Spatial analysis of dengue and the socioeconomic context of the city of Rio de Janeiro (Southeastern Brazil)

ABSTRACT

OBJECTIVE: To analyze the dengue epidemic in relation to the socioeconomic context according to geographical areas.

METHODS: An ecological study was conducted in the municipality of Rio de Janeiro (Southeastern Brazil), in areas delimited as neighborhoods, based on information about notified dengue cases concerning residents in the municipality. The average incidence rate of dengue was calculated between the epidemiological weeks: 48th of 2001 and 20th of 2002. The occurrence of dengue was correlated with socioeconomic variables through Pearson’s correlation coefficient. Moran’s global and local indexes were used to assess the spatial auto-correlation between dengue and the variables that significantly correlated with the disease. The multiple linear regression model and the conditional auto-regression spatial model were used to analyze the relationship between dengue and socioeconomic context.

RESULTS: The neighborhoods located in the west zone of the municipality presented high rates of average dengue incidence. The variables presenting significant correlation were: percentage of households connected with the general sanitary network, households with washing machines, and population density per urban area. Moran’s spatial auto-correlation index revealed spatial dependence between dengue and the selected variables. The utilized models indicated percentage of households connected with the general sanitary network as the sole variable significantly associated with the disease. The residual figures in both models revealed significant spatial auto-correlation, with a positive Moran Index (p<0.001) for linear regression model, and a negative one (p=0.005) for the conditional auto-regression one.

CONCLUSIONS: Problems related to basic sanitation contribute decisively to increase the risk of the disease.


INTRODUCTION

Dengue is the most important arbovirus in the world. Approximately 2.5 billion people are in susceptible conditions to the infection. In the last decades, its expansion has reached the tropical countries, favored by climatic, environmental and social characteristics.
Since its reintroduction in Brazil in 1976, *Aedes aegypti* has spread across the country due to the socio-environmental conditions, allied with the inefficiency of the programs that combat the vector.\(^{19,26}\)

Between 2001 and 2002 there was in the state of Rio de Janeiro the largest epidemic since the 1980s, coinciding with the isolation of serotype 3 (DEN 3):\(^{17}\) of 368,460 notified cases in the state, 177,919 were from the municipality of Rio de Janeiro.\(^{4}\)

Dengue control is one of the health services’ most difficult tasks due to the vector’s broad capacity for dispersion, to populations’ mobility, to the size of the cities’ population and to the complexity of the social and political problems that affect the quality of life and the environment.\(^{25}\)

The evaluation of different exposures to the factors involved in transmission allow to identify geographical areas with higher infection risk, and is fundamental to the development of dengue prevention and control programs.\(^{3,4}\)

In this sense, disease mapping has been a basic instrument in the public health field. Since the 1990s, the analysis techniques have been improved to generate identification maps of risk areas, resulting in differentiated attention being given by the health services.\(^{5,21,22}\)

Examples are the Brazilian studies that have used different spatial analysis techniques based on individual and aggregated data to identify areas that present higher risk of dengue.\(^{5,11,16,23,27}\)

The present study aimed to analyze the dengue epidemic in relation to the socioeconomic context.

**METHODS**

An ecological study of multiple groups was carried out in which the occurrence of dengue was analyzed according to geographical areas, delimited as neighborhoods and correlated with socioeconomic variables. The neighborhood was chosen as used as unit of analysis due to the availability of the disease data in *Sistema de Informação de Agravos de Notificação* (Sinan – Accident and Disease Reports Information System).

The municipality of Rio de Janeiro, Southeastern Brazil, had in 2000 a population of 5,857,904 inhabitants. The municipality’s total area is 1,224.56 km\(^2\), situated at 23°04’10”S and 43°47’40”W. According to physical, urbanism, administrative and planning criteria, the municipality of Rio de Janeiro is divided into ten planning areas (PA), 30 administrative regions and 158 neighborhoods.\(^{3}\) Figure 1 presents the municipality’s administrative division into neighborhoods, structured in ten PA. PA 1.0 corresponds to the city’s oldest, central region. PA 1.0 and the subdivisions of PA 2.0 and PA 3.0 occupy a smaller territorial area and have high population density and great offer of basic infrastructure services: PA 2.1 south zone, PA 2.2 north zone; PA 3.1 a suburb region of the former railroad of *Central do Brasil*, 3.2 a suburb of Leopoldina and Ilha do Governador, 3.3 the other region of the suburb of the former railroad of *Central do Brasil*. PA 4.0, PA 5.1, PA 5.2 and PA 5.3 correspond to the west zone of the city and occupy the largest territorial area of the municipality. Their neighborhoods have low population density and precarious or absent basic infrastructure.

The 125,041 notified cases concerning residents of the municipality were obtained from the Municipal Health Department through Sinan. These cases were notified in the period between the 48th epidemiological week of 2001 (November 25 to December 01) and the 20th epidemiological week of 2002 (May 12 to May 18), second curve of the epidemic process, totaling 25 weeks. This time interval is justified because it covers some important characteristics, like the identification of a single serotype (DEN 3), higher concentration of cases of the disease and a clearly defined continuous period, like the season of the year (summer).

The digital cartographic basis used in the making of the maps was provided by Rio de Janeiro’s municipal government.

A database with the municipality’s social and demographic information was created. The information was obtained at *Instituto Pereira Passos*\(^{5}\) and at *Instituto Brasileiro de Geografia e Estatística* (Brazilian Institute of Geography and Statistics).\(^{b}\)

Difficulties found in the geo-referral of the notified cases of the disease due to address problems prevented the location of the cases in census tracts. Thus, it was decided to work with the territorial unit at the neighborhood level.

The neighborhood of Joá was aggregated to Barra da Tijuca, Grumari to Recreio dos Bandeirantes, Parque Columbia to Pavuna, because these neighborhoods were recently created and many cases that occurred in these places were notified as being from the neighborhood.


\(^{3}\) In Brazil, residential areas which very frequently have precarious life conditions.

of origin. The neighborhood of Paquetá was withdrawn from the analysis because it is an island; as such, it does not present information on adjacent neighborhoods. Thus, the number of analyzed neighborhoods in the study corresponded to 154.

Subsequently, a single indicator was estimated to each neighborhood, the average incidence rate of dengue13 – new cases occurred in the 48th epidemiological week of 2001 up to the 20th epidemiological week of 2002, divided by the resident population in the year 2000.

For the characterization of the socioeconomic context, many variables presented co-linearity. Thus, those that presented higher correlation with the disease were selected:

- environmental sanitation – proportion of households with water supply coming from the general public network, proportion of households connected with the public sewage network, proportion of households with garbage collection service, proportion of households in which garbage is thrown in an empty plot or on the street;
- income and access to consumer goods – proportion of private households whose head of the family receives no income, including households whose head of the family received only benefits; proportion of households with washing machine;
- level of schooling – proportion of literate population;
- population and household density – urban population density, density of people per room in the household.

Transformations of the neperian logarithm (NL) type were employed for the outcome variable (average incidence rate of dengue) and for the independent variables. These transformations were chosen because they resulted in better approximations to the normal distribution.

To describe the association among all variables, firstly, Pearson’s correlation matrix was constructed, in which the variables with a 10% statistically significant correlation with the outcome were used in the multiple linear regression and spatial analysis.

The software used in the analysis were S-Plus 2000 for the statistical analysis and ArcView 3.2 and Terraview 3.1.4. for map making.

The spatial auto-correlation measure that was used was Moran’s global index, which tests if the connected areas present greater similarity regarding the studied indicator than expected in a random pattern, ranging from -1 to +1. The existing auto-correlation degree can be quantified, being positive for direct correlation and negative for inverse correlation.6 Besides Moran’s global index, the local index was used, which resulted in the “Moran map” of the average incidence rate of dengue, as it allows to find the spatial dependence agglomerations not observed in the global indexes, like possible clusters and outliers.

The relationships between the value of the measured attribute and of its neighbors can be observed in the graphic analysis of the local index, called Moran scatterplot, or in the two-dimensional thematic map, called Box Map. The map division is in quadrants: Q1, Q2, Q3 and Q4. The points located in Q1 and Q2 indicate zones in which the value of the measured attribute is similar to the neighbors’ mean; the first indicates positive value and positive mean and the second, negative value and negative mean. The points located in Q3 and Q4 indicate that the value of the measured attribute is not similar to its neighbors’ mean. In this case, Q3 indicates negative value and positive mean and Q4 indicates positive value and negative mean. The areas located in Q3 and Q4 can be seen as extreme or transition areas, as they do not obey the pattern observed for their neighbors.1

For the auto-correlation analysis, a neighborhood matrix was constructed for the 154 neighborhoods of the municipality, being defined by contiguous neighborhoods with at least one point in common.

The relationships between average incidence rate of dengue and socioeconomic variables were analyzed by applying the multivariate linear regression model with stepwise selection. Subsequently, for the spatial analysis a global model was applied, the Conditional Auto Regressive (CAR) spatial model, which captures the spatial structure by means of a single parameter added to the traditional regression model. CAR captures the spatial dependence of the outcome variable, and is expressed by:

\[ Y = X\beta + \varepsilon, \quad \varepsilon = \lambda W \varepsilon + \xi \]

Where \( W \) is the error component with spatial effects

\( \lambda = \) is the autoregressive coefficient

\( \xi = \) is the error component with constant, non-correlated variance.

The null hypothesis for the non-existence of auto-correlation is that \( \lambda = 0 \), that is, the error term is not spatially correlated.3 For the regression model diagnosis, maps of the residues were generated in the final multivariate linear regression and CAR models, searching for signs of a break in the independence presuppositions, that is, high concentration of positive or negative residues in one part of the map indicates the presence of spatial auto-correlation. Moran’s Index test was also used over the residues in both models to verify spatial auto-correlation quantitatively.
RESULTS

The spatial distribution of the average incidence rate of dengue indicated that neighborhoods located in PA 4.0 and in the subdivisions of PA 5.0 (west zone) were the most affected, followed by the neighborhoods located in PA 1.0 (central region) and by the neighborhoods that compose PA 2.1 (south zone) and 2.2 (north zone). Only some neighborhoods of the subdivisions of PA 3.0 presented high dengue rates (Figure 2).

At the beginning of the epidemic it was observed that the majority of the notified cases was in the city center. However, the disease propagated quickly between weeks 4-6/2002 towards the west and north zones. In this moment, the epidemic process reached more than 80% of the municipality’s neighborhoods, revealing a substantial increase in the number of cases. Up to week 12/2002, the scenario remained constant for the neighborhoods of the central region and of the west and north zones, but there was a slight reduction in the number of notified cases. In the south zone, few neighborhoods were affected by the disease until that moment. After week 14/2002, several neighborhoods of the west zone maintained a high number of cases until the end of the epidemic. In addition, the number of cases in the south zone of the city, which was affected later, started to increase. (The data were not presented)

Of the ten independent variables used in the study, three presented a statistically significant and negative correlation with the NL of the average incidence rate of dengue: percentage of households connected with the general sanitary network (-0.24); percentage of households with washing machine (-0.17); population density per urban area (-0.21).

The variables that presented a significant global spatial auto-correlation were: average incidence rate of dengue (0.21; p<0.001), percentage of households connected with the general sanitary network (0.61; p<0.001), population density per urban area (0.42; p<0.001), percentage of households with washing machine (0.27; p<0.001).

It was possible to observe through the Moran map (Figure 3) the cluster of areas in some neighborhoods located in the central region, west zone (PA 4.0 and PA 5.1) and neighborhoods in the suburbs of the Central do Brasil region of the municipality of Rio de Janeiro. The neighborhoods of the central region and west zone that compose PA 4.0 show a different pattern compared to part of the neighborhoods of the suburbs of the Central do Brasil region and of other neighborhoods of the west zone that compose PA 5.1. The first two, central region and west zone (PA 4.0), show the areas of clusters that present values of the studied attribute and the mean of their positive neighborhoods (Q1). On the other hand, in the neighborhoods of parts of the Central do Brasil region and in suburbs of the west zone (PA 5.1) the inverse occurred: areas of clusters that present values of the attribute and the mean of their negative neighbors (Q2). Few areas considered as transition areas, which can be analyzed as extremes, located in quadrants Q3 and Q4, were observed. This situation indicates that the areas belonging to these quadrants do not follow the same spatial dependence process as the others. Of the 154 analyzed areas, 113 were not significant, 16 belong to Q1, 18 belong to Q2, three belong to Q3 and four belong to Q4.

The multivariate linear regression analysis carried out with all variables revealed that the percentage of households connected with the general sanitary network was the only variable with a significant contribution to the explanation of the occurrence of the disease. However, the coefficient of determination found in the model was low.

Figure 1. Administrative distribution of the municipality according to planning areas. Municipality of Rio de Janeiro, Southeastern Brazil, 2000.

Figure 2. Average incidence rate of dengue in the 25 epidemiological weeks. Municipality of Rio de Janeiro, Southeastern Brazil, 2001-2002.
The mapping of the residues of the regression model indicated an agglomeration of positive residues in the central, south, north and west regions and in Ilha do Governador, and negative residues in the suburbs of the Central do Brasil region and of the west zone (Figure 4). Moran’s index regarding the residues was 0.19 (p<0.001), showing spatial auto-correlation.

The CAR model also showed that the percentage of households connected with the general sanitary network was the only significant variable. The coefficient found was close to the one of the multivariate linear regression model (Table). Nevertheless, the analysis of the CAR residues (Figure 4) indicated that there is not a spatial pattern as well established as in the one found in the analysis of the multiple regression residues. Moran’s index of the residues was -0.14 (p=0.005), revealing inverse auto-correlation.

The study presented limitations. For example, it did not include other possible variables related to the disease transmission, like water supply regularity,\textsuperscript{18} the vector’s spatial density, productivity of oviposition containers, building infestation rate.\textsuperscript{9,15} Due to irregular water supply, the population stores water in recipients that may become oviposition containers of the mosquito. The difficulty in accessing data on the vector limits its correlation with data on the disease.

The limitation of the used indicator refers to notification data, since a probable under-notification of dengue cases may have occurred, causing a selection bias.\textsuperscript{8}

The results obtained by means of the calculation of Moran’s global index for the studied variables expressed statistically significant auto-correlation. Besides, the analysis of the local auto-correlation indicators of the average incidence rate of dengue showed the existence of spatial dependence for 41 neighborhoods of the municipality of Rio de Janeiro. Among them, five located in the west zone and four in the central region (Q1) were considered high risk areas for the disease.

The west zone neighborhoods identified in the Moran map calculation as being the ones that had high values of the studied attribute and the mean of their neighbors (Q1) are in part the neighborhoods that were recently occupied and presented deficient and precarious basic infrastructure conditions.\textsuperscript{12} On the other hand, neighborhoods located in the central region of the city, despite the greater access to basic infrastructure services, presented high urban population density. However, the quality of these services was not assessed.

There is no consensus concerning which factors have more influence on the increase in dengue cases in a

**DISCUSSION**

The results obtained by means of the calculation of Moran’s global index for the studied variables expressed statistically significant auto-correlation. Besides, the analysis of the local auto-correlation indicators of the average incidence rate of dengue showed the existence of spatial dependence for 41 neighborhoods of the municipality of Rio de Janeiro. Among them, five located in the west zone and four in the central region (Q1) were considered high risk areas for the disease.

The west zone neighborhoods identified in the Moran map calculation as being the ones that had high values of the studied attribute and the mean of their neighbors (Q1) are in part the neighborhoods that were recently occupied and presented deficient and precarious basic infrastructure conditions.\textsuperscript{12} On the other hand, neighborhoods located in the central region of the city, despite the greater access to basic infrastructure services, presented high urban population density. However, the quality of these services was not assessed.

There is no consensus concerning which factors have more influence on the increase in dengue cases in a
Several variables have been used to identify the factors associated with dengue and, consequently, to provide a better basis for new proposals regarding the control of the disease.\textsuperscript{14,15,24}

The variable percentage of households connected with the general sanitary network explained the total variability of the average incidence rate of dengue in the municipality of Rio de Janeiro. The improvement in the sewage system has a great impact on the occurrence of infection and morbidity by many diseases. Nevertheless, even knowing that this variable is not directly connected with dengue transmission, it is possible to suggest that it is a proxy of the lack of environmental sanitation, represented by unplanned urbanization, by inadequate dwellings with precarious or inexistent basic infrastructure conditions.

According to Ault,\textsuperscript{2} unplanned urbanization and the suburban shantytowns have created not only new opportunities for the reproduction of vector populations, but also for the interaction between infected and susceptible people.

Both models were not capable of identifying the influence of the other independent variables of the present study, maybe due to the unit of analysis that was used, or because other important variables for the disease transmission were not considered, such as water supply regularity.\textsuperscript{4,9,20}

Regarding the models’ residues, spatial auto-correlation was expected in the linear regression model, as it does not presuppose the adjustment of the spatial dependence among them. However, when the spatial regression model is used, it is possible to note inverse spatial auto-correlation, when it was expected not to find spatial auto-correlation in the residues. It is possible that the variables used in the proposed spatial regression model have not captured the entire spatial pattern, probably because a global regression model was used.

Regression models with global effects start from the principle that the spatial process underlying the analyzed data is stationary, that is, the spatial auto-correlation patterns are captured in one single parameter. Nevertheless, the utilization of census data can produce diverse spatial patterns that are not captured in one single parameter. To test this hypothesis, we suggest investigating the spatial process by means of models whose parameters vary in space, like the ones that consider local spatial effects.\textsuperscript{6}

Although access to and quality of the water supply services did not present correlation with dengue, another study found that the absence of these services or irregular water supply may imply extremely favorable situations to the vector’s procreation, being determinant in the transmission.\textsuperscript{18}

Studies recognize the importance of housing and income indicators as determinants of dengue intensity, but not always do their conclusions coincide. This situation may be a reflection of the assessment of the information of collected data.\textsuperscript{9,15}

The information obtained from secondary data by means of the demographic census hides the great intrinsic variability, when aggregated to describe large regions. Thus, spatial aggregation may influence the results related to the variables, being less discriminating. In the city of Rio de Janeiro, this is even more serious, given the urban infrastructure and socioeconomic inequality problems. In view of the complexity that involves the process of becoming ill due to dengue, the obtained information regarding the disease in the municipality of Rio de Janeiro should concern smaller units of analysis, like, for example, census tracts.\textsuperscript{10,16}
Therefore, even though it was not possible to identify a multivariate model for dengue in the 2001-2002 epidemic in the municipality of Rio de Janeiro with the utilized variables, we would like to strengthen the importance of using spatial analysis tools and the methodology proposed by the present study, as it uses analysis techniques that incorporate spatial dependence in areas in the analysis of dengue occurrence.

REFERENCES


