Asthma and factory emissions in northern England: addressing public concern by combining geographical and epidemiological methods

C E Dunn, J Woodhouse, R S Bhopal, S D Acquilla

Abstract

Study objective – The prevalence of asthma was assessed to investigate respiratory health effects of airborne emissions from a factory. A geographical information system allowed flexible definition of study areas in terms of their size, distance, and location in relation to the factory. The value of the approach for this type of investigation is focussed on.

Setting – The factory is located in the south western part of a small market town in County Durham.

Measurements and main results – A total of 1573 asthma cases were identified from general practitioner computerised repeat prescribing systems. Population denominators were defined from family health services authority patient registers. The area within 1 km and immediately to the north east of the factory had an age and sex standardised asthma prevalence 24% (confidence interval 4, 44%) in excess of the expected rate (p = 0.01). The increased prevalence was confined to middle aged and elderly adults living in the area between 0.5 and 1 km to the north east of the factory.

Conclusions – The value of combining the skills of geographers and epidemiologists in addressing public health issues is shown, particularly through the use of geographical information systems which proved powerful and effective.

Department of Geography, University of Durham, South Road, Durham DH1 3LE
C E Dunn

Northern Regional Health Authority, Newcastle upon Tyne, NE6 4PY
J Woodhouse

Division of Epidemiology and Public Health, The Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH
R S Bhopal

County Durham Health Commission, Durham DH1 5XZ
S D Acquilla

Correspondence to:
Dr C E Dunn.
Accepted for publication February 1995

Interest in the effects on health of airborne pollution, particularly from point sources, has increased recently. The Small Area Health Statistics Unit (SAHSU) has a specific role in investigating these problems.1 There remains, however, a challenge to develop both the methodology and information systems for solving this type of problem. We illustrate the benefits of combining geographical information systems (GIS) technology with epidemiological concepts. We examined the respiratory health status of a population living near a single industrial point source, a factory which has been in operation since 1974 and which produces plastic coated wallpaper. Residents living near the factory had expressed concern about airborne emissions, especially in terms of unpleasant smells and the potential risk of cancer. Vapours which are emitted from the factory's chimneys and which drift across nearby housing estates initiated the local concern. Fume scrubbing apparatus was installed in 1989 to reduce emissions. A substance used in the factory processes, azodicarbonamide, is a recognised causal agent in the development of occupational asthma.23

Our null hypothesis was that after suitable standardisation no excess prevalence of asthma would be found in either children or adults in the area of the housing estate closest to the factory when compared with surrounding areas.

Methods

Setting

The factory is located in the south western part of a small market town (population approximately 36 000) in County Durham. Most residential property lies to the north east and south west of the factory with the nearest 200 m immediately to the north east of the plant (figure). The prevailing wind is from the south west, with less frequent winds from the north west or south. The plant emits a potentially complex mixture of substances from the factory ovens (principally paraffin oil and including azodicarbonamide) which are vented from stacks 20 m above ground level.

DATA SOURCES

Numerator and denominator data

A register was compiled for asthma using data from general practitioner computerised repeat prescribing systems. Patients prescribed medication for asthma were assumed to have asthma. There were 1573 such cases covering the three general practices in the study area. Denominator data were obtained from the local family health service authority’s (FHSA) computerised register: these included age, sex, and postcode of every patient registered with the three practices. Most (88%) of the population living in the study area were registered with one of these three general practices.

The unit postcode was converted to an Ordinance Survey National Grid reference by using a computerised file, the Central Postcode Directory (CPD).4 Grid references pertain to the south west corner of the 100 m grid square in which the first house in each unit postcode lies. This may mean that more than one unit postcode, or more than one patient with the same postcode, shares the same grid reference. Recent work5 has shown the implications of
the directional bias involved in using the CPD and has suggested a simple improvement to the original file, whereby a factor of 50 m is added to the northing and easting values for each grid reference. The present study makes use of both sets of grid references, although the modified file is regarded as being more reliable.

Not all patients with asthma, for example those who either never or who always attend the GP’s surgery for a prescription, will be included on the repeat prescribing register. In addition, we cannot exclude the possibilities of diagnostic uncertainty and differing prescribing thresholds by GPs, and their subsequent effects on asthma prevalence. These effects, while difficult to measure precisely, were addressed by means of standardisation of prevalence by general practice of registration as well as by age and sex. The accuracy of the denominator data may vary from practice to practice. Recent validation by the FHSA (1990) of their records against the NHS central registry has, however, led the FHSA to be confident about a much improved accuracy and reliability. The three general practices in the study area also performed checks of their own records against the FHSA register in preparation for fundholding status. Levels of discrepancy for GPs in the study area were reported as 0.6% for one practice and 6% for a second. No formal verification of records against the FHSA register was performed by the third practice, although assurance was given that the practice was diligent in informing FHSA of changes in the practice list.

Definition of study sectors and distribution of the resident population around the factory.
Geographical and epidemiological methods in asthma study

Socioeconomic data
Data from the 1981 census of population were used to examine the importance of socioeconomic factors in the prevalence of asthma: data for the 1991 census were not available at the time of the study. A set of social deprivation indicators was derived for each enumeration district (ED) to reflect wealth and income (percentage households owner occupied and percentage households with no car) and living circumstances (percentage households with more than one person per room, percentage single pensioner households and percentage single parent households).6

METHODOLOGY AND ANALYSIS
Data for both asthma patients and total background populations were imported into a GIS package, ARC/INFO.7 Descriptive (attribute) data for each patient, comprising age, sex, unit postcode, and general practice were held in the database management component of the system (INFO). For 5.2% of the total practice list and 6.1% of asthma patients, no grid reference was identifiable, either because of a missing postcode field or no grid reference match in the CPD (table 1).

A set of sectors, at 1 km intervals centred on the factory, was defined using the GIS (figure) and the age–sex standardised asthma prevalence for the population contained within each sector was calculated. This is achieved by using a procedure known as "point-in-polygon" overlay, which involves superimposing each sector with the complete set of grid reference point data (separately for asthma cases and resident population), so that those points falling in a particular sector are maintained together as an individual file together with the attribute data.

The INFO database was then used to classify the data into five-year age bands; age–sex specific prevalence rates, and standardised prevalence rates and ratios were calculated using purpose built software (SMRATE).8 Two standard populations were defined: that within a 4 km radius of the factory, in which 85% of asthma cases resided, and that within 11 km, covering 99.9% of cases. Significance levels were adjusted, depending on the number of sectors being standardised. The prevalence calculated for each sector was compared with the overall prevalence within the study area.

Table 1  Asthma case register and practice lists in relation to the general practice showing loss of data through missing postcodes or grid references

<table>
<thead>
<tr>
<th>Practice</th>
<th>No of asthma patients</th>
<th>Practice list</th>
<th>Crude prevalence (%)</th>
<th>Asthma patients grid referenced</th>
<th>Practice list grid referenced</th>
<th>Crude prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>470</td>
<td>11,980</td>
<td>3-92</td>
<td>444</td>
<td>11,374</td>
<td>3-90</td>
</tr>
<tr>
<td>2</td>
<td>691</td>
<td>13,478</td>
<td>3-13</td>
<td>663</td>
<td>12,871</td>
<td>5-15</td>
</tr>
<tr>
<td>3</td>
<td>514</td>
<td>10,769</td>
<td>4-77</td>
<td>466</td>
<td>10,107</td>
<td>4-41</td>
</tr>
<tr>
<td>Total</td>
<td>1675</td>
<td>36,227</td>
<td>4-62</td>
<td>1573</td>
<td>34,352</td>
<td>4-58</td>
</tr>
</tbody>
</table>

Table 2  Age and sex standardised rates and ratios in relation to sector based on standard population residing within a 4 km radius circle of the factory and using modified Central Postcode Directory grid references

<table>
<thead>
<tr>
<th>Sector</th>
<th>No actual cases</th>
<th>No expected cases</th>
<th>Standard rate (per 1000)</th>
<th>Standard rate (99%)</th>
<th>Confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE1000</td>
<td>199</td>
<td>160-78</td>
<td>57-5</td>
<td>1-24</td>
<td>1-04, 1-44</td>
</tr>
<tr>
<td>NE500</td>
<td>44</td>
<td>46-27</td>
<td>43-7</td>
<td>0-95</td>
<td>0-57, 1-33</td>
</tr>
<tr>
<td>NE500-1000</td>
<td>155</td>
<td>114-51</td>
<td>60-8</td>
<td>1-01</td>
<td>0-83, 1-17</td>
</tr>
<tr>
<td>NE2000</td>
<td>315</td>
<td>325-04</td>
<td>45-5</td>
<td>1-07</td>
<td>1-03, 1-10</td>
</tr>
<tr>
<td>NE3000</td>
<td>241</td>
<td>240-98</td>
<td>46-0</td>
<td>0-91</td>
<td>0-45, 1-34</td>
</tr>
<tr>
<td>NE4000</td>
<td>91</td>
<td>81-49</td>
<td>50-5</td>
<td>0-88</td>
<td>0-51, 1-24</td>
</tr>
<tr>
<td>NW1000</td>
<td>32</td>
<td>35-24</td>
<td>40-8</td>
<td>1-00</td>
<td>0-83, 1-17</td>
</tr>
<tr>
<td>NW2000</td>
<td>43</td>
<td>49-09</td>
<td>43-5</td>
<td>0-97</td>
<td>0-83, 1-11</td>
</tr>
<tr>
<td>NW3000</td>
<td>20</td>
<td>26-73</td>
<td>35-0</td>
<td>0-82</td>
<td>0-45, 1-34</td>
</tr>
<tr>
<td>NW4000</td>
<td>84</td>
<td>78-82</td>
<td>50-2</td>
<td>1-07</td>
<td>0-78, 1-36</td>
</tr>
<tr>
<td>SW1000</td>
<td>41</td>
<td>40-55</td>
<td>50-4</td>
<td>0-91</td>
<td>0-61, 1-42</td>
</tr>
<tr>
<td>SW2000</td>
<td>95</td>
<td>100-23</td>
<td>44-9</td>
<td>0-78</td>
<td>0-69, 1-21</td>
</tr>
<tr>
<td>SW3000</td>
<td>71</td>
<td>90-59</td>
<td>36-9</td>
<td>0-98</td>
<td>0-51, 1-06</td>
</tr>
<tr>
<td>SW4000</td>
<td>0</td>
<td>1-9</td>
<td>0</td>
<td>1-00</td>
<td>0-69, 1-45</td>
</tr>
<tr>
<td>SE1000</td>
<td>32</td>
<td>32-64</td>
<td>50-7</td>
<td>0-98</td>
<td>0-53, 1-43</td>
</tr>
<tr>
<td>SE2000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0-83</td>
<td>0-22, 1-43</td>
</tr>
<tr>
<td>SE3000</td>
<td>15</td>
<td>18-14</td>
<td>35-2</td>
<td>0-83</td>
<td>0-69, 1-45</td>
</tr>
<tr>
<td>SE4000</td>
<td>49</td>
<td>45-79</td>
<td>50-5</td>
<td>1-07</td>
<td>0-69, 1-45</td>
</tr>
</tbody>
</table>

* Significant.
creased prevalence was due to an increased rate among the middle aged and elderly (table 3).

**CONFOUNDING VARIABLES**

Potential explanations for the raised asthma prevalence ratio include confounding of the association between asthma and proximity to the factory by the following: socioeconomic deprivation, including housing conditions; migration, daily travel, and occupation; smoking; and variable clinical practice among general practitioners. These are considered in turn below.

**Effects of socioeconomic deprivation**

The ward which comprises almost all of sector NE1000 was ranked as having the fifth worst health and the 13th worst deprivation of all wards in the northern health region. We calculated asthma prevalence for a neighbouring area (NE2000a), part of the ward forming part of sector NE2000, which had some socioeconomic deprivation indices which were similar to those of sector NE1000 (figure). Using standard grid references, the standardised asthma prevalence for sector NE2000a was significantly raised at the 95% significance level (1.48, 95% confidence interval (CI) 1.09–1.88) and for adjusted grid references just failed to reach significance at 95% (1.35, 95% CI 0.96–1.70). Clearly socioeconomic deprivation as a confounder is an important potential explanation for our findings. Poor housing conditions, notably the presence of mould or damp and indoor air pollution from cooking or heating systems, may be implicated in asthma prevalence and deserve close attention in future studies.

**Migration**

Inward and outward migration is approximately 5–10% (1990/91 data) for the ward in sector NE1000 (personal communication, Wear Valley District Council). In addition, the potential effect of selective migration (that is, of "cases" or "healthy subjects" into or out of the study area) cannot be dismissed. Daily migration and activities may also be significant, should these involve different degrees of exposure in different areas to vehicle exhaust emissions or to asthma inducing agents in the workplace. The possibility that those living close to a factory are more likely to work there seems unlikely to account for the observed increased prevalence in the present study for three reasons. Firstly, very few (approximately 5) of the factory workforce actually live in the area with increased asthma prevalence. Secondly, only a small number of workers in the factory have been suspected of developing industrial asthma. Thirdly, the age profile indicates that the excess prevalence is present in the older age groups, and not in the 15–44 age group which is also of working age.

**Smoking**

A recent health and lifestyle study indicated that the percentage of people smoking in the population in the ward in sector NE1000 was higher than that in the ward in sector NE2000b (46% and 39%, respectively). Both wards generally have a smoking prevalence higher than that found nationally.

**Clinical practice**

If the population in the ward in NE1000 was more likely to be cared for by a general practice with a greater tendency to prescribe than general practices caring for neighbouring populations, then its asthma incidence would be raised. Using GIS to help calculate a rate standardised for age, sex, and the group general practice of registration, the prevalence rate for the NE1000 sector was 60/4/1000, compared with 57–5/1000 when the data were standardised for age and sex only. Standardising for general practice therefore strengthens the argument for increased asthma prevalence in this sector, and it seems unlikely, therefore, that variation in prescribing practice is the explanation for the spatial differences demonstrated.

**Conclusions and discussion**

Middle aged and elderly adults living 500–1000 m in a north easterly direction from the factory had a higher prevalence of asthma than expected. Is this increased prevalence real or apparent and, if real, what may be the explanatory factor(s)? There are a number of reasons why a spurious result may have occurred. Firstly, the results rely on the assumption that patients receiving anti-asthma medication on repeat prescription have asthma and vice versa. This assumes accurate and consistent use of diagnostic criteria by prescribing GPs. Spatial vari-
Geographical and epidemiological methods in asthma study

Geographical epidemiological methods for authorities areal units identifying new of descriptive mapping, from handling in invention described efficient through playing manipulating, as market in several potential applications. The role of GIS in these developments is underway to air pollution, including those from the factory. New research is underway to study the relative contributions of these factors. We now turn to the methodological aspects of the work, particularly the role of GIS in geographical research of this kind.

The study has demonstrated the value of using GIS for analysing primary care morbidity data. However, standardisation for group practice did not affect our main findings. Secondly, there is the issue of error in the grid references. Postcoded data are converted to grid references which are nominally accurate to the nearest 100 m, although in many cases accuracy will be worse than this.12,13 It is unlikely, however, that inaccuracy is confined to one or more particular geographical "sectors" of the area under study and inaccuracy is, therefore, unlikely to explain the findings.

The nature of the potential association between social class and prevalence of asthma is clear, although some studies suggest that asthma may be more prevalent in lower social groups.14 In the present study an attempt was made to compare the population for the sector of particular interest (NE1000) with a socio-economically similar population nearby. This latter population was found to show weaker evidence of raised asthma prevalence, although this result is statistically less robust because the total population is smaller. Furthermore, the area in which this population resides is geographically close to that of sector NE1000 (figure). The apparently raised prevalence found among both these populations may be due to any one of four factors: errors in the data, chance (unlikely), confounding factors, and environmental exposure to air pollutants, including those from the factory. New research is underway to study the relative contributions of these factors. We now turn to the methodological aspects of the work, particularly the role of GIS in geographical research of this kind.

The study has demonstrated the value of using GIS for analysing primary care morbidity data. GIS are being increasingly used as spatial analytical tools by geographers, planners, and others, both in health related research and in several other application areas (including emergency planning, resource management, and market analysis). A GIS may be formally defined as "a system for capturing, storing, checking, manipulating, analysing and displaying data which are spatially referenced."15 Through a combination of a database management system, graphical display, and spatial analysis, GIS provides the ability to relate together and analyse several datasets, often originating from disparate sources, in a rapid and efficient manner; these systems have been described as "the biggest step forward in the handling of geographic information since the invention of the map."15

As a tool for helping to assess spatial patterns of disease, GIS not only represents a move away from more traditional approaches based on descriptive mapping, but has also helped to initiate new insights into disease distribution.16 One of the most important features of a GIS is that users can define their own geographical areas of interest. This permits the screening of geographically related health phenomena, identifying effects that might otherwise go undetected. Thus, instead of relying on predefined areal units such as electoral wards or health authorities for geographical analyses, the user is free to identify specific spatial zones, of any size or shape (within limits) for which health status may be investigated. This flexibility allows the researcher to change the scale of investigation and focus on areas of interest which more appropriately reflect risk exposure. Thus, in the present study, sectors were centred on a single point (an industrial site) and were designed to reflect dispersal patterns of airborne emissions, as determined by wind direction. Sectors of interest were then geographically subdivided, according to underlying socioeconomic characteristics, to be used as a basis for analysing asthma prevalence. Equally, zones of interest might relate to linear features such as increasing distances from a busy road network, to test potential associations between disease and vehicle exhaust emissions.

Future studies will capitalise on recent developments in GIS research. In particular, work by Diggle and Rowlingson17 has been critical in investigating raised incidences of disease in relation to putative environmental sources which may be point sources of contamination18 or linear sources such as roads. Spatial point pattern modelling distinguishes the presence or absence of clustering in a case-control approach, which incorporates not only distance from the site of interest, but also other appropriate explanatory variables specified by the user as binary covariates, for example, smoking behaviour, occupation, or housing conditions. This approach is potentially valuable since it allows investigation of multiplicative risk factors in a single mathematical function which avoids the need to use areal units.

The present study has helped to highlight a number of issues in relation to empirical studies in geographical epidemiology. Firstly, the value of working with GIS. Secondly, the need for accurate geographically referenced data sources. Errors and inaccuracies in spatial referencing of postcoded data limit the extent to which we can be precise about the spatial patterns of morbidity. Thirdly, the value of an interdisciplinary approach to addressing a public health issue. Geographers and epidemiologists have separate skills and perspectives to bring to these issues, but only when these are combined in collaborative work can the most fruitful outcome be achieved. Fourthly, the study was initiated by residents' claims about their exposure to emissions from a specific source; we have shown that significant variations in health status may be demonstrated in responding to public concern and anxieties.

Finally, the present study has identified the need for further research, particularly on the effects of socioeconomic differences on asthma and the possibility that an industrial process emitting an agent previously identified as a cause of occupational asthma, may have affected the health of a local population. In this context, we are preparing atmospheric dispersion models to predict the spatial distribution of pollutants since our use of sectors of a circle to define exposure reflects only crudely the effects of local meteorological conditions. Modelled output in the form of contours or isolines of pollution will enable more...
realistic zones of exposure to be overlain onto health datasets in a GIS. Further, modelled values will be calibrated by relating predicted levels of air pollution to measured values from air quality monitoring in the local community.

We thank an anonymous referee for helpful criticisms.