

## Comparação entre Modelos de Regressão na Análise Espacial da Dispersão do *Triatoma Brasiliensis* no Estado do Ceará

Marcos de Meneses Rocha<sup>1</sup>  
Oscar Ricardo Vergara<sup>1</sup>  
Cassio Freitas<sup>2</sup>  
Samanta Cristina das Chagas Xavier<sup>3</sup>

<sup>1</sup>Instituto Militar de Engenharia (IME) – Seção de Engenharia Cartográfica  
Praça General Tibúrcio 80 – 20.290-270 – Urca - Rio de Janeiro – RJ, Brasil  
{mrocha,vergara}@ime.eb.br

<sup>2</sup>Escola Nacional de Ciências Estatísticas (ENCE)  
Rua André Cavalcanti 106 – 20231-050 - Santa Tereza – Rio de Janeiro – RJ, Brasil  
cassiofreitas@ibge.gov.br

<sup>3</sup>Instituto Oswaldo Cruz - IOC, FIOCRUZ - Departamento de Protozoologia, Laboratório  
de Biologia de Tripanosomatídeos  
Av. Brasil 4365 – 21045-900 - Rio de Janeiro - RJ, Brasil  
samanta@ioc.fiocruz.br

**Abstract.** This work uses techniques of spatial analysis to clarify the understanding of the disease vector's distribution, *Triatoma brasiliensis*, the most difficult specie of triatominae for controlling in Brazil. Thematic cartography and statistics are used to make an exploratory analysis of socio-economics and environmental data and the modelling uses multiple regression model and spatial autoregressive model (SAR). The results shows explanatory variables on both models explain few of the phenomenon. Future works include the use of another explanatory variables and statistics models to improve the modelling of *Triatoma brasiliensis* distribution.

**Palavras-chave:** spatial analysis, spatial statistics, geographical information systems, triatoma brasiliensis, análise espacial, estatística espacial, sistemas de informação geográfica.

### 1. Introdução

O presente trabalho emprega recursos de análise espacial e de estatística para melhor compreender a dispersão do *Triatoma brasiliensis*, com vistas a se fazer previsões sobre o número de triatomíneos presentes em uma região que possa ser empregado na vigilância epidemiológica. Busca-se, para tal, utilizar variáveis que se reconhecem como tendo influência na dispersão dos triatomíneos e verificar quais seriam as mais representativas, aplicando e comparando diversas modelagens não-espaciais e espaciais.

### 2. Justificativa

A região Nordeste apresenta grande importância nesse contexto por ter sido a segunda em número de infectados e de índices de infestação triatomínica nos inquéritos nacionais de prevalência e distribuição de vetores realizados entre 1975 e 1980 (Dias et al, 2000).

O risco de transmissão da doença persiste na região, decorrente de fatores como: baixas condições de vida da população, o fato da região ser o epicentro de dispersão do *T. brasiliensis* e do *T. pseudomaculata*, espécies de mais difícil controle pela Fundação Nacional de Saúde (FNS) e de haver baixa cobertura operativa do programa de controle da doença de Chagas pelas equipes da FNS.

### 3. Considerações sobre a doença de Chagas

A doença de Chagas é uma patologia largamente difundida no Continente Americano, sendo transmitida pelo *T. cruzi* (Kinetoplastida: Trypanosomatidae), um protozoário flagelado digenético que circula na natureza entre dezenas de espécies de mamíferos distribuídos em oito diferentes ordens e dezenas de espécies de vetores, triatomíneos da família Reduviidae (Pinto Dias, 2000). Em relação a dinâmica dos ciclos de transmissão do *T. cruzi* na natureza, ainda restam numerosos aspectos a serem esclarecidos.

Apesar da doença de Chagas estar controlada em muitas regiões do país devido, principalmente, a políticas de combate ao *Triatoma infestans*, uma espécie introduzida responsável pela transmissão intra-domiciliar, casos de emergência (ou re-emergência) da doença puderam ser vistos em algumas localidades no último ano, como por exemplo no Pará (associado ao consumo do açaí) e em Santa Catarina (associado ao consumo de caldo de cana de açúcar).

A iniciativa dos Países do Cone Sul, iniciada em 1991 e da qual o Brasil faz parte, visou eliminar a endemia chagásica na América Latina através da eliminação do principal vetor, *Triatoma infestans* e da triagem e fiscalização de rede de bancos de sangue (OPAS, 1992).

No Brasil, estas ações de controle foram implementadas de forma efetiva a partir de 1975. (Camargo et al., 1984; Silveira et al., 1984). As áreas consideradas de maior risco pelos inquéritos, que são aquela onde se encontram o *T. infestans*, foram priorizadas. Em relação às outras áreas, as ações de controle foram menos sistemáticas uma vez que os insetos vetores ali encontrados eram considerados de importância secundária (FUNASA, 1994).

A iniciativa do Cone Sul é considerada um sucesso e, atualmente, o Brasil é tido como livre da Doença de Chagas por transmissão vetorial ou transfusional. No entanto, os doentes permanecem, na medida em que a cura da Doença de Chagas é extremamente controversa. Tal fato contraria as expectativas mais otimistas de controle efetivo da endemia (Nascimento et al., 1997). Na verdade, considerar como interrompida a transmissão vetorial da Doença de Chagas a partir do controle e/ou eliminação de uma única espécie de triatomíneo vetora é bastante arriscado.

Dentre as espécies de triatomíneos, as mais relevantes, no que diz respeito à transmissão da doença de Chagas no Nordeste brasileiro, são o *T. brasiliensis*, o *T. pseudomaculata* e o *T. sordida*, sendo a primeira a espécie abordada no presente trabalho. Estudos epidemiológicos realizados por Costa et al., (1989) demonstraram que *T. brasiliensis* é o vetor de maior importância na região.

O *T. brasiliensis* apresenta uma ampla distribuição em toda a região semi-árida do nordeste brasileiro, sendo encontrado em nove estados que compõem esta região, além de ocorrer nos Estados de Minas Gerais e Tocantins. O *T. brasiliensis*, primariamente uma espécie silvestre, vem se adaptando com sucesso ao *habitat* humano, sendo encontrado no domicílio e peridomicílio. Trata-se de uma espécie que apresenta alta capacidade de reinfestação domiciliar e intradomiciliar, o que torna difícil o seu controle (Forattini, 1980; Carcavallo et al., 1998). Diotaiuti et al. (2000) e Oliveira-Lima et al. (2000) puderam comprovar o alto poder de reinfestação domiciliar do *T. brasiliensis* a partir de focos silvestres, reconstruindo-se suas colônias em um período de aproximadamente um ano.

### 4. Área de estudo

A área de estudo é o estado do Ceará, por ser a região com maior número de casos de doença de Chagas no Brasil, nele estando presente de forma endêmica e possuindo como espécie predominante o *T. brasiliensis*, segundo a FNS (Dias et al., 2000). Apresenta quase a totalidade do seu território na zona semi-árida, sendo a caatinga o principal complexo vegetacional do estado.

## 5. Materiais empregados

Utilizou-se os seguintes programas computacionais: a planilha eletrônica do OPEN OFFICE 1.0.3, o programa de geoprocessamento ARCGIS 3.2a e o programa estatístico SPLUS 6.1, da empresa *Insightful Corporation*.

## 6. Metodologia de trabalho

A metodologia aplicada constitui-se das seguintes etapas:

### 6.1. Planejamento e aquisição/coleta de dados

Verificação dos dados que possam ser explicativos para a modelagem da variável de interesse: número de triatomíneos coletados em cada município do estado do Ceará no ano de 2001. Procurou-se ter o cuidado para que o nível de precisão dos dados explicativos fosse compatível com a dos dados da variável de interesse.

### 6.2. Preparação dos dados

Retirou-se as informações redundantes ou que pudessem estar em uma escala superior à de trabalho. Os dados que estavam no meio digital foram convertidos para o formato dos programas utilizados no projeto.

### 6.3. Análise exploratória espacial

Construiu-se um mapa temático a partir dos dados de números de triatomíneos fornecidos pela FNS e cada um dos dados coletados em 2000 que se seguem: econômicos - criação de bovinos, suínos, galinhas e galos/frangos/frangas/pintos; sociais - rede geral de abastecimento de água e destino do lixo coletado; ambientais: pluviosidade anual observada e temperatura máxima anual. A seleção das variáveis baseou-se nos estudos de diversos autores, com destaque para Forattini (1980). Da comparação visual entre o primeiro e cada um dos demais, é possível ter uma primeira idéia do nível de correlação entre os mesmos. A quantificação de tal correlação ficou para a análise estatística.

### 6.4. Análise estatística de dados

Compreende a análise exploratória não-espacial e a modelagem dos dados. Na análise exploratória, visualizou-se o número de *T. brasiliensis* capturados em 2001, por meio de um sumário estatístico contendo valores da média, mediana, quartis, valores mínimo e máximo das variáveis, variância e desvio-padrão. Utilizou-se também como gráficos o histograma, o *boxplot* e o *qqplot*. O fato dos dados da variável resposta não apresentarem distribuição próxima da normal, levou à necessidade de se aplicar a transformação de Box-Cox. A modelagem dos dados foi dividida em não-espacial e espacial. Na modelagem não-espacial, empregou-se um modelo de regressão múltipla, considerando as variáveis ambientais e socio-econômicas como sendo independentes:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_k + e$$

Onde os coeficientes  $\beta_i$ ;  $i = 1, 2, \dots, k$  são parâmetros desconhecidos e o erro aleatório  $e_i$  representa todos os fatores considerados na regressão como variáveis explicativas. Nessa modelagem, acrescentou-se como variáveis explicativas a latitude e a longitude para verificar a existência ou não de correlação com a variável resposta. Adotou-se o nível de significância de 5%.

Na modelagem espacial dos dados, determinou-se primeiramente a dimensão da vizinhança a ser levada em consideração para o efeito de proximidade espacial. Em seguida,

selecionou-se a matriz de vizinhança que apresentava maior autocorrelação espacial e o tipo de modelo a ser empregado, nesse caso o modelo de regressão espacial *Spatial Auto Regressive* (SAR):

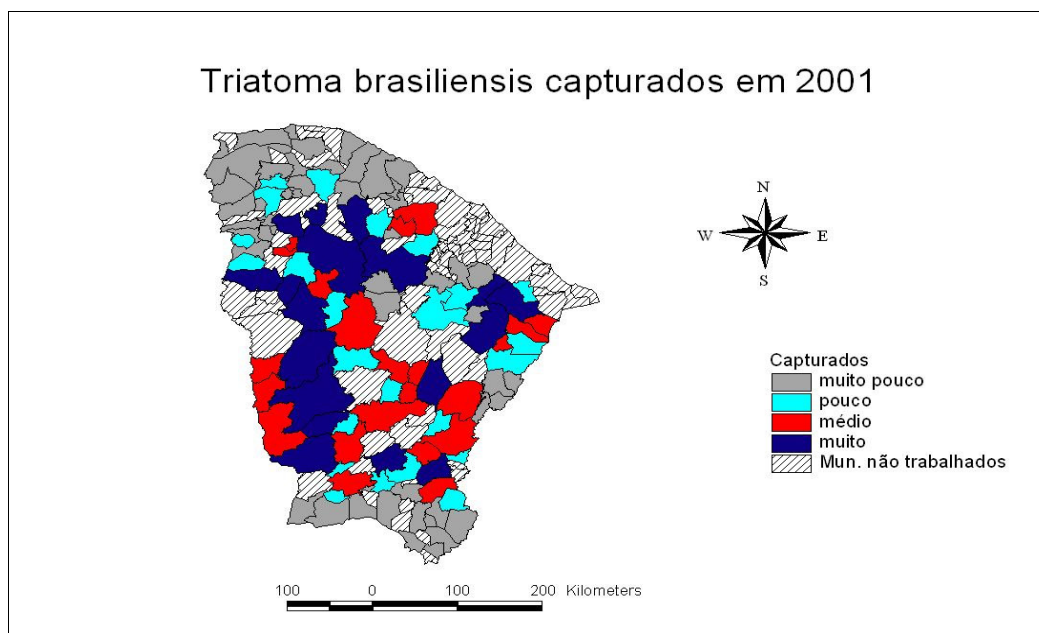
$$Y = \rho WY + \beta X + e$$

Onde  $W$  é a matriz de proximidade espacial e  $\rho$  é o coeficiente espacial autoregressivo, associado aos valores da autocorrelação espacial da variável  $Y_i$  na área  $A_i$  com os de  $Y_j$  na área  $A_j$  contida na vizinhança de  $i$ .

A comparação entre os modelos de regressão não-espacial e espacial foi feita observando-se o sumário estatístico e os gráficos empregados na análise exploratória não-espacial dos resíduos dos modelos, verificando se a distribuição dos mesmos era normal, caso contrário não poderiam ser utilizados os modelos de regressão. A estatística de *Moran* foi empregada para verificar se os resíduos apresentavam ou não correlação espacial. Caso apresentassem, o modelo espacial não teria conseguido remover tal efeito. Por fim, visualizou-se em um mapa os valores preditos da variável resposta em cada modelo, transformando-se os valores obtidos da transformação de Box-Cox inversa.

## 7. Resultados

A partir da análise exploratória espacial, observou-se que a distribuição de triatomíneos no estado do Ceará era menor na região norte, próxima ao litoral, e no sul do estado, ficando concentrada principalmente na região centro-oeste do estado (**Figura 1**). Verificou-se, ao observar os mapas temáticos, a existência de maior correlação entre a variável de interesse, triatomíneos coletados em cada município do estado do Ceará, e a criação de bovinos e galinhas, pluviosidade anual observada e temperatura máxima anual.



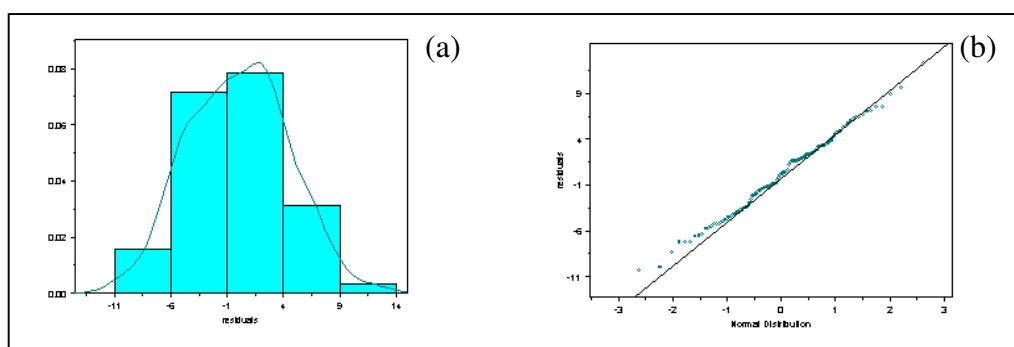
**Figura 1. Distribuição do *Triatoma brasiliensis* no Estado do Ceará**

Na modelagem não-espacial, após a aplicação da regressão linear múltipla, verificou-se que as variáveis mais explicativas seriam a criação de bovinos, a pluviosidade anual observada e as formas de abastecimento de água. O valor do  $R^2$  ajustado ficou em torno de 38% e p-valor próximo de zero, indicando que as variáveis supracitadas são pouco explicativas do fenômeno, com alto grau de confiança.

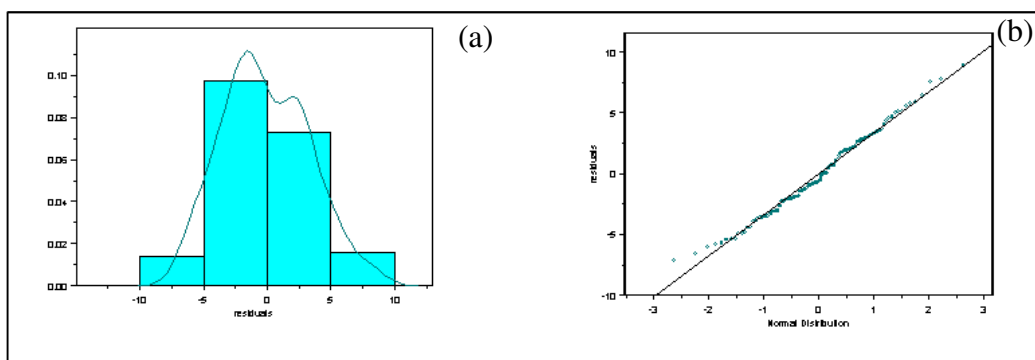
Na modelagem espacial, para efeito de proximidade espacial, empregou-se como raio de vizinhança o valor de 50 km, por apresentar um bom raio de abrangência: nem tão grande que

pucesse cobrir uma extensa área da região de trabalho, nem tão pequena a ponto de considerar um número muito pequeno de vizinhos. A matriz de vizinhança foi obtida considerando o valor 1 para os municípios que estivessem dentro do raio de 50 km e 0 para aqueles que não estivessem. Após a aplicação da regressão espacial, verificou-se que as variáveis mais explicativas da dispersão do *T. brasiliensis* foram a criação de bovinos, pluviosidade média anual e temperatura máxima anual, sendo esta última diferente da regressão não-espacial. O coeficiente espacial autoregressivo obtido foi de 0,1312, o que indica pouca dependência espacial, tendo em vista que somente valores maiores que 0,5 poderiam ser considerados relevantes para efeito de dependência espacial. O p-valor da regressão espacial em torno de 0,7 indica que não se pode rejeitar a hipótese nula. Aplicando-se o índice de Moran aos resíduos das regressões, verificou-se que o modelo não-espacial não contempla a dependência espacial, ao contrário do modelo espacial, o que torna o último mais adequado.

A **Figura 2** mostra o histograma e o *qqplot* do resíduo da regressão não-espacial e a **Figura 3** mostra o histograma e o *qqplot* do resíduo da regressão espacial.

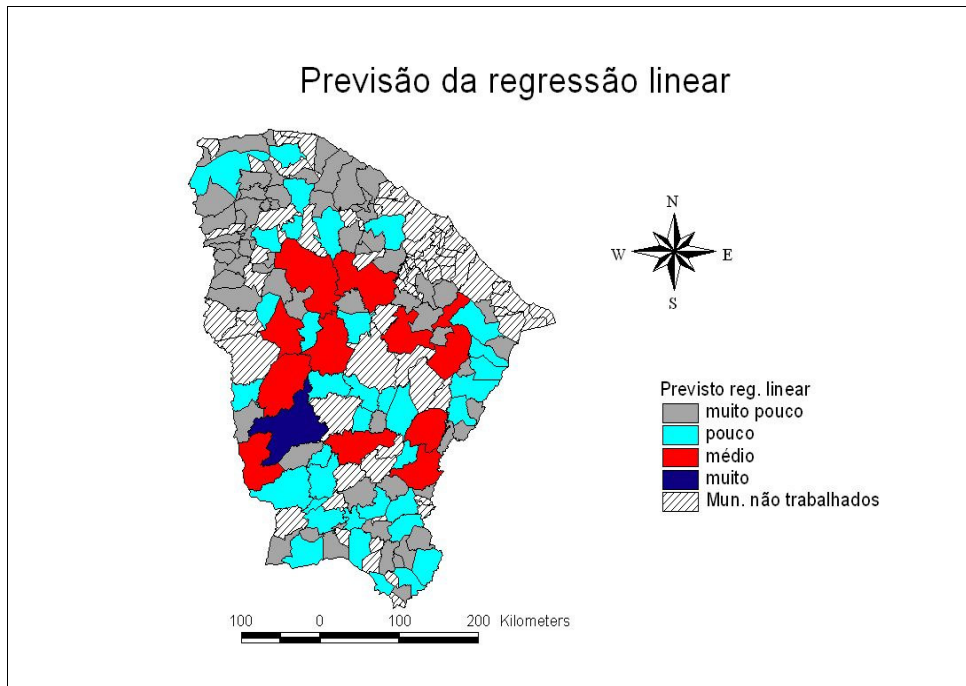


**Figura 2. (a) Histograma do resíduo da regressão não-espacial (b) Qqplot do resíduo da regressão não-espacial**

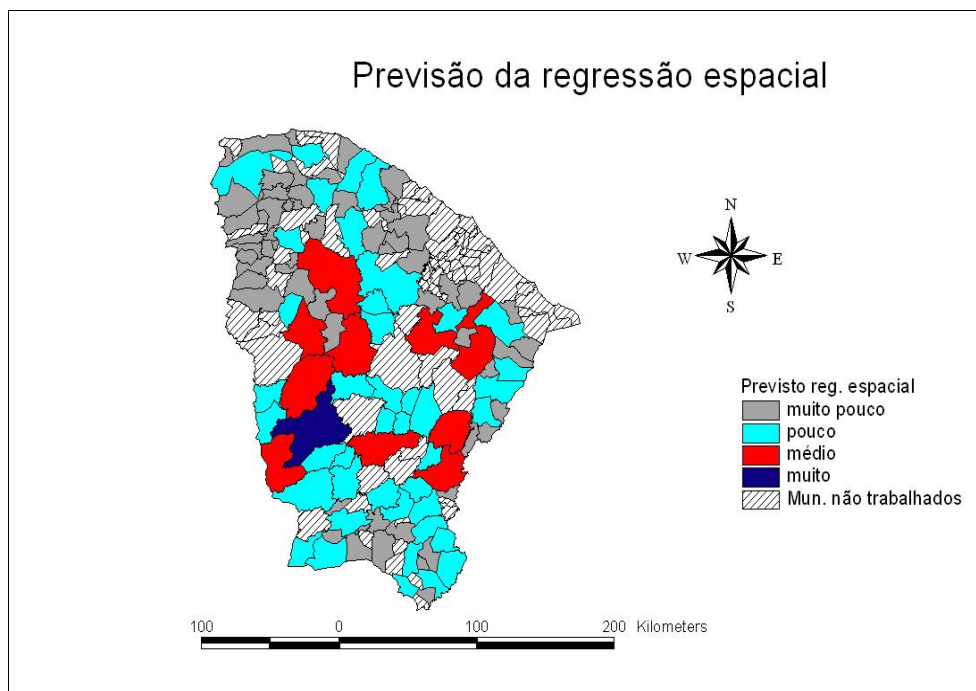


**Figura 3. (a) Histograma do resíduo da regressão espacial (b) Qqplot do resíduo da regressão espacial**

As **Figuras 4 e 5** mostram os mapas da previsão do número de triatomíneos em cada um dos municípios do estado do Ceará a partir das regressões não-espacial e espacial, respectivamente.



**Figura 4. Previsão da distribuição do *Triatoma brasiliensis* a partir da regressão não-espacial**



**Figura 5. Previsão da distribuição do *Triatoma brasiliensis* a partir da regressão espacial**

## 8. Conclusões

Conclui-se que as variáveis criação de bovinos e pluviosidade média anual foram mais explicativas na visualização dos mapas temáticos na análise exploratória espacial e nas regressões não-espacial e espacial, enquanto que a temperatura máxima anual, a criação de galinhas e a proporção de domicílios com rede geral de abastecimento de água não foram observados em todas as análises. Os valores apresentados para a criação de bovinos

confirmam o estudo de Forattini (1980) no que diz respeito à influência da ação antrópica na domiciliação dos triatomíneos, onde a devastação natural para a criação de pastagens teria provocado a migração dos triatomíneos do ambiente silvestre para o domiciliar e peridomiciliar. Quanto à pluviosidade e à temperatura, os resultados confirmam a hipótese de que o *T. brasiliensis* teria preferência por lugares secos e quentes. A temperatura não apresenta valores melhores talvez pela pouca variação das temperaturas máximas entre os municípios, as quais são valores médios para o ano de 2000. Quanto à criação de suínos e aves em geral, costuma ocorrer em um espaço físico reduzido, não gerando grandes desequilíbrios ambientais. Dependendo da extensão das criações, pode também haver grandes discrepâncias na quantidade de animais de um município para o outro. A correlação entre os domicílios com rede geral de abastecimento de água e a distribuição do *T. brasiliensis* pode ser o reflexo do critério empregado na seleção da área para captura de triatomíneos pela FNS e que poderia estar levando a uma tendenciosidade na coleta dos dados.

Os valores baixos do  $R^2$  ajustado e do coeficiente autoregressivo apontam para a baixa qualidade dos modelos de regressão não-espacial e espacial, respectivamente. Tal fato torna necessária a verificação de outros métodos estatísticos, bem como de outros indicadores que possam ser mais explicativos do modelo.

## Referências

- Camargo, M. E.; Silva, G. R.; Castilho, E. A.; Silveira, A. C. Inquérito sorológico da prevalência de infecção chagásica no Brasil. 1975/ 1980. São Paulo. **Rev. Inst. Med. Trop.**, v. 26 p.192-204, 1984.
- Carcavallo, R. U., Girón, I. G., Jurberg, J., Lent, H. **Atlas dos vetores da Doença de Chagas nas Américas**. Rio de Janeiro: Fiocruz. 3 v, 1998.
- Dias, J.C.P; Machado, E.M.M; Fernandes, A.L e Vinhaes, M.C. Esboço geral e perspectivas da doença de Chagas no Nordeste do Brasil. **Cad. Saúde Pública**, v. 14, p.19-37, 2000.
- Diotauiti, L; Filho, O.F.F; Carcneiro, F.C.F; Dias, J.C.P; Pires, H.H.R e Schofield, C.J. Aspectos operacionais do controle do *Triatoma brasiliensis*. **Cad. Saúde Pública**, v.16 (Sup. 2), p. 61-67, 2000.
- Forattini, O.P. Biogeografia, origem e distribuição da domiciliação de triatomíneos no Brasil. **Rev. Saúde Pública**, v. 14, p. 265-299, 1980.
- FUNASA (Fundação Nacional de Saúde). **Controle da Doença de Chagas: Diretrizes Técnicas**. Brasília: Fundação Nacional de Saúde, Ministério da Saúde, 1994.
- Nascimento, C.; Marassá, A. M.; Curado, I.; Piazza, R. M. E. Encontro de *Panstrongylus megistus* em ecótopo artificial: Domiciliação ou mera visitação? **Rev. Soc. Bras. Med. Trop.** v. 30, p. 333-336, 1997.
- Oliveira-Lima, J.W; Filho, OF.F; Vieira, J.B.F; Gadelha, F.V e Filho, A.M.O. Alterações do peridomicílio e suas implicações para o controle do *Triatoma brasiliensis*. **Cad. Saúde Pública**, v. 16 (Sup.2): p.75-81, 2000.
- OPAS (Organização Pan-Americana de Saúde). **Documento OPS**. PNSP/92 – 18 sobre a Iniciativa do cone sul. Washington, DC: Organização Mundial da Saúde. 1992.
- Pinto Dias, J.C. Epidemiologia. In: Brener, Z.; Andrade, Z.; Barral-Netto, M. (Eds.), **Trypanosoma cruzi e Doença de Chagas**. Rio de Janeiro: Guanabara Koogan, 2000. 48-74. p.
- Silveira, A. C.; Feitosa, V. R.; Borges, R. Distribuição de triatomíneos capturados no ambiente domiciliar no período 1975/83, Brasil. **Rev. Bras. Mal. D. Trop.** v. 36, p.15-312, 1984.