Environmental Risk Factors for the Incidence of American Cutaneous Leishmaniasis in a Sub-Andean Zone of Colombia (Chaparral, Tolima)

Carlos Valderrama-Ardila, Neal Alexander, Cristina Ferro, Horacio Cadena, Dairo Marín, Theodore R. Holford, Leonard E. Munstermann, and Clara B. Ocampo*

Biology Department, Universidad Icesi, Cali, Colombia; Centro Internacional de Entrenamiento e Investigaciones Médicas, Cali, Colombia; Laboratorio de Entomología Médica, Instituto Nacional de Salud, Bogotá, Colombia; Biology Department, Universidad del Tolima, Tolima, Colombia; Department of Public Health, Yale University, New Haven, Connecticut

Abstract. Environmental risk factors for cutaneous leishmaniasis were investigated for the largest outbreak recorded in Colombia. The outbreak began in 2003 in Chaparral, and in the following five years produced 2,313 cases in a population of 56,228. Candidate predictor variables were land use, elevation, and climatic variables such as mean temperature and precipitation. Spatial analysis showed that incidence of cutaneous leishmaniasis was higher in townships with mean temperatures in the middle of the county’s range. Incidence was independently associated with higher coverage with forest or shrubs (2.6% greater for each additional percent coverage, 95% credible interval [CI] = 0.5–4.9%), and lower population density (22% lower for each additional 100 persons/km²; 95% CI = 7–41%). The extent of forest or shrub coverage did not show major changes over time. These findings confirmed the roles of climate and land use in leishmaniasis transmission. However, environmental variables were not sufficient to explain the spatial variation in incidence.

INTRODUCTION

Advances in the ecology of vector-borne diseases are proving insightful for the study of disease transmission and its relation to environmental conditions. Techniques such as remote sensing and geographic information systems enable spatial and temporal analysis of variables such as land coverage, climate, and human settlement, all of which can influence risk of disease transmission. These methods facilitate disease surveillance and the development of models to predict future trends and possible outbreaks. For vector-borne diseases, important candidate predictors include forest fragmentation; climatic variables; including those associated with the El Niño Southern Oscillation; vegetation indices such as the normalized difference vegetation index; and land surface temperature.

Leishmaniasis provides a promising subject for spatial analysis, given the ecology of its phlebotomine sand fly vectors. Risk of American cutaneous leishmaniasis has traditionally been associated with working in or near forest. Changes in land coverage, primarily by deforestation, and shifts to other types of cultivation, have increasingly shifted the risk to rural settlements. A posteriori risk modeling for visceral leishmaniasis in Africa and Asia suggests that rainfall, altitude, and cultivation of certain plant species can predict transmission rates. Other recent studies suggest that other landscape variables (in addition to vegetation indices) are implicated, in particular land use category.

In Colombia, as in other Latin American countries, cutaneous leishmaniasis transmission has been shifting from sylvatic areas to domestic and peri-domestic habitats. Recorded cases in Colombia have increased from an average of 6,500 cases per year in the 1990s to 18,098 in 2005 and 16,098 in 2006. Cutaneous leishmaniasis comprises 95% of reported leishmaniasis cases. Colombian transmission foci are distributed widely in a variety of climates and habitats, although the Andean region has presented the highest number of cases.

To date, 147 Lutzomyia species have been recorded in Colombia, several of which (approximately 15%) have been implicated as potential leishmaniasis vectors. However, only six species have so far been confirmed as vectors, according to the criteria of Lewis and Ward and Killick-Kendrick. The parasite species most frequently associated with cutaneous leishmaniasis has been Leishmania (Viannia) panamensis, although L. (V.) brasiliensis has a broader geographic distribution. Leishmania (V.) guyanensis and the L. mexicana species complex are also found in several regions of Colombia.

In 2003, a major outbreak of cutaneous leishmaniasis, the largest recorded in Colombia in terms of incidence, occurred in the county (municipio) of Chaparral. In this outbreak, the high proportions of children ≤15 years of age (36.5% of 2,835 cases) and women (35.5% of 1,811 adults) suggested domestic transmission (Chaparral Hospital and Secretary of health Tolima). A large proportion (95% of 56 isolates) of cases were associated with Leishmania (V.) guyanensis, a species that had not been previously reported outside the Amazon and Orinoco basins of Colombia.

Four anthropophilic sand fly species were found at the time of the outbreak: Lutzomyia longiflora, Helcocyctomyia sp., Lu. colombiana, and Lu. nunezovari. Lutzomyia longiflora was the most abundant species indoors (as determined by CDC trap collections) and was considered the probable vector. Before the Chaparral outbreak, Lu. longiflora was also the most abundant anthropophilic species found in earlier outbreaks reported in mid-elevations (1,000–2,000 meters) of the Andean region near the city of Neiva (Huila Province) in 1993–1996. Planadas (Tolima Province) in 1998, and Abrego (Norte de Santander Province) in 2001–2004. Because of the intensity, apparent rapid onset, and scope of the current Chaparral outbreak, an effort was initiated to evaluate environmental factors, including remotely sensed land coverage and climatic variables, as potential risk factors for the incidence of cutaneous leishmaniasis in this region.

MATERIALS AND METHODS

Site description. The county of Chaparral is located in the Province (Departamento) of Tolima (3°43'20"N, 75°30'27"W),

* Address correspondence to Clara B. Ocampo, Centro Internacional de Entrenamiento e Investigaciones Médicas, Carrera 125 #19–225, Cali, Colombia. E-mail: claraocampo@cideim.org.co
in the inter-Andean valley of the Magdalaena River (Figure 1). Elevation in this county ranges from 372 to 4,146 meters. Its ecosystems are diverse, including tropical dry forest, mid-elevation sub-Andean moist forest, high elevation Andean mountain cloud forest and tropical high elevation grassland (paramo). The area has been occupied since pre-Hispanic times and more recently was cultivated for coffee and other crops. This cultivation has caused profound transformation of the ecosystems, particularly at lower elevations. The main agroecosystems of the area are rice in the flat valley bottoms, pasture and mixed cultivation on the lower slopes, and mostly coffee and corn at higher elevations. Above 2,000 meters, the level of transformation decreases and mixed cultivation occurs mostly along riverbanks.

**Analysis of land coverage and environmental layers.** Land use was evaluated using satellite images from LandSat TM (August 7, 1989), LandSat ETM+ (October 14, 2002), and ASTER (February 23, 2007). The sources of LandSat and ASTER images were, respectively, the Global Land Cover Facility (www.landcover.org) and the EROS Data Center (edc.usgs.gov). Coverage was analyzed with ERDAS Imagine 9.1 software (ERDAS, Atlanta, GA) by using supervised classification. The following coverage categories were chosen: forest (including primary and secondary forest), paramo, shrubs, cultivation, grassland, water bodies, exposed soil, and urban areas at a scale of 1:50,000 and 30-meters pixel size. The selected classification system followed exploratory level 1 of the Centro de Investigación en Percepción Remota. The training polygons were obtained based on visual interpretation of textures and color in different band combinations of images according to previous field experience of the classifier as suggested by Wilkie and Finn. Coverages were defined as a) forest: mature forest, stratified old secondary forest and mature planted forest; b) paramo: natural grasslands of wet soils, usually at elevations above 3,200 meters; c) shrubs: early secondary growth forest, including permanent cultivations such as coffee or fruit trees; d) cultivation: annual cultivations such as rice and corn; e) grasslands: pastures at elevations.

**Figure 1.** Elevation map of Chaparral County and location in Colombia. Chaparral is located in the western side of the Magdalaena River Valley in Tolima Province. Elevation within the county ranges from 372 meters to 4,146 meters.
Incidence. Numbers of cases per year were obtained at the level of rural township (vereda) from the Chaparral Hospital for a period of five years (2003–2007). For the denominator, population data were obtained from the national Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales located in Chaparral. The total population of the county was 56,228.

Statistical analysis. Spatial analysis was done at the rural township level using a conditional autoregressive Poisson model to represent spatial correlation. The response variable was the total number of cases over a five-year period, with logarithmic link function and log-population as an offset. The analysis therefore yielded incidence rate ratios. One township (Parque las Hermosas) was excluded because it is a nature reserve with no recorded human population at risk. A total of 2,313 cases in 142 townships of Chaparral County were included in the analysis. The overall incidence was therefore 0.82 cases per hundred person-years, with a 95% confidence interval, obtained by bootstrapping over townships, of 0.35–1.92/100 person-years. Analysis was done in the Bayesian framework in which prior beliefs about model parameters are specified by statistical distributions.

This analysis used uninformative (vague) prior distributions with high variances. For the regression slopes (log rate ratios), normal prior distributions were used with a mean of zero and variance of $10^3$. For the intercept, an improper uniform (flat) distribution was used as the prior, but, for sensitivity analysis, a normal distribution was used, with mean zero and variance $10^3$. For the precision (reciprocal variance) of the spatial random effects, a Gamma distribution with mean of $10^3$ and variance of $2 \times 10^6$ was used as the prior. Sensitivity analysis for this parameter used a Gamma prior distribution with a mean of 1 and a variance of 100. Models were fitted using Markov chain Monte Carlo techniques in the GeoBUGS component of the WinBUGS v1.4 software (Medical Research Council, Cambridge, United Kingdom). To reduce autocorrelation in the sequences of parameter values obtained, each explanatory variable was centered by subtracting its mean. Each model was fitted by an initial burn-in of at least 50,000 iterations, with parameters estimated from at least 100,000 iterations. The WinBUGS code developed herein is available online as listed in Supplemental Appendix 1 (available at www.atmhm.org). Convergence was assessed by the both the Geweke and the Heidelberger and Welch methods in the Bayesian Output Analysis library of the S-PLUS software (Insightful Corporation, Seattle, WA). Convergence was further assessed by imposing alternative initial values for parameters, and then using the Gelman-Rubin diagnostics in WinBUGS.

All explanatory (predictor) variables were averaged to provide one value per township. For example, the percent coverage of each township for each land use category was calculated. The forest and shrub categories identified from the supervised classification were joined because, separately, they showed implausibly large variation over years, whereas the combination was stable (see Discussion). Among the 19 Bioclim layers, model development concentrated on the following seven layers: mean temperature; mean temperature in wettest, driest, and warmest quarters; mean precipitation; and mean precipitation in wettest and driest quarters (layer numbers 1, 8, 9, 10, 12, 16, and 17, respectively). In later stages of analysis, the remaining 12 layers were considered (numbers 2–7, 11, 13–15, and 18–19).

RESULTS

Analysis of land coverage over time. Considered together, the coverage of forest and shrubs remained almost constant during 1989–2007 (Figure 2 and Table 1), the average over townships being 46.6% (SD = 3.1%) of the land area over an 18-year period. Areas dedicated to grasslands (21.4%, SD = 4.4%) tended to increase with time ($P < 0.001$, by chi-square test).

Spatial analysis of environmental variables. Incident cases of leishmaniasis were concentrated in a band of medium altitude (1,000–2,000 meters) are shown in Figures 3 and 4. This finding is reflected by a quadratic (inverse U-shaped) relation between cumulative incidence (over the five years) and mean temperature (Bioclim layer 1) that correlates with altitude. The peak incidence is associated with a mean temperature of 20.6°C (95% CI = 19.2–22.0°C; Table 2). Incidence was also positively associated with coverage by the combined forest-shrub category, increasing by an estimated 2.6% for each additional percent coverage (95% CI = 0.5–4.9%). Incidence was also higher in areas with lower population density, with each additional 100 persons/km² being associated with a reduction in incidence of 22% (95% CI = 7–41%). Neither average age nor sex ratio showed an association with incidence. Moreover, including these parameters in the model made little difference to the incidence rate ratios shown in Table 2. Thus, age or sex showed little confounding with the parameters shown in Table 2. None of the other Bioclim layers among those initially considered (numbers 8, 9, 10, 12, 16, and 17) showed an independent association with incidence. For some of these additional variables, e.g., mean temperature of wettest quarter (layer 8), the regression model did not efficiently separate their effect from that of mean temperature (slow mixing). This finding suggested that although temperature was a predictor of incidence, the mechanism was not necessarily causal. For example, although sand fly emergence may be associated with temperature, this relationship was not necessarily dependent on temperature, but may have been causally related to other environmental variables associated with temperature.

To assess the extent to which the models explain the spatial variation in incidence, the residual (i.e., observed/expected) rate ratio was mapped by township. For a model with no covariates, Figure 5A showed that, as expected, this result was similar to the total incidence per township (Figure 4), but was smoothed to some extent. The visual impression of spatial clustering (autocorrelation) was confirmed because the SD of the
autoregressive distribution for the spatial effects was greater than 0 (3.02 on the log-rate scale, 95% CI = 2.54–3.66). If the model with covariates (Table 2) had explained all the spatial variation, then the corresponding plot would have shown random variation with no discernible spatial pattern. However, Figure 5B indicated that this is not the case. Rather, townships to the north and west of the study area had incidences, which were higher than explained by the model’s covariates, although these were also estimated less precisely (Figure 5C). Similarly, addition of covariates reduced the SD of the spatial autoregressive distribution, but it remained far from 0 (2.32, 95% CI = 1.90–2.87). Because of this unexplained spatial variation, the remaining seven Bioclim layers (2–7, 11, 13–15, 19, and 19) were analyzed. However, none of these explained more of the variation in risk. In summary, some, but not all, of the spatial variation in incidence was explicable by environmental variables.

DISCUSSION

Risk of cutaneous and visceral leishmaniasis has previously been found to be associated with environmental factors such as land use and temperature, which are thought to act via vector density.9–11,33–36 The spatial resolution of these studies has varied from tens of meters, within one village,9 to hundreds of kilometers, with an entire country considered as a unit.36 For

<table>
<thead>
<tr>
<th>Year</th>
<th>Paramo</th>
<th>Forest</th>
<th>Shrubs</th>
<th>Cultivation</th>
<th>Grassland</th>
<th>Other</th>
<th>No data</th>
<th>Total area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>14,064 (6.7)</td>
<td>29,736 (14.2)</td>
<td>63,175 (30.1)</td>
<td>15,724 (7.5)</td>
<td>34,219 (16.3)</td>
<td>41,485 (19.8)</td>
<td>11,182 (5.3)</td>
<td>209,585 (100)</td>
</tr>
<tr>
<td>2002</td>
<td>19,272 (9.2)</td>
<td>34,023 (16.2)</td>
<td>61,298 (29.2)</td>
<td>6,086 (2.9)</td>
<td>51,582 (24.6)</td>
<td>36,078 (17.2)</td>
<td>1,244 (0.6)</td>
<td>209,585 (100)</td>
</tr>
<tr>
<td>2007</td>
<td>18,990 (9.1)</td>
<td>49,548 (23.6)</td>
<td>55,449 (26.5)</td>
<td>17,316 (8.3)</td>
<td>48,567 (23.2)</td>
<td>14,110 (6.7)</td>
<td>5,604 (2.7)</td>
<td>209,585 (100)</td>
</tr>
</tbody>
</table>

*Areas are in square kilometers. Values in parentheses are percentages of coverages.
public health purposes, it remains unclear what degree of resolution is required for useful prediction. In Colombia, at the level of county, King and others found that the optimal prediction algorithm differed between ecologic zones. The present study used a finer scale and found substantial variation in incidence of cutaneous leishmaniasis within a single county. This risk was associated with environmental factors; in particular, temperature and forest coverage.

Short-term temporal climate variation related to the El Niño phenomenon, was not considered in the current analysis. These variations have been associated with increased leishmaniasis transmission in northern and southern Colombia. However, the occurrence of this phenomenon, from March–May 2002 to February–April 2003, May–July 2004 to January–March 2005, and July–September 2006 to December–February 2007–2008, according to the changes in the Oceanic Niño Index from the National Oceanographic and Atmospheric Administration Climate Prediction Center (http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml) did not appear to correspond with the annual variation in incidence in Chaparral (Figure 4). Changes caused by El Niño probably have been secondary in magnitude to changes linked to the large altitudinal range of the county.

The classification of the remote sensing data initially showed implausibly large changes over time in the forest and shrub categories. Because deciduous tropical forest species have a similar spectral reading during dry season as shrub species, this finding was probably a misclassification related to seasonal differences in image acquisition times and differences between the LandSat and ASTER sensors. When the forest and shrub categories were combined, the classification was largely constant over time. Other significant changes in land use were not detected in particular, from forest to agricultural use or pasture.

The incidence of leishmaniasis during the outbreak was high. In 2003 and 2004, more than 20% of the populations of several townships became infected (Figure 4). The risk was highest at middling elevations (approximately 1,000–2,000 meters), with the maximum occurring at a mean temperature of 20.6°C. Higher incidence was also independently associated with lower population density and higher coverage by woodland and shrubs. These observations were consistent with other
studies carried out in Tolima and in the neighboring province of Huila, in which the most abundant vector candidate was *Lu. longiflorcosa*. This species was found largely in the dry season between 1,300 and 1,700 meters in forest and coffee plantations with shade trees.\(^2\) This pattern was also consistent with vector research linked with the current study.

Risk of cutaneous or visceral leishmaniasis has previously been associated with proximity to woodland, elevation, and temperature.\(^9\)–\(^11\), 33, 35, 36 \(^3\) The inverse association between risk and human population density may result from the sparse settlement being sparse near coffee plantations, a location that provides a habitat more favorable for vectors than other types of forest-shrub coverage. King and others\(^10\) included all of Colombia in their analysis at the county level and found that land use predicted leishmaniasis risk. Nevertheless, they estimated the probability of transmission in Chaparral to be low (<0.2). This finding proved to be a false-negative result, given the large outbreak that occurred shortly afterwards. The current study suggested that analysis with county as the unit may be too coarse-grained for effective prediction. Moreover, the Chaparral outbreak was not associated with a sudden change in environment, which indicated that such factors do not completely control leishmaniasis incidence. Within the current spatial regression analysis, this finding can be seen by looking at the component of risk, which remains unexplained in each township. For a complete model, these residual risks are expected to have a random pattern, but Figure 5B shows that systematic variation remains unexplained by environmental variables.

For transmission to reach its maximum potential, suitable environmental conditions are not sufficient. Additional factors such as population movements or cyclic fluctuations in immunity are required. In the Chaparral outbreak, the dominant parasite species was *Leishmania* (*V.*) *guyanensis*, and its novel occurrence suggested that the origin of the outbreak may have been caused by movement of persons, possibly including armed groups, from the Amazon or Orinoco basin.\(^27\)

In conclusion, spatial analysis shows that environmental factors, in particular temperature and forest coverage, are associated with the risk of cutaneous leishmaniasis. This approach is currently being extended to the national level in work, which uses maximum entropy methods and will include validation of the predictions in selected counties. Nevertheless, the current study demonstrated the limitations of environment-based prediction\(^39\) and is a reminder that effective surveillance must consider additional factors, such as movement of human populations\(^40\) or other cryptic human behaviors.

Received April 27, 2009. Accepted for publication August 17, 2009.

Acknowledgments: We thank the Chaparral Hospital staff for providing the data on human cases and location of patients; the regional environmental authority (CorTolima) for providing the township boundary data; Larry Bonneau (Yale University) and the Yale University Center for Earth Observation for providing ASTER satellite images; and

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Incidence rate ratios from spatial analysis, Colombia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variable</td>
<td>Mean (range) of the township-averaged values</td>
</tr>
<tr>
<td>Population density (hundreds of persons per km(^2))</td>
<td>0.53 (0.003–17)</td>
</tr>
<tr>
<td>Percent forest or shrub</td>
<td>57 (4–87)</td>
</tr>
<tr>
<td>Mean temperature (°C)</td>
<td>22.4 (9.8–27.2)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*These parameter estimates imply that the peak incidence occurs at a mean temperature of 20.6°C (95% confidence interval = 19.2–22.0°C).

†Bayesian equivalent of confidence intervals.
Rupert Quinnell (Leeds University, Leeds, United Kingdom), Diane McMahom Pratt (Yale University), and Nancy G. Saravia (Centro Internacional de Entrenamiento e Investigaciones Médicas) for providing important comments on the paper.

Note: Supplemental Appendix is available at www.ajtmh.org.

Financial support: This study was supported in part by the National Institutes of Health, Division of Microbiology and Infectious Diseases, International Collaboration in Infectious Disease Research Program (grant U19 AI065866).

Authors’ addresses: Carlos Valderrama-Ardila, Department of Biology, Universidad Icesi, Cali, Valle del Cauca, Colombia, E-mail: cvalderrama@icesi.edu.co. Neal Alexander, Horacio Cadena, Dairo Marín, and Clara B. Ocampo, Centro Internacional de Entrenamiento e Investigaciones Médicas, Cali, Colombia, E-mails: nalexander@ciceim.org.co, horacio.cadena@gmail.com, dmarin@cideim.org.co, and claraocampo@cideim.org.co. Cristina Ferro, Laboratorio de Entomología, Instituto Nacional de Salud, Bogotá, Colombia, E-mail: crisferrovela@gmail.co. Theodore R. Holford and Leonard E. Munstermann, School of Medicine, Yale University, New Haven, CT, E-mails: theodore.holford@yale.edu and leonard.munstermann@yale.edu.

REFERENCES


Figure 5. A. Estimates of risk by township, Chaparral County, Colombia, from model with no covariates. This panel is a smoothed version of the overall incidence from the five years (lower right panel of Figure 4). B. Unexplained (residual) spatial variation from the final model (Table 2). Some spatial pattern remains, indicating that the variables in the model do not explain all of the variation. C. Credible intervals for the rate ratios in panel B. Darker colors indicate wider credible intervals, i.e., less accurate estimates. This figure appears in color at www.ajtmh.org.


