Risk maps for Leishmaniasis in Central Tunisia

Yiannis Kamarianakis[1], Poulicos Prastacos[1], Afif Ben Salah[2], Sadok Schlif[2], Nissaf Ben Alaya[2]

[1]Regional Analysis division, Institute of Applied & Computational Mthematics, Foundation for Research and Technology-Hellas, GR 71110, Heraklion Crete, Greece

[2]Institut Pasteur Tunis, Tunisia

SUMMARY

Zoonotic cutaneous leishmaniasis (ZCL) is endemic in many rural areas of Southern and Eastern Mediterranean Region where different transmission patterns of the disease have been described. This study was carried out in a region located in Central Tunisia and aimed to investigate the spatio-temporal dynamics of the disease from 1999 to 2004. First, appropriate scan statistics were employed for the identification of statistically significant spatial, temporal and spatio-temporal clusters. A hierarchical Bayesian Poisson regression model with spatial effects was fitted to signify explanatory socio-geographic factors related to spatial rate variability. Temporal ZCL dynamics were described via a linear mixed model. Our findings of increased incidences in urban areas, support the hypothesis of increased incidences in peri-urban environments due to changes in sandfly/rodent living habits over the last years.

INTRODUCTION

Zoonotic cutaneous leishmaniasis (ZCL) is endemic in many rural areas of Southern and Eastern Mediterranean Region. It is an important health problem, with clinical manifestations varying from simple cutaneous to mucocutaneous and disseminated lesions. *Phlebotomus papatazi* is the proven vector of *L. major*, and rodents *Psammomys obesus* and *Meriones spp.* serve as animal reservoir hosts. There are several reasons behind the increased ZCL incidence in the region. The majority of them depend on human activities such as environment modifications, resettlement of non-immune populations or development of agro-industrial projects, military activities, urbanization and so on. Environmental modification, such as construction of dams, can change the temperature and humidity of the soil and vegetation, which may result in changes in the composition and density of sandfly species as well as changes in populations of rodent species. The formation of new settlements with non-immune populations facilitates the outbreak of leishmaniasis. For example, the outbreak of ZCL in the central and southern governorates of Tunisia in 1982-83 occurred following the construction of the Sidi Saad Dam. Control of reservoir hosts is an important component of the control strategy against zoonotic forms of leishmaniases. Some new approaches to control the transmission of ZCL through environmental modifications have been studied in Jordan and Tunisia with support of the EMRO/CTD/TDR Small Grants project. The destruction of *Psammomys obesus* burrows by deep ploughing, removal of chenopods and planting of trees in a 2-3 kilometer zone surrounding human settlements has resulted in a significant reduction of the incidence of cutaneous leishmaniasis among the local human population.

In this paper we present the results of a continuous selection of ZCL cases through a surveillance system carried out in the Sidi-Bouzid region, Central Tunisia. The aims of this study were to estimate the incidence rates of ZCL from 1999 to 2004, to test for spatio-temporal patterns in the rates and correlate them to socio-geographic factors. For some studies of leishmaniasis in southern America, similar in spirit to ours (although using different techniques) the interested reader may consult Assunção *et al* (2001) and Machado-Coelho *et al.* (1999).

EXPLORATORY ANALYSIS

Sidi-Bouzid is an endemic region for the disease. It is an area of 7.5 thousand km2 located in central Tunisia and lying within the boundaries of 34o30' and 35o29' S, and 8o55' and 10o04' W. The region is subdivided according to the Tunisian Census Bureau into 94 sectors (Figure 1). Its population according to the 2005 census is 395 506.

A total of 15897 ZCL cases were reported in Sidi-Bouzid from 1999 to 2004 with an average crude incidence rate of 669.7 annual cases / 100.000 inhabitants. About 63.4% of the cases were from rural areas (83 sectors), which account for the 71% of the total population (395506 inhabitants). The temporal evolution of ZCL is depicted at the left part of figure 2 and in figure 3; moreover from figure 3 one may observe a tendency of high incidence rates to occur at the eastern part of the region as time evolves. There is an obvious outbreak of the disease in 2004. The chi-square test rejects the temporal homogeneity hypothesis (χ2=1675.8, d.f.=5, P<0.0001) and Kulldorff's scan statistic reports a significant temporal cluster of high rates occurring in 2004 (observed cases are 1.66 times the expected with a log-likelihood ratio of 602 and a p-value less than 0.001). As indicated by the graph at the right part of figure 2 there is a strong association between ZCL cases per sector during the first year of the study and total ZCL cases during the six-year study-period. Thus, a sector with high incidence rates at the beginning of study period is quite likely to display high rates for the whole time period.

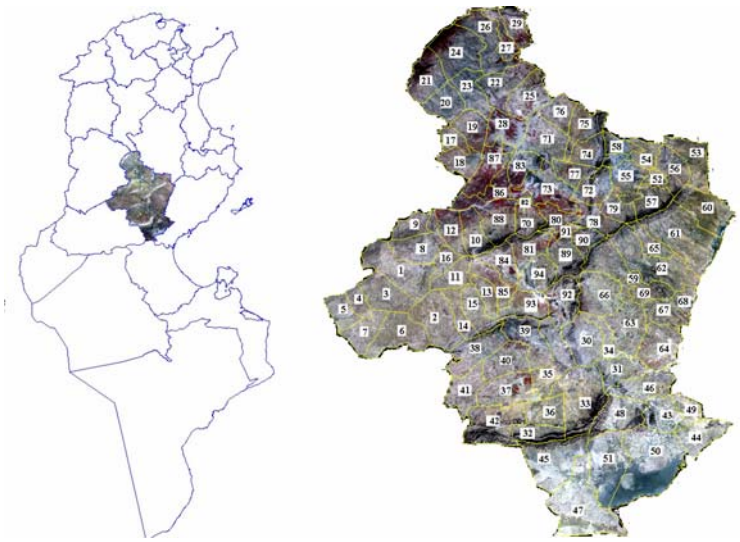

Figure 1 Left: Tunisian regions and the study area (highlighted). ***Right:*** A LandSat image of the study region with sectoral boundaries/Ids superimposed. Red tones correspond to vegetation, white-gray tones correspond to soil, green tones correspond to sparse vegetation and blue tones correspond to urban areas.

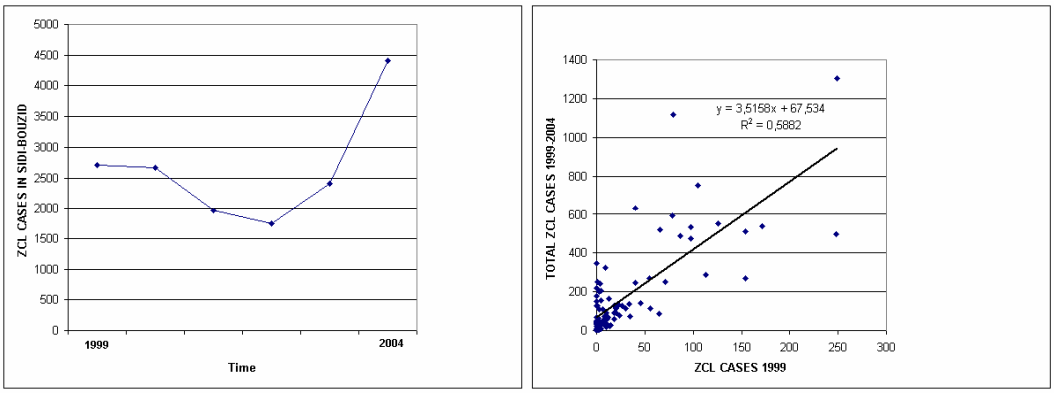

Figure 2 Left: ZCL cases in Sidi-Bouzid from 1999 to 2004. ***Right:*** Linear association between cases observed during the first year of the study and total cases from 1999 to 2004 for the 94 sectors.

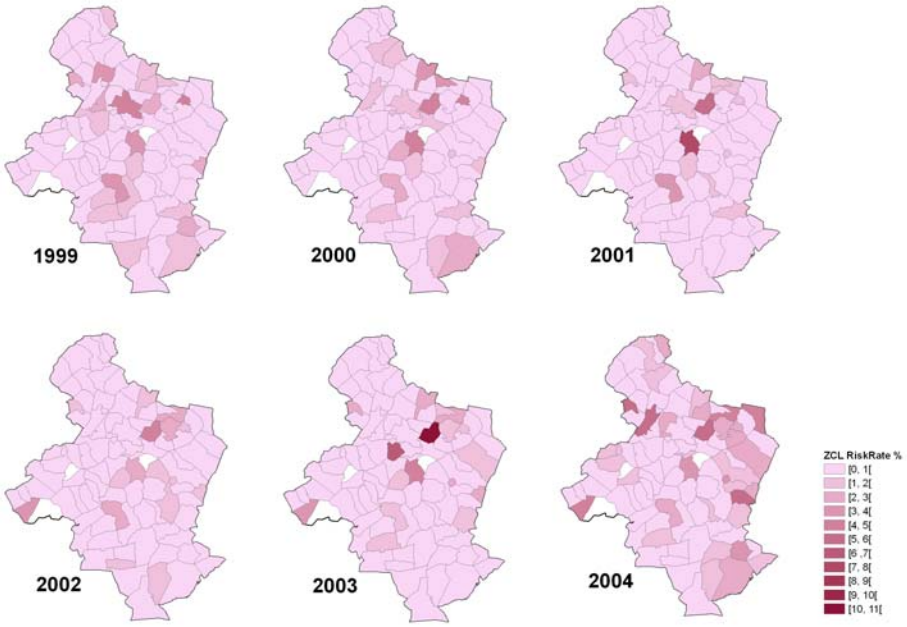

Figure 3: ZCL rates in the sectors Sidi-Bouzid from 1999 to 2004 in terms of percentage of cases in the population of each sector.

IDENTIFICATION OF SPATIAL AND SPATIO-TEMPORAL CLUSTERS

Spatial Clustering

Initially we tested for spatial homogeneity in the aggregate ZCL rates and then proceeded to separate tests for each year. Again the chi-square test was used in order to compare the number of observed ZCL cases per sector with those expected under the hypothesis of constant rate through space. To identify spatial clusters we applied the spatial scan statistic with an underlying Poisson model to the aggregate cases for the six years of the study. That is we assume the number of ZCL cases to be Poisson distributed across sectors and test the null hypothesis that the risk of ZCL is the same in all sectors. The adopted methodology, proposed by Kulldorff (1997, 1999), imposes a circular window on the map and allows the center to move over the area so that at any given position, the window includes different sets of neighboring sectors. For practical reasons, the center of the window is positioned only at the 94 sector centroids; at each position the radius of the circular window is varied continuously from zero up to a maximum radius so that the window never includes more than 50 percent of the total population. In total, the method creates a very large number of distinct circular windows, each with a different set of neighboring sectors within it, and each a possible candidate for containing a cluster of ZCL cases. For each window the method tests the null hypothesis against the alternative hypothesis that there is an elevated (or reduced) risk within, compared with outside, the window. For cluster identification, two criteria were posed: no geographical overlap between clusters and no cluster centering in another cluster. The spatial scan statistic has been used in a large number of epidemiological studies, e.g Chaput *et al.* (2002), Bakker *et al.* (2004), Jennings *et al.* (2005). Application of the spatial scan statistic is highly facilitated by the SatScan software, which is freely available from www.satscan.org.

Figures 4 and 5 present spatial clusters of high/low incidence rates for the sum of ZCL cases for each sector from 1999 to 2004. The underlying probability model for the 94 sectors is Poisson distributed and clusters were derived under two criteria: no geographical overlap between clusters and no cluster centers in another cluster. Both methods suggested a spatial cluster of high incidence rates located at the northeastern part of the region. Observed cases in this cluster are four times more than the expected ones under the hypothesis of spatial homogeneity; the cluster contains seven sectors: Sidi Lefi, Sidi Khlif, El Amra, El Hania, El Makarem, Fayedh, Ouled Haffouz. Apparently, this cluster is neighboring to the Sidi-Saad dam where the first outbreak of the disease was identified in the early eighties; the dam is located outside of the study region, very close to the borders of the spatial cluster. The population of the set of regions that comprise the cluster is 18589; except Ouled Haffouz all other sectors are rural. When we allow for overlapping between clusters (but no cluster centroids in other clusters) three more clusters of high incidence rates, located at the central part of the region, are identified; they contain twenty regions with observed cases being at least two times the expected ones (Cluster 8 in Figure 5). Spatial clusters of low incidence rates are present at the northern, western and eastern parts of the region. When we allow for regions to overlap, practically all sectors of the northwestern part of the study-region belong to at least one statistically significant cluster of low incidence rates.

Spatio-temporal Clustering

Clusters in space and time, for the six-year study period were identified via a scan statistic proposed by Kulldorff *et al.* (2005); the underlying probability model is again Poisson distributed. The space-time scan statistic is defined by a cylindrical window with a circular geographic base and with height corresponding time. The base is centered on one of several possible centroids located throughout the study region, with the radius varying continuously in size. The height reflects any possible time interval of less than or equal to half the total study period, as well as the study period as a whole. The window is then moved in space and time so that for each possible geographic location

and size, it also visits each possible time interval. In effect, we obtain an infinite number of overlapping cylinders of different size and shape, jointly covering the entire study region. Cases are assumed to be Poisson distributed with constant risk over space and time under the null hypothesis, and with different risk inside and outside at least one of the cylinders under the alternative hypothesis. For each cylinder the number of disease cases inside and outside the cylinder are noted, together with the expected number of cases reflecting the population at risk. On the basis of these numbers, the likelihood is calculated for each cylinder. The cylinder with the maximum likelihood is denoted the most likely cluster. For cluster identification, we used the two above-mentioned criteria: no geographical overlapping between clusters and no cluster centroids in other cluster.

Examination of space-time clustering reveals very similar results regarding to the spatial distribution of incidence rates; in this case though, we also have a time frame that denotes the period for which each cluster deviates significantly (figures 6 and 7). For example the secondary cluster of high incidence rates at the center of the region is prominent from 1999 to 2003 whereas for the most likely cluster the corresponding period starts in 2000 and ends in 2004. No evidence of spatial or spatio-temporal clustering exists for the southern/southwestern part of the region.

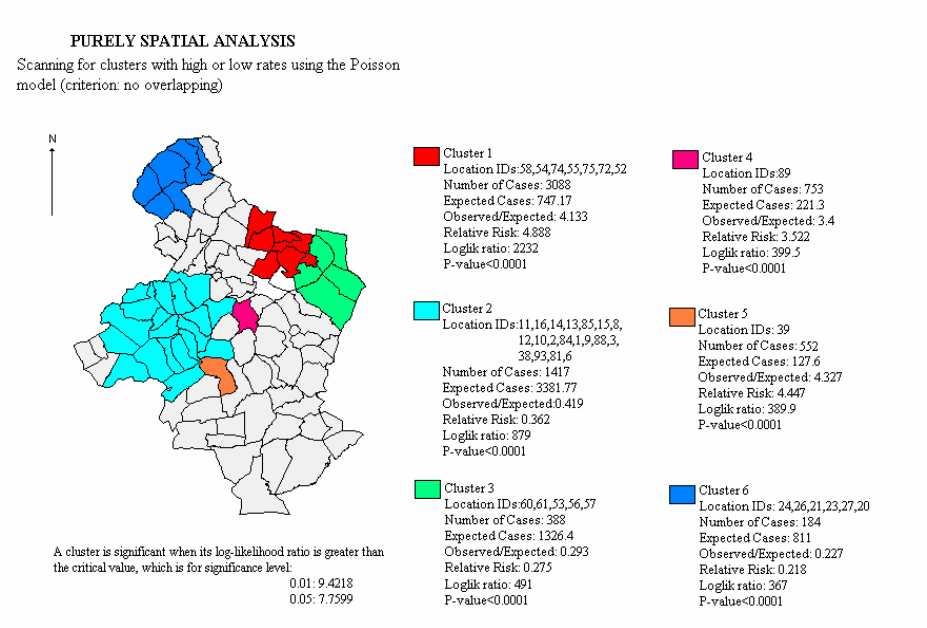

Figure 4: Spatial clusters derived from the spatial scan statistic under the Poisson model and the criterion of no geographical overlapping between clusters.

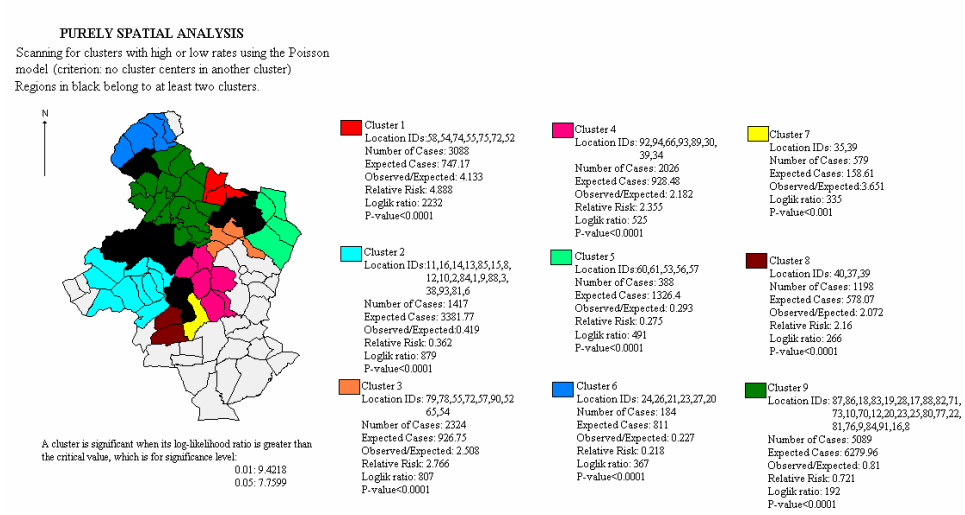

Figure 5: Spatial clusters derived from the spatial scan statistic under the Poisson model and the criterion that no cluster centers in another cluster (regions in black belong in at least two clusters).

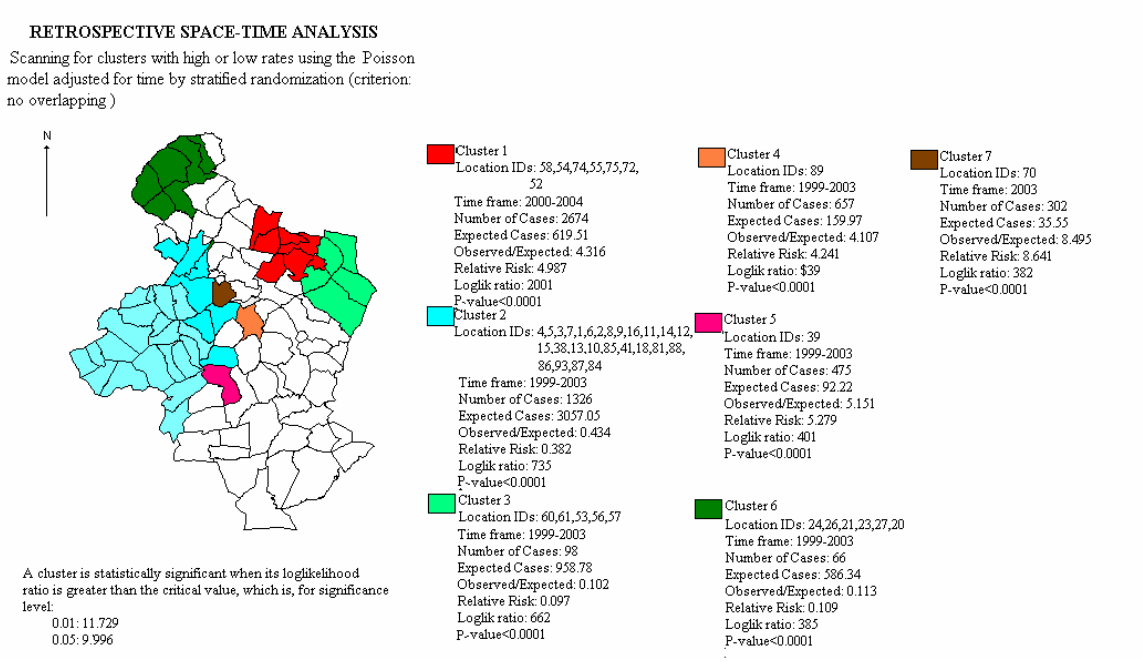

Figure 6: Clusters in space and time derived from the scan statistic under the Poisson model and the criterion of no geographical overlapping between clusters.

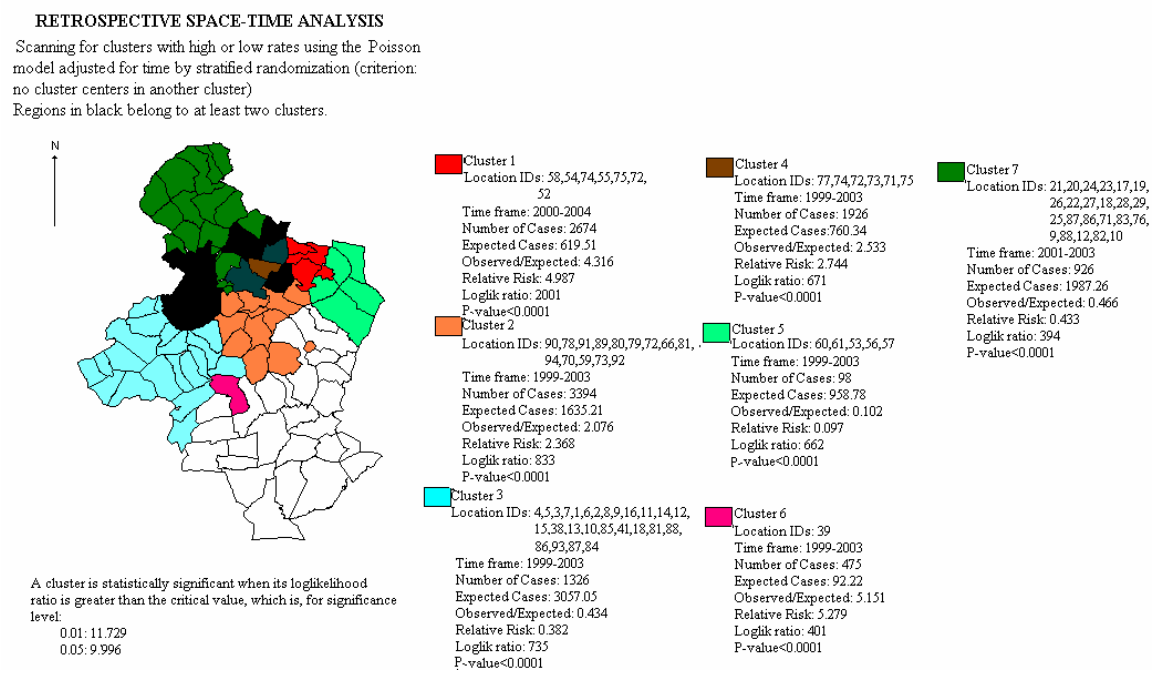

Figure 7: Clusters in space and time derived from the scan statistic under the Poisson model and the criterion that no cluster centers in another cluster (regions in black belong in at least two clusters).

CONSTRUCTING RISK MAPS VIA BAYESIAN POISSON REGRESSION

In order to explain rate variability in space, a hierarchical Bayesian Poisson Regression model was fitted to the ZCL aggregate counts. The model smoothes raw ZCL counts by fitting random effects that allow for spatial correlation using the intrinsic conditional autoregressive (CAR) prior as in Mollie (1996). These random effects represent the effect of latent (unobserved) risk factors. There is a second set of random effects in the model for which we assume an exchangeable normal prior. The random effect for each area is thus the sum of a spatially structured component and an unstructured component. This is termed a convolution prior. Under the assumed model the observed ZCL cases per sector are Poisson distributed; the logarithm of each sector's Poisson mean is dependent on the expected number of cases under the hypothesis of homogeneous spatial distribution of the disease in the region and three covariates: population density, urban/rural index and inhabitants per household.

Bayesian Poisson regression modeling revealed a significant dependence of the observed rates per sector to the urban/rural index. The negative sign of the corresponding coefficient (alpha2 in table 1 and figure 12) manifests lower rates in rural areas; this observation may seem awkward at first but it is consistent with previous findings that report high incidence rates in peri-urban areas1. It is also consistent with our findings from a large set of geo-referenced ZCL cases reported during 2004 in the northern part of the study region. Figure 8 displays a subset of this dataset that corresponds to the city of Sidi Bouzid; we plan to exploit these findings in a forthcoming publication. Population density appears to be a non-significant explanatory factor since its posterior distribution is centered to zero (see alpha1 in figure 12 and table 1) whereas the coefficient that corresponds to inhabitants per household (alpha3) tends to take negative values contrary to what one would expect. Observed

aggregate ZCL cases, expected cases under the hypothesis of homogeneous spatial distribution of ZCL, model fit and estimated relative risk, are depicted at figures 9 and 10.

The three explanatory variables of the model are not very strongly correlated. Indeed the only statistically significant correlation is between population density and inhabitants per household; Pearson's r statistic equals –0.45 indicating more inhabitants per household in rural areas. Tau and Tau.h in figure 12 and table 1 denote the precision (inverse of variance) for the spatial random effects and the exchangeable normal random effects respectively. Apparently, spatial random effects are much less dispersed around zero. The distribution of spatial effects and random error in the study area is depicted at figure 11. Unexplained log-relative risk of disease appears to be higher in the most likely spatial cluster derived from the spatial scan statistic. Moreover, high values are also present at the southern part of the region in which vegetation is very sparse (figure 1). The last two rows depict the deviance of the model with and without an exchangeable normal prior for the random effects respectively. That is in one model we assume convolution and in the other only spatial CAR priors for the random effects.

Observed values indicate superiority (in terms of short-term prediction) for the model with the convolution priors. Sampling from the posterior distribution for the parameters of interest was accomplished via Markov Chain Monte Carlo and Gibbs sampling. We run three chains and 70 000 iterations (under over-relaxation) for each one till convergence was reached; convergence was checked via the Gelman-Rubin criterion and visual inspection of the (overlapping) chains. The basic rule of thumb that MC error should be less that 5% of the standard deviation for each parameter of interest was also satisfied (table 1). After convergence was ensured 50 000 more iterations per chain were run and posterior inference was based in these final iterations.

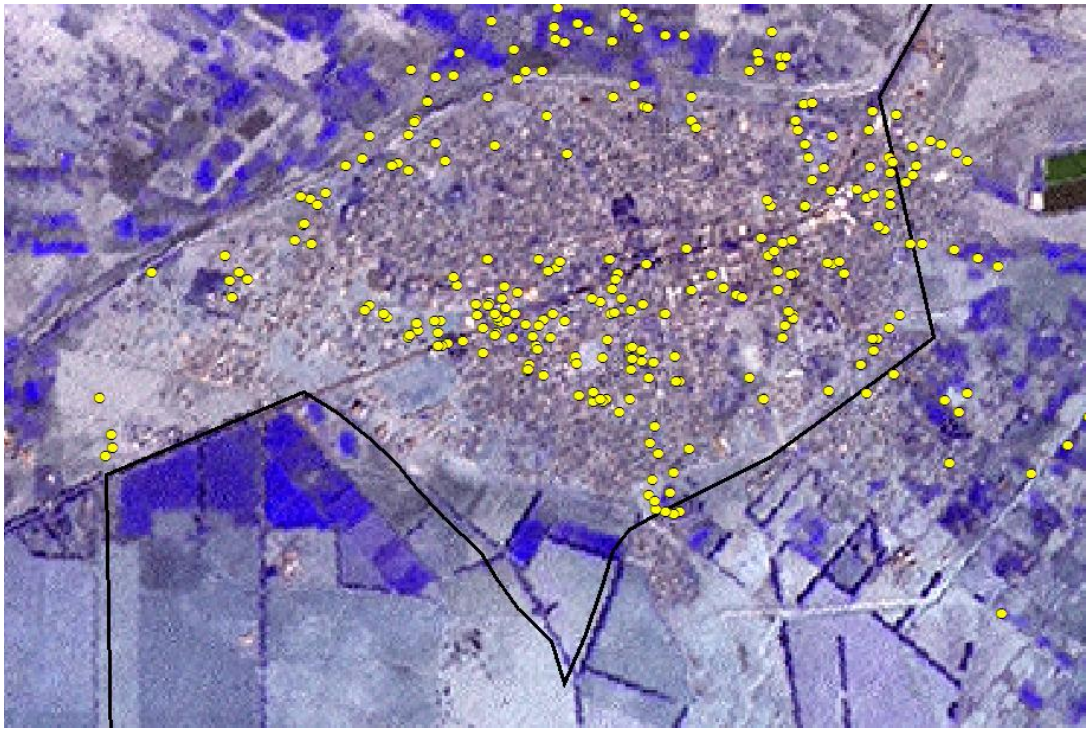

Figure 8: Geo-referenced ZCL cases reported during 2004 in the city of Sidi Bouzid.

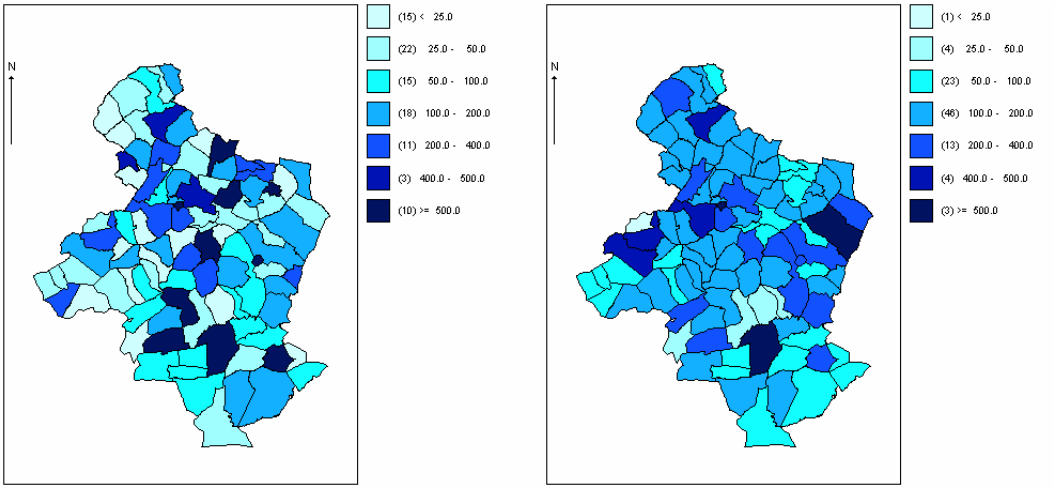

Figure 9 Left: Observed aggregate cases for the study period. ***Right:*** Expected aggregate cases under the hypothesis of homogeneous spatial ZCL distribution.

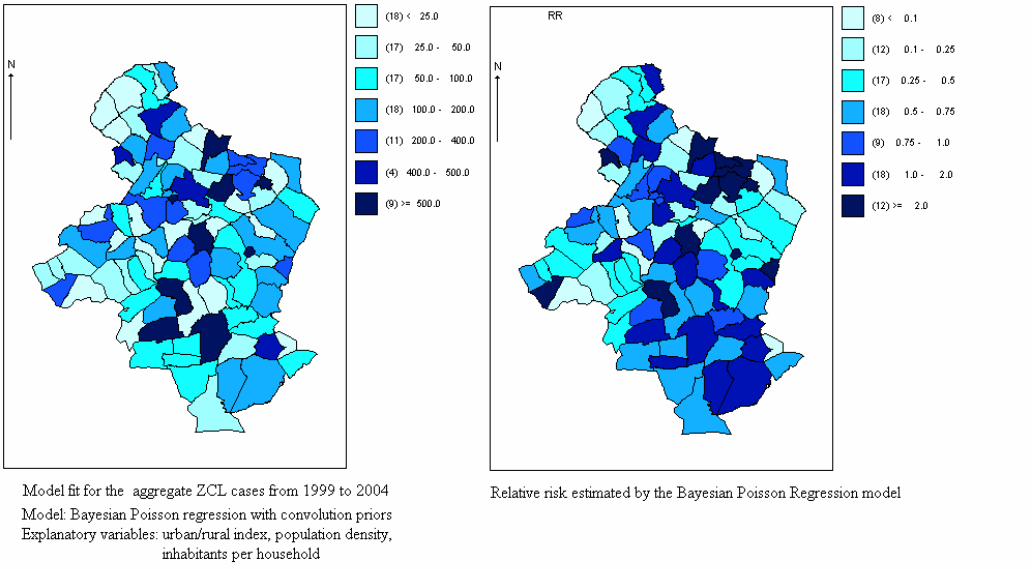

Figure 10 Left: Predicted cases from the Bayesian Poisson regression model. ***Right*:** Relative risk per sector.

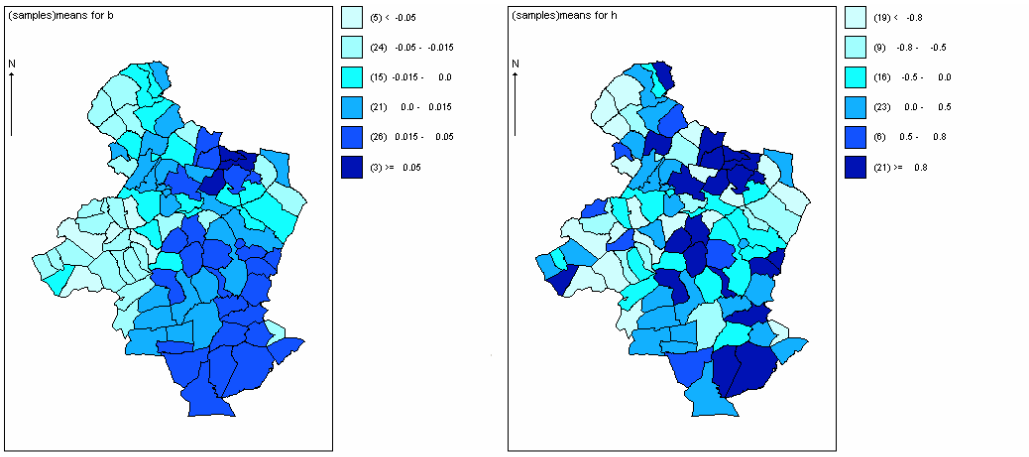

Figure 11 Left: Spatially dependent random effects. ***Right:*** Random error.

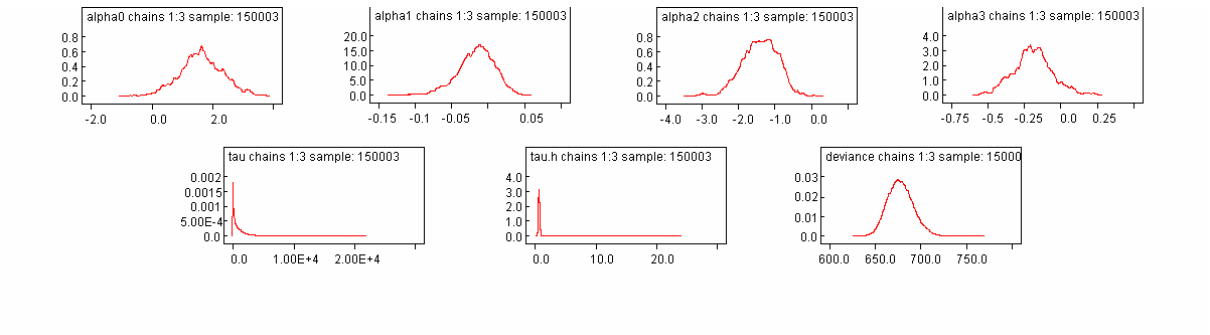

Figure 12: Simulated posterior distributions for the parameters of the Bayesian Poisson regression model.

Table 1: Summary statistics for the posterior distributions of the parameters of the Bayesian Poisson regression model.

NODE	MEAN	SD	MC ERROR	2.5%	MEDIAN	97.5%	SAMPLE
alpha0	1.62	0.7357	0.02816	0.149	1.615	3.127	150 003
alpha1	-0.017	0.0263	9.599E-4	-0.076	-0.0148	0.028	150 003
alpha2	-1.421	0.5052	0.0188	-2.447	-1.397	-0.538	150 003
alpha3	-0.191	0.1382	0.0053	-0.452	-0.1936	0.103	150 003
tau	1008	1423.0	34.51	1.045	469.6	5074	150 003
tau.h	0.791	0.3724	0.009	0.543	0.7599	1.102	150 003
deviance	677.4	14.1	0.053	651.7	676.7	707	150 003
deviance*	689.8	13.74	0.042	665.9	685.7	725.1	150 003

CONCLUSIONS

Over seventy percent of compulsory notified diseases in Tunisia are accounted for by zoonotic cutaneous leishmaniasis (ZCL), viral hepatitis and tuberculosis. Therefore, ZCL is a major public health problem for the country. Since 1982, an epidemic emerged in central Tunisia and expanded to the whole central and southern parts of the country (15/23 governorates are considered as endemic since 2002). The epidemics are cyclic and annually, two to three thousand cases are reported. The epidemic curve displays peaks and inter-epidemic periodicity of five to six years. Key factors driving spatio-temporal dynamics of the disease are presently unknown. These might include dynamics of rodent populations, dispersal of vectors, climate changes, vegetation and soil type and establishment of dense human settlements in areas where a sylvatic transmission of leishmaniasis is high (rodent-vector-rodent cycle). Currently, prediction of epidemic peaks and geographic spread is at an infant stage; thus, prevention programs are difficult to design and implement. Evaluation of the importance of such parameters will be critical for understanding the epidemiological features of ZCL and to predict future evolution.

As evident from past experience in the study region, the development of agricultural and water resources projects might enhance ZCL transmission and introduce the parasite to new areas through environmental changes in several ways. In 1982, the Sidi Saâd dam was constructed in northern Sidi Bouzid and the previously flooded shoals surrounding the dam, dried. Consequently, chenopod plants (which prefer soil of higher salinity) and are the exclusive food for P. obesus, the main reservoir of ZCL, expanded in these areas. The dramatic increase in rodent population was associated with the emergence of ZCL epidemic in man. Furthermore, drilling of hundreds of wells in other areas might have increased humidity and enhanced the vector density (P. papatasi) exposing humans to higher transmission. All these factors may have contributed to the explosion of a ZCL epidemic in 1992 in the town of Sidi Bouzid. Health authorities implemented a control program based on ploughing fields of chenopods around the town where P.obesus was very dense. This intervention that was planned and evaluated by Pasteur Institute, led to a significant reduction of the incidence among humans with a prevention fraction of disease exceeding 90% (Ben Ismaïl et al., personal communication). Consequently, in 2000, the Tunisian National Control of Parasitic Diseases Program (PDP) adopted a new multi-disciplinary strategy to intervene in ZCL transmission. It introduced ecological surveillance of areas at risk for ZCL, before the occurrence of the epidemics. Ecological surveillance consists of surveying for the emergence of rodent colonies, such as P. obesus, previously shown to harbor and increase transmission of ZCL. Indeed, previous work in the Pasteur Institute showed that ~90% of P. obesus in the governorate of Sidi Bouzid, showed evidence of anti-leishmanial antibodies at the end of the transmission cycle. Therefore, in all governorates where ZCL is endemic, PDP instituted rodent surveillance of a radius of 2 kilometers around villages with more than 5000

inhabitants. Despite this significant effort, and the analysis of transmission dynamics of the disease in other regions, control strategies remain unsatisfactory, as indicated by the number of annual cases. A global and prospective vision regarding the assessment of exposure of communities to ZCL in case of rural development is lacking. Therefore, the ability to integrate demographic, epidemiologic and ecologic factors is critically needed.

This study depicted that sectors neighboring to the Sidi Saâd dam still display statistically significant increased ZCL rates. Indeed, two alternative methods for detection of spatial clusters for the aggregate ZCL cases from 1999 to 2004 and two alternative methods for detection of spatio-temporal clusters indicated that the most likely cluster of high incidence rates is formed by seven sectors located close to the dam. Our findings coincide with previous ones reporting significant heterogeneity in the dynamics of ZCL incidence rates; ZCL occurred in outbreaks, clustering in space and time. Explanatory variables like people living per household and population density did not have significant predictive power at the sectoral level. Our findings of increased incidences in urban areas support the hypothesis of increased incidences in peri-urban environments due to changes in sandfly/rodent living habits over the last years. We plan to explore this hypothesis further in a forthcoming publication, by analyzing a dataset of geo-referenced ZCL cases, collected in the study area during 2004. For an extended version of this paper the interested reader is referred to Kamarianakis et al. (2006).

ACKNOWLEDGEMENTS

This work was financially supported by the Emphis project (www.emphis.org). The authors would like to thank L. Sallami, Z. El Ahmadi, Z. Belgacem, and A. Elfahem from the Regional Directorate of Sidi Bouzid, Tunisia and M. Bejaoui, from the Directorate of Primary Health Care, Tunisia, for their help related to data collection. Many thanks also go to K. Dellagi from the Institut Pasteur for his comments and to N. Chrysoulakis from IACM-FORTH, Greece, for his help with the satellite image.

BIBLIOGRAPHY

Assunção, R.M, Reis, I.A, Oliveira, CL., 2001 Diffusion and prediction of leishmaniasis in a large metropolitan area in Brazil with a Bayesian space-time model. Statistics in Medicine, 20, pp 2319-35.

Bakker, M.I., Hatta, M., Kwenang, A., Faber, W.R., van Beers, S.M., Klatser, P.R., Oskam, L., 2004 Population survey to determine risk factors for Mycobacterium leprae transmission and infection. International Journal of Epidemiology, 33, pp 1329-36.

Chaput, E.K., Meek, J.I., Heimer, R., 2002 Spatial analysis of human granulocytic ehrlichiosis near Lyme, Connecticut. Emerging Infectious Diseases, 8, pp 943-48.

Jennings, J.M., Curriero, F.C., Celentano, D., Ellen, J.M., 2005, Geographic identification of high gonorrhea transmission areas in Baltimore, Maryland. American Journal of Epidemiology, 161, pp 73-80.

Kamarianakis, Y., Chlif, S., Ben Salah, A., Ben Alaya, N., Prastacos, P., 2006 Zoonotic cutaneous leishmaniasis in Central Tunisia: spatio-temporal dynamics. Submitted for publication.

Kulldorff, M., 1997 A spatial scan statistic. Communications in Statistics: Theory and Methods, 26: 1481-96.

Kulldorff, M., 1999 Spatial scan statistics: Models calculations and applications. In Balakrisnan and Glaz (eds), *Recent Advances on Scan Stastics and Applications*. Boston, USA: Birkhauser, 1999.

Kulldorff, M., and Information Management Services, Inc. SaTScan™ v6.0: software for the spatial and space-time scan statistics. http://www.satscan.org/,2005.

Kulldorff M, Heffernan R, Hartman J, Assunção RM, Mostashari F., 2005 A space-time permutation scan statistic for the early detection of disease outbreaks. PLoS Medicine, 2, pp 216-24.

Machado-Coelho, G.L.L., Assunção, R,M, Mayrink, W., Caiaffa, W.T., 1999 American cutaneous leishmaniasis in southeast Brazil: space-time clustering. International Journal of Epidemiology, 28, pp 982-89.

Mollie A. 1996 Bayesian mapping of disease. In Markov Chain Monte Carlo in Practice. Gilks WR, Richardson S, Spiegelhalter DJ (eds), New York: Chapman & Hall,, pp 359-79.