SPATIAL PATTERNS OF URINARY SCHISTOSOMIASIS INFECTION IN A HIGHLY ENDEMIC AREA OF COASTAL KENYA

JULIE A. CLENNON, CHARLES H. KING, ERIC M. MUCHIRI, H. CURTIS KARIUKI, JOHN H. OUMA, PETER MUNGAI, AND URIEL KITRON

Department of Veterinary Pathobiology, University of Illinois, Urbana, Illinois; Center for Global Health and Diseases, Case Western Reserve University, Cleveland, Ohio; Division of Vector Borne Diseases Ministry of Health, Nairobi, Kenya; Kenya Medical Research Institute, Nairobi, Kenya

Abstract. Urinary schistosomiasis remains a major contributor to the disease burden along the southern coast of Kenya. Selective identification of transmission hot spots offers the potential for more effective, highly-local snail control and human chemotherapy to reduce Schistosoma haematobium transmission. In the present study, a geographic information system was used to integrate demographic, parasitologic, and household location data for an endemic village and neighboring households with the biotic, abiotic, and location data for snail collection/water contact sites. A global spatial statistic was used to detect area-wide trends of clustering for human infection at the household level. Local spatial statistics were then applied to detect specific household clusters of infection, and, as a focal spatial statistic, to evaluate clustering of infection around a putative transmission site. High infection intensities were clustered significantly around a water contact site with high numbers of snails shedding S. haematobium cercariae. When age was considered, clustering was found to be significant at different distances for different age groups.

INTRODUCTION

Urinary schistosomiasis is endemic in many sub-Saharan African countries where suitable habitats for Bulinus snails, the intermediate hosts of Schistosoma haematobium, are abundant. In Kenya, S. haematobium is highly endemic along the coast, where human exposure occurs primarily at pond and stream snail habitats. In some locations, open wells and boreholes are available, but because they typically yield hard water and are not suitable for bathing, local residents use them only for limited domestic activities, while continuing to rely on surface water sources for their water-use activities that result in long and extensive exposure (e.g., laundry, swimming). Although long-term mass chemotherapy campaigns targeting school age children have been implemented in the past, human treatment interventions have had only limited success in stopping transmission in this highly endemic region, where the reinfection rate varies between 9% and 21% per year.

To disrupt transmission more effectively and achieve prolonged disease control, there is a perceived need to develop more efficient, integrated control programs that are multifaceted, but selectively focused in space and time. Geographic information systems (GIS) and spatial analyses can be applied to consider spatial patterns of human infection, simultaneously with those of intermediate host snails, to improve efficiency of allocation for available transmission control measures.

To date, focal spatial studies of schistosomiasis have not used imagery or spatial statistics, whereas those schistosomiasis studies applying satellite imagery have used it to create national and continental schistosomiasis risk maps. In the present study, spatial aspects of S. haematobium human infection were examined in a highly endemic area on the southern coast of Kenya. Very high resolution (1–4 m²), remotely sensed imagery was used to aid in precise mapping of locations of all houses and water contact sites. Then, following integration of location data with demographic and parasitologic data in a GIS, spatial statistics were applied to elucidate the spatial patterns of infection, both among households and in relation to a transmission focal point.

MATERIALS AND METHODS

Study area and population. This study was conducted in Milalani Village, neighboring houses, and adjacent water contact/snail habitats in Nimbodze Pond, located in Msambweni Division, Kwale District, Coast Province, Kenya (Figure 1). According to a March–April 2000 demographic survey, Milalani consists of 306 households and 1,466 individuals. Additionally, some houses adjacent to Milalani from the villages of Mabatani, Nganja, and Vidungeni were also considered to analyze the effect of distance independently of Milalani Village affiliation and to create the rectangular study area required for spatial analyses. To evaluate the effect of study area size, two rectangular areas were considered, a 2,050 × 1,750-meter rectangle, including areas both east and west of Nimbodze Pond, and a smaller 1,320 × 1,720-meter rectangle comprised only of houses east of Nimbodze Pond (Figure 1).

Schistosoma haematobium infection in the division of Msambweni has been previously characterized. Water sources in the area include open wells and boreholes in addition to ponds, a stream, and a river that exhibit substantial temporal variation in water level. Milalani Village is situated between the large pond of Nimbodze and the main north-south road running from Mombasa to Tanzania (Figure 1). Compared with other open water sources in the area, Nimbodze pond has been considered a transmission hot spot because it has the highest proportion of host snail, Bulinus nasutus, shedding human cercariae (Kariuki HC and others, unpublished data).

Ethical oversight. Informed consent was obtained from area residents (or for children their parents) before determination of human infection status or water use activities. These studies were performed under human investigations protocols reviewed and approved by the Ethical Review Board of Kenya Medical Research Institute (Nairobi, Kenya) and by the Human Investigations Review Board of University Hospitals of Cleveland, Ohio.

Infection prevalence. Schistosoma haematobium infections were detected and quantified from two 10-mL urine specimens by means of Nuclepore filtration during April–May 2000. One thousand one hundred ten participants (age range...
years) from 280 households in Milalani Village were screened, and an additional 236 study participants in 46 households were surveyed from neighboring villages.

For purposes of this study, household infection levels were quantified as mean intensity and geometric mean intensity (considering only infected individuals), and as mean density and geometric mean density (mean number of eggs per person regardless of infection status) for each household.\textsuperscript{17}

**Mapping and geospatial processing.** A high resolution Ikonos satellite image was acquired for precise mapping of all households and surrounding landscape/landcover. The Ikonos satellite (Space Imaging, Atlanta, GA) is capable of generating 1-m\textsuperscript{2} panchromatic and 4-m\textsuperscript{2} multi-spectral images. The satellite rotates in a sun-synchronous orbit, passing over the same part of the Earth at roughly the same local time each day. It rotates over the Earth every 98 minutes at an altitude of approximately 680 kilometers. The size of the scene and area imaged by the sensor is 25 km\textsuperscript{2}, of which 11 km\textsuperscript{2} were used in this study. The Ikonos scene of Msambweni used in this study was acquired on March 4, 2001 at 7:45 AM, and the image is centered around 4.464°S and 39.449°E.

All houses in the study area were marked externally with unique household numbers, and hard copies of the image were used to map the households precisely. A global positioning system (GPS) (GeoExplorer II GPS; Trimble, Sunnyvale, CA) was used to confirm several house locations marked on the image, map landmarks, and map parts of Nimbdze Pond obscured by cloud cover. Given the number of houses in the Milalani area (> 300) and rest of the study area (> 3,000), the use of a GPS alone would not have allowed for such a comprehensive mapping effort. Once all houses were located, the GIS software packages ArcGIS 8.3 and ArcView 3.3 (Environmental Systems Research Institute, Redlands, CA) were used to create a digitized household level map over the Ikonos image, georectified to the Universal Transverse Mercator (UTM) zone 37S projection, 1984 datum. Demographic, parasitologic, malacologic, and environmental data were integrated into the GIS. Of the 280 households in Milalani tested, 279 houses were located and linked with the parasitologic data for 1,106 of 1,110 people. Maps presenting the distribution of human infection and snail distribution were subsequently created.

**Statistical analyses.** Statistical analyses were conducted on a variety of groupings based on sex and age. Particular emphasis was placed on very young and school age children. Since children in Kenya normally begin primary school at age six, and because it is not uncommon to have individuals as old as 21 years attending school in the study area, individuals were considered school age if they were between 6 and 21 years old.\textsuperscript{18} To further examine the effect of age, young individuals were divided into young children (0−5 years old), elementary school (6−13 years old), high school (14−17 years old), and young adult (18−21 years old). The elementary school age group was broken down further into the lower grades (6−9 years old) and upper grades (10−13 years old).

Non-spatial statistical analyses were calculated in SPSS version 11.5.0 (SPSS Inc., Chicago, IL). The homogeneity chi-square test was used to evaluate demographic group categories for differences in prevalence. The Mann-Whitney U test and the Kruskal-Wallis test were used for comparing distributions of infection intensity and density for two demographic groups and for three or more groups, respectively.

To evaluate human infection clustering by household, the software program (Point Pattern Analysis;\textsuperscript{19} San Diego State

**FIGURE 1.** Milalani Village households and neighboring houses for both the large and small analysis areas.
University, San Diego, CA) was used to calculate a global spatial statistic, the weighted K-function, and the local spatial statistics $G_i^*(d)$ and $G_i(d)$. Calculation of spatial statistics is based on giving weight to the distances between items of interest. The weighted K-function uses a distance matrix of all distances among points for analyses of the spatial distribution patterns of values (infection levels) among all locations (houses). The analysis is conducted for rectangular areas and accounts for the size of the study area, number of points (e.g., houses), distance between points, and the weight value of each point (e.g., infection intensity). The observed spatial pattern of values is compared with a confidence interval created through random allocation of observed values to all locations for a specified number of Monte Carlo iterations, which determines the $P$ value being tested. The $G_i(d)$ and $G_i^*(d)$ local spatial statistics identify local clustering or hot spots by comparing a given point’s value to all other values within specified distances, including or not including the point under consideration, respectively. To correct for multiple comparisons when using $G_i^*(d)$, significance levels were determined using Table 3 in Ord and Getis. As demonstrated by Kitron and others, some local spatial statistics can also be used as focal statistics when the weight of the point being evaluated is not included such as in calculation of $G_i(d)$. In this study, the local spatial statistic $G_i(d)$ was used as a focal statistic to assess clustering of high infection levels around a particular transmission site, Nimbodze Pond site 11, with high levels of snails shedding $S. haematobium$ cercariae.

**RESULTS**

The overall infection prevalence in Milalani Village was 53.8% with a geometric mean infection intensity of 26.2 eggs per 10 mL of urine; adjacent houses from other villages had similar prevalences (46.0–53.8%; homogeneity $\chi^2 = 4.23, P > 0.2$) and intensities (13.6–24.6 eggs per 10 mL of urine; Kruskal-Wallis $\chi^2 = 4.51; P > 0.2$). Thus, village affiliation was not associated with prevalence or intensity of infection. When houses from all villages were considered, prevalence was 52.7% and intensity was 23.5 eggs per 10 mL of urine. When only houses east of Nimbodze Pond were considered, results were similar, but with higher significance levels. This was also the case for the rest of the analyses, and to be statistically conservative, we only report the results for the large rectangular area. Infection prevalence (50.6% and 54.4%) and intensity (27.6 and 22.4 eggs/10 mL of urine) were similar for males and females, (homogeneity $\chi^2 = 1.88, P > 0.1$ and U = 57,991, $P > 0.1$), but varied significantly by age group (homogeneity $\chi^2 = 322, P < 0.001$; Kruskal-Wallis $\chi^2 = 103, P < 0.001$) (Figure 2). Infection levels were highest in individuals 6–21 years old, with infection prevalence peaking in those 10–17 years old, and intensity in those 10–13 years old. Difference in prevalence between the sexes was only significant for those more than 21 years old (homogeneity $\chi^2 = 5.63, P < 0.02$), while intensity was not significantly different between males and females for any age group.

Weighted K-function revealed clustering of household infection density with peak clustering around distances of 300 meters. When the effect of age was considered, only 6–9-year-old children exhibited global clustering. Because the household prevalence was consistently very high, analysis of local spatial clustering was performed only on infection intensity. Significant local clustering was identified only east of Nimbodze Pond. This clustering held when school age individuals were separated for analysis by sex; however, it was most prominent for those 6–13 years old regardless of sex ($G_i^*[d] > 3.71, P < 0.05$) (Figures 3 and 4), potentially indicating an increased risk of water contact at cercaria-infected water sources for 6–13-year-old children residing east of Nimbodze Pond compared with similar children residing west of Nimbodze Pond.

When focal clustering was examined, neither overall household prevalence nor average infection intensity were significantly clustered. However, mean density was significantly clustered at distances of 400–800 meters around Nimbodze Pond ($G_i[d] > 1.96, P < 0.05$). When comparing infection intensity for age groups of children (Figure 5), significant clustering was detected starting at different distances from the high-risk water source (Nimbodze Pond). Specifically, clustering of infection levels in 6–9 year-old children was significant starting at 250 meters, whereas clustering for those 10–13 years old was not significant below 500 meters; in all cases, clustering was significant up to a distance of 800 meters with a peak at 500–600 meters. The patterns of clustering seen in Figures 4 and 5 suggest a relationship between age, household distance to Nimbodze Pond, and clustering of high infection. Although overall infection prevalence for 6–21-year-old females and males were practically identical (70% and 69%, respectively), focal clustering differed by sex. For 6–9 year-old girls, clustering was significant beyond 500 meters, while elevated infection levels for boys were detected within 250 meters of Nimbodze Pond.

**DISCUSSION**

In this study, global and local spatial statistics were applied to detect household clustering of $S. haematobium$ infection by prevalence, intensity, and density. The use of the Ikonos image allowed us to generate the complete matrices of all distances among households and between households and water contact sites; such distance matrices are needed for the cal-

![Figure 2. Prevalence of infection (open bars = females, hatched bars = males) and geometric mean intensity of eggs per 10 mL of urine (▲ = females, ■ = males) in Milalani Village and surrounding houses.](image-url)
calculation of spatial statistics. Global spatial statistics can reveal a general pattern of household-level infection clustering resulting from the effect(s) of contaminated and uncontaminated water sources. It is possible that a global statistic could reveal the spatial effects of different types of water sources (e.g., ponds, boreholes), which may increase or decrease clustering at varying distance ranges. While global statistics do not detect specific clusters, local spatial clustering statistics permit identification of specific cluster locations for high or low levels of infection.

The observed differences in clustering patterns by age and sex likely reflect variations in exposure to cercariae-contaminated water between males and females, along with changes in behavior as people age, and the effects of acquired immunity. Given that susceptibility to infection with *S. haematobium* in this highly endemic population has a limited heritable component, differences in water contact behavior and acquired immunity are the probable underlying biologic determinants of infection clustering.

In contrast with both global and local statistics, focal statistics require *a priori* knowledge, in this case probable locations of transmission foci. Focal cluster statistics allow evaluation of clustering at the household-level surrounding a particular water source, but they do not allow identification of the specific households that contribute most to clustering. The results of our analyses confirm that high infection levels of urinary schistosomiasis have a significant focal distribution around a known transmission site. Of particular interest is the focal clustering pattern of infection intensity detected for 6–13 year-old children versus younger and older children, a pattern that is in agreement with changes in water contact behavior and immunity by age. This pattern may reflect that children less than six years old who live close to Nimbozde Pond have more contacts with infected water and develop immunity earlier, while children who live farther away do not have considerable water contact with infected water until they are older.

Infection levels in Msambweni were similar for females and males. While this similarity has been observed by other studies, infection is often male-biased. Observed differences in infection between males and females are most likely a result of differences in agricultural and religious practices (e.g., division of labor). Collection of water contact observation data for Msambweni is in progress, and water contact behavior will be analyzed for comparison between male and female exposures to contaminated water. Our study is the first to report the interaction with age for prevalence or intensity differences between the sexes.

In conjunction with the focal statistic, a local spatial statistic allowed us to identify that the households situated east of Nimbozde Pond contributed most to the focal clustering. Overall, focal and local spatial effects have a significant association on age-and sex-specific infection levels for *S. haematobium* infection in our study area. The introduction of alternative water sources (e.g., boreholes) and the implementation of mass community chemotherapy have failed to halt the continuing cycle of urinary schistosomiasis transmission. By identifying transmission epicenters and understanding...
spatial patterns of human infection, it may be possible to develop more effective, highly-focal snail control in conjunction with targeted chemotherapy. For future studies, to better understand the dynamics of this spatial heterogeneity, we propose using a hierarchical multi-scale approach that considers processes functioning on a range of spatial scales from snail microhabitats to whole water bodies and/or entire watersheds, while also considering human water use patterns and migration patterns that result in coarse-scale movement between water bodies or watersheds.

Received October 24, 2003. Accepted for publication December 14, 2003.

Acknowledgments: We thank the residents of Milalani, Mabatani, Nganja, and Marigiza villages for their gracious participation. We are especially thankful to Anthony Chome and Iddi Masemo as well as Malick Ndzovu, Jackson Muinde, and Joyce Bongo in helping locate households. Also, we are thankful for help from Grace Mathenge in coordinating field operations and data entry. Drs. Melinda Brady and Evelin Grijalva provided helpful comments. This paper was published with the kind permission of the Director of Medical Services, Ministry of Health, Kenya.

Financial support: This research was supported by grants from the National Institutes of Health under grants AI-45473 (National Institute of Allergy and Infectious Diseases) and TW/ES01543 (Fogarty International Center).

Authors’ addresses: Julie A. Clennon and Uriel Kitron, Division of Epidemiology and Preventive Medicine, Department of Veterinary Pathobiology, University of Illinois, 2001 South Lincoln Avenue, Urbana, IL 61802, Telephone: 217-244-5954, Fax: 217-244-7421, E-mail: jaclenno@uiuc.edu. Charles H. King, Center for Global Health and Diseases, Wolstein Research Building, Room 4126, Case Western Reserve University School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106-7286, Telephone: 216-368-4818, Fax: 216-368-4825, E-mail: chk@po.cwru.edu. Eric M. Muchiri and H. Curtis Kariuki, Division of Vector Borne Diseases, Ministry of Health, PO Box 20750, Nairobi, Kenya, Telephone: 254-20-725-833, Fax: 254-20-720-030,
REFERENCES


