Am. J. Trop. Med. Hyg., 67(5), 2002, pp. 480–489

A GEOSTATISTICAL ANALYSIS OF THE GEOGRAPHIC DISTRIBUTION OF LYMPHATIC FILARIASIS PREVALENCE IN SOUTHERN INDIA

A. SRIVIDYA, E. MICHAEL, M. PALANIYANDI, S. P. PANI AND P. K. DAS

Vector Control Research Centre (ICMR), Pondicherry, India; Department of Infectious Disease Epidemiology, Imperial College Medical School, London, United Kingdom

Abstract. Gaining a better understanding of the spatial population structure of infectious agents is increasingly recognized as being key to their more effective mapping and to improving knowledge of their overall population dynamics and control. Here, we investigate the spatial structure of bancroftian filariasis distribution using geostatistical methods in an endemic region in Southern India. Analysis of a parasite antigenemia prevalence dataset assembled by sampling 79 villages selected using a World Health Organization (WHO) proposed 25 × 25 km grid sampling procedure in a 225 × 225 km area within this region was compared with that of a corresponding microfilaraemia prevalence dataset assembled by sampling 119 randomly selected villages from a smaller subregion located within the main study area. A major finding from the analysis was that once large-scale spatial trends were removed, the antigenemia data did not show evidence for the existence of any small-scale dependency at the study sampling interval of 25 km. By contrast, analysis of the randomly sampled microfilaraemia data indicated strong spatial contagion in prevalence up to a distance of approximately 6.6 kms, suggesting the likely existence of small spatial patches or foci of transmission in the study area occurring below the sampling scale used for sampling the antigenemia data. While this could indicate differences in parasite spatial population dynamics based on antigenemia versus microfilaraemia data, the result may also suggest that the WHO recommended 25 × 25 km sampling grid for rapid filariasis mapping could have been too coarse a scale to capture and describe the likely local variation in filariasis infection in this endemic location and highlights the need for caution when applying uniform sampling schemes in diverse endemic regions for investigating the spatial pattern of this parasitic infection. The present results, on the other hand, imply that both small-scale spatial processes and large-scale factors may characterize the observed distribution of filariasis in the study region. Our preliminary analysis of a mountain range associated large-scale trend in the antigenemia data suggested that a nonlinear relationship of infection prevalence with elevation might be a factor behind such observed global spatial patterns. We conclude that geostatistic methods can provide a powerful framework for carrying out the empirical investigation and analysis of parasite spatial population structure. This study shows that their successful application, however, will crucially depend on our gaining a more thorough understanding of the appropriate geographic scales at which spatial studies should be carried out.

INTRODUCTION

Gaining a better understanding of the spatial distribution of parasitic diseases within endemic regions is increasingly recognized as being critical to the rational design and monitoring of parasitic control programs.[1–9] Not only can such knowledge lead to more precise mapping of the distribution of parasites within endemic regions, thus enabling more precise identification/estimation of the communities at risk and prioritization of target areas for control,[3,6,9,10] but improved understanding of the spatial structure of parasitic infection or disease processes (the strength and extend of focality or locality of transmission) will also be critical to making more realistic assessments of the persistence, threshold of transmission, dynamics of spread, and hence potential recovery of infection following control interventions.[11–15] This is likely to be particularly true for vector-borne diseases, which are expected to exhibit high degrees of heterogeneity in transmission patterns even at microspatial scales.[16] At a practical level, improved understanding of spatial variation will also allow an assessment of the spatial scales at which various risk factors underlying the observed disease distribution within an area operate as well as potentially facilitating the detection of the optimal spatial scale for applying a sampling framework for conducting parasitic surveys or a particular community-based control option.[14,17,18]

There have been several recent attempts to map the spatial distribution of lymphatic filariasis at various scales, from the global to national levels.[6,9,10,19,20] Based largely on aggregated spatial data (assembled from the published literature or national parasitologic survey archives) on microfilaraemia prevalence, these efforts, with the exception of Lindsay and Thomas[10] have largely tended to focus on describing the geographic distribution of filariasis infection across various areal boundaries (regional, national, state, and district) to highlight both the areal extent as well as the observed variation in the risk of infection within each studied region. While these attempts have contributed importantly to identifying and delimiting areas at risk, including target areas for prioritized treatment at global, regional, and national levels, they provide much less information regarding the spatial structure of either filariasis infection or disease. Similarly, while the work by Lindsay and Thomas,[10] which attempted to model filariasis distribution in Africa using ecologic correlates between environmental variables and the observed distribution of point prevalence data, identified several surrogate climactic factors likely to underlie the distribution of the disease on that continent, such global attempts address only one component of sample spatial variation, namely, large-scale trend or change in mean gradient in the data.[21] They provide little information on local variation or small-scale spatial dependency structure, which may often be obscured by the large-scale interpolations generated by such (mean function) methods.[14]

In this study, we use geostatistic techniques to investigate the spatial structure of observed bancroftian filariasis prevalence distribution among communities in a South Indian endemic region. Geostatistic methods (variogram analysis and kriging) facilitate the dissection of the processes underlying a given spatial dataset into both large- and small-scale components of variation and thus provide a powerful set of tools for characterizing the spatial structure of a mapped variable.[21–24]

This investigation forms part of a multicountry study initiated by the World Health Organization (WHO) Rapid Geographical Assessment of Bancroftian Filariasis (RAGFIL) program,[25] which advocated using a sampling grid of either 25 × 25 km or 50 × 50 km for selecting communities to undertake rapid filariasis prevalence surveys (by the application of rapid assessment procedures (RAP) surveys of filariasis prevalence, such as physical examination by the health workers (PEHW)[26] or estimation of parasite antigenemia prevalence by the use of the immunochromatographic card test (ICT)[27] to map and investigate filariasis distribution. While the primary analysis is based on filariasis antigeneamia prevalence data obtained from 79 communities selected from an endemic region in South India using the 25 × 25 km grid sampling framework, comparison is also made with the spatial structure of microfilaraemia prevalence data obtained from a random sample of 119 villages within the same region (used for a separate study validating different RAP methods[26]), to illustrate the importance of spatial sampling scale for the detection and quantification of spatial structure. These results are discussed in light of the WHO advocacy of a common grid dimension for grid sampling of communities for the rapid mapping and spatial analysis of filariasis distribution in diverse endemic regions.

MATERIALS AND METHODS

Study site. An area covering 41,950 square kms (pertaining approximately to a square of 225 × 225 km) lying between 78°9′25″ and 80°13′48″east, and 11°0′54″ and 13°2′31″ north in Southern India was selected for the main (analysis of antigenemia distribution) study (Figure 1). This area covers a total population of 12.5 million people distributed in 7,680 villages. The area falls partly under the tropical wet and dry zone and partly under the semiarid climatologic zone of India, bounded in the east by the Bay of Bengal. The area covers 13 districts in four South Indian states (ten districts from Tamil Nadu and one each from Andhra Pradesh, Karnataka, and Pondicherry). The districts covered in this study area are Villupuram, Chengleput, Tiruvannamalai, Vellore, Dharmapuri, Salem, Perambalur, Cuddalore, Thiruchirapalli, Thiruvarur, and Erode in Tamil Nadu, Chithoor in Andhra Pradesh, Kolar in Karnataka, and Pondicherry Urban Town. Out of 13 districts, only six districts were completely covered in the study area (Figure 1). The second dataset pertaining to microfilaraemia distribution was derived as shown in Fig.1 (insert) by random sampling of villages within a smaller subsite lying between 78°73′ and 79° 8′ East and 11° 25′ and 12° 56′ north within this region.

Infection survey methodology. Selection of villages within the study site for generating the filariasis antigenemia prevalence dataset was carried out according to the grid sampling procedure advocated by the RAGFIL program.[25] Briefly, this entailed placing grids of 25 × 25 km dimension within the chosen 225 × 225 km study area (adjusted on the right for the Bay of Bengal [Figure 1]), followed by choosing of villages nearest to each grid node (intersection point of each X [longitude] and Y [latitude] lines of the grids) for prevalence

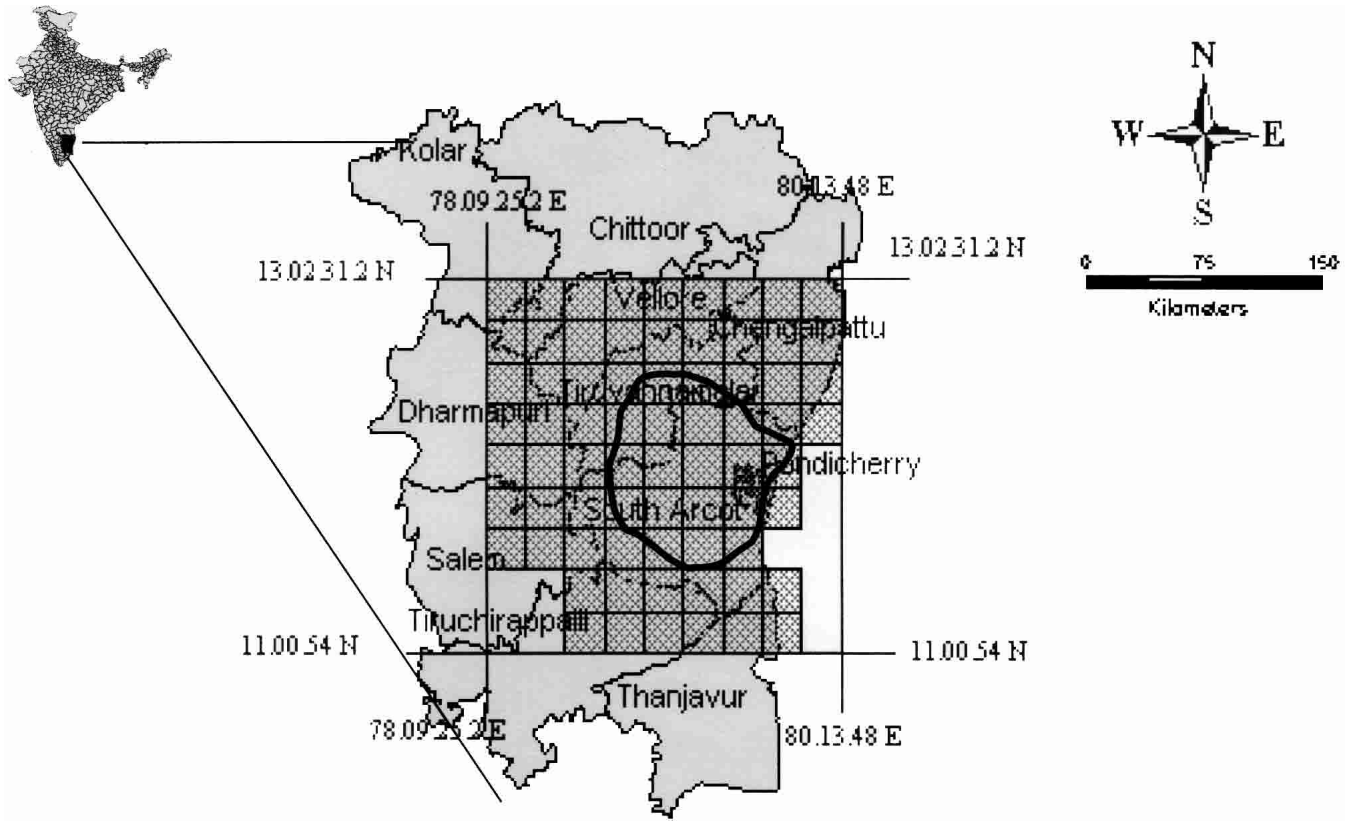

FIGURE 1. Map of the South Indian filariasis endemic region showing the site of the Rapid Geographical Assessment of Bancroftian Filariasis study (the 25 x 25 km gridded area (see text)) and the enclosed area where the microfilaraemia prevalence data used in this study were sampled.

survey by the ICT method. This resulted in the selection of 79 inhabited intersection villages for the study, with the name, geographic coordinates (collected using a geographical positioning system [GPs]), population size and area of each entered into a spreadsheet program. Within each village, 50 adult males above 15 years of age, and 40 households were selected for estimating antigenemia prevalence within the community using the ICT card test.[26,27] The number of households (40) to be sampled per village was based on the finding that on average in rural India 2 to 3 adult males would be available per household; the selection of households themselves was performed by first listing the total number of households in the village followed by estimating the sampling interval (total number of households/40) for the systematic selection of household units for the study. The first household is selected randomly (by choosing a number randomly between one and the sampling interval), and the others are chosen by the cumulative addition of the sampling interval to each previous number as described elsewhere.[28]

As far as the data for microfilaria (mf) prevalences are concerned, these villages were randomly selected from three districts under different field studies conducted by the Vector Control Research Centre(VCRC).[26,27,29] In each village, 7–10% of the population from randomly selected households were included in the microfilaraemia surveys. Microfilaria examination was undertaken using 20 ccm blood smears collected from each individual selected for these surveys.

Spatial analysis. We used basic exploratory data analysis (EDA) and visualization methods specific to the analysis of spatial data, such as data posting and contour/perspective plots (effective for initial assessments of spatial trends in the data), geostatistical methods, and large-scale trend modeling methods (fitting of trend surfaces and generalized additive modeling), to investigate the spatial structure of filariasis infection.

EDA methods. For geostatistical data, posting data methods not only allow the display of sampling locations but by varying the size of the plotting symbol proportional to the plotted attribute value also the visualization of spatial clusters and trends in the data.[24,30] Similarly, we also used two- and three-dimensional trend visualization tools, such as contour and perspective plots (derived using local spatial interpolation methods[31]), as additional tools for displaying and exploring spatial trends. Raw, untransformed data (antigenemia and microfilaraemia prevalence), were used in this initial stage of analysis.

Geostatistical analysis. Small-scale spatial dependency or structure in the present data was investigated by semivariogram analysis.[21–24,32] Because histograms of the data for both antigenemia and microfilaraemia prevalence indicated negatively skewed distributions, a logarithmic transformation, $ln(N + 1)$, was applied to the data before the analysis was performed. The semivariogram provides a measure of spatial correlation by describing how sample data are related with distance and direction. It plots $\gamma(h)$, one half the average squared difference between paired data values, versus h, the distance or lag separating the pairs.[21–24] The shape of this plot summarizes the type of spatial structure or dependence and the range (distance) over which this dependence occurs. If there is spatial dependence among the data, $\gamma(h)$ typically increases with separation distance h, may level off and even decrease after a certain distance. The value of $\gamma(h)$ when the semivariogram levels off is called the sill (c), and the distance h at which the sill is attained is called the range (a); the range represents the maximum distance at which there is spatial continuity. The semivariogram at zero lag $\gamma(0)$ must be zero, but in reality the extrapolated semivariogram usually intercepts the ordinate at a positive value known as the nugget variance (b). The nugget effect can rise from several factors including sampling error and spatially dependent variation occurring over distances much less than the sampling interval. In some cases, the sample semivariogram may be completely level. This occurs in spatially random or uniform data and implies no spatial contagion in the data at the scale being sampled. In other cases, it may be that the increase of $\gamma(h)$ seems to have no limit: the infection properties have no finite variance. Such semivariograms are said to be unbounded. If all of the semivariograms are identical regardless of the direction of h, the spatial structure of the sample data is said to be isotopic. The spatial correlation then depends only on the distance of separation, h, not on the direction. When the semivariogram is a function of the direction of h, the spatial structure is said to be anisotropic. Maximum h was set as half the maximum interpoint distance, and equal- sized distance classes were used in all cases. Finally, because semivariogram values $\gamma(h)$ are calculated only for discrete-distance classes, it is necessary to fit a model to $\gamma(h)$ as a function of h to obtain values of $\gamma(h)$ for all values of h. Because the estimated semivariogram of microfilaraemia was bounded we tested here the fits of the exponential, spherical and gaussian models.[23]

Kriging is a linear interpolation procedure that allows the prediction of unsampled values of a variate that shows spatial correlation.[21–24] It calculates predictions of unsampled values based on the model of the covariance of the observations (estimated for example by the fits of the various semivariogram models) at known locations. However, because the primary antigenemia prevalence data did not show spatial correlation, we do not perform this operation in this study.

Spatial trend analysis. Large-scale spatial trends in the data were estimated in this study by either conducting trend surface analysis using weighted least squares[31] or by fitting generalized additive models (GAMs) using sample location coordinates and any other spatial covariate as predictors.[24] Statistical evidence for spatial autocorrelation in the residuals from these analyses or indeed the raw data were evaluated using the Moran's I statistic.[33]

All the analyses above were carried out using the functions provided in the software package Splus 2000 (Mathsoft, Inc, United Kingdom) and its companion spatial analysis module S+SpatialStats.[24]

RESULTS

Table 1 summarizes the two filariasis prevalence datasets used in this study and highlights the wide range in prevalence observed among the sampled communities for each of the two measures of infection. The corresponding spatial plot of the raw antigenemia prevalence values is shown in Figure 2 and displays the spatial variation in the data. In particular, the plot indicates that a major feature of the observed geographic distribution for this variable is the occurrence of an apparent large-scale trend for high-antigenemia prevalence ranging from the south-southwest corner towards the north-northeast regions of the study area.

TABLE 1
Characteristics of the sample datasets used in the study

Infection parameter	Sampling method	No. villages sampled	Mean prevalence (%)	Range of prevalence (%)
Antigenemia*	25 × 25 grid sampling	79	8.68	0–38.0
Microfilaraemia†	Random selection	119	3.10	0–20.2

* Measured using the immunochromatographic card test card test[27]
† Based on parasite detection using 20 ccm blood samples

Given the trend shown in Figure 2, we begin our analysis of filariasis population spatial structure by calculating separate experimental semivariograms for the logarithmically transformed antigenemia data in the a) north-south (N-S), b) northeast-southwest (NE-SW), c) east-west (E-W), and d) northwest-southeast (NW-SE) directions using a tolerance of 11.25° in each direction (Figure 3). As expected given the trend, the semivariograms showed differences in spatial dependence with direction, i.e., the existence of apparent anisotropy in the spatial structure of the sampled data. In particular, while the N-S and E-W directional semivariograms were similar with an increase in $\gamma(h)$ values to a similar sill around a range of approximately 66 kms, spatial dependency in the data appeared different in the NE-SW and NW-SE directions. The NE-SW direction yielded a generally increasing semivariogram while the NW-SE direction showed a pattern of an increase followed by a slight decrease in $\gamma(h)$ values with increasing distance (Figure 3b,d). These different spatial dependencies may reflect true differences in the spatial structure of the observed antigenemia prevalence along these directions (ie., true anisotropy) or else they may reflect an artifact of spatial trends in the data (Figure 2).

Such pseudoanisotropy may be revealed by reestimating the semivariograms following the removal of large-scale spatial trend(s) from the data.[24] Note that estimating the semivariogram of a variable strictly also requires that we remove large scale spatial trends to meet the requirement of *intrinsic stationary* in the measured data over the entire sample space.[21–24] We achieved trend removal in this study by fitting a generalized additive model (GAM) to the log ICT prevalence data using locally weighted regression smooth functions of latitude and longitude as predictors[31,34] and subjecting the residuals from this model to reanalysis by variography. The fitted model (results not shown) indicated a significant nonlinear trend in the data for longitude ($F = 3.21$, df = 1,2.5, $P = .03$) but not for latitude.

The corresponding directional semivariograms using the GAM residuals are plotted in Figure 4a, and show not only similar patterns (except for a slightly lower $\gamma(h)$ values for the NE-SW direction) following the removal of large-scale trend, but also an apparent lack of small-scale spatial variation in the data along any of the four directions. This pure nugget effect is more clearly seen in the overall omnidirectional experimental semivariogram estimated for the residuals (Figure 4b), indicating that the apparent small-scale structures and anisotropy depicted in Figure 3 for the undetrended log ICT prevalence data are largely a function of the large-scale spatial trend observed for the data (Figure 2).

Available information on landscape features in the study region that may explain this coarse-scale trend was limited in this investigation to elevation and distance from the sea. A

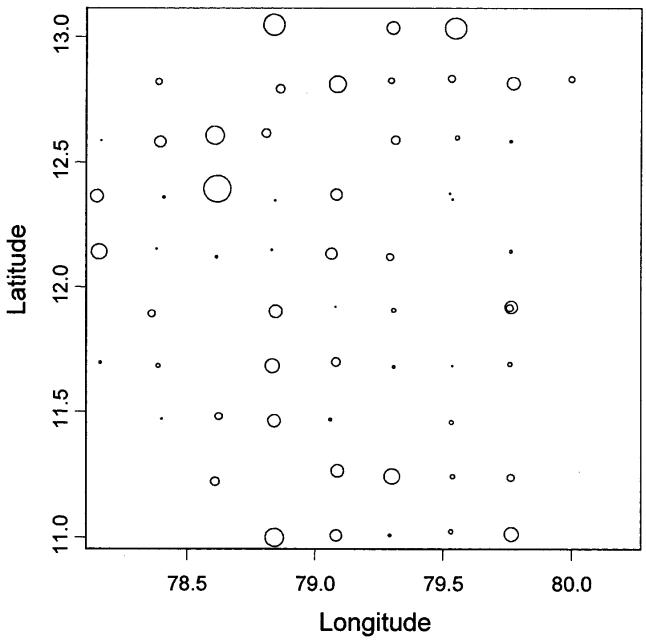

FIGURE 2. Spatial plot showing the location and magnitude of the raw immunochromatographic card test (ICT) prevalence data. Symbol sizes are proportional to the ICT prevalence measured in each study location.

GAM analysis of the distribution of log ICT prevalence as a function of these predictors suggests a likely non-linear role for elevation ($F = 3.28$, df = 1,2.4, $P = .013$) but not for distance from the sea. Figure 5 shows that the apparent association of the ICT high prevalence trend with the mountain range in the study area may partly be a function of this nonlinear relationship between prevalence and altitude. The results show that in general high prevalences of infection coincided (expect for the southeast region) with the intermediate levels of altitude (Figure 5c), which presumably occurred along the ridges of this mountain range (Figure 5a,b).

The lack of evidence for small scale structure in the ICT prevalence data does not imply that spatial dependence in the distribution of filariasis infection in this region may not occur at spatial scales smaller than that measured in this study (ie., < 25 km). Here, we address this question of the impact of sampling scale in investigations of the population spatial structure of filariasis by undertaking a spatial analysis of a separate dataset on the geographic distribution of community-level microfilaraemia prevalence available with us from the present study area (Figure 1). This dataset was constructed from a random sample of 119 villages within a smaller sub-region of the study site.[26] Figure 6a plots the sample locations of this survey and indicates the essentially random nature (although some gaps exist in the center of the sampling space) of the survey carried out. The corresponding perspective plot (Figure 6b) of the data reveals the occurrence of marked spatial heterogeneity in the distribution of this infection variable within this area, with a major cluster of high prevalence in the center, smaller local areas or hotspots of infection around the central region, and a variable trend in high microfilaraemia prevalence running along the south-north direction. As in the case of the antigenemia data, we begun the spatial analysis by first removing the apparent

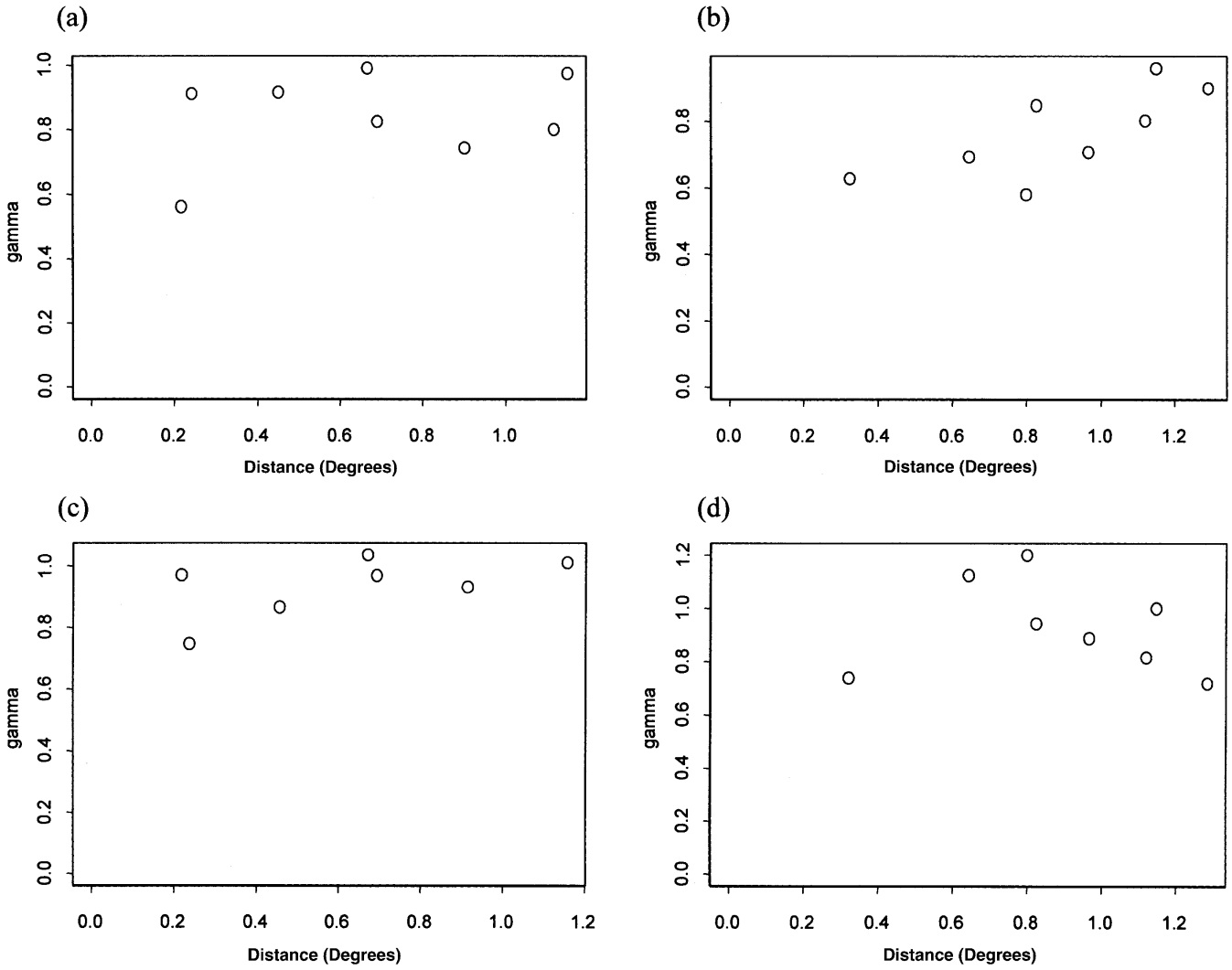

FIGURE 3. Directional experimental semivariograms for the log immunochromatographic card test prevalence ($Ln\ (N + 1)$) data in the (a) north-south, (b) northeast-southwest, (c) east-west, and (d) northwest-southeast directions before trend removal. Note that separation distance is in degrees (1° approximately equal to 112 kms).

large-scale trend observed in Figure 6b by fitting a GAM to a logarithmic transform of these data using the spatial coordinates of each point as predictors. As shown in Figure 7a, this yielded a significant effect for latitude as the major factor underlying the large-scale trend hinted at in Figure 6b. The corresponding experimental omnidirectional semivariogram for the detrended microfilaraemia data is shown in Figure 7b (directional semivariograms did not differ appreciably from each other and so not shown) and provides clear evidence for spatial dependence or autocorrelation in the data. An exponential model of the form:

$$\gamma(h) = c_0 + c\left\{1 - \exp\left(-\frac{h}{r}\right)\right. \quad (1)$$

where $\boldsymbol{c}_0$ denotes the nugget, $\boldsymbol{c}$ the sill and $\boldsymbol{r}$ the distance parameter which defines the spatial extent of the model, and portrayed by the line in Figure 7b, gave the best least squares fit to these data in comparison with the spherical and gaussian models (as assessed by the residual sum of squares between the theoretical model and the empirical semivariogram[24]). The parameter values of the fitted model suggest the existence of strong spatial continuity up to a range of ≈ 6.6km and a nugget effect (0.025) that was approximately one fifth of the sill (0.115) (i.e., approximately 80% of the average sample population variation appeared to be spatially dependent[35]).

DISCUSSION

The major finding of this study is that once large-scale spatial trends were removed, there appeared to be no indication in the present South Indian filariasis endemic site for small-scale spatial dependency in the distribution of antigenemia prevalence at the investigated sampling scale of 25 kms. Our analysis of microfilaraemia prevalence distribution from a subregion within this endemic zone suggests that this is likely to be a result of infection aggregation or local spatial variation (rather than measurement error or purely random variation) occurring in the case of infection at scales smaller than the minimum sampling interval used for the ICT survey. These results suggest that the WHO proposed sampling grid of 25 x

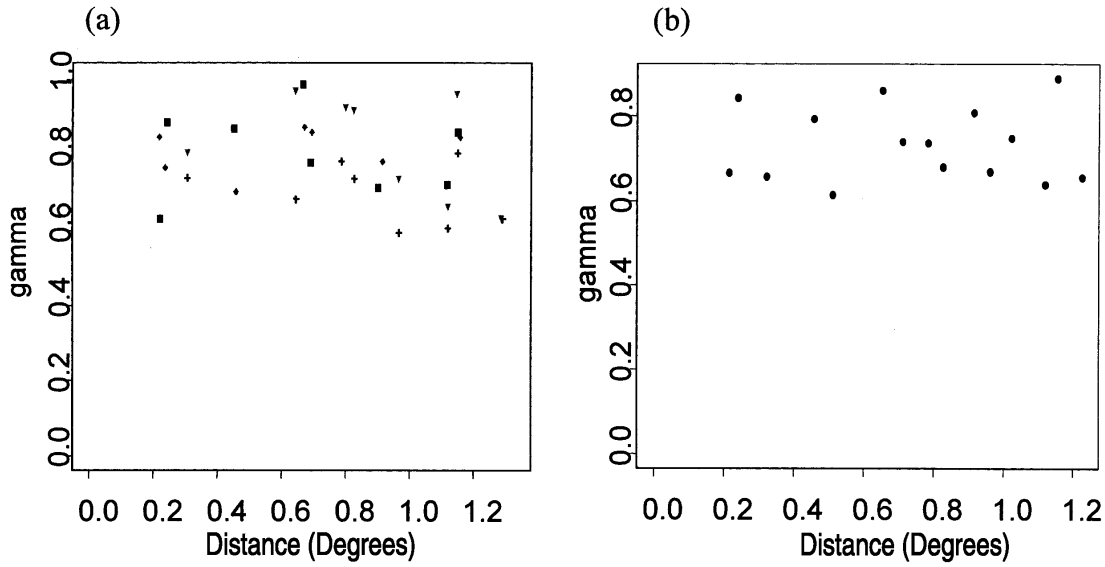

FIGURE 4. (a) Directional experimental semivariograms of the residuals of log immunochromatographic card test (ICT) prevalence in the north-south (closed squares), northeast-southwest (pluses), east-west (closed diamonds), and northwest-southeast (inverted triangles) directions after trend removal in the Rapid Geographical Assessment of Bancroftian Filariasis study area. (b) Omnidirectional experimental semivariogram for the detrended log ICT prevalence data showing constant $\gamma(h)$ values or a pure nugget effect. The lack of spatial autocorrelation in these residuals from the generalized additive models fit was further shown by a nonsignificant Moran's I statistic (Moran's I = –0.022, $P = .46$) obtained for these data.

25 kms for the rapid mapping of filariasis antigenemia prevalence (as proposed in the RAGFIL program[25]) could have been at too coarse a scale for describing and modeling the likely local spatial variability in filariasis infection in this study region, a conclusion which clearly has major implications for the execution of the multicenter RAGFIL program as follows.

First, it suggests that caution must be exercised when attempting to apply uniform sampling schemes in different endemic regions for investigating spatial patterns of filariasis infection given that the appropriate scale critical to the detection and quantification of such variation is likely to vary depending on variables as diverse as the size and shape of the study area (the sampling frame or spatial extent of

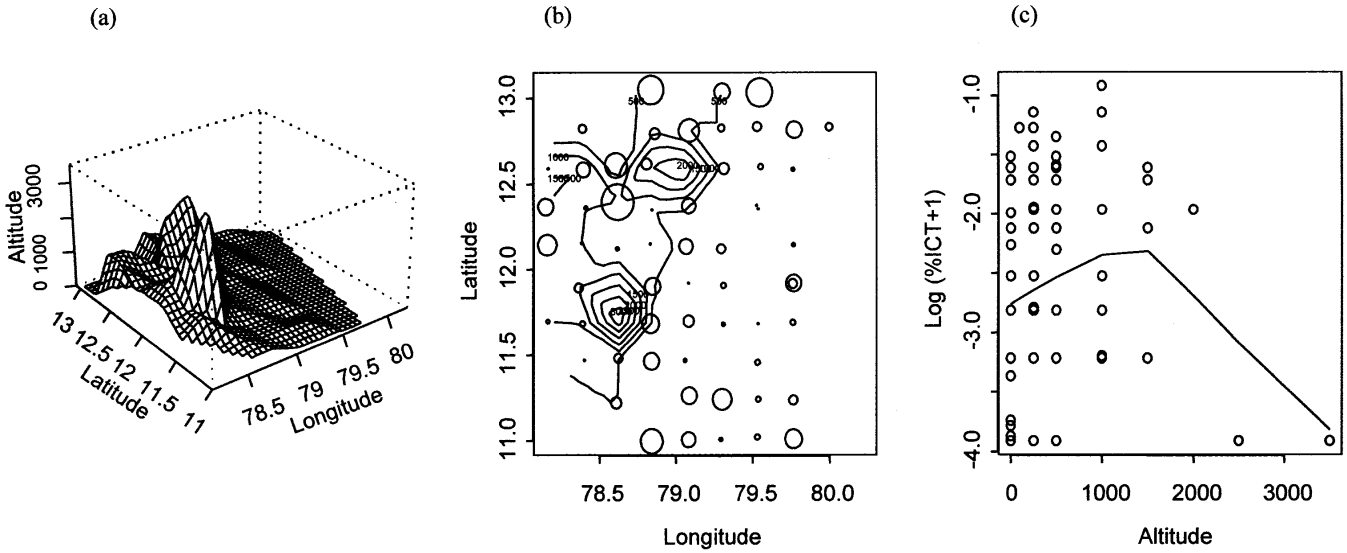

FIGURE 5. The associations between immunochromatographic card test (ICT) prevalence, elevation and the mountain range in the study area. (a) Perspective plot of the local trend surface for elevation in the Rapid Geographical Assessment of Bancroftian Filariasis area, showing the mountain range occurring in this area in the southwest to north direction. The trend surface was created using the *interp* function in Splus, which interpolates local surfaces based on fitting a fifth order trend surface within each triangle of a Delaunay triangulation of the data points.[31] (b) Spatial plot of log ICT prevalence overlaid on a contour plot of elevation indicating raised infection (expect for the southeast region) along the slopes of the mountain range. The *interp* function as described above was used as the interpolation procedure for creating the shown contours. (c) The nonlinear relationship between Log ICT prevalence and altitude. Plot shows both observed elevation values at each sampled location and the fit (solid line) of the additive model described in the text to these data. The result shows that prevalence increased with elevation up to some 1500 m before declining thereafter.

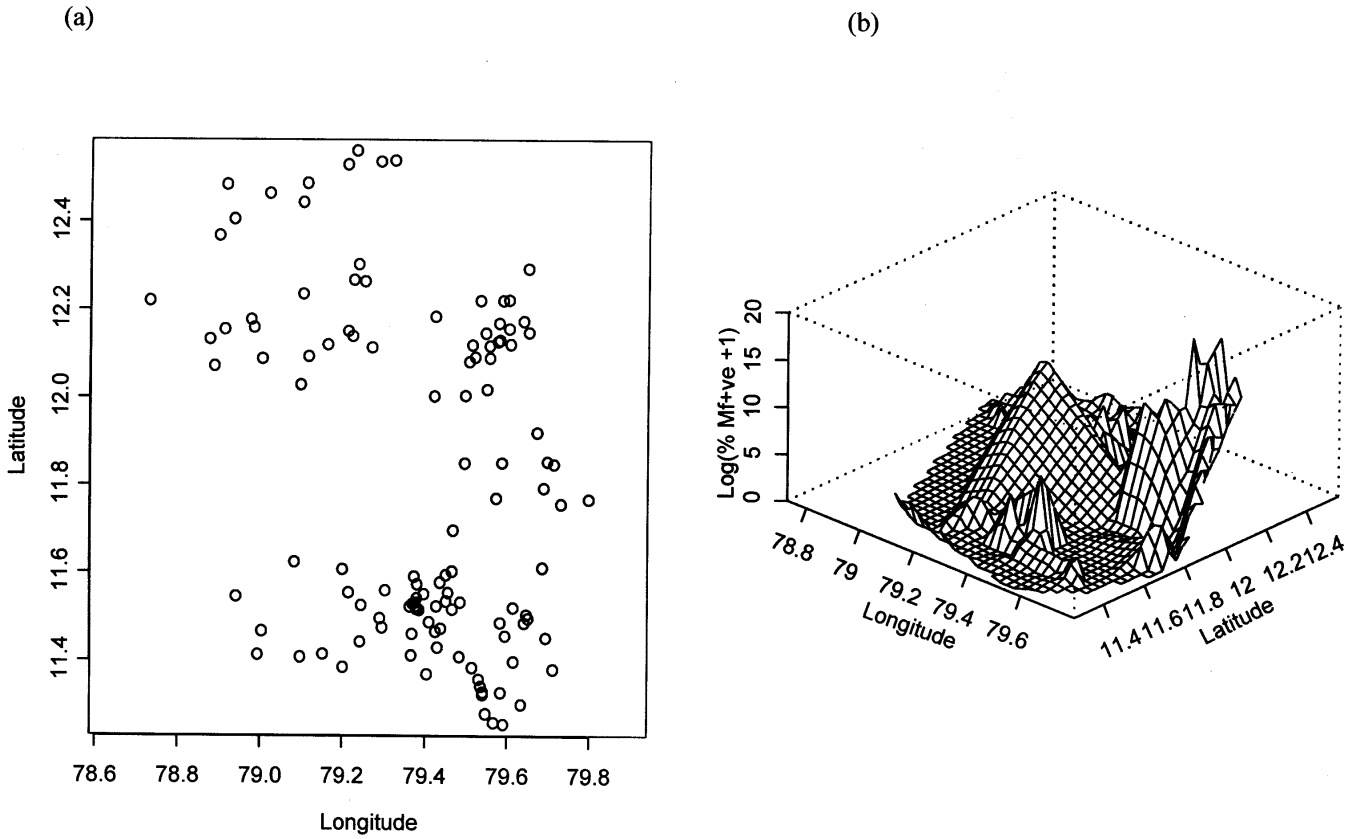

FIGURE 6. (a) Locations of 119 sites within a subregion in the Rapid Geographical Assessment of Bancroftian Filariasis area (Figure 1) where microfilaraemia prevalence data were collected. (b) Perspective plot of a local trend surface (created using the *interp* function as described in Figure 5) over the subregion for the log microfilaraemia prevalence values obtained from the study communities in (a).

study), spatial resolution and size of local sampling points (the grain), and local differences in parasite transmission dynamics and population processes including vector dispersal ranges.[11,14,17,36–39] The WHO recommendation for using a sampling grid of 25 x 25 km grid in the multicountry RAGFIL mapping project was based on preliminary analyses of spatial patterns of ICT prevalence distribution in Myanmar and microfilaraemia prevalence distribution in Gulbarga district in Karnataka state in India,[25] which indicated large filariasis transmission foci (range up to 260 kms in the case of the Myanmar dataset) in both areas. The proposal for using a grid span of 25 kms was therefore considered conservative given these results and indeed thought to represent a practical sampling scheme designed to enhance the likelihood of detecting local spatial variation occurring at smaller spatial scales in other areas. However, no analysis was made regarding the contribution of variations in the extent and grain of the study areas as well the impact of large-scale trends in the data to the finding for the existence of large filariasis foci in the above test areas. As the present results confirm, this gap in analysis means that the WHO sampling scheme should be applied cautiously when mapping filariasis distribution in disparate endemic regions. They also suggest that it may be prudent for workers to carry out supplementary field work at the outset to determine the approximate scale of investigation in their areas, perhaps via the use of multi-stage nested sampling techniques and analyses,[40] before attempts to study filariasis spatial variation using the RAGFIL approach are made.

Second, given that uncovering small-scale spatial dependence is dependent on the spatial scale of measurement, using inappropriate sampling scales have also important implications for our ability to map filariasis prevalence over larger areas from spatially stratified samples. For example, the lack of small-scale spatial dependency for the present antigenemia data at the sampling scale of 25 kms means that filariasis maps for this variable cannot be constructed for the present study area using spatial interpolation methods such as kriging.[37] This is because sample values are essentially uncorrelated at this scale, while kriging methods rely on there being spatial covariance structure in the data. Indeed, our analysis of the microfilaraemia data suggests that if these mapping methods are to be used (or alternatively if the spatial covariance structure is to be uncovered) for generating filariasis distribution maps across the present study region, sample points may well need to be located < 10 kms apart.

Nonetheless, despite these limitations the analyses described here have yielded several insights regarding the spatial structure of filariasis in the present South Indian study area. The first is that although the spatial population dynamics of microfilaraemia prevalence may well be different to those of antigenemia prevalence, the variogram results for the former suggest that the foci of filariasis transmission in this area could, on the other hand, be relatively small (with a patch diameter smaller than the sampling interval of 25 km). A small filariasis transmission patch or spatial infection cluster is also in line with theoretical expectations that vector-

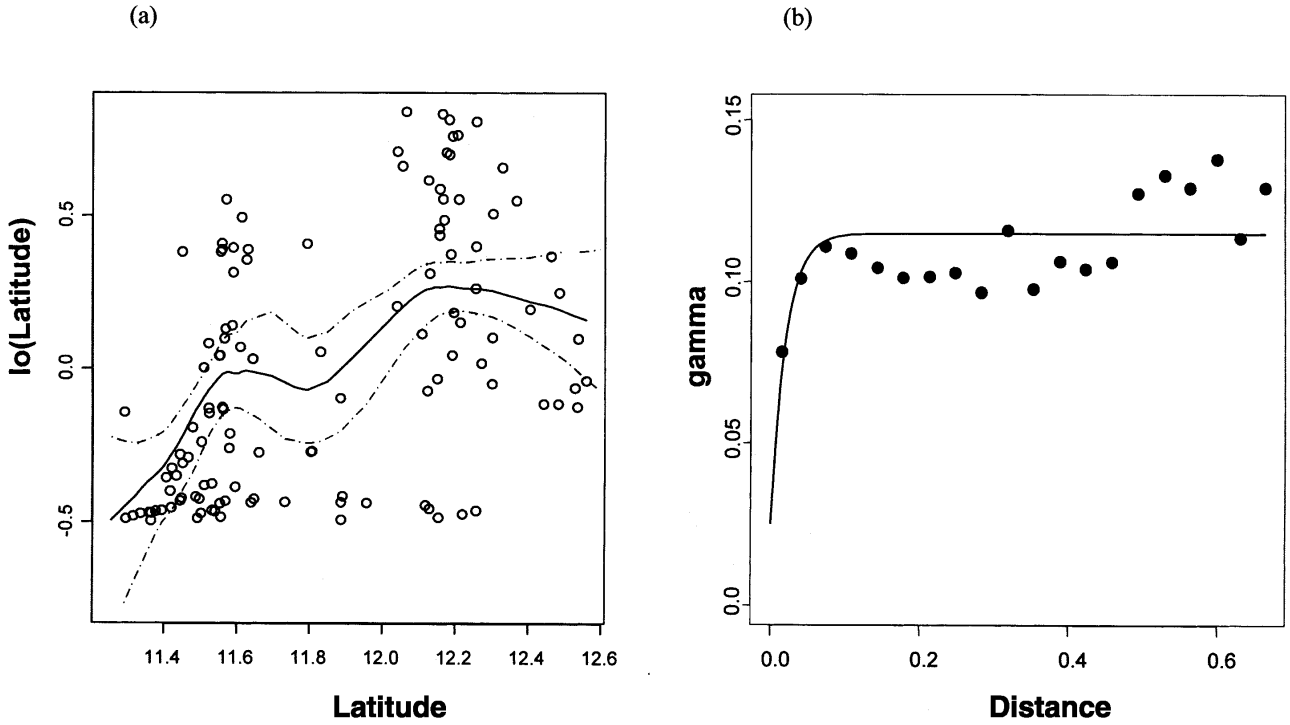

FIGURE 7. (a) Fitted function (solid line) for latitude in a generalized additive models fitted to the log microfilaraemia prevalence data (symbols) using the spatial coordinates of each point as predictors. Dashed lines show the pointwise 95% confidence intervals of the estimated function and indicate a significant nonlinear association between log microfilaraemia prevalence and latitude ($F = 6.89$, df = 1,2.9, $P < .001$). (b) Omnidirectional semivariogram for the detrended log microfilaraemia prevalence data. Symbols are experimental semivariances and the solid line is the fit of the exponential model described in the text. Parameters of the fitted model are as given in the text. Distance is in degrees.

borne parasitic diseases are more likely to exhibit high variability at local spatial scales.[16,41] The analysis of the ICT prevalence data, by contrast, suggests that parasite spatial population structure in this region is also governed by large-scale trends for high prevalence, which appeared associated with altitudinal changes from the south-southeast to the north-northeast regions of the study area (Figure 5). Taken together these results therefore imply that both small-scale spatial processes and large-scale environmental factors may characterize filariasis spatial distribution in this region.

Our study was not designed to evaluate the factors that may underlie any observed spatial variation, but preliminary analysis of the large-scale trends in the present data revealed that the association of ICT prevalence with altitude may be partially due to a non-linear relationship between prevalence and elevation such that the prevalence of infection increased with elevation up to some 1500 m after which infection prevalence declined. Such an effect of elevation on infection prevalence has been noticed previously for filariasis in India[42] and Africa[43] and has been attributed to effects of altitude on both vector survival and parasite development and survival.[44] The present results suggest, however, that there may also be an additional beneficial effect on these variables at intermediate altitude levels occurring along mountain slopes. Note that gaining a better understanding of such environmentally related correlates of infection will also allow the derivation of ecologic zones of parasite transmission, which could represent biologically valid strata for developing more realistic stratified random sampling frameworks for estimating infection burden and distribution.[7,8] We are currently examining the use of remote sensed data on environmental/climactic variables to investigate the role of this topic in mapping the distribution of lymphatic filariasis in India.

Gaining a better knowledge of both small- and large-scale components of parasite population spatial structure is increasingly being recognized as also critical to gaining a better understanding of the overall spatial population dynamics of infection.[12,14,15,40,45,46] This body of work has for example highlighted the importance of such spatial entities as population patch size, shape, and connectivity and regional landscape gradients governing dispersal (i.e.., spatial entities pertaining to both small- and large-scale spatial variation respectively), in influencing the persistence, spread and control of infectious disease agents within regions.[15,41,47–49] These findings indicate that clarifying and quantifying filariasis spatial variation at the within-regional level are likely to become increasingly important to the present global efforts for achieving control of this disease. For example, such studies may not only allow effective targeting of disease-control efforts within control areas (that will be required if the duration of treatment, say, is expected to vary with the level of parasite endemicity), but also detection of high prevalence sites that may require more intensive monitoring and management during program implementation. Characterizing parasite spatial structuring will also facilitate the development of more realistic spatially explicit population dynamic models of infection.[11,12,47] which in turn will permit more rational design of effective control programs.

The present results indicate that geostatistic approaches if carefully applied can play an important first role in the uncovering and analysis of this spatial structure for filariasis epidemiology and control. However, this study also indicates

that the successful execution of these spatial studies requires a more thorough consideration of the appropriate geographic scales at which such investigations should be carried out.[14,17,18] We suggest analyses of spatial scales, in terms of uncovering both the scales of study and the spatial scales at which different ecologic processes act, to be a priority research topic at this stage of investigations into the spatial population structure of lymphatic filariasis.

Acknowledgments: VCRC acknowledges the Indian Council of Medical Research and the WHO for support and encouragement of this study. This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, Project ID:970679. The authors thank the British Council, Chennai, India, for support under the Higher Education Link Programme that facilitated this collaborative work. EM thanks the Medical Research Council, UK, for personal support.

Authors' addresses: P. K. Das, S. P. Pani, A. Srividya, M. Palaniyandi, Vector Control Research Centre, Indira Nayar, Pondicherry-6, India, E-mail: vcrc@vsnl.com. E. Michael, Department of Infectious Disease Epidemiology, Imperial College, London, United Kingdom.

REFERENCES

1. Clarke KC, Osleeb JR, Sherry JM, Meert JP, Larson RW, 1991. The use of remote sensing and geographic information systems in UNICEF's dracunculiasis (Guinea worm) eradication effort. *Prev Vet Med 11:* 229–235.
2. Richards FO, 1993. Use of geographic information systems in control programs for onchocerciasis in Guatemala. *Bull Pan AmerHealth Organ 27:* 52–55.
3. Mott KE, Nutall I, Desjeux P, Cattand P, 1995. New geographical approaches to control of some parasitic zoonoses. *Bull World Health Organ 73:* 274–257.
4. Openshaw S, 1996. Geographical information systems and tropical diseases. *Trans R Soc Trop Med Hyg 90:* 337–339.
5. Snow RW, Marsh K, le Sueur D, 1996. The need for maps of transmission intensity to guide malaria control in Africa. *Parasit Today 12:* 455–457.
6. Michael E, Bundy DAP, 1997. Global mapping of lymphatic filariasis. *Parasit Today 13:* 472–476.
7. Robinson TP, 1998. Geographic information systems and the selection of priority areas for control of tsetse-transmitted trypanosomiasis in Africa. *Parasit Today 14:* 457–461.
8. Brooker S, Michael E, 2000. Geographical information systems (GIS) and remote sensing in the epidemiology and control of human helminth infections. *Adv in Parasit 47:* 246–288.
9. Sabesan S, Palaniyandi M, Michael E, Das PK, 2000. Mapping lymphatic filariasis at the district-level in India. *Ann Trop Med Parasit 94:* 591–606.
10. Lindsay SW, Thomas CJ, 2000. Mapping and estimating the population at risk from lymphatic filariasis in Africa. *Trans R Soc Trop Med Hyg 94:* 37–45.
11. Kareiva P, 1990. Population dynamics in spatially complex environments: theory and data. *Philo Trans R Soc Lon* ***B*** 330: 175-190.
12. Tilman D, Kareiva P, 1997. *Spatial Ecology: the Role of Space in Population Dynamics and Interspecific Interactions*, Princeton University Press, Princeton, New Jersey.
13. Kitron U, 1998. Landscape ecology and epidemiology of vector-borne diseases. *J Med Entomol 35:* 435–445.
14. Kitron U, 2000. Risk maps: transmission and burden of vector-borne diseases. *Parasit Today 16:* 324–325.
15. Grenfell BT, Harwood J, 1997. (Meta)population dynamics of infectious diseases. *Trends Evol Ecol 12:* 395–399.
16. Shope RE, 1999. Factors influencing geographic distribution and incidence of tropical infectious diseases. R.L. Guerrant, D.H. Walker and P.F. Weller, ed. *Tropical Infectious Diseases. Principles, pathogens and practice,* London, Churchill Livingston., 16-21.
17. Weins JA, 1989. Spatial scaling in ecology. *Functional Ecology 3:* 385–397.
18. Levin SA, 1992. The problem of pattern and scale in ecology. *Ecology 73:* 1943–1967.
19. Thompson DF, Malone JB, Harb M, Faris R, Huh OK, Buck AA, Cline BL, 1996. Bancroftian filariasis distribution and diurnal temperature differences in the southern Nile Delta. *Emer Infect Dis 2:* 234–235.
20. Meyrowitsch DW, Toan ND, Hao HT, Dan NT, Michael E, 1998. A review of the current status of lymphatic filariasis in Vietnam. *Acta Tropica 70:* 335–347.
21. Haining R, 1990. *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press, Cambridge, UK.
22. Isaaks EH, Srivastava RM, 1989. *Applied Geostatistics,* Oxford University Press, New York.
23. Cressie NAC, 1993. *Statistics for Spatial Data,* John Wiley & Sons, New York.
24. Kaluzny SP, Vega SC, Cardoso TP, Shelly AA, 1996. *S+SpatialStats User's Manual. Version 1.0,* Mathsoft, Inc, Seattle, Washington.
25. World Health Organization, 1998. *Research on Rapid Geographical Assessment of Bancroftian Filariasis.* TDR/TDF/COMDT/98.2.
26. Srividya A, Lall R, Ramaiah KD, Ramu K, Hoti SL, Pani SP, Das, PK, 2000. Development of Rapid Assesment Procedures for the delimiation of lymphatic filariasis endemic areas. *Trop Med Intl Health 5,:64-71*
27. Pani SP, Hoti SL, Elango A, Yuvaraj J, Lall R, Ramaiah KD, 2000. Evaluation of the ICT whole blood antigen card test to detect infection due to nocturnally periodic Wuchereria bancrofti in South India. *Trop Med Intl Health 5:* 359–363.
28. Bennett S, Woods T, Liyanage WM, Smith DL, 1991. A simplified general method for cluster-sample surveys of health in developing countries. *Rapp. Trimmest. Statist. Sanit. Mond. 44:* 98–106.
29. Das PK, Ramaiah KD, Vanamail P, Pani SP, Yuvaraj J, Balarajan K, Bundy DAP, 2001. Placebo-controlled community trial of four cycles of single - dose diethylcarbamazine or ivermectin against *Wuchereria bancrofti* infection and transmission in India. *Trans R Soc Trop Med Hyg 95:* 336–341.
30. Rossi RE, Mulla DJ, Journel AG, Franz EH, 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecol Mono 62:* 277–314.
31. Venables WN, Ripley BD, 1997. *Modern Applied Statistics with S-plus, Second Edition*. Springer-Verlag, New York.
32. Liebhold AM, Zhang X, Hohn ME, Elkinton JS, Ticehurst M, Benzon GL Campbell RW, 1991. Geostatistical analysis of gypsy moth (Lepidoptera; Lymantriidae) egg mass populations. *Environ Entomol* 20: 1407-1417.
33. Young LJ, Young JH, 1998. *Statistical Ecology. A Population Perspective,* Kluwer Academic Publishers, Boston.
34. Hastie TJ, Tibshirani RJ, 1996. *Generalized Additive Models.* London: Chapman & Hall.
35. Bailey TC, Gatrell AC, 1995. *Interactive Spatial Data Analysis,* Longman Scientific & Technical, Harlow, England.
36. Liebhold AM, Rossi RE, Kemp WP, 1993. Geostatistics and geographic information systems in applied insect ecology. *Ann Rev Entomol 38:* 303–327.
37. Sharov AA, Liebhold AM, Roberts EA, 1996. Spatial variation among counts of gypsy moths (Lepidoptera: Lymantriidae) in pheromone-baited traps at expanding population fronts. *Environ Entomol 25:* 1312–1320.
38. Bjornstad ON, Ims RA, Lambin X, 1999. Spatial population dynamics: analysing patterns and processes of population synchrony. *Trends in Eco Evol 14:* 427–432.
39. Koenig WD, 1999. Spatial autocorrelation of ecological phenomenon. *Trends Ecol Evol 14:* 22–26.
40. Webster R, Oliver MA, 2001. *Geostatistics for Environmental Scientists,* John Wiley & Sons, Chichester, England.
41. Greenwood BM, 1989. The microepidemiology of malaria and its importance to malaria control. *Trans R Soc Trop Med Hyg 83:* 25–29.
42. Raghavan NGS, 1969. Clinical manifestations and associated epidemiological factors of filariasis. *J Comm Dis 1:* 75–102.

43. Jordan P, 1956. Filariasis in the Eastern, Tanga and Northern Provinces of Tanganyika. *East African Med J 33:* 225–233.

44. Attenborough RD, Burkot TR, Gardner DS, 1997. Altitude and the risk of bites from mosquitoes infected with malaria and filariasis among the Mianmin people of Papua New Guinea. *Trans R Soc Trop Med Hyg 91:* 8–10.

45. Bolker BM, Grenfell BT, 1995. Space, persistence and dynamics of measles epidemics. *Proc R Soc Lon B*, 348, 308-320.

46. Bascompte J, Sole RV, 1998. *Modelling Spatiotemporal Dynamics in Ecology,* Springer-Verlag, New York.

47. Bolker BM, Grenfell BT, 1995. Impact of vaccination on the spatial correlation and persistence of Measles dynamics. *Proc Nat Acad Sci 93:* 12648–12653.

48. Holmes EE, 1997. Basic epidemiological concepts in a spatial context, in *Spatial Ecology: the Role of Space in Population Dynamics and Interspecific Interactions* (Tilman, D. & Kareiva, P., eds), pp111–136, Princeton University Press, Princeton, New Jersey.

49. Ferguson NM, May RM, Anderson RM, 1997. Measles: persistence and synchronicity in disease dynamics. Tilman, D. & Kareiva, P., ed. *Spatial Ecology: the Role of Space in Population Dynamics and Interspecific Interactions*, Princeton University Press, Princeton, New Jersey, 137–157.