standardise them first.

If the objective is to identify places with the greatest absolute need, then it may be appropriate to use ‘raw’ rates, without standardising for age. This will often highlight localities with high numbers of older people as being in greatest need. If we want to study inequalities over and above the effects of age, it would be better to age-standardise the rates. This was the chosen approach in Case Study (vi).

Case Study (vii) – Smoking in Nottingham

This Health Equity Audit into the city’s ‘New Leaf’ Stop Smoking Service began by linking the Mosaic classification nationally to TGI survey data on smoking prevalence. Some of the groups with the highest smoking index were substantially over-represented in Nottingham. ‘New Leaf’ clients were then matched to Mosaic groups on the basis of their postcode, and a reassuringly high proportion were found to come from the groups most in need of the service. There were a few ‘cold spots’, principally in central Nottingham, where uptake of the service was low even in those groups. Work is now under way with the GP practices in those areas to address this gap.

3. Targeting your intervention

Having identified the geodemographic make-up of the population you wish to target, the information which the segmentation tool provides about where and how those groups live, work, shop and play can be used to refine strategies for consultation, health promotion and service delivery. There are increasing opportunities to incorporate ‘insight’ from studies other than your own (see Box 4).

Box 4. Generating ‘insight’ – linkage of datasets

When attempting to tailor a social marketing approach to a particular geodemographic group or groups, it is of great benefit to be able to ‘unlock’ the body of research already coded to that same segmentation tool. For example, a Mosaic user can cross-reference their own findings about each group against a growing database of characteristics, covering everything from fear of the dark, to viewing habits, to participation in Bingo. This can help to answer questions such as whether readers of the Daily Telegraph are more or less likely than average to have diabetes (or vice versa). Access to such a library of data does not usually come free, and the skills to link it with health statistics in a meaningful way are still in short supply. Analysts must be careful to avoid the ‘ecological fallacy’ – the unfounded assumption that the general characteristics of an area will apply to any particular individual within it. However the National Social Marketing Centre is working together with APHO to ensure that the potential for insight generation is fully realised.
Case Study (i) cont – Access to Healthy Foods in Great Yarmouth & Waveney

Analysis of the local maps of Mosaic groups likely to consume the least fruit and vegetables showed that these areas were not particularly deficient in fresh food outlets. Various possible explanations for the low uptake were floated, including high cost, inconvenience, and lack of food preparation knowledge. The ‘pen portrait’ supplied for each geodemographic group helped the team to choose consultation methods to which the population was likely to respond, and ultimately to design a mobile ‘fruit & veg van’ service suited to their needs.

Case Study (viii) – Snack Right

The Snack Right social marketing campaign was undertaken by the Cheshire and Merseyside Public Health Group (CHAMPS), to address a perceived need to reduce consumption of unhealthy snacks by pre-school children during the ‘danger period’ between the end of organised day care and teatime. Mosaic was used to build up a picture of the target population of families with young children in deprived areas. This knowledge of the relevant Mosaic segments informed the design and conduct of focus group work, and also helped to identify potential retail partners to involve in the campaign.

Case Study (ix) – Spatial Targeting of Area Based Initiatives

Regeneration programmes often take the form of Area Based Initiatives, confined to a boundary drawn on a map. Geodemographics can be used to assess how well this boundary reflects the intended target population. In a study of Sure Start boundaries, a judgement was made as to which P² categories the scheme should be aiming to reach. The ideal scenario would be for everybody living inside the Sure Start boundary, and nobody outside it, to belong to these priority groups. However this would be unlikely to produce a sensible, contiguous boundary. In Nottingham, the breakdown was as shown in Table 2:

Table 2 – Distribution of priority and non-priority groups relative to Sure Start boundary in Nottingham.

<table>
<thead>
<tr>
<th></th>
<th>Within SureStart boundary</th>
<th>Not within boundary</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority P² groups</td>
<td>42,648</td>
<td>40,294</td>
<td>82,942</td>
</tr>
<tr>
<td>Non-priority groups</td>
<td>31,090</td>
<td>141,956</td>
<td>173,046</td>
</tr>
<tr>
<td>Total</td>
<td>73,738</td>
<td>182,250</td>
<td>255,988</td>
</tr>
</tbody>
</table>

4. Evaluation/reach analysis of targeted interventions

After our targeted intervention has been running for some time, we can evaluate whether it has reached its intended audience, using the techniques already discussed for analysis of service utilisation.

Case Study (x) – Reach Analysis of Sure Start Programmes

As part of its evaluation of Sure Start programmes, Kirklees made use of geodemographics to identify whether those using Sure Start programmes (as opposed to just living within the boundary) were families who may be most in need. Sure Start programmes are not necessarily used by all the families within the defined area, and are also used by families living outside it.

Data matching of Child Health records with the records of those using the Sure Start programmes enabled comparisons to be made between the profile of children living in a Sure Start area and those making use of Sure Start services. This type of analysis allowed Kirklees to see where families were using services and where they were not. The value of geodemographic analysis was that it allowed managers to see at very local levels where families who were equally deprived but with very different cultural backgrounds were missing out. Further qualitative work could then be undertaken with these separate communities to gain a better understanding of how services could be developed to be more attractive to non-users.

5. Further reading

The APHO website (www.apho.org.uk) contains links to a range of further resources relating to geodemographic tools and their applications in public health.
Summary

Public sector use of geodemographic segmentation tools has lagged behind that of the commercial sector, but is now beginning to catch up as analysts in health and related fields overcome their suspicion of market research techniques. The methodology behind many commercial models remains largely confidential, but free and transparent tools are now available from the ONS. In public health, the use of geodemographic segmentation is still in its relative infancy, but is attracting an increasing amount of interest.

There is a temptation to regard any new approach as a panacea, but as the case studies in this Technical Briefing will have illustrated, the result of applying a segmentation tool to a local or regional issue is rarely an end in itself. Rather it serves as a means of helping us to formulate relevant hypotheses, and focus our subsequent rounds of research, consultation and service design on the most appropriate target groups.

The novelty of geodemographics means that it is also an area of skills shortage, particularly when it comes to extracting maximum value by using it as a key to the growing repositories of ‘insight’ based upon each rival classification. This is where the National Social Marketing Centre and APHO can offer valuable advice and support, helping geodemographic segmentation to realise its true potential in the effort to improve health and tackle health inequalities.

Glossary and Abbreviations

**Algorithm**: A mathematical or statistical procedure, often consisting of a repetitive sequence of steps.

**Cluster Analysis**: Methodology for grouping similar units (e.g. neighbourhoods) together into groups or ‘clusters’, which are as different as possible from each other.

**COPD = Chronic Obstructive Pulmonary Disease.**

**Correlation**: Statistical measure of the strength of the relationship between two variables.

**Data Reduction**: Use of statistical techniques such as Factor Analysis or Principal Components Analysis to summarise data using as few variables as possible.

**Discrimination**: In this context, the extent to which a few geodemographic categories capture the bulk of cases of a particular characteristic of interest, as opposed to it being spread evenly across them all.

**Ecological Fallacy**: The assumption that the characteristics typifying a geographical area apply equally to every individual living in that area.

**Health Equity Audit**: An assessment of how fairly health services and resources are distributed in relation to the needs of different groups.

**Homogeneous**: Similar to each other.

**Multi-level Modelling**: Statistical technique for analysing situations where individuals belong to groups, which in turn belong to bigger groups.

**ONS = Office for National Statistics.**

**OPCS = Office for Population Censuses and Surveys (forerunner of ONS).**

**Output Area (OA)**: Smallest area for which 2001 Census results are released. OAs are based on postcodes and fit within 2003 ward boundaries. There are 165,665 OAs in England, each containing on average 300 people.

**Reach Analysis**: An assessment of whether a targeted intervention is being taken advantage of by the people it was intended to help.

**Standardisation**: Manipulation of figures to remove the influence of differing units of measurement, or factors such as age (age standardisation). For advice on methods of age standardisation, see APHO Technical Briefing 3: Commonly used public health statistics and their confidence intervals.

**Super Output Area (SOA)**: Statistical areas built from groups of OAs. There are 32,482 Lower Super Output Areas (LSOAs) in England, each containing on average 1500 people, which combine to form 6780 Middle Super Output Areas (MSOAs) with an average population of 7200. The ONS Beginners’ Guide to Geography can be found at http://www.statistics.gov.uk/geography/beginners_guide.asp (accessed 9/4/09)

**Targeted Intervention**: An intervention which is intended to benefit a particular group of people.
References


All links accessed 9/4/09

About the Association of Public Health Observatories (APHO)

The Association of Public Health Observatories (APHO) represents and co-ordinates a network of 12 public health observatories (PHOs) working across the five nations of England, Scotland, Wales, Northern Ireland and the Republic of Ireland.

APHO facilitates joint working across the PHOs to produce information, data and intelligence on people’s health and health care for practitioners, policy makers and the public.

APHO is the:

- single point of contact for external partners
- learning network for members and participants
- advocate for users of public health information and intelligence

Further information about APHO and the work of PHOs can be found at http://www.apho.org.uk

Updates and more material, including methods and tools to support our Technical Briefing series will be made available through our website at http://www.apho.org.uk

About the National Social Marketing Centre

The National Social Marketing Centre is a strategic partnership between the Department of Health and Consumer Focus (previously National Consumer Council). More information about social marketing including a case study database and other resources can be found at: http://www.nsmcentre.org.uk

Contact details:
Dominic McVey
Director of Research and Development
National Social Marketing Centre
20 Grosvenor Gardens
London, SW1W 0DH

For further information contact:
Association of Public Health Observatories
Alcuin Research & Resource Centre, University of York, Heslington, York, YO10 5DD

Telephone: 01904 328216
Fax: 01904 321870
http://www.apho.org.uk