

Utilização de dados censitários em substituição a informações socioeconômicas obtidas no nível individual: uma avaliação empírica

Use of Census Information to Proxy Individual Socioeconomic Characteristics: an Empirical Evaluation

Guilherme L. Werneck

Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro-RJ
Núcleo de Estudos de Saúde Coletiva, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ

Carlos H. N. Costa

Instituto de Doenças Tropicais Nathan Portella, Universidade Federal do Piauí, Teresina-PI

Resumo

A obtenção de informação socioeconômica em nível individual nem sempre é uma tarefa de simples execução. Muitos pesquisadores optam por substituir a informação individual por dados socioeconômicos derivados de recenseamentos censitários. O objetivo deste artigo é avaliar a validade do uso de dados censitários em substituição às informações geradas em nível individual. Com base em estudo de caso-controle de base populacional para investigar o impacto da ocupação urbana e degradação ambiental na incidência de leishmaniose visceral no Município de Teresina, Estado do Piauí, Brasil, são comparadas estimativas de efeito para diferentes fatores de risco quando controladas, ora para variáveis censitárias, ora para variáveis obtidas mediante entrevistas individuais. Considerando-se uma discrepância relativa de 10% ou menos como critério de similaridade entre as razões de chance [odds ratios (OR)] obtidas com dados individuais (ORi) e dados censitários (ORc), percebe-se que 47% das estimativas de ORc poderiam ser consideradas medidas diferentes daquelas obtidas pela ORi. Tomando-se a ORi como “padrão-ouro”, a especificidade e a sensibilidade da ORc foram de apenas 60%. O uso de variáveis “censitárias” tende a não prover correção efetiva para fatores de confusão socioeconômicos. Dados censitários não devem ser considerados, por si só, bons substitutos de dados individuais, mas expressões de conceitos que operam em outro nível hierárquico de determinação.

Palavras-chave: fatores socioeconômicos; censos; fatores de confusão; viés; falácia ecológica.

Summary

Gathering socioeconomic information at the individual level is not a simple task. Many researchers use census data as a proxy for individual characteristics. The objective of this report was to evaluate the validity of the use of census information as a proxy for individual socioeconomic characteristics. Using data from a population-based case-control study to investigate the impact of land occupation and environmental degradation on the incidence of visceral leishmaniasis in the Municipality of Teresina, Piauí State, Brazil, this report compares measures of effect for different risk factors, controlling for census variables and individual-level variables. Considering a relative discrepancy of 10% or less as a criterion of similarity between the odds ratios (OR) obtained by using individual level socioeconomic variables (ORi) and census data (ORc), some 47% of the estimates of ORc could be considered as a different estimate, compared to the ORi. Considering the ORi as the gold-standard, the specificity and sensitivity of the ORc were both approximately 60%. The use of census variables does not necessarily provide effective control for socioeconomic confounding. Census data should not be considered by themselves good substitutes for individual-level data as they represent different concepts that operate in another level of determination.

Keywords: socioeconomic factors; censuses; confounding factors; bias; ecological fallacy.

Endereço para correspondência:

Rua São Francisco Xavier, 524, 7º andar, Bloco D, Maracanã, Rio de Janeiro-RJ. CEP: 20559-900
E-mail: gwerneck@nesc.ufrj.br

Introdução

A utilização de informações socioeconômicas na pesquisa epidemiológica é de importância capital. Sua necessidade sucede das substanciais evidências de conexões entre estrato social, emprego, educação e renda e eventos ligados à saúde.¹⁻¹⁰ O interesse nesse tipo de informação não é privilégio dos estudos que exploram o papel de fatores socioeconômicos na determinação do adoecer. Mesmo naqueles estudos que não focalizam essas relações, é praticamente inevitável que algumas dessas variáveis cumpram os requisitos básicos para serem consideradas fatores de confusão e devam, portanto, ser incluídas na análise dos dados.¹¹

A obtenção de dados socioeconômicos em nível individual nem sempre é uma tarefa de simples execução. Entrevistas individuais associadas à inspeção detalhada da habitação são estratégias comumente utilizadas na pesquisa epidemiológica, mas esses procedimentos apresentam limitações inerentes à dificuldade de operacionalização de conceitos complexos como os de classe social, poder de compra, mobilidade social, entre outros.¹²⁻¹⁵

Mesmo que se ignorem tais restrições conceituais, alto custo e as dificuldades logísticas na coleta de dados, ainda assim, nem sempre há garantia de que sejam obtidas informações válidas. O problema é mais evidente na medida em que se amplia, progressivamente, a disponibilidade de grandes bancos de dados informatizados sobre morbidade e mortalidade e a sua utilização como fonte de dados para estudos epidemiológicos.¹⁶ Essas bases de dados podem contribuir para ampliar nossa compreensão acerca da qualidade da assistência à saúde e do papel desempenhado por diferentes fatores na produção do adoecimento. Infelizmente, informações socioeconômicas válidas nem sempre estão disponíveis nos instrumentos utilizados para a coleta desses dados.

Para superar esses problemas, muitos pesquisadores optam pela utilização de dados socioeconômicos derivados de recenseamentos censitários.^{11,17-19} Na abordagem deste estudo, os indivíduos são caracterizados pelo perfil socioeconômico do setor censitário onde está localizada a sua residência. Se essa solução não é a ideal, pode-se aventar que, pelo menos do ponto de vista do controle do confundimento em estudos epidemiológicos, é possível que algumas variáveis

obtidas no nível agregado sejam boas aproximações daquelas que seriam obtidas mediante entrevista individual. Como consequência, sua inclusão em modelos de regressão permitiria a estimação acurada das medidas de associação de interesse.

Resultados de alguns estudos sobre a validade do uso de informações socioeconômicas derivadas do censo em substituição à informação individual são conflitantes.²⁰⁻²³ Até o momento da conclusão deste relato, não foram detectados estudos desse tipo no contexto brasileiro. Nesse sentido, torna-se relevante a elaboração de trabalhos sobre o tema, gerando informações que possam ampliar as oportunidades de desenvolvimento de estudos epidemiológicos em nosso meio.

Aqui, são relatados os resultados de uma avaliação empírica da validade do uso de dados censitários, em substituição às informações geradas em nível individual, sobre um estudo de base populacional para investigar fatores de risco para leishmaniose visceral no Município de Teresina, Estado do Piauí.

Metodologia

Entre 1995 e 1996, um estudo de caso-controle de base populacional para investigar o impacto da ocupação urbana e degradação ambiental na incidência de leishmaniose visceral (LV) foi desenvolvido em Teresina, Piauí. Entre julho de 1995 e fevereiro de 1996, todos os novos casos de leishmaniose visceral notificados à representação da Fundação Nacional de Saúde (Funasa) em Teresina foram selecionados. O diagnóstico baseou-se no quadro clínico, na identificação de *Leishmania chagasi* em aspirado de medula óssea ou na presença de testes sorológicos positivos para *L. chagasi*. Foram consideradas elegíveis para participar do estudo apenas as pessoas acima de um ano de idade e residentes em Teresina à época do diagnóstico. Entre os 62 pacientes elegíveis, seis (10%) morreram imediatamente após o diagnóstico e 12 (19%) não puderam ser localizados. Este estudo incluiu resultados para 44 casos de leishmaniose visceral.

Entre junho de 1995 e maio de 1996, 200 residências foram amostradas aleatoriamente, a partir de uma lista de endereços registrada na companhia elétrica do Estado do Piauí (Cepisa). Naquela época, essa lista cobria cerca de 93% das habitações urbanas de Teresina. Uma pessoa em cada residência acima

de um ano de idade e sem evidência clínica de IV foi selecionada como controle. Um total de 176 (88%) controles elegíveis ou adultos responsáveis completou as entrevistas. Oito pessoas (4%) recusaram-se a participar do estudo e 16 (8%) não puderam ser localizadas.

Um questionário, estruturado com perguntas pré-codificadas, foi utilizado para entrevista, incluindo diversos fatores relacionados, potencialmente, à ocorrência de leishmaniose visceral, como idade, sexo, presença de cão doméstico e outros animais, tipo de moradia, entre outros. Informações detalhadas sobre condições sanitárias e sociais também foram levantadas em inspeção direta do interior da habitação e do peridomicílio.

A localização dos domicílios foi determinada utilizando-se um sistema de posicionamento global. Dados socioeconômicos em nível de setor censitário, provenientes do Censo de 1991, foram vinculados às residências dos indivíduos mediante sistema de informação geográfico.²⁴

Com o objetivo de avaliar a aplicabilidade de dados censitários em substituição às informações obtidas em nível individual para fins de controle de confundimento, procedeu-se da seguinte maneira: (1) seleção das variáveis socioeconômicas disponíveis em nível individual e censitário; (2) estimativa dos riscos relativos associados a cada uma das variáveis socioeconômicas; (3) seleção de variáveis de interesse primário do estudo – ou seja, aquelas para as quais se desejam obter estimativas de risco relativo, controladas para confundimento socioeconômico –; (4) estimativa dos riscos relativos associados às variáveis de interesse primário, controlando, ora para variáveis socioeconômicas no nível individual, ora para aquelas derivadas do censo; e (5) comparação dos resultados entre os modelos.

As seguintes variáveis socioeconômicas estavam disponíveis nos níveis individual e censitário: escolaridade do chefe da família; abastecimento de água com canalização interna; esgotamento sanitário ligado à rede geral ou fossa séptica; e recolhimento de lixo. As variáveis socioeconômicas disponíveis em nível individual foram dicotomizadas em sim/não (água e esgoto), regular/irregular (recolhimento de lixo) e até 4ª série/>4ª série (escolaridade). As variáveis socioeconômicas disponíveis em nível censitário, originalmente expressas como percentuais, foram dicotomizadas segundo critérios apontados por mo-

delos de árvores de classificação,²⁵ da seguinte forma: abastecimento de água (< ou >18% dos domicílios); esgotamento sanitário (< ou >60%); recolhimento de lixo (< ou >2%); e chefes de família sem escolaridade (< ou >35%).²⁶

As variáveis de interesse primário e selecionadas para avaliação foram: relato de recolhimento de cão pela Funasa, em razão da infecção por *L. chagasi* nos 12 meses que antecederam a entrevista (sim/não); relato de presença de raposas nas redondezas da habitação (sim/não); e grau de exposição domiciliar ao vetor *L. longipalpis* (alto/baixo). Tanto a presença de cães infectados como a de raposas no peridomicílio são consideradas fatores de risco para leishmaniose visceral, por serem esses animais reservatórios domésticos e selvagens da infecção, respectivamente.²⁷ O grau de exposição domiciliar também pode ser considerado um fator de risco, na medida em que reflete a probabilidade de o flebotômio invadir a residência para se alimentar em humanos.²⁸ Um domicílio seria classificado como de “alto” grau de exposição ao flebotômio se possuísse pelo menos duas das três seguintes características: habitação não completamente coberta por laje e/ou telha; ausência de forro completo na sala e quartos; e mais de quatro moradores permanentes.

Todas as variáveis socioeconômicas incluídas neste estudo são, teoricamente, potenciais fatores carreadores de confusão. Por um lado, elas têm sido apontadas como fatores de risco para a ocorrência de leishmaniose visceral,²⁹ por outro lado, estão associadas, na base populacional, às três exposições em questão.

A utilização de informações socioeconômicas na pesquisa epidemiológica, como estrato social, emprego, nível de educação e renda, é de importância capital.

Utilizou-se regressão logística não condicional para obter razões de chance como estimativas dos riscos relativos associados a cada variável de interesse primário.³⁰ Para cada variável socioeconômica, dois modelos foram ajustados: um com a variável “individual” e o outro com a respectiva contraparte “censitária”. Para cada variável de interesse primário, dois grupos de modelos de regressão foram ajustados (um incluindo

a variável socioeconômica “individual”; e outro, a respectiva contraparte “censitária”). Todos os modelos consideraram idade como co-variável. Comparando-se as razões de chance [*odds ratios* (OR)] obtidas mediante a utilização de modelos ajustados com a variável censitária (OR_c) e com a variável individual (OR_i), pode-se obter uma medida-resumo para quantificar o viés induzido pela substituição da variável “individual” pela “censitária”, que foi denominada de discrepância relativa, definida como:³¹

$$\frac{OR_c - OR_i}{OR_i}$$

Resultados

A Tabela 1 mostra os resultados das associações entre variáveis socioeconômicas e ocorrência de IV. Houve grande variação nas razões de chance, particularmente para as variáveis abastecimento de água e esgotamento sanitário, devendo-se salientar que apenas a primeira dessas duas, além do recolhimento de lixo,

mostrou-se associada com a ocorrência de IV de maneira significativa. Nota-se, também, que as razões de chance associadas às variáveis censitárias tenderam, sistematicamente, para o valor nulo (OR=1), quando comparadas com aquelas estimadas utilizando-se modelos com variáveis individuais.

Os resultados das associações entre as variáveis de interesse primário e a ocorrência de IV estão apresentados na Tabela 2. Adotando-se, como critério para existência de confundimento, a modificação em mais de 10% da razão de chances estimada com ajustamento apenas para idade, pode-se observar que, em nível individual, a variável abastecimento de água confunde as relações entre ocorrência de IV e recolhimento de cão e presença de raposa; escolaridade confunde a relação entre IV e grau de exposição do domicílio; e a variável recolhimento de lixo confunde a relação entre IV e presença de raposa no peridomicílio. Quando são ajustadas todas as quatro variáveis, simultaneamente, há evidência de confundimento apenas para a relação entre ocorrência de IV e presença de raposa.

Tabela 1 - Razões de chance (OR) e intervalos de confiança de 95% (IC95%) para leishmaniose visceral associados com variáveis socioeconômicas obtidas em nível individual e censitário, no Município de Teresina, Estado do Piauí, Brasil, 1995 - 1996

Característica	OR individual ^a	IC95%	OR censitário ^a	IC95%	Discrepância relativa
Abastecimento de água					
Melhor categoria	1,00		1,00		
Pior categoria	6,99	1,32-37,00	3,24	1,19-8,78	-0,54
Esgotamento sanitário					
Melhor categoria	1,00		1,00		
Pior categoria	2,80	0,39-19,90	1,13	0,41-3,07	-0,60
Escolaridade					
Melhor categoria	1,00		1,00		
Pior categoria	2,42	0,86-6,78	2,18	0,89-5,33	-0,10
Recolhimento de lixo					
Melhor categoria	1,00		1,00		
Pior categoria	5,40	1,88-15,50	4,92	1,89-12,90	-0,09

a) Ajustado por idade

Tabela 2 - Razões de chance (OR) e intervalos de confiança de 95% (IC95%) para leishmaniose visceral associados com recolhimento de cão pela Fundação Nacional de Saúde em razão da infecção por *L. chagasi* nos 12 meses que antecederam a entrevista, relato de presença de raposas na redondeza do domicílio e grau de exposição domiciliar ao vetor *L. longipalpis*, ajustados para variáveis socioeconômicas (individual e censitária), no Município de Teresina, Estado do Piauí, Brasil, 1995 - 1996

Variável socioeconômica de ajuste	Recolhimento de cão pela Funasa nos 12 meses anteriores (sim versus não)		Discrepância relativa	Presença de raposa (sim versus não)		Discrepância relativa	Grau de exposição do domicílio ao flebotômico (alto versus baixo)		Discrepância relativa
	individual	censitária		individual	censitária		individual	censitária	
Ajuste apenas por idade	4,67 ^a (1,16-18,67) ^a			9,90 ^a (1,89-51,84) ^a			3,40 ^a (1,32-8,75) ^a		
Abastecimento de água	5,32 (1,28-22,1)	4,82 (1,23-18,9)	-0,09	7,78 (1,21-49,9)	6,51 (1,11-38,2)	-0,16	3,63 (1,33-9,93)	3,38 (1,30-8,77)	-0,07
Esgotamento sanitário	4,74 (1,16-19,3)	5,02 (1,23-20,5)	0,06	9,90 (1,88-52,3)	10,1 (1,88-53,8)	0,02	3,35 (1,30-8,62)	3,41 (1,32-8,78)	0,02
Escolaridade	4,30 (1,06-17,4)	5,01 (1,27-19,8)	0,17	8,97 (1,67-48,1)	7,98 (1,44-44,1)	-0,11	3,05 (1,17-7,94)	3,26 (1,26-8,42)	0,07
Recolhimento de lixo	4,46 (1,12-17,8)	5,77 (1,43-23,2)	0,29	3,86 (0,57-25,9)	5,21 (0,87-31,3)	0,35	3,48 (1,32-9,16)	3,49 (1,32-9,19)	0,00
TOTAL	4,70 (1,12-19,7)	5,94 (1,43-24,6)	0,26	4,83 (0,60-38,7)	5,48 (0,85-35,1)	0,13	3,62 (1,29-10,2)	3,79 (1,41-10,2)	0,05

a) Razão de chances e intervalos de confiança de 95% ajustados por idade; todas as outras razões de chance e respectivos intervalos de confiança de 95% apresentados são ajustados por idade e variável socioeconômica.

Os resultados fornecidos pelo ajuste via variável censitária são, até certo ponto, decepcionantes. Se considerarmos um critério de similaridade entre as OR_i e OR_c baseado em uma discrepância relativa de 10% ou menos, pode-se perceber que, entre as 15 medidas

calculadas, 47% delas estariam fora desse critério, ou seja, poderiam ser consideradas medidas diferentes daquelas obtidas pela OR_i . Tomando-se a OR_i como "padrão-ouro", a especificidade e a sensibilidade da OR_c são de apenas 60% (6/10 e 3/5, respectivamente).

Ademais, nas três situações em que a OR_c indica a existência de confundimento em concordância com o OR_p , as duas medidas divergem em mais de 10%, levando a conclusões bastante diferentes sobre os efeitos postulados.

Discussão

Os resultados deste estudo indicam que os efeitos de variáveis socioeconômicas tendem a ser atenuados quando se utilizam dados obtidos em nível censitário, em substituição à informação individual, sugerindo um possível viés de má classificação não diferencial.

Os resultados apontam para a necessidade de uma reabilitação de abordagens mais integradas, onde o foco de investigação inclua não só fatores sociais e econômicos obtidos em nível individual, mas também as influências de características ambientais, culturais, sociais, grupais, na saúde das populações.

Um segundo aspecto evidenciado é o de que, tomando-se a variável definida em nível individual como a mais adequada para fins de controle de confundimento, observou-se que o uso de variáveis “censitárias” tende a não prover correção efetiva para o efeito de confundimento socioeconômico.

Os resultados aqui apresentados não podem ser generalizados facilmente, em função de, pelo menos, dois aspectos: (1) derivam de uma única experiência empírica, que, como tal, pode não traduzir as relações de ocorrência estudadas de maneira adequada, em função

de erros sistemáticos e/ou aleatórios; (2) mesmo que esses resultados sejam válidos e confiáveis, a experiência empírica captada é bastante particular, tanto no que diz respeito à população observada quanto às relações de ocorrência estudadas. Tendo em vista esses limites, os resultados apresentados podem ser úteis, como mínimo, para subsidiar reflexões sobre o tema.

Poder-se-ia supor, indutivamente, que a substituição da informação individual pela censitária seria uma estratégia inadequada, seja para estimar riscos associados aos próprios fatores socioeconômicos, seja para permitir um conveniente ajuste para confundimento. Considerando os limites de inferência mencionados no parágrafo anterior, talvez seja o caso de se utilizarem esses resultados para questionar até que ponto a substituição de um tipo de informação por outro configura a questão de fundo mais relevante. Pode-se argumentar que, como princípio básico, dados censitários não devem ser considerados, por si, apenas bons substitutos de dados individuais. Eles são, também, expressões de conceitos que operam em um outro nível de determinação e que, em certas situações, talvez possam ser utilizados para tal fim. A situação inversa – dados individuais substituindo dados coletivos – deve ser considerada da mesma forma.

Conclui-se, dos resultados apresentados, que, mais do que rejeitar a hipótese geral de que os dados censitários servem como bons substitutos para informações individuais, eles apontam para a necessidade de uma reabilitação de abordagens mais integradas, onde o foco de investigação inclua não só fatores sociais e econômicos obtidos em nível individual, mas também as influências de características ambientais, culturais, sociais, grupais, na saúde das populações. Atualmente, o uso de modelos de múltiplos níveis (multiníveis), em que variáveis individuais e contextuais são simultaneamente especificadas, parece ser a opção preferencial para a análise de dados socioeconômicos em epidemiologia.

Referências bibliográficas

1. Bronfman M, Tuirán RA. La Desigualdad social ante la muerte: clases sociales y mortalidad en la niñez. Cuadernos Medico Sociales 1984;29/30:53-75.
2. Marmot MG, Kongevinas M, Elston MA. Social/economic status and disease. Annual Review of Public Health 1987;8:111-135.

3. Williams DR. Socioeconomic differentials in health: a review and redirection. *Social Psychology Quarterly* 1990;53:81-99.
4. Link BG, Phelan J. Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior* 1995;Spec No:80-94.
5. Turner JB. Economic context and the health effects of unemployment. *Journal of Health and Social Behavior* 1995;36:213-229.
6. Lynch J. Social position and health (editorial). *Annals of Epidemiology* 1996;6:21-23.
7. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality, and mortality. *American Journal of Public Health* 1997;87:1491-1498.
8. Kennedy BP, Kawachi I, Prothrow-Stith D, Lochner K, Gupta V. Social capital, income inequality, and firearm violent crime. *Social Science and Medicine* 1998;47:7-17.
9. Muntaner C, Eaton WW, Diala C, Kessler RC, Sorlie PD. Social class, assets, organizational control and the prevalence of common groups of psychiatric disorders. *Social Science and Medicine* 1998;47:2043-2053.
10. Berkman LE, Kawachi I, editors. *Social epidemiology*. Oxford: Oxford University Press; 2000.
11. Liberatos P, Link BG, Kelsey JL. The Measurement of social class in epidemiology. *Epidemiologic Reviews* 1988;10:87-121.
12. Barros MBA. A Utilização do conceito de classe social nos estudos dos perfis epidemiológicos: uma proposta. *Revista de Saúde Pública* 1986;20:269-273.
13. Lombardi C, Bronfman M, Facchini LA, Victora CG, Barros FC, Béria JU, Teixeira AMB. Operacionalização do conceito de classe social em estudos epidemiológicos. *Revista de Saúde Pública* 1988;22:253-265.
14. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annual Review of Public Health* 1997;18:341-378.
15. Lynch J, Kaplan G. Socioeconomic position. In: Berkman LE, Kawachi I, editors. *Social epidemiology*. Oxford: Oxford University Press; 2000. p. 13-35.
16. Associação Brasileira de Pós-graduação em Saúde Coletiva. III Plano Diretor para o desenvolvimento da epidemiologia no Brasil 2000-2004. Rio de Janeiro: Abrasco; 2000. p. 35.
17. Krieger N. Social class and the black/white crossover in the age-specific incidence of breast cancer: a study linking census-derived data to population-based registry records. *American Journal of Epidemiology* 1990;131:804-814.
18. Chen FM, Breiman RF, Farley M, Plikaytis B, Deaver K, Cetron MS. Geocoding and linking data from population-based surveillance and the US census to evaluate the impact of median household income on the epidemiology of invasive *Streptococcus pneumoniae* infections. *American Journal of Epidemiology* 1998;148:1212-1218.
19. Kawachi I, Kennedy BP, Glass R. Social capital and self-rated health: a contextual analysis. *American Journal of Public Health* 1999;89:1187-1193.
20. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *American Journal of Public Health* 1992; 92:703-710.
21. Geronimus AT, Bound J, Neidert LJ. On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. *Journal of the American Statistical Association* 1996;91:529-537.
22. Woodward M. Small area statistics as markers for personal social status in the Scottish heart health study. *Journal of Epidemiology and Community Health* 1996;50:570-576.
23. Geronimus AT, Bound J. Use of census-based aggregate variables to proxy for socioeconomic group: evidence from national samples. *American Journal of Epidemiology* 1998;148:475-486.
24. IDRISI [computer program]. Version 2.007. Worcester, MA: Clark Labs; 1997.
25. Clark LA, Pregibon D. Tree-based models. In: Chambers JM, Hastie TJ, editors. *Statistical Models in S*. New York: Chapman & Hall; 1993. p. 377-419.
26. S-PLUS [computer program]. Version 4.5. Seattle, WA: Mathsoft; 1998.
27. Marzochi MCA, Marzochi KBE. Tegumentary and visceral leishmaniasis in Brazil - emerging anthro-pozoonosis and possibilities for their control. *Cadernos de Saúde Pública* 1994;10(Supl.2):359-375.

28. Quinnell RJ, Dye C. Correlates of the peridomestic abundance of *Lutzomyia longipalpis* (Diptera: Psychodidae) in Amazonian Brazil. *Medical and Veterinary Entomology* 1994; 8:219-224.
29. Wijeyaratne PM, Jones-Arsenault LK, Murphy CJ. Endemic disease and development: the leishmaniases. *Acta Tropica* 1994;56:349-364.
30. Intercooled Stata [computer program]. Version 6.0. College Station, TX: Stata Corporation; 1999.
31. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research: principles and quantitative methods*. New York: Van Nostrand Reinhold; 1982. p. 195.