An analysis of the geographical distribution of severe malaria in children in Kilifi District, Kenya

Joanna Armstrong Schellenberg, James N Newell, Robert W Snow, Victor Mung’ala, Kevin Marsh, Peter G Smith, and Richard J Hayes

Background
Although malaria is known to be a major cause of child mortality and morbidity throughout sub-Saharan Africa there are few detailed studies of malaria mortality rates and incidence of severe malarial disease in defined communities. We have studied the geographical pattern of admissions to hospital with severe malaria and the stability of this pattern over time in Kilifi District on the Kenyan Coast.

Methods
Over a 2-year period all children under 5 years of age with severe malaria admitted to the district hospital and living in a rural study population of about 50,000 people were identified. Annual censuses were carried out in the study area, and all households were mapped using a hand-held satellite navigation system. The resulting databases were linked using a geographical information system (GIS).

Results
Using methods originally developed for the study of the geographical distribution of childhood leukaemia we assessed the spatial pattern of hospital admission rates for severe malaria. As expected, admission rates were significantly higher in children with easier access to the hospital. For example, those living more than 25 km from the hospital had admission rates which were about one-fifth of those for children living within 5 km of the hospital. Those living more than 2.5 km from the nearest road had admission rates that were about half of those for children living within 0.5 km of a road. We also investigated short-term local fluctuations in severe malaria and found evidence of space-time clustering of severe malaria.

Conclusions
Hospital admission rates for severe malaria are higher in households with better access to hospital than in those further away. The finding of space-time clusters of severe malaria suggests that it would be of value to conduct case-control studies of environmental, genetic and human behavioural factors involved in the aetiology of the disease.

Keywords
Severe malaria, Kenya, geographical information system, mapping, spatial pattern, space-time clustering

Accepted 23 May 1997
unclear why a small proportion of those infected suffer severe disease or die while others have only a series of relatively benign episodes before gaining protective immunity. As there are no reliable methods for the determination of malaria-specific mortality in the community, we have studied the epidemiology of severe malaria through hospital admission rates for severe life-threatening disease.

A bewildering array of methods are available to display disease distributions and to analyse the spatial pattern. Here we have used two methods, both originally proposed for the study of the pattern of childhood leukaemia in the UK, to study the spatial distribution of severe malaria in children. We have also investigated the stability of the spatial pattern, both from year to year and over shorter time intervals. We draw conclusions concerning the use of hospital data for such work and discuss how related studies may give clues as to the aetiology of severe disease.

Data sources
The study included all cases of severe malaria diagnosed in Kilifi District Hospital (KDH) between 1 May 1991 and 30 April 1993 in children aged under 5 years living in a defined rural study population. The hospital is 60 km north of Mombasa on the Kenya coast, and is the only hospital serving Kilifi District, which includes a rural population of around 50,000 people who live in about 5000 widely scattered homesteads. The study area and population are described in detail elsewhere. Using preliminary hand-drawn maps to locate the households on foot, the latitude and longitude of each household were found using a hand-held satellite tracking device (SatNav: Trimble Navigational Europe Ltd, UK). Co-ordinates were checked by comparison with the hand-drawn maps, and households were re-visited to correct errors. Where discrepancies were still apparent after checking, households' co-ordinates were taken as missing. The town of Kilifi, immediately surrounding the hospital, and three other small urban areas with many migrant workers were omitted from the study population and the mapping exercise. De jure censuses were conducted in May 1992 and May 1993. Between May 1991 and April 1993 all paediatric admissions to KDH were examined on admission and during their stay in hospital by one of the Kenya Medical Research Institute (KEMRI) clinicians, who provided 24-hour cover for the 30-bed paediatric ward. Each child admitted underwent a full clinical, haematological and parasitological investigation. Details were recorded on a standard form, which included information to enable the child's household to be identified from the map database. Children with a primary diagnosis of malaria were defined as having severe malaria if they exhibited any one of the following: cerebral malaria (in coma; unable to localize pain); anaemia (haemoglobin < 5.1 g/dL with a peripheral parasitaemia > 10,000/μL); two or more generalized convulsions within the 24th period before admission; hyperparasitaemia (> 20% infected red cells); prostrated (unable to sit or stand unaided); or death with a confirmed diagnosis of malaria. Within each year (May 1991 to April 1992, and May 1992 to April 1993) second and subsequent admissions for severe malaria in the same child were omitted from the database (n = 6).

Data were compiled using dBase IV (Ashton-Tate). We also used Map-Info (Map-Info, Troy, NY, USA), SAS (SAS Inc, Cary, NC, USA) and EGRET (SERC, Seattle, USA) software for analysis, on a 486 colour notebook computer (Compaq Contura 4/25cx).

Analytical methods and results
Complete demographic information and household co-ordinates were available for 10,247 children under 5 years in May 1992 living in 3598 households, and 9868 children in 3618 households in May 1993. This represents over 98% of children in the study area. Between May 1991 and April 1993, 358 cases of severe malaria under 5 years old were admitted to KDH from the study area. Five children were admitted with severe malaria
in both years of the study. Analysis was restricted to 343 (96%) children for whom household data were complete. A map of the study area is given in Figure 1, which shows the location of all households with children under 5 years at the May 1992 census. Also shown on this map are households with a case of severe malaria presenting to KDH between May 1991 and April 1993. Overall incidence of severe malaria was 18.2/1000 (186/10 247) between May 1991 and April 1992, and 15.9/1000 (157/9868) in the following year. The distribution of the disease by age showed that incidence was highest in those under 3 years, with a peak in 1–2 year olds, as discussed elsewhere. Severe malaria showed strong seasonality, with peaks in June and July in both years, during or shortly after the main rains and a further small peak in January–February 1993, after the short rains. Year-to-year variations in the timing and incidence of severe disease are expected in this population. We therefore assessed the spatial pattern of disease within each year separately.

Equal-population-areas (EPA)

The study area consisted of scattered households, with no natural population units such as villages. In order to map disease incidence rates, households were grouped into contiguous areas. There was no obvious choice for the size of such areas: larger areas enable broad variations in disease incidence to be assessed whereas smaller areas facilitate examination of local variations in incidence. However, larger areas are subject to less random variation than smaller areas. The analysis was repeated, therefore, for a range of different-sized areas.

Neighbouring households were combined to form arbitrary geographical areas of roughly equal population size using a simple algorithm. For example, starting with the most southerly household, the nearest households to the first household were added one by one until the population of the first area reached 100 children under five (from the May 1992 census). The most southerly unallocated household was then allocated to the second area and the process repeated until all households were allocated to an area. The last area (area 101) consisted of fewer than 100 children living at the northermost part of the study area and was omitted from the spatial analysis.

This approach for investigating spatial pattern of disease is well suited for application with household-level data on population and disease incidence, but has been relatively little used to our knowledge. Spatial patterns were investigated using areas of different sizes. Thus, the algorithm was used to create (a) 10 areas of approximately 1000 children; (b) 20 areas of approximately 500 children; (c) 50 areas of approximately 200 children; and (d) 100 areas of approximately 100 children. For each set of areas (a) to (d), rates of admission to KDH with severe malaria were calculated and directly standardized for age (in one-year age groups) and sex, using the population of the whole area as reference. Standardized rates were calculated for each of the 2 years of the study and were plotted on maps shaded according to the quintiles of the distribution in each year. For example, Figure 2 shows the data for severe malaria in 1991–1992 for 10, 20, 50 and 100 EPA. If disease were distributed evenly throughout the study area, these would have 18.6, 9.3, 3.7 and 1.86 cases expected in each area, respectively. Statistical significance of departure of the crude rates from homogeneity in all areas was investigated using the approach of Pothoff and Whittinghill. This is a more powerful test of ‘contagion’ than the standard χ² test as it is based on the number of pairs of cases that occur in the same area relative to the total number of pairs of cases. Results are shown in Table 1 for both 1991–1992 and 1992–1993. In the first year of the study the spatial distribution of severe malaria in the area was not homogeneous (P < 0.001 when it was divided into 10, 20, 50 or 100 EPA). In the second year statistically significant evidence of heterogeneity was only apparent when the area was divided into 10 (P = 0.03). We repeated the significance testing stratified by age and sex but this made no material difference to the findings (data not shown).

Poisson regression

Visual inspection of the maps drawn for the EPA analysis (Figure 2) suggested, as expected, a relationship between access to hospital and severe malaria admission rates, with higher rates of admission to hospital in those living nearer to the hospital and in those living nearer to a road. The former pattern was apparent for all four different area sizes, but the latter was only apparent for the smaller areas. The hypothesis was investigated for both years separately using Poisson regression methods. Incidence rates of disease for each household were related to the distance of that household from the nearest road and from the hospital, as the crow flies. Both distance variables were categorized into approximate quintiles. There was a clear tendency for admission rates for severe malaria to be lower in those most distant from the hospital (Table 2). For example, those living more than 25 km from the hospital had admission rates of about one-fifth of those for children living within 5 km of the hospital. Also, those living more than 2 km from the nearest road had admission rates that were less than half of those for children living within 0.5 km of a road. Regression analysis of the rates in the second year gave very similar results (data not shown).

Rank correlations for EPA

In order to assess whether the spatial pattern of disease changed from year to year, we ranked areas according to the incidence of severe malaria in 1991–1992 and in 1992–1993 and compared the ranking in the 2 years using the Spearman rank correlation test (Table 1). Area-specific rates showed evidence of association between the 2 years, suggesting that the pattern of disease was to some extent stable over this time period.

Local short-term fluctuations in severe malaria

We have previously reported evidence of space-time clustering of severe malaria. Analysing data on severe malaria diagnosed in the period from May 1989 to April 1992, we found a significant excess of pairs of cases residing within 1 km of each other and presenting to the hospital within 3 days of each other. The analytical method used, originally proposed by Knox, does not lead directly to a visual representation of the ‘clusters’ and nor does it make direct use of population data. Here we have adopted a different approach to investigate further the phenomenon of micro-epidemics of severe malaria within broad seasonal variations.

Besag and Newell have described a method for analysing spatial data which is based on small areas not of equal population but of equal case numbers. Their original application was to study the spatial distribution of childhood leukaemia in the UK and they used population data from the census at
Figure 2  Equal-population-area (EPA) maps of the study area, shaded by quintiles of the rate of severe malaria presenting to KDH between May 1991 and April 1992. Ten EPA of approximately 1000 children; 20 of approximately 500 children; 50 of approximately 200 children; and 100 of approximately 100 children are shown. Solid lines show all-weather roads used as bus or matatu (shared taxi) routes.
small-area level. Although their method was developed to assess spatial anomalies in disease we have applied it here to assess anomalies in both space and time, or apparent space-time ‘clusters’.

The method is based on choosing a small positive integer \( k \) (e.g. \( k = 3 \)) and a time interval \( t \) (e.g. \( t < 7 \) days) and for each case of disease in turn drawing a circle centred on that case which passes through the household of the \( k \)th nearest case presenting to hospital within the specified time interval \( t \). In each of the circles the total number of children under 5 years is counted and special note is made if this number is below some specified threshold. In our analyses the under-5 population in each circle was estimated by summing the census populations of the individual households. Circles with the smallest populations represent the most anomalous areas for a particular choice of \( k \) and \( t \).

We compared the distribution of population size in each circle with the distribution that would be expected if cases of disease were randomly distributed among children at risk, based on the Poisson distribution. We used Monte Carlo simulation to decide whether there was evidence of overall clustering which should be investigated further using maps. We considered the dates of onset (admission) as fixed, and randomly permuted these dates of onset among the locations of children at risk, computing circles to include the \( k \)th neighbouring case within \( t \) days around each case. The number of significant circles observed \( (P < 0.05) \) was compared with the numbers found in each of 999 simulations. If the rank of the observed number of circles among the 1000 realizations (999 simulations plus one observed value) was less than 50 (i.e. \( < 0.05 \times 1000 \)) then the observed set was considered to show ‘clustering’ at the 5% level. The results of the Monte Carlo simulations are summarized in Table 3. It can be seen that there was evidence of space-time clustering of severe disease in both years, but that this was more marked in the first year than the second. Figure 1 illustrates the clustering for \( k = 2 \) and \( t = 7 \) days in August, September and October of 1991: the circles marked have a Poisson probability of two or more events less than 0.05. These 11 circles together make up three separate aggregations or ‘clusters’. Although Table 3 suggests that overall there is evidence of clustering, it should be kept in mind that some of the anomalies shown in Figure 1 may have occurred by chance.

### Discussion

We have presented here the first application of several statistical techniques developed for the study of chronic disease to examine the spatial distribution of a major parasitic disease in Africa. This was made possible using a geographical information system (GIS), whereby data relating to disease incidence and household population were assigned spatial co-ordinates and linked together.

### Table 1 Significance testing for spatial pattern in severe malaria admission rates and stability of rates over time

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( z = 8.7 )</td>
<td>( z = 2.2 )</td>
<td>0.52</td>
</tr>
<tr>
<td>10</td>
<td>( P &lt; 0.001 )</td>
<td>( P = 0.03 )</td>
<td>( P = 0.12 )</td>
</tr>
<tr>
<td></td>
<td>( z = 6.2 )</td>
<td>( z = 1.1 )</td>
<td>0.64</td>
</tr>
<tr>
<td>20</td>
<td>( P &lt; 0.001 )</td>
<td>( P = 0.29 )</td>
<td>( P &lt; 0.001 )</td>
</tr>
<tr>
<td></td>
<td>( z = 8.7 )</td>
<td>( z = 0.9 )</td>
<td>0.24</td>
</tr>
<tr>
<td>50</td>
<td>( P &lt; 0.001 )</td>
<td>( P = 0.39 )</td>
<td>( P = 0.10 )</td>
</tr>
<tr>
<td></td>
<td>( z = 4.5 )</td>
<td>( z = 0.29 )</td>
<td>0.36</td>
</tr>
</tbody>
</table>
| 100             | \( P < 0.001 \) | \( P = 0.77 \) | \( P < 0.0001 \)

*Pothoff and Whittinghill's test.*

---

### Table 2 Rates of admission to hospital with severe malaria from May 1991 to April 1992 by distance to the hospital and to the nearest road

<table>
<thead>
<tr>
<th>Distance to hospital</th>
<th>Cases/CYAR</th>
<th>Rate</th>
<th>Rate ratio (95% CI)</th>
<th>Adjusted*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Under 5 km</strong></td>
<td>37/1170</td>
<td>31.6</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>5-10 km</strong></td>
<td>53/2342</td>
<td>22.6</td>
<td>0.72 (0.5-1.1)</td>
<td>0.78 (0.5-1.2)</td>
</tr>
<tr>
<td><strong>10-15 km</strong></td>
<td>54/2573</td>
<td>21.0</td>
<td>0.66 (0.4-1.0)</td>
<td>0.70 (0.5-1.1)</td>
</tr>
<tr>
<td><strong>15-25 km</strong></td>
<td>32/2158</td>
<td>14.8</td>
<td>0.47 (0.3-0.8)</td>
<td>0.57 (0.4-0.9)</td>
</tr>
<tr>
<td><strong>Over 25 km</strong></td>
<td>10/2004</td>
<td>5.0</td>
<td>0.16 (0.1-0.3)</td>
<td>0.19 (0.1-0.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Under 0.5 km</strong></td>
<td>61/2339</td>
<td>26.1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>0.5-1 km</strong></td>
<td>39/2103</td>
<td>18.5</td>
<td>0.71 (0.5-1.1)</td>
<td>0.70 (0.5-1.0)</td>
</tr>
<tr>
<td><strong>1-1.5 km</strong></td>
<td>29/1630</td>
<td>17.8</td>
<td>0.68 (0.4-1.1)</td>
<td>0.68 (0.4-1.1)</td>
</tr>
<tr>
<td><strong>1.5-2.5 km</strong></td>
<td>43/2506</td>
<td>17.2</td>
<td>0.66 (0.5-1.0)</td>
<td>0.76 (0.5-1.3)</td>
</tr>
<tr>
<td><strong>Over 2.5 km</strong></td>
<td>14/1669</td>
<td>8.4</td>
<td>0.32 (0.2-0.6)</td>
<td>0.47 (0.3-0.9)</td>
</tr>
</tbody>
</table>

*Case per 1000 child-years at risk.*

b *Child-years at risk (CYAR), estimated from May 1992 census.*

\( t \) *Confidence interval.*

\( d \) *Adjusted for age (in one-year groups), sex, and hospital/road distance.*

\( k \) *Likelihood ratio test for significance of association between disease rate and distance to hospital, adjusted for age, sex and distance to road: LRT \( \chi^2 \) for trend = 23.9 on 1 d.f. \( P < 0.001 \).*

\( t \) *Likelihood ratio test for significance of association between disease rate and distance to road, adjusted for age, sex and distance to hospital: LRT \( \chi^2 \) for trend = 5.7 on 1 d.f. \( P = 0.02 \).*
We found that the rate of admission to KDH with severe malaria was strongly associated with the distance a child lived from the hospital and to the distance from a road. For reference we also assessed the spatial pattern of admission rates for all causes other than severe malaria and found both from visual inspection of maps and from Poisson regression models a similar pattern of lower admission rates for children with poorer access (data not shown). These findings were not unexpected and a likely explanation is that only a fraction of cases present to hospital and the probability of presentation is strongly related to ease of access to hospital.

A more surprising finding was that of apparent space-time clustering of cases of severe malaria. Clinicians working in malaria-endemic areas often provide anecdotal evidence of simultaneous presentation of cases of severe malaria from neighbouring households. In formal analysis, defining what is meant by a cluster has always been problematic, and the two methods we used are based on quite different approaches. Knox’s method identifies groups of cases that are anomalous after allowing for the usual rate of disease in that place at that time—a certain type of ‘cluster’. The Besag and Newell method detects rates that are anomalous on a more absolute scale: if seasonal and geographical rate fluctuations are thought of as part of the clustering process, this may be a more natural definition. Our study area is the first in which space-time clustering of severe malaria has been substantiated using both methods.

Is it possible that our apparent space-time clusters are due not to local short-term increases in transmission or pathogenicity of malaria but due simply to the deficiencies of our population database with respect to local short-term population movement? Exactly how this would arise is unclear, but if it were the case we might expect to see similar anomalies for child mortality rates in the same population. We therefore repeated the analyses for all-cause child mortality (data not shown), and found no space-time anomalies in either year. This supports the view that there is a tendency for severe malaria to occur in localized micro-epidemics.

Our findings must be viewed with reference to the limitations of the available data, both in terms of the number of events (numerator) and the population in each household (denominator). Anomalies in either of these may lead to apparent heterogeneity in disease: such anomalies are particularly likely to arise with spatial analysis of disease in developing countries because of the lack of reliable event and denominator data. Undoubtedly, many children die from malaria without reaching health facilities or hospitals. This may be due in part to rapid onset of symptoms, poor accessibility of health facilities, and use of traditional healers rather than Western medical care. Researchers who have tried to estimate the proportion of deaths in the community due to malaria by interviewing bereaved relatives (verbal autopsies) have had only limited success. Our chosen alternative was to study the epidemiology of severe malaria in hospitals themselves, accepting the underestimation of disease incidence that results.

Deficiencies in the denominator are also likely. Although we had annual census data, the population is highly mobile and a child may have acquired infection outside their own household. Mothers and children make frequent visits away from home for several days at a time, often to visit relatives, for agricultural work, because of illness, or to attend funerals, which are major social events in Giriama society (CS Molyneux and RW Snow, unpublished data).

We have presented the results of the first in-depth investigation of the spatial pattern of severe, life-threatening malaria in sub-Saharan Africa. We found evidence relating hospital admission rates strongly to access to hospital. Furthermore we have shown that severe malaria occurs in space-time clusters and physically identified such clusters on a map. The methods used allow for the overall spatial and temporal variation. The spatial pattern of severe malaria suggests that using hospital-based incidence is likely to be an underestimate of the true incidence in the community, but our finding of space-time clustering suggests that despite such limitations we may be able to identify factors which are associated with severe disease clusters. For example, research might now focus on discordant ‘virulence’ factors in parasite isolates from clustered severe cases and cases that did not occur in a cluster but had equally good access to hospital.

Acknowledgements

We thank Dayo Forster, and the staff of the KEMRI unit at Kilifi, especially Rogers Chisengwa and Lewis Mitsanze. This investigation received financial support from KEMRI and the Wellcome Trust (reference number 037547/Z/93/Z/1.5). RWS is a Senior Wellcome Fellow in Basic Biomedical Sciences and KM is a Senior Wellcome Fellow in Clinical Science. JAS is supported by the UK Medical Research Council. This paper is published with the permission of the Director of KEMRI.

References


