Chapter 4: Geographic Information Systems in Spatial Epidemiology and Public Health

R.Haining and R.Maheswaran.

1. Introduction: what is a Geographical Information System?

Definitions of Geographical Information Systems (GISs) usually fall into one of three categories: toolbox-, database- and organization-based definitions (Burrough and McDonnell, 2000). Geographical or spatial data are data about entities in the real world (physically represented as point, line or area objects) that define their location and the attributes recorded at each location. Knowing where entities are located allows spatial relationships between them to be defined such as distances between points, adjacency or otherwise of areas, the proximity of one object to another. A GIS database is distinguished from most other databases by knowing where in geographical space objects (points, lines and polygons) are located in relation to one another since most other kinds of databases capture only entities and their attributes. Database definitions of GISs emphasize this difference: "any… computer based set of procedures used to *store and manipulate geographically referenced data*" (quoted in Burrough and McDonnell, 2000, p.11, italics added). Geography is stored in the form of a series of discrete layers in the database (a layer for the roads, a layer for the waste sites, a layer for the forested areas, a layer for the deprivation scores by census area for example) enabling different features to be switched on or off when visualizing or mapping an area.

Toolbox-based definitions, on the other hand, emphasize system functionality and the place of GISs within information technology: "*a powerful set of tools* for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes" (quoted in Burrough and McDonnell, 2000, p.11, italics added). Burrough and Mc Donnell (2000, p.15) list some of the basic operational requirements for a GIS. These include being able to: show the locations of entities individually and in relation to others ("identify all the areas within a 15 minute travel distance of location x"); compute the physical size of areas; show the result of intersecting or overlaying different layers of spatial data (for example air pollution data, socio-economic data, facilities data); count up the number of cases of an entity within a given distance; determine paths of least cost or least resistance over a surface or network. The different layers in the database can be overlaid and/or buffered in order to respond to such queries which may be thought of as cartographic modelling. Certain forms of statistical modelling may also be available inside a GIS. For example, the ArcGIS software includes Geostatistical Analyst which enables various forms of spatial statistical analysis to be executed within the GIS.

GIS capabilities, in terms of database management and toolbox functionality, have been exploited in many application fields (see for example Burrough and McDonnell, 2000, p.9). In spatial epidemiology GIS capabilities have been used to capture the geographical distribution of disease (and how that distribution changes over discrete periods of time) and the relationship between the occurrence of a disease and various environmental as well as social and economic factors. In public health research however, interest often focuses on the ways GIS can also contribute to: planning public health services and interventions; improving access to health care services; assessing the locational impacts of health policy; facilitating community participation (by groups or by individuals) in addressing local health concerns. When applied to these sorts of questions then a GIS meets the organization-based definition of a GIS: "*a decision support system* involving the integration of spatially referenced data in a problem solving environment" (quoted in Burrough and McDonnell, 2000, p.11, italics added). Cromley and McLafferty (2012, p.14) remark: "GIS, as a means of exploring health problems and finding ways to address them, has taken its place in the conceptual and methodological foundations of public health." GIS literacy and GIS capability has become important for spatial epidemiology and public health research and practice.

In the context of this handbook however there is another definition of a GIS: as an important enabling technology in the implementation of Geographic Information *Science* (GISc). Here we take GISc to refer to all those areas of scientific enquiry in which *where* events occur matters, perhaps for the purposes of description, explanation or prediction of those events. It follows that the acquisition and processing of spatially referenced data are of fundamental importance to the progress of those areas of science and a GIS facilitates key aspects of the handling and processing of that data. Both spatial epidemiology and public health research draw on three key aspects of GISc: spatial database management, spatial data visualization and mapping and finally spatial analysis. The latter includes cartographic and topological analysis (e.g. measuring distances and areas; spatial relationships amongst observations including overlay and buffering operations), mathematical modelling (e.g. network and surface analysis) and spatial statistics. Particular GIS products, of which there are a large number, are differentiated by how they manage spatial data and the functionality they possess.

A GIS is more than a computerized or automated mapping system and moreover, as emphasized by Cromley and McLafferty (2012, p.16), should be seen as part of a "larger constellation of computer technologies for capturing and processing geographic data" which includes the global positioning system (GPS), satellite data collection systems and digital scanners. This constellation of computer technologies used to process geographical data is therefore wider than proprietary GIS software packages (e.g. ArcGIS, MapInfo) and includes databases and statistical packages (e.g. Microsoft Access, SAS, R, Stata, WinBUGS). Database functions within such packages may be used to link health datasets with datasets that already contain geographically located information (e.g. grid coordinates, census output areas, environmental exposure values). Statistical modelling of data may be undertaken within statistical packages and while this functionality may be considered to be part of what is involved when working in a GIS or GISc environment, standard GIS packages at best have only limited spatial statistical modelling capability. In several of the examples quoted in Sections 4 and 5, various computer software packages are used to process and analyse data and these could all be regarded as "using GIS within the field of GISc".

Physically a GIS comprises: computer hardware (e.g. networked computers; large format scanners and printers), software (e.g. that enables data input and output, storage and database management) and an organizational structure that includes skilled people who are able to operate the system. But growth of the internet means users of GIS do not need to have their own in-house, physical, GIS in order to access geographical data or the results of spatial data processing or even to undertake data analysis. Distributed GIS services have made it possible for many more users to take advantage of GIS capability. Peng and Tsou (2003) identified applications supported by distributed GIS: data sharing (both the original data and "data about data" - metadata); information sharing (online publishing of the results of data analyses and online accessing of servers running a GIS application that can process a client's request for information and return the results); data processing (providing online access to GIS analysis tools that can be applied to a client's data); location-based services (providing information on a client's local environment or where they can best access particular health services mirroring developments in, for example, retailing). More recently it has become possible for users to post their own map data and to annotate them online. Mashups allow data from multiple sources to be integrated and mapping mashups using Google Maps is an example of this development (Cho, 2007). Cromley and McLafferty (2012, p38-41) provide a brief overview of this evolving area of GIS. These developments have the potential to contribute to the development of public participation GIS (PPGIS) – systems that promote the participation of communities and individuals in raising and tackling issues of local concern (see for example Sheppard et al. 1999, and for an example of PPGIS in health research and planning see Beyer and Rushton, 2009). Chapter 1 of this volume contains detailed discussion of the important integrating contribution geospatial methods and GIS in particular can make to population level studies of disease risk and public health.

2. The world as captured in a Geographical Information System database: abstracting reality

The database sits at the heart of a GIS and for this reason much attention has been given in the scientific literature on GIS to the relationship between geographical reality and what is stored in the spatial database. An understanding of the processes that take us from the limitless complexity of the real world to a database with a finite number of spatially referenced bits of information about that world enables the user to exploit GIS and implement GISc in a critical and hence rigorous way. In this section we describe the two processes of conceptualisation and representation which take us from geographical reality to a model of that reality that becomes an essential part of the model for the GIS database. The other two ingredients of that model are the attributes that are stored and the topological relationships. Subsequently measurements are taken so that the database can be populated with data.

Figure 1 shows a two-step process in which at the first step observed phenomena in the world are conceptualised in terms of either a field (or continuous surface) view or an object (or discrete space) view. Environmental phenomena (air quality, soil type, temperature) are usually conceptualised as fields. For example ground level air quality is conceptualised as a field because it is possible to go to any location on the earth's surface and measure air quality. On the other hand a house or a hospital, a river or a road, a reservoir or a waste site,

Figure 1 here

are all conceptualised as objects. Depending on the map scale they would be conceptualised as point, line or polygon (area) objects. For such phenomena the real world is conceptualised as an empty space populated by these discrete, identifiable objects to which attributes are attached. These two models are how the geographical world is conceptualised and with few exceptions all geographical phenomena are modelled in terms of one or other of these two views. One important exception is population data which are sometimes conceptualised as fields (e.g. when constructing a population density map) or as objects (e.g. when producing a dot map). The next question is how these two conceptual models are represented for the purpose of constructing a model for the spatial dimension of the database which of necessity must comprise a finite number of bits of spatial data.

In the case of the object view, each object (whether point, line or polygon) is located on the earth's surface and attribute values attached to them and topological relationships defined for them. The objects, of which there are a finite number, underpin the model for the spatial database. But other issues can affect the representational choice. For example, although household data refer to point objects, for confidentiality reasons much of this data is available only at the aggregated level – that is for (irregular) census areas.

Field data present a different challenge. In order to capture field data in a (finite) spatial database the surface needs to be sampled. This sampling may take the form of a regular or irregular sample of points (point data), a series of contour lines linking locations with the same value of the attribute (line data) or a set of polygons that partition the space. These polygons may be regular or irregular in shape. Regular polygons are constructed independently of the attribute as in the case of a matrix of regular pixels ("picture elements") of given size from a remotely sensed image. Associated with each pixel value is a single value which provides a measure of the attribute at that location. Landsat Thematic Mapper data for example, which provide reflectance values for each pixel which are then classified in terms of, say, land cover features, cover a ground dimension of 30 metres by 30 metres. This defines the data's spatial resolution and any landscape or environmental feature smaller than that will not be detected. The larger the pixel size the larger the spatial filter and the smoother any landscape will appear. The location of each pixel is given by its position in the row and column of the matrix from which topological relations can be defined (Haining, 2003, p.78-79). Irregular polygons that partition an area arise when following attribute boundaries such as the edge of a forested area or the edge of a built-up area. The TIN (triangulated irregular network) is also an example of irregular polygons that partition a study area. TINs are constructed from a set of sample points on a surface. The TIN is used to capture surface variation and is often used to describe topography capturing both surface slope and aspect. Because field data are often captured by sampling this raises the question as to what sort of sample plan to adopt (including what type of sampling design and sample point density). It also introduces sampling error into any analysis based on such data as well as the scale and partition effects noted above in relation to aggregated object data (Haining, 2003, p.100-113).

Inside any GIS, map data are stored and processed in one of two main ways – the tessellation data model and the vector data model. The interested reader is referred to any standard text on GIS for a description of these two data models. At this point we merely note that both the field and object views of geographical reality can be captured by either of these two GIS data models. That said, much but not all environmental data are captured using the tessellation data model and most commonly within that set of data models the raster model structured as a regular array of square pixels. This is because much environmental data are obtained via remote sensing methods (aerial photographs and satellite imagery). Point and area based population data and network data (e.g. lines of transportation) are usually captured using the vector data model. In the case of population data this is because reporting takes place through irregular census areas.

We now give a few examples to illustrate datasets that could be stored in a GIS database each geographically referenced attribute forming a layer within the GIS database.

- We have data on the set of GP clinics for a region, represented as points with fixed locations from which can be derived topological relations. Attached to each clinic there are a series of attributes (patient lists with addresses and who have recently had tests for high blood pressure or been screened for a form of cancer).
- We have data on the size of the susceptible population in a region represented as counts by census areas. Attached to each census area is also the count of the number of cases of the disease during a given interval of time. There are also data recording social, economic and demographic data for the resident population of each census area. Finally there are data obtained from remotely sensed imagery that measure environmental (e.g. land cover) conditions across the area. Topological relationships across the set of census areas might be defined in terms of which census areas share a common boundary.
- As part of an emergency response analysis and establishing a regions vulnerability to a disaster, we have data on the location and capacity of hospitals, the regional road network (with travel times) and a map of environmental risk (arising from say the risk of flooding or fire) which partitions the map into areas classified by whether they have high to low levels of vulnerability. So the map objects in this case are points, lines and areas and topological relations between the hospitals and between the hospitals and the areas with defined levels of environmental risk are specified through the road network.

Assigning locations to objects is central to the creation of a geographical database. This creates challenges for the researcher. Whilst the location of a hospital or a road may be fixed for the period of a study, in studies of environmental exposure even the daily movement patterns of populations can affect exposure levels (e.g. to air pollutants) and in studies of the distribution of cases of a chronic disease the resident histories of populations becomes important (see chapter 34 for an extended discussion). In both cases the analyst is forced to consider the validity of assigning populations to locations based on, for example, their current residential address. Measuring location may be problematic in other ways. Map projection issues do not need to be considered when undertaking large scale studies (analysing small areas in high detail) but will be an issue when undertaking small scale studies (analysing large areas, say at the scale of Europe or the continental USA, in limited detail) particularly where different data sets need to be integrated or overlaid.

In this section we have focused on how the geographical world is conceptualised and represented for purposes of storage in a GIS database. But the database also comprises the attributes that are attached to each location. These attributes go through their own processes of conceptualisation and representation. Many studies in epidemiology need to control for deprivation since deprivation often acts as a confounder in environmental exposure-response studies. There are many forms of deprivation such as material (economic) deprivation and social deprivation. Before such terms can be used in any study it is necessary to consider how deprivation is to be conceptualised and then how that conceptualisation is best represented – that is, how it will be measured. Deprivation may be a confounder at the level of the individual but it may also be a confounder at the area level. Some area level attributes represent aggregates of the resident population (proportion of the population unemployed or living below the poverty line) whilst some attributes are only defined at the area level – for example social cohesion or social capital both of which need to be conceptualised but when measured are attributes of groups of people and communities, not of individuals.

The model for the GIS database is defined by how geography is captured, how spatial relationships are defined between the finite elements representing that geography, what attributes are included in the database and how they are to be measured. Although a GIS database cannot be used to model attributes over time, attributes will refer to points in time or intervals of time and so time needs to be made explicit. Subsequent use of the database may produce maps comparing disease rates in a population between one time period and another.

All models are simplifications of the reality they describe and involve a trade-off between descriptive power and model complexity with diminishing returns (see figure 2). We can speak of the quality of the model for the GIS database but this cannot be assessed independently of the planned use of the database. The model is chosen on the basis of whether it is fit for purpose recognizing that the quality of the model may not be uniform across all parts of the study area (for example, across both urban and rural areas of a study region). Model quality is assessed by such criteria as: resolution (is the spatial detail adequate?); completeness (what are all the attributes necessary for undertaking a robust study into, for example, cancer incidence by small area and are these data available? (Swerdlow, 1992, Wakefield and Elliott, 1999)); as well as clarity, precision and consistency. For extensive discussion of these terms the reader is referred to Guptill and Morrison (1995).

Figure 2 here

We now turn from considering the model for the spatial database to issues surrounding the quality of the data that populates the spatial database.

3. The Geographical Information System database and data quality.

Figure 3 describes the set of relationships in the construction of the GIS database and the terms often used to describe those relationships. In the case of data quality there is a different type of trade-off to that in the case of model quality. Now the trade–off is between data quality and cost. Whilst the costs of data acquisition may be falling on a per unit basis as more methods are routinely employed to capture data, it is still the case that if data quality demands for a project grow the cost of acquiring that data, particularly if it has to be primary data, grows.

Fig. 3

A single item of geographical data comprises the triple: "what, where and when" – the attribute, the location to which the measurement refers and the time period to which it refers. It is sometimes referred to as the geographical space-time data cube. Veregin and Hargitai (1995) identify four primary dimensions of data quality: accuracy, resolution, consistency and completeness thus giving rise to a 4-by-3 data quality matrix. Rather than try to address the full matrix, discussion will focus on the four primary dimensions choosing examples to illustrate the point. It should be noted that, as with model quality, data quality may not be uniform across the study area particularly if the study area is large. There can be many reasons for this. In the case of mortality data it can be due to diagnostic 'fads' and other biases in specifying the cause of death (Lopez, 1992). We briefly discuss each of the four dimensions.

All measurements contain error due to the inevitable imprecision associated with the process of taking a measurement. Other types of error (inaccuracy) associated with spatial data include: sampling error; operator error and even deliberate error (as in the case of Barnardisation of UK census data). A particular concern with GIS databases is locational error when, for example, a patient is assigned to the wrong postal code or zip code. Underlying the idea of "data error" is the assumption that there is a true value. When drawing on socio-economic data it is not always clear that this is so – what is the "true" level of social capital or deprivation in an area? In large databases it is often necessary to resort to automated methods to try to flush out possible data errors including screening for both distributional and spatial outliers – values that are very different from their neighbours. Errors not only infect data values but can also propagate as a result of such GIS operations as map overlaying and buffering (Haining, 2003, p.124-7; Burrough and McDonnell, 2000, p.237-9).

Resolution refers to the amount of spatial and/or temporal detail provided. Spatial aggregates act as filters on the real underlying variability. Large spatial units suppress variation but can make it easier to detect patterns; small spatial units retain more variation but the resulting statistics are more affected by small errors. An important consequence of working with aggregated data is that results arising from analysing such data are conditional on the size of the census tracts (the scale effect) and their specific boundaries (the partition effect). Geographers and GI scientists refer to this as the modifiable areal unit problem (MAUP). There are other challenges arising from resolution issues. Regression modelling in spatial epidemiology requires that all data are reported on the same spatial framework but whilst much population data are reported by irregular census areas, environmental data are usually reported by grid squares. The challenge is to find a common spatial framework or undertake statistical modelling that recognizes the uncertainty arising when data are transferred from one reporting framework to another. Geographical data today are being collected at finer and finer spatial scales but whilst improved spatial precision is to be welcomed, this raises issues of statistical precision when calculating statistics for small areas. The small number problem arises when working with small census units comparing and analysing rates of a rare disease. Elsewhere in this book, small area estimation is discussed in chapters 6 and 28, whilst chapter 5 considers aggregation effects.

Consistency refers to the absence of contradictions in a database. For example cases of a disease may be reported in a postcode where no one lives This may be due to a geocoding error or perhaps due to the use of two databases (one health one demographic) which do not correspond in time. Lack of completeness refers to the situation where there are missing data or there has been undercounting. The use of internet methods that seek to engage the public in addressing public health issues may suffer from a lack of completeness. These and other issues associated with addressing GIS data quality are reviewed in amongst other sources Haining, 2003, p.61-74, and in greater detail with reference to health data by Cromley and McLafferty, 2012, p43-74 who also discuss spatial databases with particular reference to US data sources in their chapter 3. Many of the issues raised in these early sections are also discussed in Maheswaran and Craglia (2004).

4. Geographical Information Systems in the study of disease

In this section, we illustrate the use of GIS and GISc in the study of disease. We have grouped the studies into five subsections – Environmental epidemiology; Communicable diseases; Geographical epidemiology; Exposure assessment; and Disease clusters and environmental sources – but recognise that there is overlap between the subsections. We use a small number of examples but describe them in some detail in order to provide understanding of the epidemiological and public health contexts in which the scientific investigations were carried out. We have mainly used examples of work we have been involved in, supplemented by examples drawn from the work of others.

4.1.Environmental epidemiology

Air pollution and cardiovascular diseases in Sheffield - Existing and routinely collected data held by local and health authorities were used to investigate associations between outdoor air pollution and cardiovascular disease at the small area level. (For further discussion of pollution fields see chapters 15 and 16 in this volume.) The project demonstrates the advantages and limitations of working with routinely available data. The project used air pollution estimates for three outdoor air pollutants – particulate matter, nitrogen oxides, and carbon monoxide – which had been generated by the city council using an air pollution model (Indic AirViro). The modelling process took into account a range of emission sources (represented by points, lines and polygons) and meteorological conditions and generated a pollution surface for each pollutant at a 200 metre grid square resolution.

These data were imported into a GIS (ArcGIS) in order to validate model outputs and link with health outcome data. Abnormal patterns in the particulate matter pollution surface were apparent in a few quite localised areas of the city. Further investigation revealed that these were due to errors in the emissions database where old coal fired burners which had been replaced years ago had not been updated in the city council's emissions database (Brindley et al., 2004). These areas were excluded from the analysis. We also found that pollution concentrations for nitrogen oxides were overestimated by the model when compared with measured values from monitoring stations. However, the concentrations were all overestimated at a broadly comparable level, suggesting that relative measures of pollution were probably valid (Brindley et al., 2004). Analyses therefore used relative categories by quintile and not absolute values.

The outcome data comprised deaths and hospital admissions for coronary heart disease and stroke (1994-1998), which were provided at the enumeration district level by the health authority. Enumeration district level population estimates were based on the 1991 UK census and scaled using health authority mid-year estimates for 1994-1998.

As the outcome and exposure data were available at different spatial frameworks with 1030 enumeration districts and 10,847 pollution concentration grids, GIS was used to integrate both sets of data. We calculated the average pollution level within each enumeration district using a procedure based on postcode centroids. We first assigned to each domestic postcode centroid the value of the 200 metre pollution grid in which it lay. We then took the average of the values for all postcode centroids which lay within the enumeration district polygon. There are varying numbers of households within postcodes. We used the number of domestic delivery points at each postcode to weight the average value calculated for each enumeration district (Brindley et al., 2005). In order to take some account of daily local population movements, we also assigned to each postcode centroid the pollution value of the average of grid squares which fell within a 1km buffer of each postcode centroid based on surveys indicating that 1km was the average walking journey length to see if this improved associations with health outcomes.

We undertook analyses using methods with increasing levels of complexity. Standard Poisson regression carried out using SAS showed some associations between air pollutants and coronary heart disease and stroke mortality and to a lesser extent hospital admissions (Maheswaran et al., 2005a and 2005b). We found that using the 1 km buffer made little difference to the results. Bayesian analyses carried out using WinBUGS and, taking into account errors in variables and within area variation were subsequently used and these continued to show associations between air pollutants and mortality from coronary heart disease and stroke (Maheswaran et al. 2006a, Haining et al. 2007 and Haining et al. 2010).

Air pollution and stroke in South London – This example describes a subsequent project carried out to further investigate the association between outdoor air pollution and stroke. Point location case data and grid point resolution pollution data were used in this work. The work was carried out in an area in South London which was covered by the South London Stroke Register, a population based stroke register set up in 1995 to capture all cases of first ever stroke occurring amongst the resident population living within a defined geographical area.

The air pollution modelling for this study had been carried out at a very fine spatial resolution (20 metre grid point resolution) using bespoke air pollution modelling and was available for particulate matter and nitrogen dioxide. The modelling process took into account a wide range of pollution sources and emissions including major and minor road networks modelled with detailed information on vehicle stock, traffic flows and speed for each road segment, pollution sources in the London Atmospheric Emissions Inventory including large and small regulated industrial processes, boiler plants, domestic and commercial combustion sources, agriculture, rail, ships and airports, and pollution carried into the area by prevailing winds. Model validation was carried out by comparing modelled values with measured pollution values from background monitoring stations and by visually inspecting maps showing pollution values overlaid on road networks and other pollution sources within ArcGIS. The model validated well in terms of absolute values when compared with monitored pollution values (Maheswaran et al., 2010).

We investigated two outcomes, the incidence of stroke and survival after stroke. Stroke incidence was examined using a small area level ecological study design and survival after stroke investigated using a cohort study design.

To examine incidence, a denominator population (population at risk) was needed and we used population counts in census output areas from the UK 2001 census for this purpose. This was the smallest geographical unit at which census population counts by five-year age band and sex were available with approximately 300 people per output area. Stroke cases were assigned to output areas using the point in polygon method within ArcGIS, with a case assigned to the output area in which the postcode centroid of residence of the case was located. Observed and expected counts were calculated for each output area.

Linkage of grid point resolution pollution data to residential postcodes was also carried out within the GIS. All residential postcodes in the study area were assigned the pollution value of the grid point closest to the residential postcode centroid. Where there were equidistant points, the average value was taken. For the ecological study, an average value was then calculated for each output area, taking the average of values assigned to all postcode centroids which fell within the output area polygon, using point in polygon to link postcodes to output areas. A total population count by postcode was available from the 2001 census and this was used to weight the average pollution value calculated for each output area.

The survival analysis was carried out at the individual level, using Cox regression modelling. Follow-up was for up to 11 years. Patients were assigned the pollution value attached to their residential postcode centroid. For patients who moved, we took the average of the values at the start and end of their contribution to the study. The start value was from their postcode of residence when they had the stroke. The end value was either the value from their postcode of residence at the time of death or the value from their postcode of residence at the end of the study.

We found that living in a more polluted area was associated with decreased survival after stroke (Maheswaran et al., 2010). In the ecological analysis, we found that there was no spatial structure to the output area level incidence rates when assessed using WinBUGS, Moran's I and visual inspection of maps and we therefore used standard Poisson regression methods in SAS. There was no clear evidence of association between outdoor air pollutants and the incidence of stroke, although there was a suggestion of association in the 65-79 year age group in relation to ischemic stroke (Maheswaran et al., 2012). There was also a suggestion that the association was stronger for mild ischemic stroke (Maheswaran et al., 2014a).

4.2 Communicable diseases

Communicable or contagious diseases are infectious diseases involving some causative disease agent such as a virus, bacterium or parasite that is transmitted either from person to person or via some vector and/or intermediate host such as an animal. In the case of person to person transmission the disease is spread by contact and the geography of that spread will depend on the geography of human interactions and may take any of several forms including local clustering, spread following the urban hierarchy and mixtures of the two (Cliff and Haggett 2004). Some non-vector communicable diseases are spread through socially induced exposure to risk as in the case of HIV/AIDS (Rhodes et al., 2005) others are spread through environmental exposure so that the geography of cases will be a function of the geography of the environmental risk – for example the river network or water distribution system in the case of a water borne disease (Lake et al., 2007). Modelling such communicable diseases involves the use of epidemic models that partition the population into those who are susceptible (S), those who are infected (I) and those who have recovered or been removed (R) – known as SIR models. Models may be aggregated in terms of population groups defined by their locations which may be small areas, regions or urban places such as the STEM model (The Eclipse Foundation, 2011) or based on interactions amongst individuals as in the case of agent based models (Lee et al., 2008, Perez and Dragicevic, 2009). GISs are used to map such disease spread in space and time, to help identify disease clusters or concentrations, to map risk, to try to predict disease spread (see for example Oppong et al. 2012). Typically GIS provides, integrates and updates the data inputs and data layers that are used by epidemic and interaction models which then return outputs for mapping or animation by the GIS. The GIS is loosely coupled with these models. Databases may be at many different scales but there is growing interest in global scale databases for communicable diseases reflecting the global nature of threats to human health arising from population mobility and other aspects of globalisation. Studies that cover a large portion of the earth's surface drawing data from different countries (who may have different mapping conventions) raise map projection and other issues and here the ability of GISs to integrate spatial data is particularly valuable.

In the case of vector borne diseases, GISs are used in the ecological study of agent-vector-host relationships and their links to human populations (for an overview see Cromley and McLafferty 2012 p263-302). Habitat modelling can be used to assess exposure risk whilst land cover changes brought about by urban development or climate change can be used to assess whether the exposure risk is increasing or decreasing. GISs have also been used to assess the environmental characteristics of Lyme disease (Glass et al. 1992) and West Nile virus (Ruiz et al. 2007) case locations. The GIS operation of point-in-polygon can be used to look for case clusters in relation to different ecological characteristics.

The study of communicable diseases spans many disciplines involving researchers in virology, molecular biology, geography, epidemiology and public health so it is important to make connections. Formally integrating knowledge from different disciplines is challenging but as Ge et al. (2012) demonstrate, GIS can help mitigate the disciplinary gaps by providing a platform for updating data inputs and layers from different disciplines, facilitating analysis and integrating outputs for mapping (see also chapter 26). Their study into the spatial-temporal dynamics of avian influenza H5N1 in East and South-East Asia used GIS-based knowledge fusion. Genetic sequences were used to create phylogenetic trees to estimate and map the H5N1 virus' ability to survive and spread. Adding information about virus location together with spatial interpolation techniques produce maps of H5N1 risk. Maps of risk can also be produced by modelling social, economic, environmental and other data (Gilbert et al. 2008) and can also be obtained by analysing large concentrations of outbreaks using spatial statistics. Ge et al. (2012) used the Dempster-Shafer Inference Theory of Evidence to integrate the three raster layer probability maps, mapping the resulting output in a GIS.

4.3 Geographical Epidemiology

Migration and health inequalities in Sheffield - Socioeconomic gradients in mortality at the geographical level exist across many cities and regions worldwide. These patterns may endure despite efforts by health and local government authorities to reduce inequalities by targeting appropriate interventions at deprived areas. One potential explanation for enduring inequalities at the geographical level is selective migration. This is the situation where people in poor health, or those with the socioeconomic determinants of poor health e.g. unemployment, move from affluent to deprived areas while those in good health or with the socioeconomic determinants of good health e.g. high income, move from deprived to affluent areas. Thus, although interventions may benefit individuals in deprived areas, this selective migration may perpetuate inequalities when examined at the geographical level. In this example, we describe the use of GISc in epidemiological investigation of migration and area level mortality patterns.

We examined for evidence of selective migration and investigated the impact of selective migration on geographical inequalities in health in Sheffield, a city where there is a striking East-West gradient in area level deprivation which is closely mirrored by gradients in life expectancy (Maheswaran et al. 2014b). The project was carried out because the local authority wanted to know if selective migration contributed to the enduring gradient in health inequalities across the city.

We used a total population cohort dataset which was provided by the local health authority in anonymised format. The dataset was created from the general practice database the health authority held. This was a continually updated register of people resident in Sheffield who were registered with a general practice. The health authority kept regular "snapshots" taken from this register, including the residential location of people at the snapshot time point. For this project, the health authority provided a dataset with a record for each individual who was resident in Sheffield at any point within the cohort time frame. If an individual was present in a snapshot, their census area (lower super-output area from the UK 2001 census) and electoral ward of residence were provided for that snapshot time point. These census areas typically contain approximately 1500 people. Death records were also linked into this dataset, and if an individual had died the census area and ward of residence at death were provided.

Analysis of migration can be a very complex undertaking, depending on the number of time points analysed, the number of geographical units used and the length of time people resided in each geographical unit (see also chapter 34 in this Handbook). We started out with a simplified analysis where we used two time points (residential location at the start of the study and at death or the end of the study) and divided census areas into two categories (high and low deprivation). We found clear evidence of selective migration. People moving from low to high deprivation areas had higher mortality than those remaining in low deprivation areas. Conversely, people moving from high to low deprivation areas had lower mortality than those remaining in high deprivation areas. The magnitude of these differentials in mortality risk diminished with increasing age. We were also provided with data on health status and socioeconomic circumstances for a sample of the population. These data had been obtained in a survey carried out before the start point of the migration analysis. Analysis of these data showed that people tended to carry their pre-existing risks with them (Maheswaran et al., 2014b).

We examined the impact of migration on geographical gradients in mortality by putting people back to where they were at the start of the cohort time frame and comparing the mortality gradient across the city based on this location with the gradient based on residence at time of death. We found that selective migration made little contribution to existing socioeconomic gradients in mortality across the city (Maheswaran et al., 2014b).

The mapping, database manipulation and analysis for this project was carried out in R and included use of the GIS functionalities in R.

Alcohol-related mortality in England - Lifestyle related factors which include smoking, alcohol consumption, diet and physical activity are key determinants of health. These lifestyle related determinants of health are potentially modifiable and are therefore of significant public health concern. In this example, we illustrate examining the geographical epidemiology of alcohol related mortality in relation to socioeconomic deprivation at the small area level.

Alcohol consumption data from UK surveys suggest that alcohol consumption is marginally higher in more affluent socioeconomic groups. However, this does not appear consistent with alcohol related mortality, which appears to be higher in lower socioeconomic groups. We investigated the association with mortality at a national scale using a small area level ecological correlation study (Erskine et al., 2010). We used electoral wards as the units of analysis, of which there were 8797, using an existing dataset on alcohol related mortality which had been compiled by the Office for National Statistics for surveillance of alcohol related mortality. The deaths had been assigned to wards using a postcode to ward look-up table. The deaths included in the dataset were those considered to be most likely to be directly attributable to alcohol. The predominant condition in this group was liver cirrhosis.

In addition to examining associations between alcohol-related mortality and socioeconomic deprivation at the small area level, we also examined associations in relation to gender, age and urban-rural location. The dataset supplied included the Carstairs Index as the indicator of socioeconomic deprivation at the ward level. This is a standardised combination of four variables from the 2001 census – male unemployment, overcrowding, low social class and lack of car ownership.

The analysis was based on 18,716 male and 10,123 female deaths over a five year period (1999-2003). We found a strong association between socioeconomic deprivation at the electoral ward level and alcohol related mortality. The differential in relative risk was most pronounced in the 25-44 year age band. Mortality rates were higher in men than women and also higher in urban areas (Erskine et al., 2010).

The main analysis was carried out using standard Poisson regression methods in SAS due to the number of wards in the dataset and substantive analytical detail required. We also carried out Bayesian analysis on a small subset of the data to explore gender variation in the spatial pattern of alcohol-related deaths using WinBUGS (Strong et al., 2012). The adjacency matrix for this analysis was generated in ArcGIS. We initially fitted separate models for men and women and subsequently modelled male and female deaths jointly using a shared component for random effects. We investigated a range of different unstructured and spatially structured specifications for the gender specific and shared random effects. We found significant spatial variation in ward-level alcohol-related mortality for men but this was much less marked for women. After accounting for deprivation, there was significant unexplained elevated risk in a very small number of wards.

4.4 Exposure assessment

Improving estimates of air pollution exposure - Most studies examining the association between air pollution and health outcomes have used either monitored or modelled air pollution values to estimate exposure (see also chapter 15 in this Handbook). Monitored values are from fixed site air quality monitoring stations. These epidemiological studies have generally not taken daily population movements and time spent in different locations into account. Most have generally used outdoor monitored or modelled estimates and the indoor vs outdoor concentrations have generally not been taken into account. An important element determining the dose of pollution taken in by people is the activity being undertaken, as higher energy expenditure is associated with an increased respiratory rate and depth of breathing, resulting in higher doses of pollution. These aspects are all challenging to incorporate in large scale epidemiological studies which are needed to examine associations with health outcomes.

The example described here is a detailed study undertaken by de Nazelle and co-workers on a small number of subjects, 36 healthy young volunteers, undertaken to accurately assess exposure (de Nazelle, 2013). The methodology used global positioning systems (GPS) and accelerometer functions in smart phones for exposure assessment. A computer programme was developed which used accelerometer readings to estimate energy expenditure.

Estimation of air pollution exposure used modelled annual mean pollution estimates at a fine spatial scale as the start point. The estimation subsequently took into account temporal variation, both within day and by day of the week, and the microenvironment, by sampling indoor and outdoor concentrations, and measuring exposure whilst using different modes of transport.

A GIS platform was used to integrate the air pollution, GPS location and activity data. Various comparisons were carried out. There was substantial variation when pollution estimates incorporating all the refinements including energy expenditure were compared with estimates based on home location address only, with little correlation between the two.

This example illustrates the potential for substantial exposure misclassification and also bias, with a tendency to underestimate exposure, in standard epidemiological studies. The methods described in this work are very involved and the challenge will be to use such methods in large scale epidemiological studies.

Assessing environmental influences on diet and exercise - Public Health is concerned with the influence of environmental factors such as parks and green spaces and fast food outlets on health. Parks and green spaces provide places for physical activity while fast food outlets may promote the consumption of foods high in saturated fat and low in fibre. Several studies have been carried out to examine the potential influence of these environmental factors on diet and physical activity and most studies have examined exposure around the residential location. However, exposures around activity spaces away from these residential locations have been much less well studied.

The example described here is work carried out by Zenk and co-workers in which exposure in activity spaces was examined using a combination of GPS and accelerometers (Zenk et al., 2011). GPS were programmed to record participants' position every 30 seconds over a 7-day period and data were obtained on 120 participants. Two measures of activity space were created using the GPS information downloaded into a GIS.

The first measure was referred to as a one standard deviation ellipse. The central location of all GPS points for a participant was calculated. An ellipse was then created around this central point and the one standard deviation limit meant that approximately 68% of all GPS points were included within the ellipse. The long axis of the ellipse was in the direction of maximum dispersion, while the short axis was in the direction of minimum dispersion. This measure was calculated using the spatial statistics toolbox in ArcGIS.

The second measure was referred to as the daily path area. This was created by first buffering around every GPS point for the participant using a 0.5 mile radius and then dissolving the boundaries between these buffers to create the daily path area. This measure was also created using ArcGIS.

For comparison with the standard residential location methods, a 0.5 mile street network buffer was created around the census block centroid of each participant's residential location.

The density of fast food outlets and the percentage of land that was designated as municipal park land were calculated for each of the three exposure areas. The study found very low correlations between exposures based on neighbourhood location and exposures based on activity space, especially with activity space defined as the daily path area.

The study found no association between residential neighbourhood fast food outlet density and diet. However, a positive association between fast food outlet density and an unhealthy diet was found when the daily path area was used to calculate exposure. No associations were found between physical activity and park land use in analyses using each of the three exposure space definitions.

Although the study found that fast food density in the daily path area was associated with an unhealthy diet at the individual level, there is a potential problem with interpreting this association as causal, i.e. that increased exposure to fast food outlets is the cause of people eating unhealthily. This is because the daily path area is defined by the participant choosing to go along particular routes and they may have gone along those particular routes in order to access fast food outlets.

4.5 Disease clusters and environmental sources

Rapid initial assessment of apparent disease clusters – Concerns about apparent clusters of disease and potentially elevated risks of disease around environmental sources of pollution such as factories frequently arise. These clusters, real or apparent, have the potential to cause substantial public anxiety and media interest and can result in substantial public health resources being spent in addressing these concerns if they are not handled in a timely and effective manner (Maheswaran and Staines, 1997).

Identifying disease clusters is somewhat different from examining if diseases have the general propensity for clustering. Clusters may occur in areas where there is no obvious cause, or may occur around environmental sources, typically around point sources but also, to a lesser extent, around line and area sources. The statistical issues around clusters and clustering are covered in other chapters in this book (see chapters 8, 9, 14 and 28). Here we describe an example, in which we were involved, of a facility set up to investigate apparent clusters of disease to support public health investigation (Aylinet al., 1999).

The Rapid Inquiry Facility was set up within the Small Area Health Statistics Unit in the UK to carry out a rapid initial assessment of apparent disease clusters. The facility is a system which combines three technical elements – a database, a GIS and automated statistical analytical methodology.

The database integrates datasets on health outcomes including deaths, hospital admissions, congenital malformations and cancer registrations. The geographical identifier for these outcome data is the postcode centroid. The database includes population denominator counts by age and sex at the small area level. These are counts from population censuses, with population estimates for inter-censal years. The database also includes information on socio-economic deprivation at the small area level. These are basic minimum requirements for adjusting for potential confounding variables as disease incidence and mortality can vary substantially by age, gender and socioeconomic status.

A GIS platform is fundamental to this system and allows integration of data from different spatial frameworks. The system offers flexibility regarding the spatial resolution at which diseases can be investigated, with the smallest unit being electoral wards and census output areas. Point source locations of environmental pollutants can be specified and different buffers created around these sources. The GIS also allows an adjacency matrix to be generated for use in smoothing risk maps.

The statistical analysis allows for automated calculation of absolute and relative risks for different buffer zones around point sources. The calculation includes confidence intervals and significance testing. The statistical methodology also includes the production of smoothed maps displaying a risk surface using Bayesian methodology utilising the adjacency matrix created within the GIS. Areas with significantly higher or lower risks are also identified on these risk maps.

The system allows the rapid initial assessment of apparent disease clusters which have caused concern to members of the general public, media, politicians or public health staff. It does not provide definitive answers but can rule out clusters which do not exist statistically. If clusters or high rates in some areas are found, further investigation is needed. A key first step is to examine for artefacts and errors in the data. Incomplete data capture in some areas, inaccuracies in the data recorded and relevance of the conditions being examined, e.g. cancer "clusters" which comprise conditions which are aetiologically unrelated, all need to be considered.

The Rapid Inquiry Facility has been acquired for use elsewhere e.g. in Utah (Ball et al., 2008), and has undergone further enhancements (Beale et al. 2010). Chapter 2 in this Handbook describes environmental exposure research in detail whilst cluster detection and modelling are discussed in chapters 3, 8, 9, 14 and 28.

5. Examples of Geographical Information Systems in the provision of public health services

In this section we give examples where GIS and GISc have been used to examine and inform the provision of public health services. We have grouped the studies into four subsections – access to services; needs assessment and health equity; variation in utilisation; and planning the location of services – recognising the overlap between sections. Health services are also discussed in chapter 29 of this Handbook.

5.1 Access to services

Uptake of breast cancer screening in North Derbyshire - In this example work was undertaken to inform local planning decisions (Maheswaran et al., 2006b). An increase in capacity in the provision of screening for breast cancer was needed in North Derbyshire. This was because a change in national policy meant that the age range of women invited for screening was to be increased from 50-64 years to 50-70 years. In addition, two view mammography, which was then being undertaken only at the initial screen, was to be instituted at all screening rounds. The health authority was interested to know if there was still an issue with distance from screening site, i.e. if uptake was lower amongst women living further away, and if uptake was lower amongst women living in more socio-economically deprived areas, in order to take these factors into consideration when reorganising services.

Data were provided at the individual level for women invited for screening. This dataset contained the postcode of women invited for screening, whether or not they attended, and the screening location to which they were invited. A postcode to census enumeration district look-up table was used, which also contained eastings and northings for postcode centroids.

Road travel distance to a screening location was calculated from the postcode centroid to the grid location of the screening centre using 1:10,000 resolution road network data within a GIS (MapInfo). Screening was provided at a fixed site (the main district general hospital in the area) and at 12 locations throughout the health district using a mobile screening unit.

Socio-economic deprivation was assessed using the Townsend score of the census enumeration district in which the postcode was situated. This area level derivation indicator was a standardised combination of four 1991 census variables (unemployment, no car ownership, non-home ownership, overcrowding). Data were analysed on 34,868 women. Overall uptake of screening was 78%.

As this was an individual level dataset, the analysis was carried out at the individual level using logistic regression in SAS, modelling the binary outcome of attendance or non-attendance. We found a small decrease in uptake with increasing distance from the screening location. The effect of distance on uptake, although still detectable, was likely to have been largely ameliorated through the use of the mobile unit. Deprivation however, did have a clear effect, with lower uptake amongst women living in more deprived areas (Maheswaran et al., 2006b).

Walk-in centres and primary care access - This next example relates to work carried out to inform government policy on providing access to primary health care. General practice surgeries were under increasing pressure due to increasing demands for their services and walk-in centres were seen as one option for relieving pressure on these surgeries. A wave of walk-in centres had been set up in England and the purpose of this work was to evaluate if these walk-in centres had reduced waiting times at general practices (Maheswaran et al., 2007).

Waiting times for a general practice appointment were monitored by the primary care access survey, a regular monthly survey carried out nationally to assess the waiting time measured in days to the next available surgery appointment with a general practitioner. The survey was carried out on all NHS general practices in England and there was a 48-hour target set by government. We obtained these monthly survey data from the Department of Health in England.

We used two approaches to calculate exposure of general practices to walk-in centres. The first was the straight line distance from each general practice postcode centroid to the postcode centroid of the nearest walk-in centre which was already in operation that month. This approach took into account the phased opening of walk-in centres.

The second approach used a function based on walk-in centre attendance rates by distance. For this second approach we used attendance data, which were available for four walk-in centres, to create the function. Attendance data and population denominator counts were available by census output area. These output areas were assigned to 1 km concentric rings around the walk-in centres and attendance rates calculated for these distance bands. An exponential distance decay function was fitted to these rates. We used this function to calculate distance decay values for each general practice by month on the basis of its distance to each walk-in centre. We then summed the values for each general practice by month which in effect took into account the effect of multiple walk-in centres in the vicinity of a general practice.

We analysed data on 2509 general practices in 56 health authority areas in England and included 32 walk-in centres in the analysis. We found no evidence to suggest that walk-in centres shortened waiting times for access to primary care. As part of the project, we also examined the effect of area level deprivation on waiting times and found clear evidence that the waiting time target was less likely to be achieved in more deprived areas.

ArcGIS was used to visualise locations of walk-in centres, general practices and health authority boundaries. Straight line distances were calculated using Pythagoras' theorem in a Microsoft Access database. Statistical analyses were carried out in SAS.

Renal replacement therapy in the Trent Region - The purpose of this example is to illustrate the use of GISc in relation to access to health services and health outcomes (Maheswaran 2003). End stage renal failure typically results from chronic renal disease caused by a variety of medical conditions. When patients are in end stage renal failure, they require some form of renal replacement therapy in order to survive. The three options are haemodialysis, peritoneal dialysis and renal transplantation.

This example describes work we carried out in collaboration with the Trent Public Health Observatory in order to inform planning decisions in the Trent Region. The need for more renal services was being reviewed. One of the considerations was the provision of more renal units which would improve access. Renal units are typically classified into main units and satellite units, with the latter providing mainly haemodialysis.

The Observatory assembled the data and carried out descriptive analyses. Renal units within and surrounding Trent Region were identified and their locations geo-referenced. Prevalence data on all patients residing in the Region who were receiving any of the three forms of renal replacement therapy were obtained. Patients were assigned to census enumeration districts. Denominator populations for enumeration districts were obtained from the 1991 census and scaled to subsequent mid-year estimates for health authorities within the region. The Townsend score was used as an indicator of socioeconomic deprivation at the enumeration district level. The percentage of the population of African and Asian origin at the enumeration district level was also obtained from the census. These factors were taken into account because renal disease is commoner in more deprived communities and also has a higher prevalence amongst people of African and Asian origin. Access to renal units was assessed by calculating road travel distances from census enumeration district population centroids to the nearest renal unit.

We used Poisson regression to assess associations between travel distance, deprivation and renal replacement therapy rates in the region. Renal replacement therapy rates were higher in more deprived areas. However, when the individual modalities of renal replacement therapy were examined, rates were higher for haemodialysis but not for transplantation in more deprived areas. This raises the issue of inequalities in health care, as transplantation is the preferred option for end-stage renal disease and it would be expected that transplantation rates would also be higher in more deprived areas.

With regard to geographical access, haemodialysis rates were lower in places further from renal units. This might be expected to some extent because distance from a renal unit might be taken into consideration when decisions are taken regarding whether to use haemodialysis or peritoneal dialysis. There may also be the issue of "reverse causality" for the association. The need for haemodialysis might cause people to move to live closer to renal units.

MapInfo was used to calculate road travel distances and to map and visualise locations of renal units in relation to regional geography. Assembly of the dataset was carried out using Microsoft Access and statistical analysis was carried out in SAS.

5.2 Needs assessment and health equity

Health equity profiles - This example describes the use of GIS and GISc in health needs assessment and assessment of health equity. Health needs assessment may be carried out to assess the health needs of a population, people with a particular condition or the need for a specific intervention. There is overlap with the process of health equity auditing, the first step of which is a health equity profile. The health equity profile may be assessed from spatial and social perspectives, with the latter sub-classified by age, gender, class and ethnicity.

This example is of a health equity profile that was undertaken to support planning and reconfiguration of services in a health authority in North West England. The work used existing and routinely collected information to assess equity. The range of indicators used included mortality data, hospital admissions data and general practice level data including data from the Quality and Outcomes Framework. The conditions for which the equity work was to be undertaken were predetermined by the health authority and included cardiovascular disease, diabetes, chronic obstructive pulmonary disease and alcohol related conditions.

The spatial frameworks at which the data were assembled and analysed were electoral ward level, census based lower super output area level and general practice population level. The latter is not clearly defined geographically and patients registered with a particular general practice could come from a wide area. Nevertheless, the majority of patients registered with a practice live in the local area close to the practice. From the planning perspective, primary health care services are organised around practices and this is therefore a useful level at which to investigate health needs and equity.

Data were assigned to wards and output areas by the health authority using postcode to area geography look-up tables. General Practice level data were generated at this level. The data were used to produce a range of choropleth maps using ArcGIS. Scatterplots and other graphs were used to carry out exploratory spatial data analysis. For GP practice level information, we produced a graphic which was able to show a range of practice level indices in the same figure (**Figure 4**). This form of visualisation was useful for identifying outliers and considering a range of related indicators together. This bespoke graphic was created in R and brings together a range of geographical information.

Figure 4 here

The synthesis of data from different sources allowed variations in need and equity to be identified. Outlying practices were investigated further by the health authority but it should be noted that there may be good reasons for variations in practice. Being an outlier does not automatically indicate unusual or substandard practice.

Physical activity in socioeconomically deprived areas - This example relates to a physical activity intervention that was offered as part of a multistage intervention leading up to a trial of an intervention to maintain increased physical activity (Ying et al., 2014). Preventative services are being increasingly recognised as an important element of public health offered to communities. The prevalence of cardiovascular diseases is higher in more socio-economically deprived communities and attempts to reduce inequalities in health have led to services being targeted at more deprived communities.

In the initial phase of the intervention, middle aged people living in deprived neighbourhoods in Sheffield were offered a physical activity intervention to increase their physical activity levels. The intervention comprised a motivational DVD to be sent by post along with additional information. The DVD was sent to those who responded to an initial invitation letter asking them if they would like to receive the intervention. The health authority for Sheffield had previously characterised neighbourhoods, and people aged 40-64 in these selected most deprived neighbourhoods were to receive the intervention. This first phase was offered to all people invited, unlike the subsequent phase in which a further intervention was given on a randomised basis to some participants.

The overall uptake was extremely low, with an overall mean of 7%. Investigation into factors associated with this low uptake included examining small area level factors which might be associated with low uptake. These factors were investigated using postcode areas as the units of analysis. Deprivation was assessed using the proportion of households in a postcode receiving housing benefit. This variable was not available for all postcodes as the data were not provided for postcodes with small numbers of households or for postcodes where most or all of the households were receiving benefits. An alternative indicator, the Index of Multiple Deprivation available at lower super-output area level, was also used and assigned to all postcodes within the super-output area. Other factors investigated included walking distance to the nearest gym, walking distance to the nearest swimming pool and walking distance to the nearest municipal green space. The network distance analysis was undertaken by linking and using datasets with a fine spatial resolution within ArcGIS. A detailed description of the datasets and methodology used is provided in Goyder et al. 2014.

The spatial analysis was complicated by a number of factors. There were 2455 postcode areas analysed in the study. The postcode areas were not contiguous due to the way areas were selected. Only postcodes in selected deprived areas with one or more residents aged 40-64 years were included in the study. There were very low counts for most postcode areas. In 66 postcode areas, only one postal invitation was sent out, with no responses in sixty of the postcodes. In addition, in postcodes where more than one invitation was sent, there was a zero response from 996 postcodes.

We developed and used Bayesian hierarchical Bernoulli-binomial spatial mixture zero-inflated Binomial models to model over-dispersion and to separate the systematic and random variations in the noisy and mostly low response rates (Ying et al., 2014). The models allowed for investigation of variations in patterns of mail outs, zero responses and response rates. We found that response rates were lower in postcodes in which a higher proportion of households received housing benefit. There was little evidence of association with the other variables examined. The postcode polygon adjacency matrix was created using ArcGIS. Spatial analysis was carried out in WinBUGS and the statistical outputs were visualised using R.

5.3Variation in utilisation

Geographical variation in potentially avoidable admissions - Hospital services are coming under increasing pressure from the increasing demand for emergency hospital admissions.

Not all hospital admissions are essential. Whilst some emergency hospital admissions, e.g. for a heart attack or meningococcal meningitis, are clearly urgent and essential, at the other end of the spectrum there are admissions which could have been avoided for a number of reasons, including better social and preventative community care. In this example, we describe a project examining geographical variation in potentially avoidable admissions (O'Cathain et al., 2014).

A group of hospital admissions were used which were considered to be likely to contain a substantial number of avoidable admissions. This group of admissions included non-specific chest pain, non-specific abdominal pain and chronic obstructive pulmonary disease, and were selected through a consensus process with specialists in the field. It is important to note though that not all admissions with these conditions would have been avoidable.

Using this list of defined conditions, a standardised avoidable admissions rate was calculated for health authority areas using direct standardisation. The population for the whole study area (England) was used as the standard population.

The spatial framework used for the geographical areas was primary care trust boundaries. These trusts were health authorities responsible for the health of the population resident within their defined geographical boundaries. They received money from central government and are part of the NHS structure within England. The rationale for using these primary care trust boundaries was that these trusts commissioned emergency and urgent medical care (along with primary and other levels of care) for their residents and therefore the care within a primary care trust area could be considered to comprise an emergency and urgent care system.

We analysed 152 primary care trusts in this project. There were 3.3 million admissions over a three year period which came under the category of defined conditions for this project, accounting for 22% of all emergency admissions. There was a 3.4 fold variation in potentially avoidable admission rates across the primary care trust areas examined.

Geographical variation was investigated by thematically mapping standardised avoidable admission rates and there was clustering of primary care trusts with high and low rates, with noticeable clusters of high rates in NW and NE England. We investigated associations between a range of primary care trust level factors and avoidable admission rates using general linear modelling with primary care trusts as the units of analysis. The list of factors examined were those which had been previously associated with geographical variation in admission rates. The large area ecological level regression was considered appropriate because the level of interest was systems operating at primary care trust level. Primary care trust level deprivation explained 72% of the observed variation across trusts. Factors related to emergency departments, ambulance services and general practice also explained some of the variation (O'Cathain et al., 2014).

A subsequent phase of the project identified trusts that were outliers, that is with variation not explained by the factors used in the regression, and carried out in-depth case studies of a selection of these trusts using qualitative methods. Data manipulation and analysis were carried out using R and SPSS and mapping was carried out in R.

Geographical variation in use of computed tomography (CT) scanning services – Nixon and co-workers (Nixon et al., 2014) carried out a study to examine geographical variation in the use of CT scanning in a region in New Zealand. CT scanning can be carried out on an emergency or a routine basis. Emergency scans may be carried out for inpatients or for patients in emergency departments before a decision to admit has been taken. Routine scans may be carried out for inpatients or for outpatients. CT scanning may be used purely for diagnostic purposes or for carrying out procedures under CT guidance.

The study region comprised a mix of urban and rural areas, including remote rural areas. The study area contained two large urban areas, each of which had a CT scanner at the main hospital in the area. Most of the rural areas had a rural hospital which served the local population. The spatial framework used for the analysis was the catchment area. Catchment areas were geographically defined using census based units. The catchment areas were defined as areas from which most of the patients using the local hospital came from.

The study found that there was large variation in age adjusted CT utilisation rates. Urban areas had 63% higher CT utilisation rates compared with remote rural areas.

There are a number of possible explanations for the variation in use. These include availability of services, variation in clinical practice, differences in population morbidity and demography. Availability of services is most likely to be the key factor driving utilisation rates. Overall utilisation rates will be the result of a mix of appropriate use (although appropriate use can be difficult to accurately define for a range of clinical situations) and potentially borderline or inappropriate use including supply-induced demand.

Variation in clinical practice is another potentially important factor responsible for geographic variations in utilisation rates. Experienced doctors may be more likely to have a higher threshold for using CT scanning, and be more likely to be able to identify patients for whom the scanning is appropriate. However, in one rural area where CT was subsequently introduced, a clear increase in utilisation rates was observed, suggesting that availability of CT is the overriding factor driving utilisation rates.

Referral rates were also lower for outpatient specialist CT diagnostic and procedure purposes in rural areas, even though these areas were served by specialists running clinics in the rural hospitals. A possible explanation is variation in clinical practice, with the specialists for example taking travel distances and inconvenience for rural patients into account if they arranged CT scanning for these patients.

This study highlights the importance of geography when investigating variations in utilisation. Mapping was used to visually display hospital catchment areas by type and the objectives of the study in relation to geography were achieved without the need to resort to more complex GIS functionality. The study was designed to be descriptive and further investigation is needed to identify potential explanations. In addition, it was not designed to examine outcomes, and variation in utilisation needs to be accompanied by subsequent study of outcomes, as lower utilisation does not automatically mean poorer outcomes.

5.4 Planning the location of services

Optimising access to stroke care

Ischemic stroke is a condition caused by blockage of an artery supplying blood to the brain. This acute event is most commonly caused by a blood clot. Treatment in the acute phase includes administration of a treatment (recombinant tissue-type plasminogen activator), which breaks down the clot to re-establish the blood supply. This treatment needs to be given as soon as possible after the onset of ischemic stroke but the diagnosis needs to be confirmed first using imaging techniques.

In the USA, a hospital which has the facilities for emergency investigation and treatment of stroke patients may be certified as a primary stroke centre. The process of achieving certification is relatively onerous and is initiated by hospitals, typically larger hospitals in urban areas. This voluntary self-initiation process has led to an uneven distribution of these primary stroke centres and substantial proportions of the population are not covered by these centres.

Leira et al. (2012) set out to quantify the percentage of the population not covered by primary stroke centres in the state of Iowa. The state comprises a mix of urban and rural areas and had 12 certified primary stroke centres. The authors used a location-allocation model to examine the percentage of the population that would have been covered in a hypothetical situation where the 12 centres were allocated de novo. They also examined how many additional centres would be needed to achieve 75% population coverage using the location-allocation model compared with a weighted random selection of additional centres. The weighted random selection was set up to mimic the current situation where larger hospitals were more likely to self-initiate the certification process. The additional sites were selected from 108 hospitals in the state which had the requirements to be designated as primary stroke centres.

There is a range of location-allocation models (see Cromley and McLafferty (2012) for a description of models) that can be used in different situations and Leira e al (2012) used the maximal coverage model for their investigation. They constructed a time-distance matrix from ZIP code tabulation area postcode centroids to potential locations, used population counts in ZIP code tabulation areas, and used pre-specified maximum time distance thresholds (15, 30 and 45 min) in their calculations. GIS tools used in their work included mapping software for visualisation purposes, Microsoft's Bing Maps API for calculating travel times and a web-based maximal coverage model calculator implemented using Java and PHP.

The authors found that the 12 existing centres only covered 37% of the Iowa population when a 30 minute maximum threshold was used for defining access. The hypothetical assignment of the 12 centres starting de novo would have covered 47.5% of the population using the maximal coverage model. A further 54 primary stroke centres would be needed to reach 75%

population coverage if selected using the weighted random selection process but only a further 31 would be needed if selected using the maximal coverage location-allocation model.

Whilst the authors acknowledge a number of limitations to the theoretical optimisation modelling approach, it nevertheless is useful in informing the debate about extending population coverage for rapid access to thrombolysis following ischemic stroke.

6. Concluding remarks

Many of the applications discussed in this chapter have used regression together with other statistical modelling tools. More can be found on these methods in chapters 5, 6, 7 and 13. In this overview we have adopted a broad conceptualisation of GIS, embedding it within GISc thereby stressing the wider contribution it makes to how we work with geographically referenced data. The definition of what constitutes a GIS and its functionality has evolved and will continue to do so. As the domain of spatial analysis has expanded (particularly the field of spatial statistics) it is no longer reasonable (if it ever was) to expect that any single GIS product will include within it all the tools a spatial epidemiologist or public health analyst working in the field of GISc might wish to call on. Interaction between a GIS and other software systems has become increasingly important as has the interaction between GIS and the internet.

GIS and GISc make important contributions to all those areas of scientific investigation and policy making where elements of geography are integral – where place (from the global to the neighbourhood and community scales) and spatial relationships matter (see chapter 33 for other examples). As will be evident from this overview and other chapters in this volume, whilst advanced spatial statistical methodology has an important role to play in research, even basic GIS functionality such as data integration, mapping and the implementation of simple spatial queries often provide important insights, placing the study of population disease and the delivery of health services in their broader social and environmental contexts.

Appendix 1
At the time of writing (March 2015) there is an entry in Wikipedia: "List of geographic information system software" giving both open source and commercial GIS products. Notable in the former category are GRASS GIS and QGIS. Notable in the latter category are ERDAS IMAGINE, ESRI (which includes ArcMap, ArcGIS, ArcIMS), Intergraph and Mapinfo. However there are many more. See: http://en.wikipedia.org/wiki/List_of_geographic_information_systems_software#cite_note-sstfoss4g-4

GIS functionality falls into the following broad categories (Cromley and McLafferty, 2012, p. 30):

(i) Measurement (e.g. distance, length, perimeter, area, centroid, buffering, volume, shape);
(ii) Topology (e.g. adjacency, polygon overlay, point and line in polygon, dissolve, merge);
(iii) Network and location analysis (e.g. connectivity, shortest path, routing, service areas, location-allocation modelling, accessibility modelling);
(iv) Surface analysis (e.g. slope, aspect, filtering, line of sight, viewsheds, contours, watersheds)
(v) Statistical analysis (e.g. spatial sampling, spatial weights, exploratory data analysis, nearest neighbour analysis, spatial autocorrelation, spatial interpolation, geostatistics, trend surface analysis).

GeoDa is free software for undertaking some forms of (frequentist) spatial statistical analysis including exploratory spatial data analysis. It is user friendly and employs a drop down menu style. It also has some normal spatial regression modelling capability (the so-called spatial error and spatial lag regression models with likelihood-based diagnostics to help the user select). The software has some nice features including a linked windows capability that allows the user to link database spreadsheet rows with the corresponding locations on a map and on graphs. It contains both "global" and "local" statistics such as local and global measures of spatial autocorrelation. The software is available from: https://geodacenter.asu.edu/ where tutorials can also be found. It is particularly useful for teaching (especially at the undergraduate level) and could be used in laboratory classes for a course in spatial epidemiology. The software was developed by Luc Anselin. For more advanced spatial modelling the researcher needs to investigate amongst others the following: WinBUGS and GeoBUGS (Bayesian modelling), STATA, S-PLUS and the R library. See other chapters in this Handbook. In this era of "big data" potential users need to be aware that many of these softwares encounter difficulties when used to fit models to large data sets.

References

Aylin P, Maheswaran R, Wakefield J, Cockings S, Jarup L, Arnold R, Wheeler G, Elliott P. (1999). A national facility for small area disease mapping and rapid initial assessment of apparent disease clusters around a point source: the UK Small Area Health Statistics Unit. *Journal of Public Health Medicine* 21:289-93.

Ball W, LeFevre S, Jarup L, Beale L. (2008). Comparison of different methods for spatial analysis of cancer data in Utah. *Environ. Health Perspect.* 116:1120-4.

Beale L, Hodgson S, Abellan J, LeFevre S, Jarup L. (2010). Evaluation of Spatial Relationships between Health and the Environment: The Rapid Inquiry Facility. *Environ. Health Perspect.* 118:1306-1312.

Beyer, K and G.Rushton (2009). Mapping cancer for community engagement. *Public Health Research, Practice and Policy,* 6(1), 1-8.

Brindley P, Maheswaran R, Pearson T, Wise S, Haining RP. (2004). Using modelled outdoor air pollution data for health surveillance. In: Maheswaran R, Craglia M (eds.) *GIS in Public Health Practice*. Boca Raton: CRC Press, 125-49.

Brindley P, Wise SM, Maheswaran R, Haining RP. (2005). The effect of alternative representations of population location on the areal interpolation of air pollution exposure. *Computers, Environment and Urban Systems* 29:455-9.

Burrough, P.A. and R.A. McDonnell (2000). *Principles of Geographical Information Systems*, Oxford University Press, Oxford.

Cho, G (2007). *Geographic Information Science: Mastering the legal issues*, Wiley, Chichester, UK.

Cliff, A. and Haggett, P. (2004). Time, travel and infection. *British Medical Bulletin*, 69(1), 87-99.

Cowen, D.J. (1988). GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing*, 54, 1551-4.

Cromley, E.K. and S.L.McLafferty (2012). *GIS and Public Health*, 2nd edition, The Guilford Press, New York.

de Lepper, M.J.C., H.J.Scholten and R.M.Stern Eds. (1995). *The Added Value of Geographic Information Systems in Public and Environmental Health*, Kluwer Academic Publishers, Dordrecht, Netherlands.

de Nazelle A, Seto E, Donaire-Gonzalez D, Mendez M, Matamala J, Nieuwenhuijsen MJ, Jerrett M. (2013). Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environmental Pollution* 176:92-9.

The Eclipse Foundation. (2011). The spatiotemporal epidemiological modeller (STEM) Project. Retrieved August 11th 2014 from *www.eclipse.org/stem*.

Erskine S, Maheswaran R, Pearson T, Gleeson D. (2010). Socioeconomic deprivation, urban-rural location and alcohol-related mortality in England and Wales. *BMC Public Health* 10:99.

Ge E-J, Haining R, Li C-P, Yu Z-G, Waye MMY, K-H Chu and Leung Y. (2012). Using knowledge fusion to analyze avian influenza H5N1 in East and Southeast Asia. *PLoS ONE*, doi: 10.1371/journal.pone.0029617.

Gilbert M, Xiao X, Pfeiffer DU, Epprecht M, Boles S, Czarnecki C, Chaitaweesub P, Kalpravidh W, Minh PQ, Otte MJ, Martin and Slingenbergh J. (2008). Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. *Proceedings of the National Academy of Sciences* 105(12): 4769-4774.

Glass, G., Morgan, J., Johnson, D., Noy, P., Israel, E. and Schwartz, B. (1992). Infectious disease epidemiology and GIS: a case study of Lyme disease. *Geographical Information Systems* 3(3), 65-69.

Goyder E, Hind D, Breckon J, Dimairo M, Minton J, Everson-Hock E, Read S, Copeland R, Crank H, Horspool K, Humphreys L, Hutchison A, Kesterton S, Latimer N, Scott E, Swaile P, Walters SJ, Wood R, Collins K, Cooper C. (2014). A randomised controlled trial and cost-effectiveness evaluation of 'booster' interventions to sustain increases in physical activity in middle-aged adults in deprived urban neighbourhoods. *Health Technol. Assess.* 18:1-210.

Guptill, S. and J.Morrison (1995). *Elements of Spatial Data Quality*. Elsevier Science, Oxford.

Haining, R. (2003). *Spatial Data Analysis: Theory and Practice*. Cambridge University Press, Cambridge.

Haining R, Law J, Maheswaran R, Pearson T, Brindley P. (2007). Bayesian modelling of environmental risk: example using a small area ecological study of coronary heart disease mortality in relation to modelled outdoor nitrogen oxide levels. *Stochastic Environmental Research and Risk Assessment* 21:501-9.

Haining R, Li G, Maheswaran R, Blangiardo M, Law J, Best N, Richardson S. (2010). Inference from ecological models: estimating the relative risk of stroke from air pollution exposure using small area data. *Spatial and Spatio-Temporal Epidemiology* 1:123-31.

Lake, I., Harrison, F., Chalmers, R., Bentham, G., Nichols, G., Hunter, P. et al. (2007). Case control study of environmental and social factors influencing cryptosporidiosis. *European Journal of Epidemiology*, 22(11), 805-811.

Lee, B., Bedford, V., Roberts, M. and Carley, K. (2008). Virtual epidemic in a virtual city: simulating the spread of influenza in a U.S. metropolitan area. *Translational Research*, 151(6), 275-287.

Leira EC, Fairchild G, Segre AM, Rushton G, Froehler MT, Polgreen PM. (2012). Primary stroke centers should be located using maximal coverage models for optimal access. *Stroke* 43:2417-22.

Lopez, A. (1992). Mortality data. In P.J.Elliott, J.Cuzick, D.English and R.Stern (eds.), *Geographical and Environmental Epidemiology: Methods for small area studies*, pp.37-50. Oxford University Press, Oxford.

MacNab YC, Read S, Strong M, Pearson T, Maheswaran R, Goyder E. (2014). Bayesian hierarchical modelling of noisy spatial rates on a modestly large and discontinuous irregular lattice. *Statistical Methods in Medical Research* (PMID: 24671659).

Maheswaran R, Staines A. (1997). Cancer clusters and their origins. *Chemistry and Industry* 7:254-6.

Maheswaran R, Payne N, Meechan D, Burden RP, Fryers PR, Wight J, Hutchinson A. (2003). Socioeconomic deprivation, travel distance and renal replacement therapy in the Trent Region, United Kingdom 2000: an ecological study. *Journal of Epidemiology and Community Health* 57:523-4.

Maheswaran, R. and M. Craglia (2004). *GIS in Public Health Practice*, CRC Press, Boca Raton.

Maheswaran R, Haining RP, Brindley P, Law J, Pearson T, Fryers PR, Wise S, Campbell MJ. (2005). Outdoor air pollution and stroke in Sheffield, United Kingdom – a small-area level geographical study. *Stroke* 36:239-43. (a)

Maheswaran R, Haining RP, Brindley P, Law J, Pearson T, Fryers PR, Wise S, Campbell MJ. (2005). Outdoor air pollution, mortality and hospital admissions from coronary heart disease in Sheffield, United Kingdom – a small-area level ecological study. *European Heart Journal* 26:2543-9. (b)

Maheswaran R, Haining RP, Pearson T, Law J, Brindley P, Best NG. (2006). Outdoor NOx and stroke mortality – adjusting for small area level smoking prevalence using a Bayesian approach. *Statistical Methods in Medical Research* 15:499-516. (a)

Maheswaran R, Pearson T, Jordan H, Black D. (2006). Socioeconomic deprivation, travel distance, location of service and uptake of breast cancer screening in North Derbyshire, UK. *Journal of Epidemiology and Community Health* 60:208-12. (b)

Maheswaran R, Pearson T, Munro J, Jiwa M, Campbell MJ, Nicholl J. (2007). Impact of NHS walk-in centres on primary care access times: ecological study. *British Medical Journal* 334:838-41.

Maheswaran R, Pearson T, Smeeton NC, Beevers SD, Campbell MJ, Wolfe CD. (2010). Impact of outdoor air pollution on survival after stroke: population-based cohort study. *Stroke* 41:869-77.

Maheswaran R, Pearson T, Smeeton NC, Beevers SD, Campbell MJ, Wolfe CD. (2012). Outdoor air pollution and incidence of ischemic and hemorrhagic stroke: a small-area level ecological study. *Stroke* 43:22-7.

Maheswaran R, Pearson T, Beevers SD, Campbell MJ, Wolfe CD. (2014). Outdoor air pollution, subtypes and severity of ischemic stroke - a small-area level ecological study. *International Journal of Health Geographics* 13:23. (a)

Maheswaran R, Pearson T, Strong M, Clifford P, Brewins L, Wight J. (2014). Assessing the impact of selective migration and care homes on geographical inequalities in health - a total population cohort study in Sheffield. *Spatial and Spatio-Temporal Epidemiology* 10:85-97. (b)

Nixon G, Samaranayaka A, de Graaf B, McKechnie R, Blattner K, Dovey S. (2014). Geographic disparities in the utilisation of computed tomography scanning services in southern *New Zealand. Health Policy* doi: 10.1016/j.healthpol.2014.05.002.

O'Cathain A, Knowles E, Maheswaran R, Pearson T, Turner J, Hirst E, Goodacre S, Nicholl J. (2014). A system-wide approach to explaining variation in potentially avoidable emergency admissions: national ecological study. *BMJ Quality and Safety* 23:47-55.

Oppong JR, Tiwarri C, Rucktongsook W, Huddleston J and Arbona S. (2012). Mapping late testers for HIV in Texas. *Health & Place* 18: 568-575.

Peng, Z-R. and M-H. Tsou (2003). *Internet GIS: Distributed geographic information services for the Internet and wireless networks.* Wiley, Hoboken, New Jersey.

Perez, L. and Dragicevic, S (2009). An agent based approach for modelling dynamics of contagious disease spread. *International Journal of Health Geographics*, 8:50. doi:10.1186/1476-072X-8-50.

Rhodes, T., Singer, M., Bourgois, P., Friedman, S. and Strathdee, S. (2005). The social structural production of HIV risk among injecting drug users. *Social Science and Medicine*, 61(5), 1026-1044.

Ruiz, M., Walker, E., Forster, E., Haramis, L. and Kitron, U. (2007). Association of West Nile virus illness and urban landscapes in Chicago and Detroit. *International Journal of Health Geographics*, 5, 10. doi:10.1186/1476-072X-6-10.

Sheppard, E., H. Couclelis, S.Graham, J.W.Harrington and H.Onsrud (1999). Geographies of the Information Society. *International Journal of Geographical Information Science*, 13(8), 797-823.

Strong M, Pearson T, MacNab YC, Maheswaran R. (2012). Mapping gender variation in the spatial pattern of alcohol-related mortality: A Bayesian analysis using data from South Yorkshire, United Kingdom. *Spatial and Spatio-Temporal Epidemiology* 3:141-9.

Swerdlow, A. (1992). Cancer incidence data for adults. In P.J.Elliott, J.Cuzick, D.English and R.Stern (eds.) ,*Geographical and Environmental Epidemiology: Methods for small area studies*, pp.51-62. Oxford University Press, Oxford.

Veregin, H. and P. Hargitai (1995). An evaluation matrix for geographical data quality. In S. Guptill, and J.Morrison (eds.), *Elements of Spatial Data Quality*, pp. 167-188. Elsevier Science, Oxford.

Wakefield, J. and P.Elliott (1999). Issues in the statistical analysis of small area health data. *Statistics in Medicine*, 18, 2377-99.

Zenk SN, Schulz AJ, Matthews SA, Odoms-Young A, Wilbur J, Wegrzyn L, Gibbs K, Braunschweig C, Stokes C.(2011). Activity space environment and dietary and physical activity behaviors: a pilot study. *Health and Place* 17:1150-61.