Distribution of cases of systemic lupus erythematosus at time of first symptom in an urban area

N D Hopkinson, K R Muir, M A Oliver, M Doherty, R J Powell

Abstract

**Objectives**—To determine the geographical distribution of cases of systemic lupus erythematosus (SLE) in a defined geographical area in the East Midlands, UK, and, in particular, to search for spatial variation in cases that may implicate the role of environmental factors in SLE aetiology.

**Methods**—Six methods of case ascertainment were used. The postcode of the patient’s domicile at time of first definite symptom of SLE was used for analysis which included case mapping, probability mapping by electoral ward, and variogram analysis.

**Results**—The study area population of 613,700 contained 200 SLE patients, 188 of whom experienced their first symptom whilst residing in the area. Case mapping revealed 12 SLE patients residing within an area of one square mile, including four men and six patients with RNP antibodies. The use of probability mapping showed five wards in close proximity to each other to have a greater number of SLE cases than would be expected by chance (p < 0.1). The ‘cluster’ of patients seen on the case map fell into two wards which showed a significant excess of cases only when combined (p = 0.006). The variogram of the incidence rates for each ward did not confirm any structure or pattern to the distribution of cases for the whole area.

**Conclusions**—Some areas have a greater than expected prevalence of SLE. The normal result from variogram analysis suggests that the cause(s) for these excess numbers of cases does not have an effect across the whole study area.


There is marked variation in the geographic distribution of many human diseases, and Snow’s investigation of the clustering of cases of cholera around a water pump in London in 1848 led to the subsequent identification of its causative agent. Similarly, Burkitt’s geographical analysis of the lymphoma that now bears his name has given insight into its possible causes. More recently, the distribution of other cancers, particularly leukaemia, has been scrutinised using a variety of techniques and such studies have led to hypotheses relating environmental factors to the occurrence of a disease.

Systemic lupus erythematosus (SLE) is an uncommon disease (prevalence, 24.7/100,000) in which genetic and environmental factors are likely to be important. Genetic influences are notable in complement deficiencies and in twin concordance studies. The importance of environmental effects has been argued on the basis of geographical differences in prevalence rates between so-called common genetic stock, for example black people in Africa and those in North America and, similarly, Chinese groups living in Malaysia, Hawaii, and North America. It is well established also that a number of drugs, for example hydralazine and procainamide, can induce a disease which is not dissimilar to idiopathic SLE. However, the spatial distribution of SLE has received little attention.

The geographical distribution of SLE in the Nottingham Health District, UK, has been studied to assess associations with environmental factors. Methods of data analysis applied to cases of SLE included probability mapping, and variogram analysis.

**Methods**

The study was approved by the local ethics committee. The study area was that of the Nottingham Health District, which is served by two main hospitals, the Queen’s Medical Centre (QMC) and Nottingham City Hospital. An attempt was made to identify all patients with SLE residing in this area during the period 1 May 1989–30 April 1990.

**Identification of Patients**

Six principal methods of ascertainment were used, details of which have been published previously. From these sources, patients with definite or possible SLE who resided in the study area were contacted and invited to attend the hospital to consult one of the investigators (NH).

**Patient Interview**

At interview the diagnosis of SLE was confirmed, or refuted, based on the principles of Fries and Holman, using evidence of multisystem disease, immunological abnormalities (for example positive ANA with or
without antibodies to dsDNA) and in the absence of a more appropriate diagnosis. For this study patients were included regardless of the number of American Rheumatism Association (ARA) criteria that they fulfilled. The time of the first definite symptom of SLE was identified retrospectively and the patient's address and postcode at that time were recorded. If the patient had lived there for less than six months, their previous address was used for analysis.

The expected number of cases of SLE in each electoral ward was calculated assuming a uniform distribution of SLE throughout the area. The incidence rates for each ward were tested against their expected number using a cumulative Poisson probability test and the wards ranked according to techniques described previously. To highlight those wards at the top of this table, a p value of 0.1 was arbitrarily taken as being significant. This is superior to a simple observed:expected ratio because it allows for ward population. The p values quoted are one sided.

Geostatistical analysis

Theory: variogram analysis. The variogram is the central tool of geostatistics, and is a set of analytical procedures based on the regionalised variable theory, developed by Matheron. It has been used widely for describing and estimating the spatial distribution of metals, diamonds, soil properties, and atmospheric pollutants. Geostatistics provides a statistical tool for describing variation across the Earth's surface. A recent development of the method has been applied to geographical patterns of rare diseases such as the analysis of the risk of childhood cancer, in which it has proved valuable. However, the technique is relatively insensitive to point clusters.

The variogram. Following geostatistical convention, the rate of incidence is treated as a variable, Z(x), distributed in space in which x denotes the coordinates, xl and x2, for the two dimensions at the centre (or centroid) of each ward. Where a variable is spatially dependent, the variance of any pair of values depends on the separation in both distance and direction between them—the 'lag', denoted by the vector h. The variance of the differences between any two centroids a given distance apart is known as the semivariance, g, because it is the variance per site when the sites are considered in pairs. It is defined by:

\[ g(h) = 1/2E[(Z(x) - Z(x + h))^2] \]

where Z(x) and Z(x + h) are the value of the property at any two places, x and x + h, separated by h, and E denotes the expectation. The function g(h) that relates the variance to the lag is the variogram.

Results

On 1 May 1989, 163 people with SLE were residing in the area and a further 37 presented between that date and 30 April 1990. Of these 200 cases (177 women, 23 men), 188 experienced their first symptom while residing in Nottingham and were therefore suitable for inclusion in the geographical analysis.

Case map of Nottingham

The map of cases for Nottingham, plotted according to the individuals' domicile at the time of first symptom of SLE, independent of ARA criteria, showed that they were not distributed uniformly (fig 1). One area to the

The study area

The defined area approximated to the metropolitan community of Greater Nottingham, which consists of the districts of Nottingham, Broxtowe, Gedling, Rushcliffe, and the Hucknall wards of Ashfield. Based on 1988 mid-year estimates, the population of the area was 613 700 (300 600 males, 313 100 females). The age and sex distribution of this local population accords with that of the UK. All the major social classes are represented, with a distribution similar to that in the whole of England and Wales. The main industries in the area include mining, quarrying, and the manufacture of textiles, chemicals, and tobacco.

Geographical distribution of SLE

The home of each patient with SLE was referenced by the postcode, according to the above criteria, and from this the geographical coordinates were derived. The electoral ward to which each case belonged was also derived from the postcode and this was referenced by the coordinates of the ward centroid (the population weighted centre). The cases were plotted on a map of Nottingham Health District to examine the spatial distribution. The rate of incidence of SLE for each ward was defined as the number of cases divided by the population in that ward:

\[ F(x_i) = L(x_i)/n(x_i) \]

where \( F(x_i) \) = rate of incidence or frequency, \( L(x_i) \) = number of cases in ward i, and \( x_i \) denotes the coordinates of the centroids of the wards.

We used two analytical approaches to assess potential patterns in the distribution of cases. These were chosen from a range of techniques in order to identify different types of pattern that may be present, as described further below.

Poisson probability test

The Poisson analysis is based on the concept that if cases are distributed independently of one another in geographical space with an intensity related only to population density, then the difference between the observed and expected number of cases will follow a Poisson distribution. The aim is to identify anomalies in the distribution of cases that might indicate real differences in the risk of developing the disease from area to area.

The study area

The defined area approximated to the metropolitan community of Greater Nottingham, which consists of the districts of Nottingham, Broxtowe, Gedling, Rushcliffe, and the Hucknall wards of Ashfield. Based on 1988 mid-year estimates, the population of the area was 613 700 (300 600 males, 313 100 females). Based on 1988 mid-year estimates, the population of the area was 613 700 (300 600 males, 313 100 females). The age and sex distribution of this local population accords with that of the UK. All the major social classes are represented, with a distribution similar to that in the whole of England and Wales. The main industries in the area include mining, quarrying, and the manufacture of textiles, chemicals, and tobacco.
north of Nottingham yet within the Borough of Nottingham City, was notable for having an apparent 'cluster' of SLE cases. This area included five patients who were living within an area less than 0-5 mile\(^2\) at the time of their first symptom. Three of these patients were men, and three had high titre antibodies against RNP. Interestingly, the domiciles of two of the men patients were in the same road when their symptoms arose. A further seven patients, including three with antibodies to RNP, were living within one mile of this central area. None of these patients was related, knew each other, or had the same occupation. Onset of the patients’ symptoms occurred over the period 1975–89. Table 1 shows the clinical details of the 12 patients in this area, who were referred by different General Practitioners.

**Table 1 Clinical features of the cluster of 12 patients with systemic lupus erythematosus observed in the wards of Bulwell West and East**

<table>
<thead>
<tr>
<th>No</th>
<th>Sex</th>
<th>Age (yr)</th>
<th>Main clinical features</th>
<th>Autoantibodies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>46</td>
<td>Arthritis</td>
<td>ANA 1:400, ddsDNA</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>45</td>
<td>Malar rash, arthritis, pericarditis, migraine</td>
<td>ANA 1:1600, Ro, dsDNA</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>60</td>
<td>Subacute cutaneous LE</td>
<td>ANA 1:400, Ro</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>20</td>
<td>Arthritis, photosensitivity</td>
<td>ANA 1:6400, dsDNA, RNP</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>43</td>
<td>Arthritis, oral ulcers</td>
<td>ANA 1:200, dsDNA, RNP</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>44</td>
<td>Arthritis, seizures</td>
<td>ANA 1:1600, Ro, La</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>37</td>
<td>Arthritis, oral ulcers</td>
<td>ANA 1:800</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>43</td>
<td>Arthritis</td>
<td>ANA 1:1600, dsDNA</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>16</td>
<td>Severe Raynaud’s</td>
<td>ANA 1:1600, RNP</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>35</td>
<td>Raynaud’s, arthritis, myositis</td>
<td>ANA 1:3200, RNP</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>37</td>
<td>Raynaud’s, arthritis</td>
<td>ANA 1:6400, dsDNA, RNP</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>51</td>
<td>Arthritis, pleurisy</td>
<td>ANA 1:3200, Ro, RNP</td>
</tr>
</tbody>
</table>

ANA = Antinuclear antibodies; dsDNA = double stranded DNA; RNP = ribonucleoprotein.

Using the study area population of 613,700 the rate of SLE patients presenting with their first symptom of SLE was 188/613,700 × 100,000 = 30.63/100,000 population.

Assuming an even spread of SLE across the population of the study area, the expected and observed number of cases per electoral ward was derived. Of 104 wards, 34 had no cases of SLE, 23 had one case and 47 had more than one case. Table 2 shows the wards with the smallest values of p. Clearly, this represents the top of a long list of electoral wards, the majority of which show a rate of occurrence of SLE similar to that expected by chance. At the bottom of the list are some wards that showed no cases of SLE despite being well populated. Using the incidence rate of 30.63/100,000, all wards with a population of 3265 or more would be expected to have at least one case of SLE. Table 3 shows the 16 wards in the study area with larger populations than this, but with no case of SLE. Figure 1 shows the ratios of the observed to the expected number of cases of SLE for each electoral ward.

Using a Poisson probability of less than 0·01 as being significant, two wards had a greater than expected number of SLE cases: Carlton Hill in the Borough of Gedling and Lenton in the Borough of Nottingham. In figure 1, the geographical distribution of wards with a Poisson probability of p < 0·1 is highlighted by a circular symbol. Two such wards, Trent and Bridge, are adjacent to each other; and a further two, Lenton and Carlton Hill, are within two wards of these. Further away is the ward of Bulwell West. The cluster of 12 cases detailed above was contained within the wards of Bulwell West and Bulwell East. Neither ward on its own (table 2) showed an excess of cases using conventional statistical significance, but a significant excess in the number of SLE cases (observed 12, expected 6·94; p = 0·006) was apparent when these two wards were combined.

Population pyramids for the five areas with a p value less than 0·1 compared with the study area as a whole, have been constructed (data not shown). Bulwell West has a considerably higher proportion of children aged 0–14 compared with the whole study population, which would reduce the expected at risk population for SLE, with consequent enhancement of the significance of the p value. In this SLE cohort, the greatest incidence rates were observed in the 40–59 year age group, which was in fact under represented in these high incidence wards, again increasing the significance of their observation.
Table 2  Mapping of cases of systemic lupus erythematosus (SLE) by place of residence at first symptom, according to electoral ward

<table>
<thead>
<tr>
<th>Electoral ward</th>
<th>Popn</th>
<th>Obs cases</th>
<th>Rate (1/10)</th>
<th>×Area rate</th>
<th>Exp cases</th>
<th>Excess p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carlton</td>
<td>6255</td>
<td>7</td>
<td>1·12</td>
<td>3·65</td>
<td>1·92</td>
<td>5·08</td>
</tr>
<tr>
<td>Lenton</td>
<td>7535</td>
<td>7</td>
<td>0·93</td>
<td>3·03</td>
<td>2·31</td>
<td>4·69</td>
</tr>
<tr>
<td>Trent</td>
<td>9300</td>
<td>6</td>
<td>0·65</td>
<td>2·11</td>
<td>2·85</td>
<td>3·15</td>
</tr>
<tr>
<td>Bridge</td>
<td>9765</td>
<td>6</td>
<td>0·61</td>
<td>2·01</td>
<td>2·99</td>
<td>3·01</td>
</tr>
<tr>
<td>Bulwell W</td>
<td>12965</td>
<td>7</td>
<td>0·54</td>
<td>1·76</td>
<td>3·97</td>
<td>3·03</td>
</tr>
<tr>
<td>Oxclose</td>
<td>3705</td>
<td>3</td>
<td>0·81</td>
<td>2·64</td>
<td>1·13</td>
<td>1·87</td>
</tr>
<tr>
<td>Robin Hood</td>
<td>10735</td>
<td>6</td>
<td>0·56</td>
<td>1·82</td>
<td>3·29</td>
<td>2·71</td>
</tr>
<tr>
<td>Beeston C</td>
<td>6310</td>
<td>4</td>
<td>0·63</td>
<td>2·07</td>
<td>1·93</td>
<td>2·07</td>
</tr>
<tr>
<td>Bulwell E</td>
<td>9690</td>
<td>5</td>
<td>0·52</td>
<td>1·66</td>
<td>2·97</td>
<td>2·03</td>
</tr>
<tr>
<td>Hucknall W</td>
<td>9820</td>
<td>5</td>
<td>0·51</td>
<td>1·66</td>
<td>3·01</td>
<td>1·99</td>
</tr>
</tbody>
</table>

Popn = Population; Obs cases = observed number of SLE cases; Rate = ward rate of SLE cases per 1000 population; ×Area rate = ward rate at study area rate; Exp cases = expected number of SLE cases assuming an even spread of SLE per head across the study area; Excess = excess of observed cases over expected cases.

Table 3  Wards in the study area without any case of systemic lupus erythematosus (SLE) despite having an expected number > 1

<table>
<thead>
<tr>
<th>Electoral ward</th>
<th>Popn</th>
<th>Obs cases</th>
<th>Exp cases</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basford</td>
<td>9655</td>
<td>0</td>
<td>2·96</td>
<td>0·052</td>
</tr>
<tr>
<td>St Anns</td>
<td>8875</td>
<td>0</td>
<td>2·72</td>
<td>0·066</td>
</tr>
<tr>
<td>Bingham</td>
<td>7490</td>
<td>0</td>
<td>2·29</td>
<td>0·10</td>
</tr>
<tr>
<td>Calverton</td>
<td>6635</td>
<td>0</td>
<td>2·03</td>
<td>0·13</td>
</tr>
<tr>
<td>Kingwell</td>
<td>5925</td>
<td>0</td>
<td>1·81</td>
<td>0·16</td>
</tr>
<tr>
<td>Stapleford W</td>
<td>5785</td>
<td>0</td>
<td>1·77</td>
<td>0·17</td>
</tr>
<tr>
<td>Stapleford E</td>
<td>5775</td>
<td>0</td>
<td>1·77</td>
<td>0·17</td>
</tr>
<tr>
<td>Hucknall E</td>
<td>5665</td>
<td>0</td>
<td>1·74</td>
<td>0·18</td>
</tr>
<tr>
<td>Toten</td>
<td>5570</td>
<td>0</td>
<td>1·71</td>
<td>0·18</td>
</tr>
<tr>
<td>Bonington</td>
<td>4910</td>
<td>0</td>
<td>1·50</td>
<td>0·22</td>
</tr>
<tr>
<td>Melton</td>
<td>4845</td>
<td>0</td>
<td>1·48</td>
<td>0·23</td>
</tr>
<tr>
<td>Eastwood E</td>
<td>4835</td>
<td>0</td>
<td>1·48</td>
<td>0·23</td>
</tr>
<tr>
<td>Bestwood Park</td>
<td>4535</td>
<td>0</td>
<td>1·39</td>
<td>0·25</td>
</tr>
<tr>
<td>Marham</td>
<td>4430</td>
<td>0</td>
<td>1·36</td>
<td>0·26</td>
</tr>
<tr>
<td>Burton Joyce</td>
<td>3795</td>
<td>0</td>
<td>1·16</td>
<td>0·31</td>
</tr>
<tr>
<td>Beeston C</td>
<td>3540</td>
<td>0</td>
<td>1·08</td>
<td>0·34</td>
</tr>
</tbody>
</table>

Popn = Population; Obs cases = observed number of SLE cases; Rate = ward rate of SLE cases per 1000 population; Exp cases = expected number of SLE cases assuming an even spread of SLE per head across the study area.

Figure 2  Variogram of frequency of cases of systemic lupus erythematosus in Nottinghamshire.

VARIOGRAM ANALYSIS

The variogram of the incidence rates for each ward suggested that there was no apparent structure or pattern for the defined study area (fig 2).

Discussion

When the place of residence of the patient at the time of their first symptom of SLE was mapped, the large number of cases in the Bulwell area and particularly the close proximity of the individuals led us to consider complex geographical analyses. The cumulative Poisson probability test for the wards of Bulwell West and Bulwell East taken together showed that there was a greater number of SLE cases than would be expected by chance.

The Poisson probabilities for the individual wards of Bulwell West and East were not statistically significant because of the arbitrary subdivision of large areas by the ward boundaries. This highlights one of the main disadvantages of probability mapping: some areas of high incidence may be missed because areas are divided haphazardly. The technique is relatively insensitive and will only detect substantial aggregations of SLE. It has previously been emphasised in an investigation of childhood cancer in the West Midlands, that probability mapping is a ‘first level’ tool of analysis and that other methods should be applied to data before final conclusions can be drawn. Nevertheless, it has the advantage of being a simple and quick technique that will highlight a very ‘abnormal’ pattern such as that seen with SLE cases in Nottingham.

With the use of a Poisson probability test, a number of false positive results may be expected by chance. However, five wards with a Poisson probability of p < 0·1 seemed to aggregate into two areas (fig 1), the first comprising Carlton Hill, Lenton, Trent, and Bridge, and the second Bulwell West. Such aggregation of wards with a ‘significant’ Poisson probability makes it less likely that they have occurred by chance.

Non-environmental factors could be responsible for an unusually large or small number of cases. There may be variation of General Practitioner referral patterns across the study area, which encompasses both recognition of diagnosis and demographic differences in the ward populations. However, comparison of population pyramids between individual wards and the whole study area suggests that age distribution does not account for the excess of cases seen in those wards with a significant p value (<0·1). A number of wards demonstrated an absence of SLE cases despite having a population size that would have been predictive of at least one case, as calculated from the expected ward prevalence rates. Conversely, however, if environmental factors alone were responsible for SLE, one might argue that no cases should be the norm. Using Poisson probability testing, no statistical significance was demonstrated for any ward; however, it should be noted that the Poisson distribution curve is skewed to the left, and the resulting power to detect statistical significance in the wards with no cases is thus weak.

The calculated variogram of SLE incidence in Nottingham failed to confirm the discrete cluster suggested visually by the case map and by subsequent Poisson analysis. Variogram analysis may appear superior to probability mapping because it is not constrained by the borders of geographical wards and avoids sudden transitions between neighbouring areas. However, this technique is best suited for detecting spatial variation across a large area, rather than a local concentration of cases such as the
apparent SLE cluster. Consequently, the cause(s) of the excess cases seen in Bulwell West, Bulwell East and other wards do not appear to have a discernible effect across the whole area.

The development of SLE may depend on a genetically susceptible individual being exposed to some environmental factor(s). However, as little is known about the incubation period of SLE, it is possible that such environmental factors may promote the expression of existing but previously dormant disease.

In conclusion, some areas in Nottingham have a higher than expected prevalence of SLE; however, current statistical methodologies, especially those adapted from non-medical spatial analyses, do not support the clustering seen on case mapping. Further investigations are in progress to define relevant local environmental factors and their possible role in disease causation.

The study was financed by Lupus UK, to whom we are indebted. We also wish to thank all our colleagues for their help with this study, and Mrs S Todd for assistance with computer analysis.

Distribution of cases of systemic lupus erythematosus at time of first symptom in an urban area.
N D Hopkinson, K R Muir, M A Oliver, M Doherty and R J Powell

Ann Rheum Dis 1995 54: 891-895
doi: 10.1136/ard.54.11.891

Updated information and services can be found at:
http://ard.bmj.com/content/54/11/891

These include:

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

Notes

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

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

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