NIH Public Access
Author Manuscript

Social and spatial processes associated with childhood diarrheal disease in Matlab, Bangladesh

Carolina Perez-Heydrich[a,b], **Jill M. Furgurson**[a,c], **Sophia Giebultowicz**[a,c], **Jennifer J. Winston**[a,c], **Mohammad Yunus**[d], **Peter Kim Streatfield**[d], and **Michael Emch**[a,c,*]

[a]Carolina Population Center, University of North Carolina at Chapel Hill, CB# 8120, University Square, 123 West Franklin Street, Chapel Hill, NC 27516-2524, USA

[b]Department of Biostatistics, University of North Carolina at Chapel Hill, 3101 McGavran-Greenberg, CB#7420, Chapel Hill, NC 27599, USA

[c]Department of Geography, University of North Carolina at Chapel Hill, Saunders Hall, Campus Box 3220, Chapel Hill, NC 27599-3220, USA

[d]International Centre for Diarrhoeal Disease Research, GPO Box 128, Dhaka 1000, Bangladesh

Abstract

We develop novel methods for conceptualizing geographic space and social networks to evaluate their respective and combined contributions to childhood diarrheal incidence. After defining maternal networks according to direct familial linkages between females, and road networks using satellite imagery of the study area, we use a spatial econometrics model to evaluate the significance of correlation terms relating childhood diarrheal incidence to the incidence observed within respective networks. Disease was significantly clustered within road networks across time, but only inconsistently correlated within maternal networks. These methods could be widely applied to systems in which both social and spatial processes jointly influence health outcomes.

Keywords

1. Introduction

Spatial and social network structures influence infectious disease transmission within populations. Across the globe, diarrheal disease contributes to significant morbidity and mortality due to a combination of complex and dynamic biological, environmental and social factors that vary over time and space. Spatial analysis can be used to identify clusters of disease incidence and to correlate observed patterns to measurable variables (Douven and Scholten, 1994). However, without considering the underlying causal structure and relationships, the use of spatial analysis runs the risk of over simplifying the construction of space and place in which transmission operates (Radil et al., 2010). It is especially important

Corresponding author: Michael Emch, Department of Geography, University of North Carolina at Chapel Hill, Saunders Hall, Campus Box 3220, Chapel Hill, NC 27599-3220, USA, Phone: (919) 843-1010; Fax: (919) 962-1537; emch@email.unc.edu.

to consider demographic and spatial heterogeneity for disease microorganisms, such as diarrheal pathogens, which are transmitted via multiple pathways (Savill et al., 2006).

This case study integrates spatial and social network analysis methods to evaluate the relative contributions of the local neighborhood and a maternal-based social network to diarrheal disease transmission dynamics. Diarrheal diseases cause approximately one in five childhood deaths around the world (Boschi-Pinto et al., 2008). Globally, South Asia and Africa carry the highest burden of diarrheal morbidity and mortality, and children bear the majority of that burden (Deen et al., 2008; Nair et al., 2010). Common pathogens such as cholera, rotavirus and shigellosis, are transmitted to humans via multiple pathways, including the aquatic environment or through contact with infected persons, fomites, or contaminated food and water (Craig, 1988; Islam et al., 2011). Transmission is therefore a product of both spatial processes, which affect exposure to pathogens in the physical and social environments, as well as social processes, which affect sanitation and hygiene behavior, access, storage and use of water, and food storage, preparation, and distribution (Bates et al., 2007). Models that combine spatial and social network analysis are rare but have the potential to enhance understanding of the underlying transmission processes.

Social networks have been the focus of numerous studies in recent years due to their influence on various social processes, such as business activity (Danis et al., 2011), internet use (Rosen et al., 2011), and terrorism (Knoke, 2012). Smith and Christakis (2008) outline some of the mechanisms through which social networks affect health: (a) the provision of social support (both perceived and actual), (b) social influence (e.g., norms, social control), (c) social engagement, (d) person-to-person contacts (e.g., pathogen exposure, secondhand cigarette smoke), and (e) access to resources (e.g., money, jobs, information). Analysis of social networks facilitates the study of the relationships between individual actors in a social environment, and the patterns and implications of the relational ties between actors (Wasserman and Faust, 1994). Within a social network framework, the linkages among collections of individuals become the units of analysis and can be used to investigate theories about relational processes. In the fields of infectious disease and public health there has been growing attention to the importance of social networks as researchers recognize that disease does not spread uniformly through populations or space (Carrington et al., 2005). Conventional disease transmission models often assume homogenous mixing within the host population, meaning all members of the population are equally likely to come into contact with one another. However, individuals tend to interact with only a subset of the total population and the clustered nature of social networks leads to a non-random probability that two individuals will make contact (Ames et al., 2011). Thus, underlying social network structure will impact temporal and spatio-social patterns of disease transmission, decreasing the utility of models that assume homogenous mixing (Watts and Strogatz, 1998; Ames et al., 2011).

Spatial data analysis is used in a variety of disciplines to address relationships between observed spatial processes and certain variables of interest (Bailey and Gatrell, 1995). The spatial approach towards the study of disease and health focuses on mapping spatial patterns and correlating those patterns to measureable variables (Douven and Scholten, 1994). Although such analyses prove useful in detecting non-random disease patterns across physical space, a more comprehensive approach could incorporate information from non-spatial networks, and thus consider how the two could interact to drive observed disease processes. Through the integration of spatial and social network analyses, hypotheses associated with disease dynamics can be formulated to simultaneously address how geographical and social contexts influence observed disease patterns.

There has been growing attention in the last decade to the importance of social networks on health (Klovdahl et al., 2001; Smith and Christakis, 2008), but few studies combine spatial and social network approaches (Verdery et al., 2012) and even fewer focus on health outcomes. In a study of nine communities in rural Ecuador, Bates et al. (2007) found that social network and geographical space measures predicted diarrheal disease prevalence. However, in this study social networks had a positive effect on disease risk, suggesting that social connectedness was a measure of social cohesion rather than pathogen transmission. Emch et al. (2012) compared spatial and social clustering of cholera and shigellosis through a model that included kinship-based networks and variable spatial buffers to address clustering across different neighborhood scales. Previous studies using road networks are rare and demonstrate that roads operate on disease transmission via a variety of mechanisms (Eisenberg et al., 2006). In a study on the transmission of foot and mouth disease in livestock in the United Kingdom, Savill et al. (2006) found that Euclidean distance between infectious and susceptible farms was a better predictor of transmission risk than via the road network except when major intervening geographical features were present. However, the authors suggest that one reason for this finding may be that infection through social networks is a potential uncontrolled confounding factor.

In this study, we develop novel methods for conceptualizing geographic space and social networks to evaluate the relative contributions of local environment and kinship-based social networks to diarrheal disease transmission dynamics in Matlab, Bangladesh. Infant mortality in Bangladesh has declined in recent years but remains high with one out of 15 children dying before they reach the age of five (NIPORT, 2009). In the Matlab HDSS area, the two leading causes of infant mortality are diarrheal disease and pneumonia and in 2010. Infant mortality rate was 30 deaths per 1000 infants (icddr,b, 2012), which was lower than the national rate of 38 deaths per 1000 infants (World Bank, 2010). Diarrheal diseases are endemic and occasionally expand into epidemics (Ali et al., 2002; Islam et al., 2011). Using the WHO-standard world population age structure, the age-standardized rate of mortality due to diarrheal disease in Matlab is 13.14 deaths per 1000 people (icddr,b, 2012). Bangladesh is among the most densely populated countries in the world, with 1,142 inhabitants per square kilometer, and 72% of those inhabitants live in rural areas (World Bank, 2010). Density within villages promotes intimate social relations among kin, with households serving as the basic unit of kinship in the rural parts of the country (Chowdhury, 1995). In Matlab, as in other rural areas of Bangladesh, households are linked within baris (compounds of patrilineally-related households). Thus, people have well defined obligations to their respective bari and weaker ties to their extended, larger kin group that reside outside of the bari.

The study area in Matlab is characteristic of many rural Bangladeshi communities. A largely traditional social structure governs the formation of social networks. Sexual divisions of labor exist between women and men, and traditional institutions and customs regulate the lives of most women. The social system of patrilocality dictates that a married couple resides near the family of the husband. A recently married woman will leave her natal family and household and move near her husband's parents. Their children will replicate this pattern with sons staying in the family bari of residence and daughters moving to their husband's bari. The practice of purdah restricts women's mobility outside the bari and creates a social barrier between women and the public sphere (Nosaka and Bairagi, 2008). When viewed as a broader set of cultural norms, purdah can be seen as limiting female interactions with men through exclusionary customs and labor division (Amin, 2002). To accommodate the constraints of purdah, women have completely separate responsibilities from men, whether confined to their bari or working in the fields. However, there is evidence that the restriction on women's mobility has been changing in the rural areas

during the last 20 years as more women work outside the home and female education within villages is often seen as advantageous (Arends-Kuenning and Amin, 2001).

The purpose of this study was to characterize the influence of spatial and social networks on childhood diarrheal incidence in Matlab, Bangladesh. In order to more accurately represent the underlying social network structure, as relevant to childhood diarrheal disease incidence in Matlab, we defined maternal networks by capturing the kinship links between households of related females. We then defined the spatial neighborhood according to road networks, and thus addressed the role of geographic space in new ways to provide a more accurate representation of spatial relationships. Roads can influence disease transmission in various ways (Eisenberg et al., 2006) but this study assumes roads facilitate social interactions as well as contact patterns relevant to the spread of infectious disease. The network enables the construction of spatial matrices that measure geographic distance between households along roads. While the road network represents physical (spatial) connections between people, the maternal network embodies non-spatial connections through which, for example, information between related women could be transferred. The main goal of this analysis is to address how kinship influences impact childhood disease incidence above and beyond the influence of interactions among people within spatial neighborhoods. Because most direct, personal interactions are expected to be spatially constrained, we must simultaneously control for these spatial dependencies when addressing the impacts of non-spatial, social neighborhoods on infectious disease outcomes. Together, maternal and transportation networks were used to evaluate the relative contributions of local environment and kinship-based social networks to diarrheal disease incidence in children under five.

2. Methods

2.1 Study area and data

This study was conducted in Matlab, Bangladesh, which is located about 55 kilometers southeast of Dhaka (Fig. 1). The study population was sampled from a Health and Demographic Surveillance System (HDSS) maintained by the International Center for Diarrheal Disease Research, Bangladesh (icddr,b). Since 1966, the HDSS has monitored the approximately 220,000 residents of Matlab, who live in patrilineally-related household clusters called baris Households within the same bari share resources such as wells, latrines, cooking facilities, and food although they sleep in separate structures within the bari compound. Baris are situated within villages, of which there are approximately 140 throughout Matlab. Settlement is in close proximity to water bodies, including ponds, canals, and rivers, with bari residents often using these as communal water sources (Carrel et al., 2005). The northern half of Matlab is somewhat less densely populated, with much of the settlements clustered around primary waterways. The southern region, where the main hospital and administrative centers are located, features a higher density of baris and more villages situated along smaller channels of the main river. Land used for agriculture and regions that are more susceptible to flooding act as natural barriers amongst communities. Approximately 40% of Matlab is located within a flood-control embankment that was built in 1988–1989 as part of the Meghna-Dhonagoda Irrigation Project (Emch, 2000). While residents in this area have a higher level of protection from seasonal flooding, residential distribution does not differ significantly between the two areas.

Once a month, baris are visited by community health workers who gather individual-level data on demographic events and disease incidence. The definition of diarrhea in this study is 3 or more bloody or watery stools during the previous 24 hours. Community health research workers asked mothers whether their children under 5 had a diarrheal event during a 24 hour recall period which was the dependent variable for this study. To create socioeconomic

covariates, we aggregated socio-demographic data to the bari level to generate a mean maternal education index and a wealth index based on ownership of household assets.

Five years of demographic and health data, from 1999 to 2003, were used for this analysis. A subset of 500 spatially referenced baris was randomly selected from the total number of baris with at least one child under five years of age during the study period. All recorded childhood diarrheal disease incidence was integrated with the bari-level data (Fig. 1). Spatial data was available through the Matlab geographic information system (GIS), which contains information on locations of all baris, health facilities, and environmental features. These data were used to generate environmental covariates, including distance to nearest river, pond density within 200 meters, and tubewell density. Existing GIS data were updated for the purposes of this study, as described in the Methods section.

2.2 Generation of bari-level road networks

While means of transportation and mobility in Matlab vary, it is probable that simple straight-line distance measures between households will not accurately reflect true proximity based on individual travel practices. The presence of roads and footpaths in the area serves as a more efficient means of getting from point A to point B; using distances via these routes may be a more realistic representation of movement and connectivity. Residents of households located further away from a road may still utilize these resources, especially if the final destination is accessible by road.

To accurately calculate travel distances, a transportation network of all major and secondary roads in Matlab was created based on existing spatial data and high-resolution satellite imagery (Emch, 1999; Ali et al., 2001). The Matlab GIS includes local roads, which were generated using aerial photographs and GPS ground-truth data collection in the 1990s. Additional roads have since been built and connectivity in the study area improved, allowing individuals greater access to travel by foot or other means. The new transportation network thus improved upon the earlier data by integrating additional and recent features. These were digitized using current aerial and satellite imagery available through Google Earth software (Google, 2011). The software allows users to import imagery directly into an existing GIS. It also contains functions that allow the user to generate line and polygon features within the software. Roads were primarily digitized using this method, with reference shapefiles imported from the GIS into Google Earth for accuracy assessment (Esri, 2010). The resulting network provided an improved system of connected roads, which could then be used to measure transportation distances between baris in Matlab (Fig. 1). While certain roads are more heavily traveled than others, the road segments were not weighted due to the fact that individuals living near less-used roads would still most likely use that route if it was closest to their bari.

The Network Analyst extension in ArcGIS (Esri, 2010) was used to calculate the shortest possible distances via road between all baris used in the analysis. The extension contains an Origin-Distance (OD) Cost Matrix function, which identifies and measures the least-cost paths along a user-defined network. The output produced is a matrix of "costs" associated with travel from an origin location to a set of destination locations. The maximum distance search radius may be specified but was not set in this case, as the sample of baris measured was randomly distributed throughout the study area. Each bari was identified as both an origin and a destination within the cost-distance matrix. Based on the road network, distances in meters between each origin and all other destinations were measured.

2.3 Generation of bari-level maternal networks

Maternal networks were generated according to data from a historic demographic surveillance system (D'Souza, 1981), which provided information on cohabitation of individuals within households from 1983–2003. As mentioned previously, we restricted our analysis to a random subset of 500 baris with children under five years of age documented within the demographic surveillance system. For each household in the surveillance system, we determined which individuals lived together at any point within this period, and defined these cohabitating individuals as a unique family. We created unique identifiers for each family, which were based on the identification number of the eldest female in the household. When individuals moved from one household to another, in order to start a family of their own, we assumed that they maintained ties to the members of their original family. Using this approach, we were able to capture kinship linkages between females even after they moved from their households, and thus individuals could be linked to their grandmothers, mothers, sisters and daughters despite occupying separate households.

After defining maternal networks (i.e., linkages between individuals according to familial ties), we then defined bari networks according to the familial ties of bari residents. In other words, bari C and bari D had a tie if at least one individual from bari C was related to at least one individual in bari D. We used matrix operations to define bari-to-bari connections. First, we defined familial linkages among individuals in the sample. Specifically, let $\mathbf{M}$ be an $n \times m$ matrix of individual-to-family ties, where n and m correspond to the number of individuals and families, respectively, included in this study. Let

$$\mathbf{F} = \mathbf{M}\mathbf{M}',$$

so that $\mathbf{F}$ represents the $n \times n$ matrix of individual-to-individual ties based on familial relationships. Elements of $\mathbf{F}$ (f_{ij}) are binary, with 1 representing a familial connection between individual i and individual j. Now let $\mathbf{B}$ be an $n \times b$ matrix of individual-to-bari ties, where b corresponds to the number of baris in the sample and n is defined as above. Finally, the familial linkages among baris were defined as:

$$\mathbf{T} = \mathbf{B}'\mathbf{F}\mathbf{B}.$$

Figure 2 depicts the ties corresponding to a subset of the network generated using these methods.

2.4 Model formulation

We used a simultaneous autoregressive lag model with a spatially correlated error term (LeSage and Pace, 2009) to determine how childhood diarrheal incidence within a bari was associated with diarrheal incidence among individuals within maternal networks, while controlling for diarrheal incidence among individuals connected spatially through road networks. Specifically, the model was of the form:

$$Y = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}Y + u$$
$$u = \lambda\mathbf{V}Y + \varepsilon$$

where Y represents the bari-level incidence rate of childhood diarrhea, $\mathbf{X}$ is a design matrix of bari-level covariates (i.e., distance to nearest river, pond density within 200 m, tube well density, mean wealth index, and mean maternal education), β is the vector of corresponding regression coefficients, $\mathbf{W}$ is the row-standardized maternal network matrix ($\mathbf{T}$), ρ is an

autocorrelation coefficient defining the relationship between incidence within a bari's maternal network and its own observed incidence, $\mathbf{V}$ is the distance-weighted road network matrix, λ is a spatial autoregressive coefficient used to correct for spatial dependence of childhood diarrhea incidence within a bari's road network, and ε is the residual normally distributed random error term, $\varepsilon \sim N(0, \sigma^2)$. The model was fit in R (Team, 2010) using the *spdep* package (Bivand and Gebhardt, 2000).

3. Results

Overall, bari-level maternal networks were relatively sparse. The average number of links observed was 0.435. Of the 497 baris included in the sample, 345 had no familial links to other baris, 105 had one link, 34 had two links, nine had three links, and four had four links. Summary statistics for bari-level covariates included in the analysis are listed in Table 1. Mean diarrheal disease incidence among children under five ranged between 264 and 434 cases per 1,000 people, with the lowest and highest incidence rates observed in 1999 and 2003, respectively. Median road network distance between baris was 22.2 km, with an interquartile range between 8.6 and 38.8 km.

Moran's index values plotted in Figure 3 range between +1.0 and −1.0, where values near +1.0 are indicative of clustering, values near −1.0 indicate dispersion, and values near 0 indicate no autocorrelation. Moran's I tests for autocorrelation demonstrated that childhood diarrheal incidence significantly clustered within spatial road networks across most years (Fig. 3). After correcting for the spatial dependence of disease through the use of a social lag model with spatially correlated errors, the impact of kinship networks on childhood diarrheal disease incidence was significant for only two of the five years (1999 and 2001; Fig. 4). Estimates of the social autoregressive coefficient ρ indicate how extensively diarrhea incidence within a bari's network correlates with a bari's own disease incidence. As with Moran's index values, ρ is constrained to lie between +1.0 and −1.0, with values near the former indicating strong positive autocorrelation and values near the latter corresponding to strong negative autocorrelation. The estimates for ρ in 1999 and 2001 were 0.169 (SE = 0.0672) and 0.168 (SE = 0.0577), respectively. Although these estimates were found to be statistically different from zero, they are suggestive of only a slight positive correlation between bari-level diarrhea incidence and disease incidence within a bari's network.

4. Discussion

Overall, we found that childhood diarrheal incidence consistently clustered in space, while the influence of maternal networks on disease incidence was inconsistent across time. Although network structure remained constant across all years, the influence of maternal networks on childhood diarrheal incidence varied. This reflects findings from previous work by Giebultowicz et al. (2011) and Emch et al. (2012) who also found temporal variability in the effects of kinship networks on enteric disease incidence. Potential reasons for this may be related to changes in information diffusion between related females or fluctuations in disease incidence brought about through other unaccounted environmental processes. For instance, Emch et al. (2012) found temporal fluctuations in the clustering of cholera and shigellosis among households throughout the study area, and suggested that under outbreak, rather than endemic disease conditions, the relative importance of transmission mechanisms on disease incidence could vary (Tien and Earn, 2010). The years in which the influence of maternal networks was statistically significant corresponded to the years of lowest disease incidence. In this case, perhaps, maternal networks may contribute more strongly to disease prevention than to increased transmission, as observed by Bates et al. (2007). For example, shared hygienic and/or vaccination practices among kin could reduce disease risk within kinship networks. Sparse network structure could have also contributed to observed

inconsistencies in significance. The low density of maternal networks suggested that familial linkages among individuals were more prevalent within rather than between baris. Thus, most related women included in the study sample resided in the same bari, and bari-to-bari familial ties were rare. This is contrary to our expectations. We had anticipated there to be more spatial distance between related females than related men, since the latter typically remain near their fathers and the former leave their households to live with their husbands.

Because specific mechanisms driving these patterns are largely speculative, the time-varying significance of the social lag coefficient (ρ) warrants further investigation. Possible non-social factors influencing this pattern could also be related to changes in baseline incidence due to climatic factors (Ruiz-Moreno et al., 2010), which would influence the incidence of disease within social networks, and could contribute to increased transmission among members of a network component. Although we did control for environmental factors known to be associated with diarrheal disease incidence (i.e., tube well density, pond density, distance to nearest river), the effects of social and/or spatial networks could still have been confounded by time-dependent environmental processes that were not considered in the analysis. Future work should therefore regard the influence of temporal processes when characterizing social and spatial risk factors associated with disease incidence.

The most novel aspect of the approach we present in this paper corresponds to the implementation of Google Earth software for the construction of road networks. Road networks were expected to more accurately describe the spatial connections between baris in comparison to networks generated using Euclidean distances. We formally tested this assumption by comparing the model fits via Akaiki's Information Criterion (AIC) of two models that defined the spatial weight matrix **V** according to a: (1) road network, and (2) Euclidean distance network. A difference in AIC (ΔAIC) between models of less than 2.0 indicates that both models are equally supported by the data. Otherwise, the lower the AIC value, the greater the relative support attributed to a model. Upon comparing the two models we found that for three of the five years the two approaches resulted in similar model fits (ΔAIC < 2.0), but for the other two years, the approach using road networks was best supported (ΔAIC > 4.0). Thus, overall, the use of road networks was associated with better model support compared to the use of a Euclidean distance network.

Although the use of road rather than Euclidean networks was better supported by the data, there are certain limitations regarding the generation and use of these networks. For instance, while all visible roads were integrated into the GIS and used to build an interconnected physical network, there were likely routes such as footpaths that were undetectable from aerial imagery due to tree cover or other features. Additionally, using a road network to measure connectivity does not account for other means of transportation, such as canals, where transport occurs by boat. Road networks could also be generated to reflect costs associated with travel, which was disregarded in this analysis. When constructing the cost-distance matrix for the generation of road networks, we assumed an equal cost of travel among all roads so that they were not weighted in terms of the most prominent mode of transportation (i.e., bicycle, car, bus, foot). There was no way to determine the type of transportation available to household members from DHSS data. Thus, we were limited in assuming that mode of transportation within road networks had no influence in the spatial connectivity of individuals.

Our use of maternal networks was also somewhat limited in that relationships with non-kin were disregarded. Our motivation behind generating and using a maternal kin network was to investigate how ties to family members can influence childhood diarrhea incidence through mechanisms associated with information transfer and/or common hygienic/ vaccination practices. A priori, we believed the diffusion of common practices and

knowledge to be strongest among female relatives. For instance, women of child bearing age are known to routinely visit their families. This is especially the case for child birthing, when they may even go back and live at home with their mothers for some time. In our analysis, social interactions among non-kin neighbors, which are likely to be common, would have subsequently been captured through spatial networks. The significance of these social interactions would have therefore been muted by the control of spatial dependencies. Thus, although our results suggest that female kin relationships have a minor influence on childhood diarrheal disease, these findings do not suggest that social networks are unimportant. Rather, we found that bari-level familial linkages, specifically, fail to consistently influence disease incidence.

Overall, with this paper we provide novel methods for (1) defining social and spatial structure, and (2) integrating their effects on population health outcomes. This work demonstrates the utility of considering how both spatial and social networks may impact disease transmission, and improves upon network definitions of prior studies. As Bates et al. (2007) describe, the relationship between spatial and social configurations needs to be understood as either a dynamic joint effect or one that approaches collinearity in order to best target prevention efforts. Few studies have implemented a joint spatial-social approach to address research questions (Verdery et al., 2012); however, numerous studies have used spatial models to identify and characterize observed patterns of disease clustering (Riley, 2007), and there are an increasing number of studies that utilize social network methods to understand contact patterns that affect transmission (Klovdahl et al., 1994; Eubank et al., 2004). Using an integrated spatial and social network framework to understand the transmission of infectious pathogens within a population allows insights into both underlying patterns of disease and potential mechanisms associated with observed patterns (Radil et al., 2010). Specifically, we used road networks rather than spatial networks based on Euclidian distance, to consider spatial clustering of individuals. This likely provides a more accurate representation of the spatial interactions among people between baris, and could be widely applied in other settings where spatial context is influenced by travel and/or commuting behavior. Additionally, we define social networks in terms of maternal linkages, whereby females are linked to one another based on direct family ties. The finding that childhood diarrheal disease clusters spatially across all years affirms the importance of environmental factors in disease transmission, while the sporadic significance of social effects in some years, above and beyond the spatial clustering of disease, suggests that maternal networks may also play a role under certain circumstances. The methods presented here could be easily and widely implemented to other study systems where both spatial and social processes are likely to influence outcomes of interest.

Acknowledgments

References

Ali M, Emch M, Ashley C, Streatfield PK. Implementation of a Medical Geographic Information System: Concepts and Uses. Journal of Health, Population, and Nutrition. 2001; 19:100–110.

Ali M, Emch M, Donnay JP, Yunus M, Sack RB. The spatial epidemiology of cholera in an endemic area of Bangladesh. Social science & medicine. 2002; 55:1015–1024. [PubMed: 12220086]

Ali M, Emch M, Von Seidlein L, Yunus M, Sack DA, Rao M, Holmgren J, Clemens JD. Herd immunity conferred by killed oral cholera vaccines in Bangladesh: a reanalysis. The Lancet. 2005; 366:44–49.

Ames GM, George DB, Hampson CP, Kanarek AR, Mcbee CD, Lockwood DR, Achter JD, Webb CT. Using network properties to predict disease dynamics on human contact networks. Proceedings of the Royal Society B: Biological Sciences. 2011; 278:3544–3550.

Amin S. The poverty–purdah trap in rural Bangladesh: implications for women's roles in the family. Development and Change. 2002; 28:213–233.

Arends-Kuenning M, Amin S. Women's capabilites and the right to education in Bangladesh. International Journal of Politics, Culture, and Society. 2001; 15:125–142.

Bailey, T.; Gatrell, T. Interactive Spatial Data Analysis. New York: J. Wiley; 1995.

Bates SJ, Trostle J, Cevallos WT, Hubbard A, Eisenberg JNS. Relating Diarrheal Disease to Social Networks and the Geographic Configuration of Communities in Rural Ecuador. American Journal of Epidemiology. 2007; 166:1088–1095. [PubMed: 17690221]

Bivand R, Gebhardt A. Implementing functions for spatial statistical analysis using the R language. Journal of Geographical Systems. 2000; 2:307–317.

Boschi-Pinto C, Velebit L, Shibuya K. Estimating child mortality due to diarrhoea in developing countries. Bulletin of the World Health Organization. 2008; 86:710–717. [PubMed: 18797647]

Carrel M, Emch M, Streatfield PK, Yunus M. Spatio-temporal clustering of cholera: The impact of flood control in Matlab, Bangladesh, 1983–2003. Health and Place. 2009; 15:741–752. [PubMed: 19217821]

Carrington, PJ.; Scott, J.; Wasserman, S. Models and methods in social network analysis. Cambridge University Press; 2005.

Chowdhury A. Families in Bangladesh. Journal of Comparative Family Studies. 1995; 26:27–42.

Craig M. Time-space clustering of Vibrio cholerae 01 in Matlab, Bangladesh, 1970–1982. Social science & medicine. 1988; 26:5–13. [PubMed: 3353754]

D'souza S. A population laboratory for studying disease processes and mortality--the Demographic Surveillance System, Matlab Comilla, Bangladesh. Rural Demography. 1981; 8:29–51. [PubMed: 12338517]

Danis WM, De Clercq D, Petricevic O. Are social networks more important for new business activity in emerging than developed economies? An empirical extension. International Business Review. 2011; 20:394–408.

Deen JL, Von Seidlein L, Sur D, Agtini M, Lucas MES, Lopez AL, Kim DR, Ali M, Clemens JD. The high burden of cholera in children: comparison of incidence from endemicareas in Asia and Africa. PLoS Neglected Tropical Diseases. 2008; 2

Douven, W.; Scholten, H. Spatial analysis in health research. In: De Lepper, M.; Scholten, H.; Stern, RM., editors. The added value of geographical information systems in public and environmental health. GeoJournal Library. 1994. p. 117-133.

Eisenberg JNS, Cevallos W, Ponce K, Levy K, Bates SJ, Scott JC, Hubbard A, Vieira N, Endara P, Espinel M, Trueba G, Riley LW, Trostle J. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. Proceedings of the National Academy of Sciences. 2006; 103:19460–19465.

Emch M. Diarrheal disease risk in Matlab, Bangladesh. Social science & medicine. 1999; 49:519–530. [PubMed: 10414811]

Emch M. Relationships between flood control, kala-azar, and diarrheal disease in Bangladesh. Environment and Planning. 2000; 32:1051–1063.

Emch M, Root ED, Giebultowicz S, Ali M, Perez-Heydrich C, Yunus M. Integration of Spatial and Social Network Analysis in Disease Transmission Studies. Annals of the Association of American Geographers. 2012; 102(5)

Esri. ArcGIS Desktop. Redlands, CA: Environmental Research Systems Institute; 2010.

Eubank S, Guclu H, Kumar A, Marathe MV, Srinivasan A, Toroczkai Z, Wang N. Modellig disease outbreaks in realistic urban social networks. Nature. 2004; 429:184–184. [PubMed: 15107864]

Giebultowicz S, Ali M, Yunus M, Emch M. A comparison of spatial and social clustering of cholera in Matlab, Bangladesh. Health & Place. 2011; 17:490–497. [PubMed: 21195655]

Google. Google Earth. Google; 2011.

icddr, b. Registration of health and demographic events 2010, Scientific Report No. 117. Dhaka: icddr,b; 2012. Health and Demographic Surveillance System – Matlab, v.44.

Islam MS, Hossain MA, Khan SI, Khan MNH, Sack RB, Albert MJ, Hug A, Colwell RR. Survival of Shigella dysenteriae Type 1 on Fomites. Journal of Health, Population, and Nutrition. 2011; 19:177–182.

Klovdahl AS, Graviss EA, Yaganehdoost A, Ross MW, Wanger A, Adams GJ, Musser JM. Networks and tuberculosis: an undetected community outbreak involving public places. Social science & medicine. 2001; 52:681–694. [PubMed: 11218173]

Klovdahl AS, Potterat JJ, Woodhouse DE, Muth JB, Muth SQ, Darrow WW. Social networks and infectious disease: The Colorado Springs study. Social science & medicine. 1994; 38:79–88. [PubMed: 8146718]

Knoke, D. Social Networks and Terrorism. In: Safar, M.; Mahdi, K., editors. Social networking and community behvior modeling: qualitiative and quantitativ measures. 2012. p. 232-246.

Lesage, JP.; Pace, RK. Introduction to Spatial Econometrics. Boca Raton, FL: CRC Press; 2009.

Nair GB, Thandavarayan R, Bhattacharya MK, Krishnan T, Ganguly S, Saha DR, Rajendran K, Byomkesh M, Ghosh M, Okamoto K, Takeda Y. Emerging trends in the entiology of enteric pathogens as evidenced from an active surveillance of hospitalized diarrhoeal patients in Kolkota, India. Gut Pathogens. 2010; 2

National Institute of Population Research and Training (NIPORT), Mitra and Associates, Macro International. Bangladesh Demographic and Health Survey 2007. Dhaka, Bangladesh and Calverton, Maryland, USA: National Institute of Population Research and Training, Mitra and Associates, and Macro International; 2009.

Nosaka A, Bairagi R. Traditional roles, modern behavior: Intergenerational intervention and contraception in rural Bangladesh. Human Organization. 2008; 67:407–416.

R Core Development Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010.

Radil SM, Flint C, Tita GE. Spatializing social networks: using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. Annals of the Association of American Geographers. 2010; 100:307–326.

Riley S. Large-Scale Spatial-Transmission Models of Infectious Disease. Science. 2007; 316:1298–1301. [PubMed: 17540894]

Rosen D, Barnett GA, Kim JH. Social networks and online environments: when science and practice co-evolve. Social Network Analysis and Mining. 2011; 1:27–42.

Ruiz-Moreno D, Pascual M, Emch M, Yunus M. Spatial clustering in the spatiotemporal dynamics of endemic cholera. BMC Infectious Diseases. 2010; 10

Savill NJ, Shaw DJ, Deardon R, Tildesley MJ, Keeling MJ, Woolhouse ME, Brooks SP, Grenfell BT. Topographic determinants of foot and mouth disease transmission in the UK 2001 epidemic. BMC Veterinary Research. 2006; 2

Smith KP, Christakis NA. Social Networks and Health. Annual Review of Sociology. 2008; 34:405–429.

Verdery AM, Entwisle B, Faust K, Rindfuss RR. Social and Spatial networks: Kinship distance and dwelling unit proximity in rural Thailand. Social Networks. 2012; 34:112–127.

Wasserman, S.; Faust, K. Social network analysis: methods and applications. Cambridge University Press; 1994.

Watts DJ, Strogatz HS. Collective dynamics of 'small-world' networks. Nature. 1998; 393:440–442. [PubMed: 9623998]

World Bank. [Accessed July, 2012] World Development Indicators. 2010. Online at http://www.worldbank.org/data/dataquery.html

Highlights

- We examine the influence of social and spatial networks on childhood diarrhea.

- We generate social-maternal and spatial-road networks.

- For the analysis, we implement a spatial econometric model.

- Road networks influence diarrheal incidence; maternal networks have mild influence.
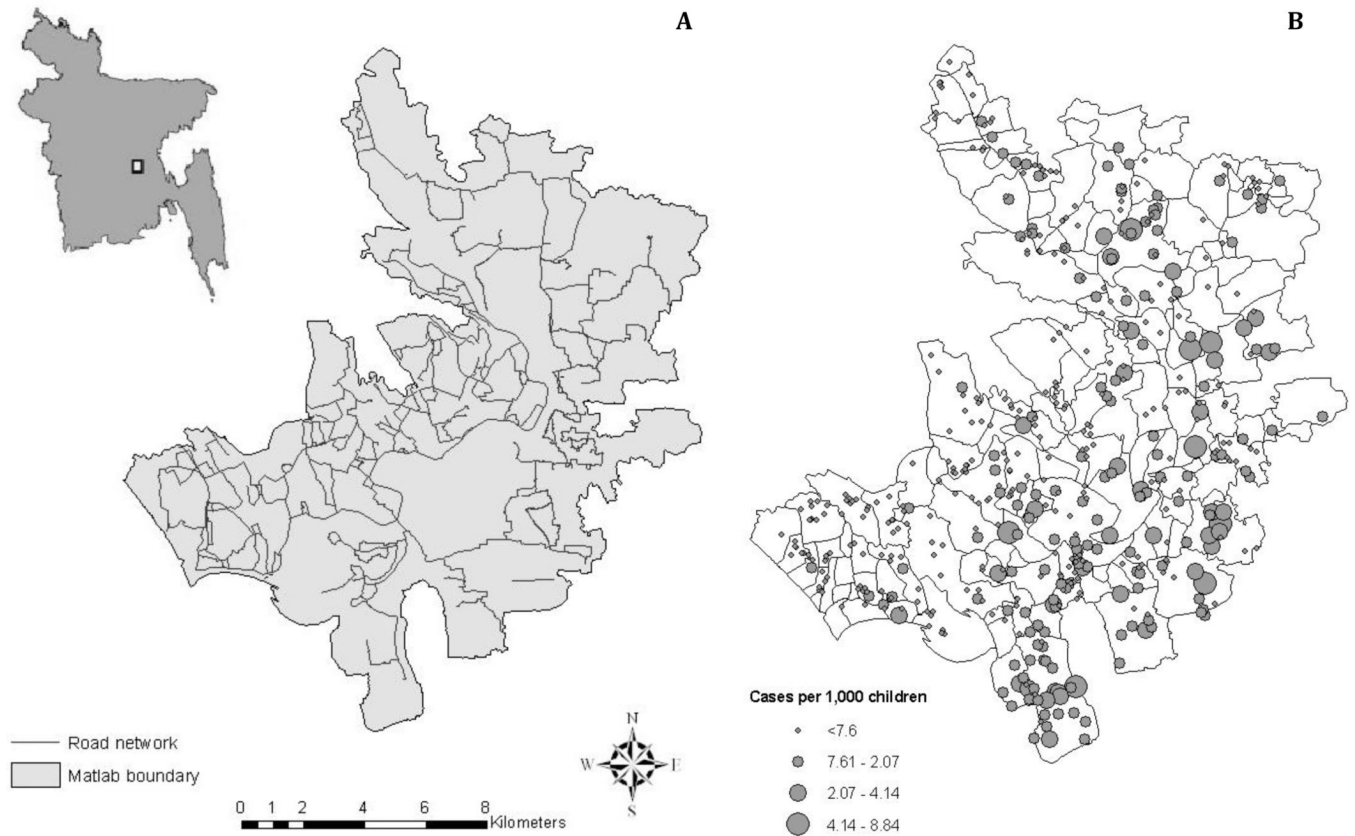
- The methods presented here are widely applicable.

Figure 1.
Map of (A) road network and (B) diarrheal disease incidence within the study area, Matlab, Bangladesh, located about 55 kilometers southeast of Dhaka. Roads were digitized using satellite imagery provided by Google Earth (Google, 2012). Total bari-level incidence per 1,000 people between 1999 and 2003 are plotted in (B) using graduated symbols, whereby smaller points represent baris with lower disease incidence and larger points represent baris with high disease incidence during the study period. Village boundaries are also overlaid in (B).
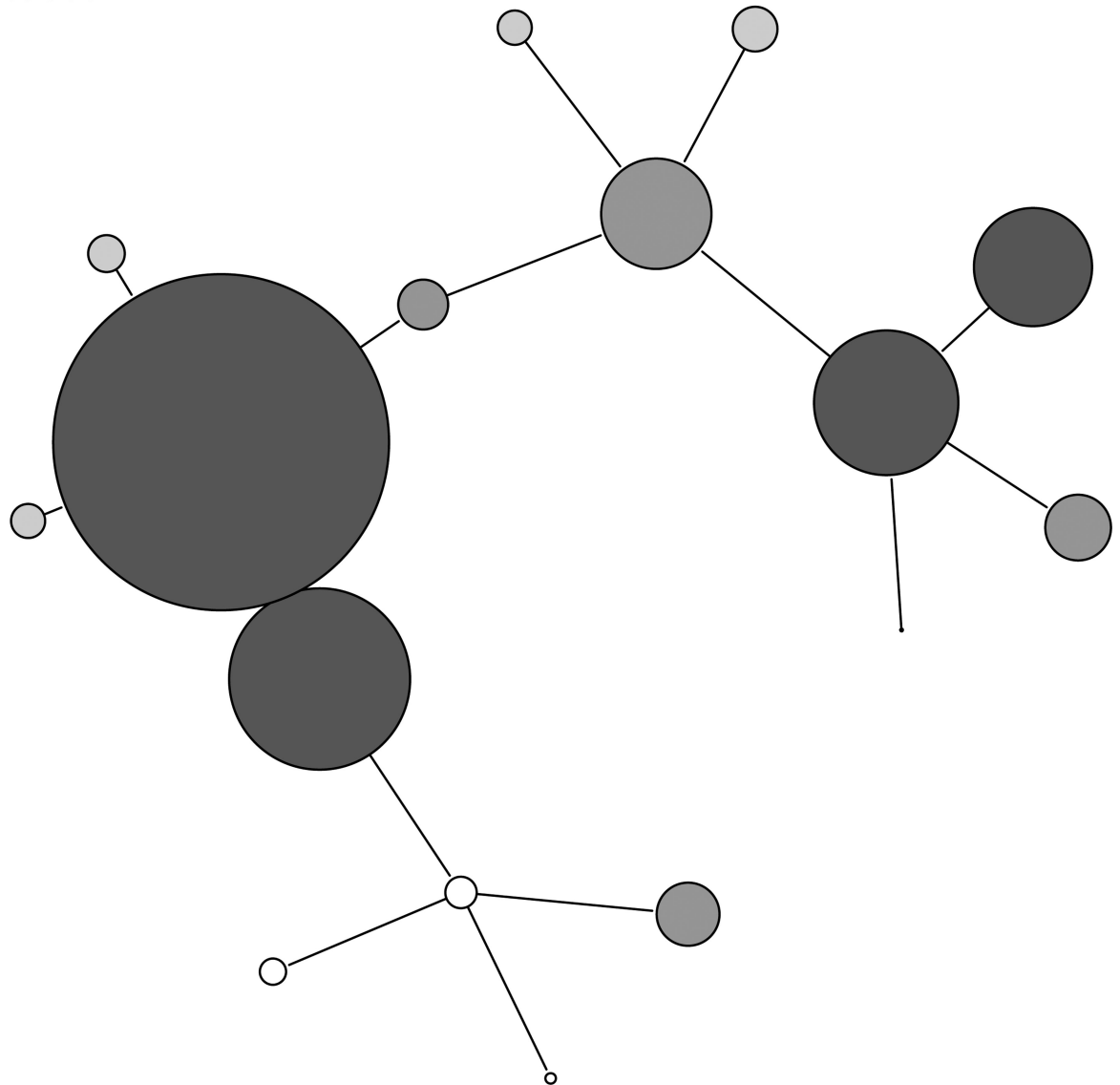
Figure 2.
Maternal network graph indicating the structure of family linkages within sample baris. The above graph depicts a subset of **T**, the maternal network adjacency matrix. Colors and sizes of nodes correspond to total number of childhood diarrheal cases observed in a given bari. Nodes are connected by edges that depict which individual baris are linked according to family ties.
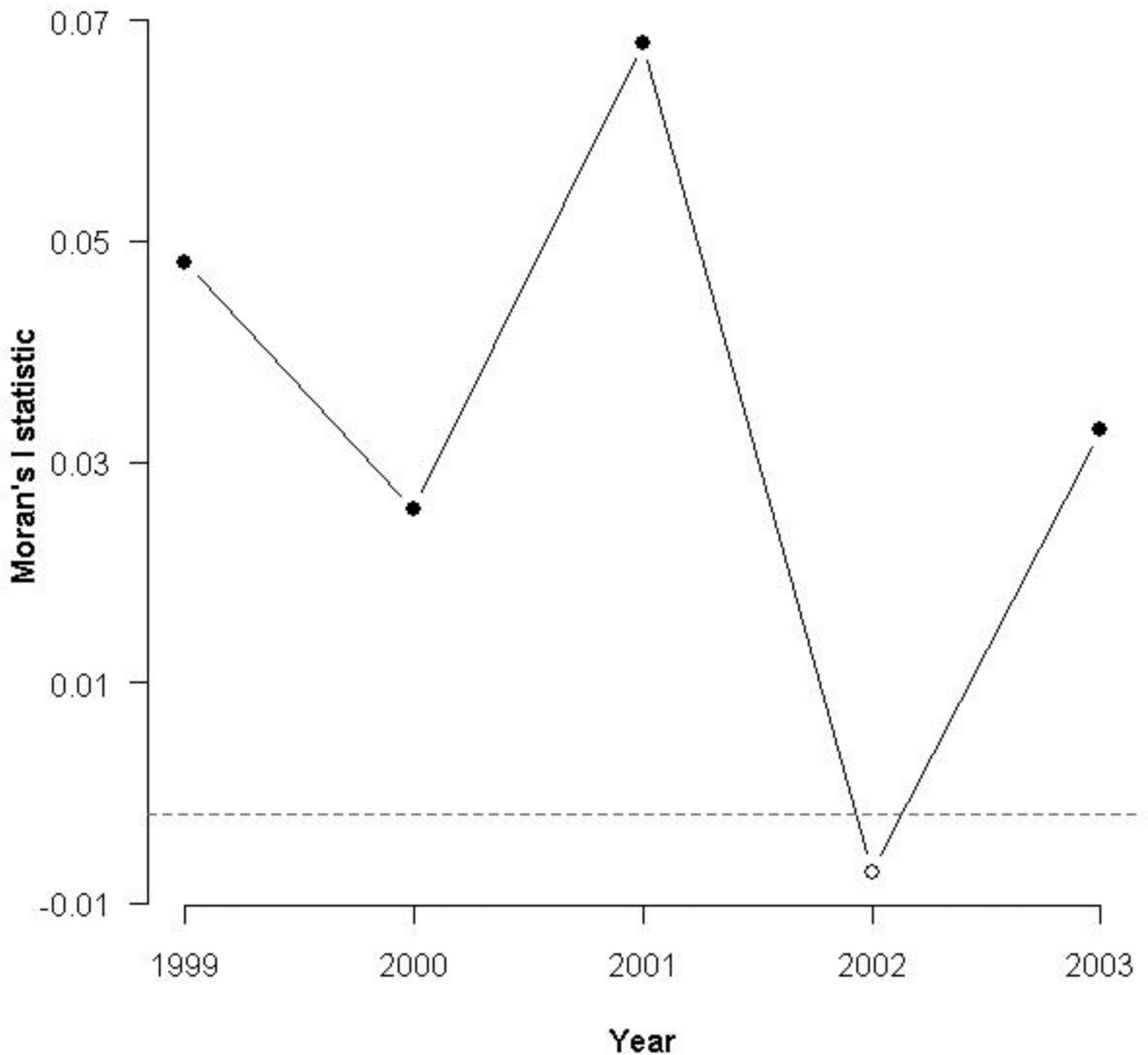
Figure 3.
Moran's I statistics used to determine the significance of spatial clustering of diarrheal cases road networks between 1999–2003. Points highlighted in black indicate significant clustering effects. The horizontal dashed line represents the expected value of the Moran coefficient under the null hypothesis of no autocorrelation in geographic space. Index values that are greater than that expected by random chance (assuming a significance level of 0.05) are represented by filled dots. Childhood diarrheal disease incidence significantly clustered within road networks across four of five years.
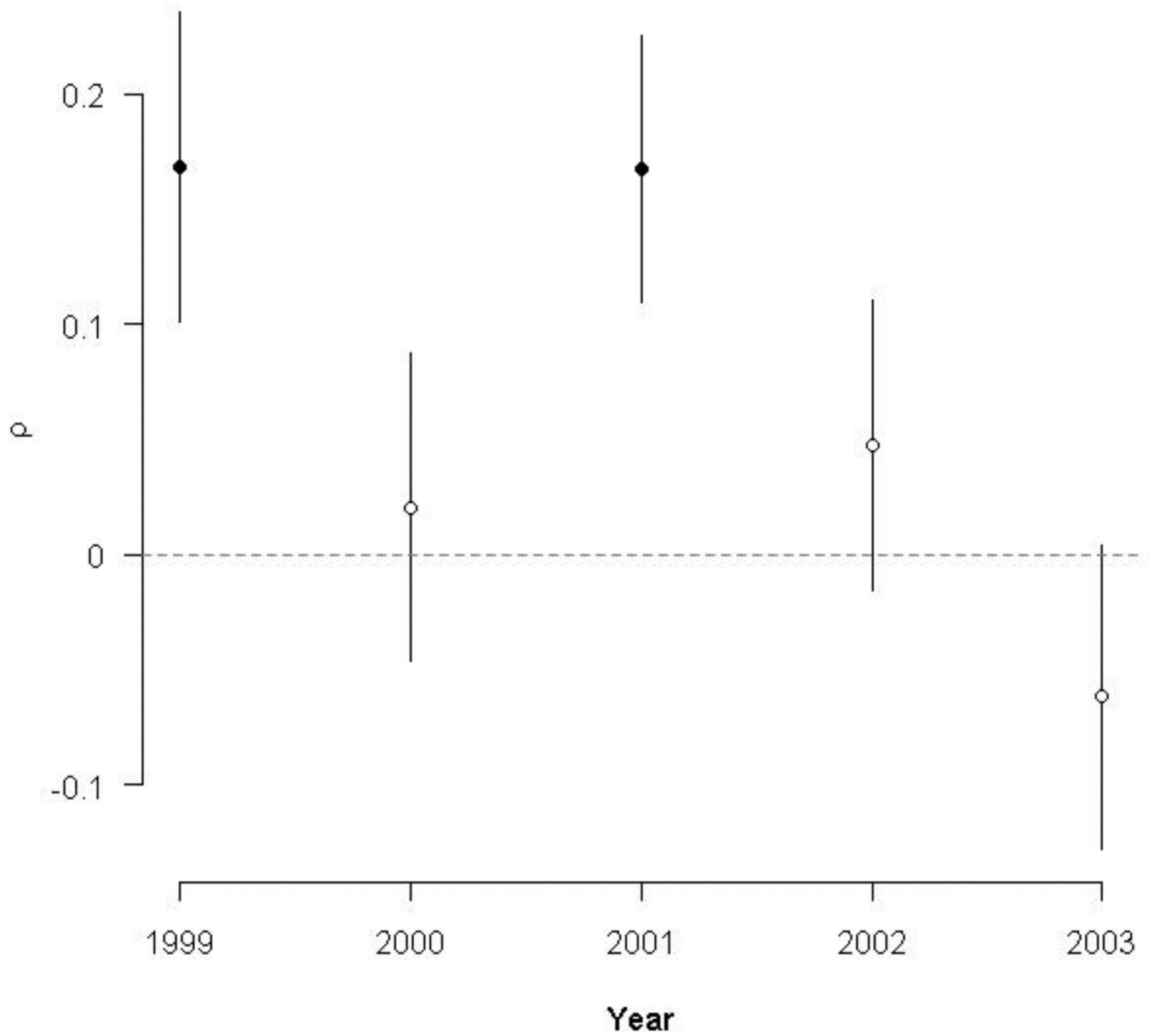
Figure 4.
Estimates of the lag simultaneous autoregressive lag coefficient (ρ) and standard errors.
Points highlighted in black indicate estimates that were significantly different from zero.
Overall, social structure was only significantly associated with childhood diarrheal incidence
for two of the five years.

Table 1

Distribution of covariate values across all baris included in the analysis. Pond and tubewell density represent counts within a 1000 m circular neighborhood. Wealth index represents a composite index variable associated with socioeconomic status. Maternal education represents the mean number of years of schooling for mothers in a bari. Childhood diarrhea incidence represents the number of new cases per 1,000 children per year.

	Median	25th percentile	75th percentile
Distance to River (m)	165.74	61.19	333.67
Pond Density	2.00	0.00	5.00
Tubewell Density	12.00	7.00	19.00
Wealth Index	3.00	2.20	3.67
Maternal Education	3.71	2.57	5.05
Childhood Diarrhea Incidence (1999)	0.00	0.00	400.00
Childhood Diarrhea Incidence (2000)	200.00	0.00	500.00
Childhood Diarrhea Incidence (2001)	166.67	0.00	428.57
Childhood Diarrhea Incidence (2002)	200.00	0.00	500.00
Childhood Diarrhea Incidence (2003)	285.71	0.00	666.67