Residential Setting as a Risk Factor for Lyme Disease in a Hyperendemic Region


The hypothesis that residence in a uniform medium-density residential development is associated with lower incidence of Lyme disease is tested with data from a rural, 12-town region of south-central Connecticut where the disease is hyperendemic. The residential setting for 424 cases identified by active surveillance from 1993 through 1995 was determined. Cases located within the Eastern Coastal ecologic region, where tick densities are known to be higher than inland and where most of the population in the region resides, were selected for further analysis. Within this region, residence in a homogeneous area of medium-density development at least 30 acres (12 ha) in size was associated with a two- to 10-fold lower level of risk than residence in surrounding less developed areas, depending on the estimate of residential population. Type of residential development may be an important factor to consider, in addition to other environmental variables, in studies of peridomestic vector-borne disease in human populations.

MATERIALS AND METHODS

Study area

The study area is approximately 325 square miles (842 km²) (figure 1), has a population of 83,614 (17), and consists of two ecologic regions: the Eastern Coastal Hardwoods and the Southeast Hills Central Hardwoods (18). The Eastern Coastal Hardwoods region (figure 2) is a seaboard region lying within 5–7 miles (8.1–11.3 km) of eastern Long Island Sound and extending up the Connecticut River Valley. Elevations range from sea level to 300–400 feet (92–122 m). The...
mean annual temperature of this ecoregion is the highest in the state (51°F (11°C)) and the average length of the frost-free period (195 days) is the longest in the state. The major forest vegetation is central hardwoods, particularly white, red, and black oak. The Southeast Hills ecoregion lies within 30 miles (48 km) of eastern Long Island Sound. Elevations range from 150 to 500 feet (46–153 m). The major forest vegetation is hardwoods-hemlock. The mean annual temperature is about 2°F (1.1°C) colder than in the Eastern Coastal ecoregion, and the frost-free season is about a month shorter.

Surveillance and geocoding

The Connecticut Department of Public Health began active surveillance for Lyme disease in the 12-town region in 1992. Disease reports meeting the national surveillance case definition for Lyme disease were counted as cases (19). Cases were geocoded based on the reported residential street address using address-matching procedures. More than 84 percent of cases identified by active surveillance from 1993 through 1995 (424 of 503) were successfully geocoded. Cases that could not be geocoded because the residential address was incomplete or given as a post office box were dropped from subsequent analyses.

The Geographic Information System (GIS) software package Transcad (20) was used to compile the street network and case databases for the project and to perform the address-match geocoding. Address-matching against a database of address-ranged street segments yielded a latitude/longitude for each case (21). These coordinates were converted to the Connecticut Coordinate System, a system of state plane coordinates, using the GIS software package pcARC/INFO (22). This made it possible to integrate the case database with other digital spatial databases maintained by the Connecticut Department of Environmental Protection.

Distribution of Lyme disease by ecologic region

A GIS coverage of the ecoregions was digitized from a 1:250,000 map obtained from the Connecticut Department of Environmental Protection (23). The boundary between the two regions is represented as a transition zone with an arbitrary 1,000-foot (305 m) buffer of the boundary shown on the map. The GIS coverage of Lyme disease case locations was overlaid with the ecoregion coverage so that the number of cases within each ecoregion could be ascertained. The 1990 population of each ecoregion was estimated by areal interpolation from 1990 Census block groups using a method that assigns population within a block group to the area within a 100-foot (30.5 m) buffer on each side of the street network in the block group (24).

Distribution of Lyme disease cases in the Eastern Coastal Region by residential setting

The structure of land use in the region was modeled in a 1:24,000 land use/land cover GIS coverage in ARC/INFO format obtained from the Natural Resources Center of the Connecticut Department of Environmental Protection (25). This data layer is a representation of LANDSAT Thematic Mapper Satellite Imagery of the state from 1987, 1988, and 1989.

Given the town structure of local government in the region, villages could not be identified based on political boundaries because they are not incorporated and not all are considered census-defined places. For this research, a village was defined as a compact region of at least 30 acres (12 ha) uniformly classified as medium-density residential in the Connecticut Department of Environmental Protection land use/land cover GIS coverage. All of these villages are located...
in either the Eastern Coastal Hardwoods ecoregion or the transition zone. The case locations from the 1993, 1994, and 1995 active surveillance periods were overlaid with the village locations, and each case was identified as being strictly within or strictly outside a village.

Two estimates of village population were made. First, the villages were overlaid with a GIS coverage of 1990 Census block groups. The 1990 populations of block groups containing villages were summed. This procedure overestimates village population because the villages occupy only a portion of the whole block group area. Second, village population was estimated by areal interpolation using the areal weighting method (24). The block group population was multiplied by the proportion of the block group area accounted for by the village area. The village area was weighted by a factor of 2 to reflect the finding that village residential lot sizes are half the size of residential lots outside the village. This estimate can be considered a lower bound on village population.

RESULTS

Land use structure

The study area is predominantly rural and has no major industrial areas. The dominant land use/land cover is deciduous forest (72 percent). Other important categories include cleared areas of soil, grass, or pasture (5 percent), the deep water area of the Connecticut River and estuary (4 percent), and medium-density residential (5 percent). High-density residential/commercial land accounts for less than 0.5 percent of the total.

Residential development occurs mainly at medium or very low densities. Approximately 85 percent of the 39,226 housing units in the study area are single-family detached structures (not including mobile homes), and only 4 percent of the housing units are in structures with five or more units (17). About 14 percent of the total units are seasonal/recreational; most of these are found in the towns located on Long Island Sound. These units, however, are generally occupied during the peak risk periods for Lyme disease in May and June and in early autumn.

There are two types of medium-density residential development in the study area. In the historic town centers and villages, the natural landscape has been entirely modified and structures are clustered together on small lots in developments of medium density (two to four structures per acre (0.4 ha)), as indicated on property maps used by town assessors. Based on the GIS land use analysis, 35 villages were identified in the study area (figure 2). The area within villages

FIGURE 2. Lyme disease cases and population concentrations by major ecologic region in the southern Connecticut study area. Top, Lyme disease cases from 1993 through 1995 by major ecologic region. Bottom, village population centers by major ecologic region.
equaled 27 percent of the medium-/high-density residential area in the study area as a whole.

Newer medium-density developments are less compact and houses occupy 1–2 acre (0.4–0.8 ha) lots with more woodlands interspersed. The rest of the housing units in the region have been built on lots ranging in size from 2 to 5 acres (0.8–2.0 ha) or more and residential density is low. These housing units are regularly spaced along the road network of the region in woodland or cleared areas.

**Distribution of cases by geographic setting**

The distribution of cases and population by ecoregion indicates that most cases occur in the Eastern Coastal region (table 1). The relative risk of Lyme disease in the coastal region is 2.4. The results also indicate that the risk is not uniform within the coastal region.

There is a significant relative risk associated with living outside a village area. The risk ranges from 1.9 to 10.4 depending on the estimate of village population. GIS aided in the visualization of this relation. At the individual town level, very few cases appear in the villages (figure 3); this effect was seen for all villages, those on the coast and those inland.

To investigate the size of medium-density residential development that might be required to protect against Lyme disease, the number of cases observed in medium-density residential areas was plotted against the size of the homogeneous residential development (figure 4). The number of cases drops as an area of medium-density development reaches around 8 acres (3.2 ha) in size.

To the extent that some cases identified by active surveillance cannot be geocoded, the analysis of environmental conditions associated with high numbers of Lyme disease cases may be biased. Although these cases cannot be included in the analysis, some analysis of their locations can be made based on the limited locational information available. Of the 79 cases observed during the 1993 through 1995 surveillance years that could not be geocoded, 13 were reported for towns that contained no villages, and these cases can be assumed, by definition, not to have occurred in villages.

Of the remaining 66 cases located in towns containing villages, no information that could be used to determine whether or not a case was in a village was available for 39 cases. No address information at all was provided for 17 of these cases. The reported street name could not be found in the street network database for 13 cases, and only a post office box was reported for nine cases.

For the remaining 27 cases, no street number was reported, but the street name could be found in the street network database. These cases were evaluated to determine whether or not the entire street was outside a village; 23 of the 27 cases were found to be on streets that were completely outside villages. Based on this finding, the cases would be outside villages, magnifying the effect observed in this study.

**DISCUSSION**

GIS analysis in the investigation of Lyme disease has been useful for analyzing the local environmental conditions where people live, and, within those areas, where cases have and have not been observed. A number of studies have used ticks harvested from deer carcasses and hunter-reported location of kill to generate maps of tick and infected tick distribution (6, 26, 475).
An obvious problem with this approach is that the derived distribution of ticks is affected by where deer are hunted and killed, the fact that most hunting occurs in the fall, and the accuracy of the reported kill location.

Other studies have mapped the environmental variables believed to be associated with high tick density to identify the tick collection sites. In New Jersey, ticks were collected at 22 public school or recreational areas in one township to evaluate a risk index based on a GIS assessment of habitat suitability for ticks (28). In Rhode Island, a more systematic sampling site identification procedure used GIS analysis to identify forested areas and to compare the distribution of risk to the distribution of reported Lyme disease cases (29). These studies did not explicitly consider the characteristics of residential areas where people might be coming into contact with infected ticks. The relations between 127 environmental variables (not including empirical data on tick density and infection) and Lyme disease incidence were explored in Maryland (10). Four environmental variables, including location in one of two watersheds, on loamy soils, and in forested areas, were associated with increased risk. Residence in "highly developed" areas was associated with a threefold decrease in risk, but an operational definition of high development was not given.

Regardless of whether new residential development takes place in local areas where tick density is observed or modeled to be high or low, exurban residential development in the ecologic regions where the disease cycle functions will itself create new habitat for the mammals involved, including new sources of food supply in the herbaceous shrubs, lawns, and other landscape modifications associated with such development in the northeastern United States. That this type of residential development is currently a preferred form of development in the region is clear. While it is unlikely that land use patterns can be altered in the short term, the increased cost of higher relative risk for Lyme and other vector-borne diseases may now be added to the already noted environmental and public service provision costs of this kind of development.
How people live in environments should not be a neglected factor in the revived interest in vector-borne disease and the GIS applications to study it, which to date have emphasized the physical environmental attributes of regions. Aside from the implications for intervention, the distribution of population across various residential settings has important implications for spatial sampling and the design of case-control studies. In the short-term, GIS can be used to identify residential areas at higher relative risk for Lyme disease, and tick control methods that have been shown to be effective in residential areas (including timed application of acaricides and yard debris removal) can be implemented in those places (30–32). In the long-term, GIS can be used in land use planning to design clustered residential developments and redevelopments that more effectively separate human populations from potential vectors of disease.

ACKNOWLEDGMENTS

The authors thank Pat Mshar for her surveillance efforts, Kevin P. Joy for his work on the 1992 and 1993 surveillance data, and Drs. David Dennis, James Hadler, and Kirby Stafford for their helpful comments on the manuscript.

This research was supported by the Centers for Disease Control and Prevention cooperative agreement no. U50/CCU106598.

REFERENCES

25. GIS data catalog. Hartford, CT: Natural Resources Center, Department of Environmental Protection, 1996.