

Estudo caso-controle com resposta multinomial: uma proposta de análise

Case-control studies with multinomial responses: a proposal for analysis

Ana Carolina Cintra Nunes Mafra ¹
Luciana Bertoldi Nucci ¹
Ricardo Cordeiro ¹
Celso Stephan ¹

¹ Universidade Estadual de
Campinas, Campinas, Brasil.

Correspondência
A. C. C. N. Mafra
Universidade Estadual de
Campinas.
Rua Tessália Vieira de
Camargo 126, Campinas, SP
13083-970, Brasil.
anacarol.nunes@gmail.com

Abstract

This study reviews articles on case-control studies in which the cases were classified in two or more types. Application of multinomial models and their adequacy for case-control studies are discussed. Among the available multinomial adjustments, we argue that the polytomous logistic model is the most suitable for obtaining epidemiological measures of risk and association in case-control studies. By way of illustration, we present an application of this model in a population-based case-control study, comparing the results with those obtained in a binomial logistic model. The multinomial approach allows investigating, in a single analysis, the occurrence of associations between covariates and more or more subclasses of cases, thus providing the epidemiologically relevant possibility of identifying individualized risk and protective factors for each subclass.

Case-Control Studies; Logistic Models; Statistical Analysis

Introdução

O entendimento de relações causais que poderiam explicar o surgimento de casos de uma doença é buscado comumente por meio da execução de estudos caso-controle. Neles, a distribuição de exposições associadas aos doentes analisados é comparada àquela existente na população onde habitam estes doentes ¹. Esse tipo de investigação propicia uma maneira eficiente de se obter estimativas populacionais de risco relativo e fração atribuível. A razão dessa eficiência advém da capacidade que o estudo caso-controle tem, a um custo relativamente baixo, de estimar a distribuição populacional de exposições a partir do seu grupo controle. Essa propriedade dispensa o consumo de tempo, bem como recursos financeiros e administrativos, na enumeração e seguimento de um grande conjunto de indivíduos, classificados de acordo com níveis de exposição ².

Em decorrência do caráter dual dos estudos caso-controle (que classifica indivíduos como doentes ou não doentes), neles, muito frequentemente, a relação entre exposição e adoecimento é modelada pela distribuição binomial, por meio de ajustes de modelos lineares generalizados, tendo como função de ligação o *logito* da probabilidade de doença. Entretanto, nem sempre é conveniente classificar dicotomicamente a população estudada. Pode ser interessante, para os propósitos do estudo, subdividir os casos analisa-

dos de acordo com características tais como gravidade, tipo celular, forma clínica etc., obtendo-se maior compreensão do desfecho e sua relação com as exposições estudadas. Essa subdivisão resulta em subconjuntos dos indivíduos estudados, formados pelos diferentes tipos de casos e pelos controles. Tal reestruturação requer um novo método de análise, pois nessas situações, modelos que utilizam a distribuição binomial não são adequados. Apesar de intuitiva, a análise de vários modelos binomiais distintos não leva em consideração a correta composição do grupo controle, que representa a população fonte de todos os tipos de casos; isso só seria adequado se houvesse um grupo controle correspondente a cada tipo de caso. Uma opção natural para a análise são modelos lineares generalizados com resposta multinomial, uma extensão da distribuição binomial para mais de dois resultados possíveis³. Entretanto, a escolha e aplicação do modelo multinomial deve ser adequada ao delineamento do estudo.

Este artigo revê o uso de modelos multinomiais com foco em estudos caso-controle. Como ilustração foram analisados dados de um estudo de dengue.

Estudos caso-controle com resposta multinomial

Foi feita uma revisão bibliográfica na base PubMed com os termos *multinomial case control*; *ordinal case control* e *nominal case control*, e na base SciELO com os mesmos termos em português no mês de setembro de 2009. Tal revisão resultou em 448 artigos, sendo 445 na base PubMed e três artigos na base SciELO, após a exclusão das duplicações. Desses, focou-se a pesquisa em publicações com delineamento do estudo tipo caso-controle em que os casos são subdivididos em dois ou mais grupos por tipo ou gravidade do caso. Em razão disso, 247 foram excluídos por não apresentar delineamento do tipo caso-controle; 102 pelo desfecho não ter sido subdividido, ou seja, a resposta permaneceu binária; 57 pela resposta ter sido estratificada por variáveis que não dividiam os casos em subtipos, como variáveis relacionadas à exposição, ou grupos de controles e; em 11 casos não se pôde acessar as publicações. Dos 31 trabalhos restantes, dois eram metodológicos^{4,5} e serão comentados mais à frente; dois, por serem estudos caso-controle emparelhados, fizeram uso de métodos Bayesianos para a análise^{6,7}; dois fizeram apenas análises descritivas^{8,9}; três consideraram o desfecho ordinal, e destes, um utilizou para a análise o modelo de *Odds* Proporcionais¹⁰ e os outros dois

analisaram apenas os casos por meio do modelo politômico^{11,12}; quatro trabalhos ajustaram diversos modelos binomiais (um para cada tipo de caso)^{13,14,15,16}; um utilizou um modelo denominado *modified polytomous logit model*¹⁷; e 13 utilizaram o modelo politômico com todos os dados^{18,19,20,21,22,23,24,25,26,27,28,29,30}, porém dois deles não usaram a categoria de controles para as comparações^{25,28}.

Em toda a revisão feita observa-se que, apesar de grande parte dos artigos ter optado por utilizar em suas análises o modelo politômico (48%), não há na literatura discussão sobre esta escolha e sobre a adequação destas análises ao delineamento dos estudos caso-controle e aos pressupostos epidemiológicos relacionados a este.

A análise desse tipo de dados pressupõe a representação da população fonte de casos por meio do grupo controle. As medidas epidemiológicas de risco e associação resultantes, tais como razão de proporção de incidência, razão de taxa de incidência e fração atribuível sempre comparam os casos com o grupo de controles². Entretanto, quando se tem mais de um grupo de casos, a pressuposição de que os controles representam a população contínua a mesma, devendo este grupo ser a categoria de referência em todas as comparações realizadas na análise. Caso contrário, as estimativas obtidas com o ajuste do modelo não podem ser utilizadas para a obtenção das medidas epidemiológicas citadas. Nota-se, portanto, que o melhor ajuste deve considerar esse fato.

Modelos multinomiais

Nos ajustes multinomiais existe uma distinção importante entre as abordagens nominal e ordinal. Essa pressupõe uma ordem ou uma gradação entre as categorias do fenômeno estudado^{31,32}. São exemplos de modelos multinomiais ordinais os modelos Categorias Adjacentes³³, Razão Contínua³⁴, *Odds* Proporcionais³⁵, entre outros. O modelo logístico de Categorias Adjacentes compara entre si duas categorias adjacentes ao longo de uma escala ordinal. O modelo de Razão Contínua compara entre si uma categoria com todas as outras superiores na escala ordinal. O modelo de *Odds* Proporcionais compara entre si categorias abaixo e acima de pontos de corte ao longo de uma escala ordinal e pressupõe proporcionalidade entre os *odds* das categorias. Embora seja crescente o uso desses modelos em estudos epidemiológicos, sua aplicação especificamente em estudos caso-controle não é adequada, pois em nenhum deles consegue-se comparar a distribuição de exposições entre casos e a população fon-

te, propriedade fundamental para a estimação de riscos em estudos caso-controle. A escolha do ajuste que define o modelo multinomial é determinante para a adequabilidade das análises.

Muito embora seja freqüente categorizarmos casos de acordo com uma escala ordinal (por exemplo, hipertenso leve, moderado e grave), a melhor escolha para a modelagem da relação entre exposições e ocorrência de casos assim classificados é a abordagem nominal, tendo os controles como categoria de base. Epidemiologicamente, essa é a única opção, dentre as possibilidades de análise utilizando a distribuição multinomial, que mantém a lógica do estudo caso-controle, em que a distribuição de exposições dos casos incidentes (agora, classificados em diferentes grupos) é comparada à distribuição de exposições na população que gerou estes casos, estimada por meio do grupo controle. Na abordagem nominal isso é feito principalmente por meio do modelo logístico Politômico³⁶, no qual, escolhendo-se como referência a categoria de controles, compara-se esta, um a um, com todos os grupos de casos. Aliás, esse modelo não possui restrição quanto à ordinalidade da resposta, podendo ser usado tanto em modelos multinomiais ordinais quanto nominais³².

Ahn et al.⁴ discutem o modelo Estereótipo, que pode ser aplicado tanto para desfechos categóricos ordinais quanto nominais, pois independente das restrições de ordem e tem a vantagem de permitir inferências com amostragem dependente do desfecho, como nos estudos de caso-controle. Apesar da crescente discussão acerca de inferências, tanto paramétrica quanto não-paramétrica (Bayesiana), para dados ordinais, o modelo Estereótipo não é tão usado, principalmente por sua complexidade computacional e seus rígidos pressupostos.

No caso específico de estudos caso-controle emparelhado com resposta ordinal, Mukherjee et al.³⁷ discutem o uso de modelos logísticos quando se tem um grupo controle e os casos subdivididos de acordo com a ordem natural de gravidade da doença. Os autores propõem o modelo condicional de Categorias Adjacentes comparando-o com o modelo Politômico, que ignora a ordem natural de gravidade da doença, e com o modelo de regressão logística condicional, que agrupa os diferentes casos em uma única categoria. O foco da comparação é o tratamento do pareamento nos ajustes, e o artigo não aborda as especificidades do delineamento epidemiologicamente.

Ilustração: dengue no Distrito Sul de Campinas

A seguir, apresentamos uma ilustração da aplicação do modelo multinomial politômico, anteriormente descrito, na análise de um estudo caso-controle no qual os casos foram subdivididos em dois grupos. Tal aplicação se estende trivialmente para situações em que os casos são classificados em três ou mais classes.

Realizou-se um estudo caso-controle de base populacional no Distrito Sul da cidade de Campinas, Estado de São Paulo, Brasil³⁸. Os casos foram indivíduos residentes nessa área, com diagnóstico de dengue confirmado laboratorialmente, cujo início dos sintomas ocorreu entre outubro de 2006 e maio de 2007. Os controles foram moradores da região, sorteados com base em um cadastro universal, que não apresentaram a doença no período de estudo. Casos e controles foram entrevistados, indagando-se sobre diversas covariáveis clínicas, biológicas, sociais e ambientais. A amostra total resultou em 727 casos e 746 controles. Para o presente estudo, foram excluídos os menores de 20 anos, resultando em 555 casos e 736 controles. Em razão de algumas perdas de informação, foram analisados 540 casos e 730 controles (98,4%). Com base nas entrevistas e em dados clínicos e laboratoriais, os casos foram classificados em leves e graves (não leves), resultando em 379 casos leves e 161 casos graves.

A Tabela 1 descreve alguns aspectos da população estudada, salientando características sócio-demográficas e variáveis relacionadas ao dengue.

Ajustes dos modelos lineares generalizados

Inicialmente, foram ajustados modelos logísticos binomiais, tendo como variável resposta o *status* do indivíduo (caso ou controle) e como variável preditora, uma a uma, as co-variáveis de interesse para o estudo. Baseando-se nesses modelos simples foi ajustado um modelo binomial múltiplo, contendo como variáveis preditoras aquelas cujos valores de *p* das estimativas de “efeito” foram menores que 0,25 nos ajustes simples e que, após o ajuste múltiplo, obtiveram valor de *p* inferior a 0,05.

A seguir, ajustou-se um modelo multinomial politômico com três níveis, tendo como variável resposta o *status* do indivíduo (controle, caso leve, caso grave) e como variáveis preditoras aquelas utilizadas no ajuste logístico binomial anteriormente descrito. A categoria de base foi o grupo de controles.

Tabela 1

Proporção (%) de indivíduos em cada classe das variáveis analisadas para casos e controles. Estudo de dengue no Distrito Sul de Campinas, São Paulo, Brasil, 2007.

Variável	Casos (n = 540)	Controles (n = 730)	Total (N = 1.270)
Sexo			
Masculino	45,00	44,80	44,90
Feminino	55,00	55,20	55,10
Idade (anos)			
≤ 40	53,10	38,50	44,70
40-60	36,70	45,60	41,80
> 60	10,20	15,90	13,50
Escolaridade (anos de estudo)			
≤ 4	39,10	30,10	33,90
4-8	24,40	21,90	23,00
> 8	36,50	47,90	43,10
Trabalha			
Sim	65,90	51,90	57,90
Não	34,10	48,10	42,10
Já teve dengue			
Sim	2,60	2,70	2,70
Não	97,40	97,30	97,30
Enchente/Lama			
Sim	27,40	7,40	15,90
Não	72,60	92,60	84,10
Saneamento básico			
Rede	56,90	81,90	71,30
Fossa	18,90	13,30	15,70
Céu aberto	24,30	4,80	13,10

Para a análise da qualidade do ajuste dos modelos ajustados foram analisados o logaritmo da função de verossimilhança, o desvio (ou *deviance*) e o Critério de Informação de Akaike (AIC).

Resultados

A Tabela 2 mostra a análise binomial, descrevendo estatísticas obtidas em um ajuste logístico para a população estudada classificada dicotomicamente como casos e controles, em que foram testadas as co-variáveis selecionadas nos ajustes binomiais simples.

A Tabela 3 mostra as estatísticas obtidas no ajuste multinomial múltiplo. As variáveis selecionadas no ajuste binomial múltiplo (Tabela 2) mantiveram-se no modelo.

As estimativas nas abordagens binomial e multinomial são semelhantes, exceto para a variável escolaridade do chefe, que no ajuste binomial foi significativa como fator de proteção e no ajuste multinomial também se mostrou fator de

proteção; entretanto, não foi significativa para casos graves, demonstrando importância apenas para os casos leves.

No ajuste multinomial observa-se que as variáveis período de trabalho diurno, armazenamento de água, presença de mosquito picando, falta de saneamento básico (céu aberto) e presença de criadouros implicavam riscos mais elevados para casos graves do que para casos leves. Ao passo que para a variável ocorrência de enchente e/ou lama isso se inverte, o risco é maior para os casos leves do que para os casos graves.

As medidas de qualidade do ajuste, tanto do modelo binomial múltiplo quanto do multinomial múltiplo, estão na Tabela 4. Podemos notar que ambos os modelos múltiplos estão bem ajustados.

Tabela 2

Medidas obtidas através de ajuste de Modelo Binomial Múltiplo. Resposta: casos vs. controles (n = 1.270). Estudo de dengue no Distrito Sul de Campinas, São Paulo, Brasil, 2007.

Variável	Estimativa β	Erro padrão	Valor de p	OR	IC95%	
					Limite inferior	Limite superior
Idade	-0,016	0,005	0,001	0,985	0,975	0,994
Escolaridade do chefe	-0,036	0,017	0,031	0,965	0,933	0,997
Período de trabalho						
Não trabalha						
Diurno	0,837	0,141	< 0,001	2,309	1,751	3,046
Noturno	-0,276	0,392	0,482	0,759	0,352	1,637
Armazena água						
Não						
Sim	1,002	0,266	< 0,001	2,724	1,618	4,587
Freqüência da coleta de lixo (por semana)						
< 3 vezes						
3-5 vezes	-1,098	0,252	< 0,001	0,334	0,203	0,547
> 5 vezes	-0,930	0,277	0,001	0,394	0,229	0,679
Enchente/Lama						
Não						
Sim	0,764	0,203	< 0,001	2,146	1,441	3,198
Presença de mosquito picando						
Não						
Sim	0,355	0,137	0,010	1,426	1,091	1,865
Saneamento básico						
Rede						
Fossa	0,172	0,195	0,376	1,188	0,811	1,739
Céu aberto	0,879	0,262	0,001	2,408	1,440	4,027
Criadouro de mosquito						
Não						
Sim	0,578	0,136	< 0,001	1,783	1,367	2,327

OR: odds ratio; IC95%: intervalo de 95% de confiança.

Discussão

A análise multinomial permite a identificação diferenciada entre os fatores de risco para cada categoria de casos, o que não pode ser visto quando se realiza apenas uma análise binomial. A opção de se realizar várias análises binomiais não pode ser aplicada a estudos caso-controle por inadequação conceitual, conforme comentado anteriormente.

Não foi encontrada estatística para comparação quantitativa desses dois ajustes. Isso ocorre em razão de diferentes variáveis respostas para cada um dos modelos. Essa mudança impede a utilização de quaisquer das estatísticas comumente utilizadas para comparação de modelos ou critérios de seleção de modelos. Podemos observar apenas as medidas de qualidade do ajuste

de cada um, sem compará-las por meio de teste estatístico. Portanto, a avaliação dessas medidas deve ser realizada e a escolha será aquela que for mais conveniente às análises epidemiológicas.

É importante ressaltar que, na ilustração apresentada, ao se realizar o ajuste que considera a gravidade do caso de dengue, observaram-se fatores nos quais a intensidade de risco/proteção alterou-se para tipos específicos de casos e em um deles até deixou de ser significativo estatisticamente.

Outra análise a ser considerada é a utilização do modelo Estereótipo, que não pressupõe independência do desfecho no processo de amostragem, pois considera tanto desfechos nominais quanto ordinais e permite que se escolha a categoria de controles para a comparação com as demais. Com essas características, apesar de fac-

Tabela 3

Medidas obtidas através de ajuste de Modelo Multinomial Múltiplo. Resposta: casos leves, casos graves e controles (n = 1.270). Estudo de dengue no Distrito Sul de Campinas, São Paulo, Brasil, 2007.

Variável	PARTE A: COMPARAÇÃO ENTRE CASOS LEVES E CONTROLES					
	Estimativa β	Erro padrão	Valor de p	OR	IC95%	
					Limite inferior	Limite superior
Idade	-0,010	0,005	0,047	0,990	0,980	1,000
Escolaridade do chefe	-0,037	0,018	0,045	0,964	0,930	0,999
Período de trabalho						
Não trabalha						
Diurno	0,796	0,153	< 0,001	2,217	1,642	2,993
Noturno	-0,291	0,435	0,503	0,747	0,319	1,752
Armazena água						
Não						
Sim	0,925	0,282	0,001	2,523	1,451	4,386
Frequência da coleta de lixo (por semana)						
< 3 vezes						
3-5 vezes	-1,186	0,265	< 0,001	0,306	0,182	0,514
> 5 vezes	-0,925	0,293	0,002	0,306	0,182	0,514
Enchente/Lama						
Não						
Sim	0,813	0,215	< 0,001	2,254	1,479	3,436
Presença de mosquito picando						
Não						
Sim	0,296	0,148	0,046	1,345	1,005	1,799
Saneamento básico						
Rede						
Fossa	0,180	0,214	0,400	1,197	0,787	1,821
Céu aberto	0,839	0,280	0,003	2,314	1,338	4,004
Criadouro de mosquito						
Não						
Sim	0,511	0,148	0,001	1,667	1,246	2,230

(continua)

tível ao delineamento de estudos caso-controle, esse modelo ainda tem sua aplicabilidade restrita. A não linearidade dos parâmetros adicionais torna inviável a estimação computacional por modelos lineares generalizados já implementados nos programas estatísticos, exigindo algoritmos complexos³⁹ ou estimativas Bayesianas recentemente propostas⁴.

No caso particular de estudo caso-controle emparelhado com resposta ordinal, o modelo condicional de Categorias Adjacentes pode ser aplicado desde que a pressuposição de proporcionalidade dos *odds* seja mantida para evitar perda de poder do modelo proposto, o que não é necessário no modelo Político³⁷.

Concluindo, nos marcos de um estudo caso-controle, a análise mais genérica – utilizando-se

modelos lineares generalizados – que estima relações entre adoecimento e co-variáveis de interesse é o clássico modelo logístico binomial. Tal análise classifica a população estudada dicotomicamente em doentes e não doentes, o que não permite identificar eventuais diferenças nos efeitos de fatores de risco associadas aos diferentes grupos de doentes. Quando os casos estudados são subdivididos em diferentes classes, a abordagem multinomial politômica abre a possibilidade de investigar, em um único modelo, a ocorrência de associações entre co-variáveis e uma ou algumas das classes de casos particularmente, permitindo a identificação de fatores de risco individualizados para cada subgrupo de casos. Além disso, nesse contexto, entendendo-se a doença estudada como um processo único, que

Tabela 3 (continuação)

Variável	PARTE B: COMPARAÇÃO ENTRE CASOS GRAVES E CONTROLES					IC95%	
	Estimativa β	Erro padrão	Valor de p	OR	Limite inferior	Limite superior	
Idade	-0,030	0,007	< 0,001	0,970	0,956	0,985	
Escolaridade do chefe	-0,034	0,025	0,183	0,967	0,920	1,016	
Período de trabalho							
Não trabalha							
Diurno	0,973	0,210	< 0,001	2,646	1,753	3,994	
Noturno	-0,211	0,597	0,723	0,810	0,251	2,607	
Armazena água							
Não							
Sim	1,1176	0,322	< 0,001	3,240	1,722	6,097	
Freqüência da coleta de lixo (por semana)							
< 3 vezes							
3-5 vezes	-0,891	0,328	0,007	0,410	0,216	0,780	
> 5 vezes	-0,984	0,389	0,011	0,374	0,174	0,802	
Enchente/Lama							
Não							
Sim	0,637	0,269	0,018	1,890	1,115	3,205	
Presença de mosquito picando							
Não							
Sim	0,508	0,206	0,014	1,662	1,110	2,489	
Saneamento básico							
Rede							
Fossa	0,153	0,276	0,579	1,165	0,679	2,001	
Céu aberto	0,991	0,347	0,004	2,695	1,364	5,324	
Criadouro de mosquito							
Não							
Sim	0,754	0,205	< 0,001	2,125	1,422	3,177	

OR: odds ratio; IC95%: intervalo de 95% de confiança.

Tabela 4

Medidas de qualidade do ajuste dos modelos Binomial Múltiplo e Multinomial Múltiplo. Estudo de dengue no Distrito Sul de Campinas, São Paulo, Brasil, 2007.

Modelo Múltiplo	Log-verossimilhança	AIC	Desvio (modelo nulo-completo)	Valor de p (desvio - χ^2)
Binomial	-722,55	1471,1	282,4	< 0,001
Multinomial	-1042,534	2137,068	305,024	< 0,001

AIC: Critério de Informação de Akaike.

se manifesta de diferentes formas, a realização de uma análise multinomial que apreenda todo o processo parece ser, do ponto de vista epide-

miológico, mais adequada do que a realização da análise binomial para interpretar o mesmo processo.

Resumo

Este estudo revisa alguns artigos com delineamento do tipo caso-controle em que se tem os casos classificados em dois ou mais tipos. A aplicação de modelos para respostas multinomiais e a adequabilidade destes aos estudos caso-controle é discutida. Dentre os ajustes possíveis, argumenta-se que o modelo logístico Político é o mais indicado para a obtenção de medidas epidemiológicas de risco e associação neste tipo de delineamento. Como ilustração, apresenta-se uma aplicação desse modelo em um estudo caso-controle de base populacional, comparando os resultados com aqueles obtidos em um ajuste logístico binomial. A abordagem multinomial possibilita investigar, em uma única análise, a ocorrência de associações entre co-variáveis e uma ou algumas das classes de casos estudados, abrindo a possibilidade, de grande interesse epidemiológico, de identificação de fatores de risco e de proteção individualizados para cada subgrupo de interesse.

Estudos de Casos e Controles; Modelos Logísticos; Análise Estatística

Colaboradores

A. C. C. N. Mafra, L. B. Nucci e R. Cordeiro participaram da revisão e sintetização da literatura sobre modelos multinomiais, discussão sobre a concepção do artigo e redação do texto. C. Stephan auxiliou na atualização da revisão bibliográfica, discussão e respostas aos pareceristas.

Agradecimentos

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento deste estudo por meio dos processos 2006/05920-7, 2007/01892-1 e 2007/07435-1.

Referências

- Rothman KJ, Greenland S. Modern epidemiology. Philadelphia: Lippincott Raven; 1998.
- Cordeiro R. O mito da doença rara. Rev Bras Epidemiol 2005; 8:111-6.
- Dubin N, Pasternack BS. Risk assessment for case-control subgroups by polychotomous logistic regression. Am J Epidemiol 1986; 123:1101-17.
- Ahn J, Mukherjee B, Banerjee M, Cooney KA. Bayesian inference for the stereotype regression model: application to a case-control study of prostate cancer. Stat Med 2009; 28:3139-57.
- Mueller JC, Riemenschneider M, Schoepfer-Wendels A, Gohlke H, Konta L, Friedrich P, et al. Weak independent association signals between IDE polymorphisms, Alzheimer's disease and cognitive measures. Neurobiol Aging 2007; 28:727-34.
- Liu J, Gustafson P, Cherry N, Burstyn I. Bayesian analysis of a matched case-control study with expert prior information on both the misclassification of exposure and the exposure-disease association. Stat Med 2009; 28:3411-23.
- Sinha S, Mukherjee B, Ghosh M. Bayesian semi-parametric modeling for matched case-control studies with multiple disease states. Biometrics 2004; 60:41-9.
- De La Paz MA, Zhang J, Fridovich I. Red blood cell antioxidant enzymes in age-related macular degeneration. Br J Ophthalmol 1996; 80:445-50.
- Hattori E, Ebihara M, Yamada K, Ohba H, Shibuya H, Yoshikawa T. Identification of a compound short tandem repeat stretch in the 5'-upstream region of the cholecystokinin gene, and its association with panic disorder but not with schizophrenia. Mol Psychiatry 2001; 6:465-70.
- Tishler PV, Larkin EK, Schluchter MD, Redline S. Incidence of sleep-disordered breathing in an urban adult population: the relative importance of risk factors in the development of sleep-disordered breathing. JAMA 2003; 289:2230-7.
- Lee BH, Schofer JL, Koppelman FS. Bicycle safety helmet legislation and bicycle-related non-fatal injuries in California. Accid Anal Prev 2005; 37: 93-102.

12. Soran A, D'Angelo G, Begovic M, Ardic F, Harlak A, Samuel Wieand H, et al. Breast cancer-related lymphedema: what are the significant predictors and how they affect the severity of lymphedema? *Breast J* 2006; 12:536-43.
13. Hasselmann MH, Reichenheim ME. Parental violence and the occurrence of severe and acute malnutrition in childhood. *Paediatr Perinat Epidemiol* 2006; 20:299-311.
14. Hindorf LA, Lemaitre RN, Smith NL, Bis JC, Marciano KD, Rice KM, et al. Common genetic variation in six lipid-related and statin-related genes, statin use and risk of incident nonfatal myocardial infarction and stroke. *Pharmacogenet Genomics* 2008; 18:677-82.
15. Mavaddat N, Dunning AM, Ponder BA, Easton DF, Pharoah PD. Common genetic variation in candidate genes and susceptibility to subtypes of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2009; 18:255-9.
16. Zammit W, Mtiraoui N, Kallel C, Mercier E, Al-mawi WY, Mahjoub T. A case-control study on the association of idiopathic recurrent pregnancy loss with autoantibodies against beta2-glycoprotein I and annexin V. *Reproduction* 2006; 131:817-22.
17. Studdert DM, Bhattacharya J, Schoenbaum M, Warren B, Escarce JJ. Personal choices of health plans by managed care experts. *Med Care* 2002; 40:375-86.
18. Balestrieri M, Rucci P, Sbrana A, Ravani L, Benvenuti A, Gonnelli C, et al. Lifetime rhythmicity and mania as correlates of suicidal ideation and attempts in mood disorders. *Compr Psychiatry* 2006; 47:334-41.
19. Bonner MR, McCann SE, Moysich KB. Dietary factors and the risk of testicular cancer. *Nutr Cancer* 2002; 44:35-43.
20. Conner KR, Beautrais AL, Conwell Y. Moderators of the relationship between alcohol dependence and suicide and medically serious suicide attempts: analyses of Canterbury Suicide Project data. *Alcohol Clin Exp Res* 2003; 27:1156-61.
21. del Rincon I, Haas RW, Pogolian S, Escalante A. Lower limb arterial incompressibility and obstruction in rheumatoid arthritis. *Ann Rheum Dis* 2005; 64:425-32.
22. Kagen J, Lautenbach E, Bilker WB, Matro J, Bell LM, Dominguez TE, et al. Risk factors for mediastinitis following median sternotomy in children. *Pediatr Infect Dis J* 2007; 26:613-8.
23. Le Moual N, Siroux V, Pin I, Kauffmann F, Kennedy SM. Asthma severity and exposure to occupational asthrogens. *Am J Respir Crit Care Med* 2005; 172:440-5.
24. Palli D, Russo A, Ottini L, Masala G, Saieva C, Amorusi A, et al. Red meat, family history, and increased risk of gastric cancer with microsatellite instability. *Cancer Res* 2001; 61:5415-9.
25. Pezzini A, Del Zotto E, Magoni M, Costa A, Archetti S, Grassi M, et al. Inherited thrombophilic disorders in young adults with ischemic stroke and patent foramen ovale. *Stroke* 2003; 34:28-33.
26. Schroeder JC, Weinberg CR. Use of missing-data methods to correct bias and improve precision in case-control studies in which cases are subtyped but subtype information is incomplete. *Am J Epidemiol* 2001; 154:954-62.
27. Tan Q, Christiansen L, Christensen K, Kruse TA, Bathum L. Apolipoprotein E genotype frequency patterns in aged Danes as revealed by logistic regression models. *Eur J Epidemiol* 2004; 19:651-6.
28. Unicomb LE, O'Reilly LC, Kirk MD, Stafford RJ, Smith HV, Becker NG, et al. Risk factors for infection with *Campylobacter jejuni* flaA genotypes. *Epidemiol Infect* 2008; 136:1480-91.
29. Yang YF, Li H, Xu XQ, Diao YT, Fang XQ, Wang Y, et al. An expression of squamous cell carcinoma antigen 2 in peripheral blood within the different stages of esophageal carcinogenesis. *Dis Esophagus* 2008; 21:395-401.
30. Zanetti R, Rosso S, Martinez C, Nieto A, Miranda A, Mercier M, et al. Comparison of risk patterns in carcinoma and melanoma of the skin in men: a multi-centre case-case-control study. *Br J Cancer* 2006; 94:743-51.
31. Abreu MNS, Siqueira AL, Cardoso CS, Caiiffa WT. Ordinal logistic regression models: application in quality of life studies. *Cad Saúde Pública* 2008; 24 Suppl 4:S581-91.
32. Ananth CV, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. *Int J Epidemiol* 1997; 26:1323-33.
33. Agresti A. Categorical data analysis. 2nd Ed. New York: John Wiley & Sons; 2002.
34. Feinberg B. Analysis of cross-classified data. 2nd Ed. Cambridge: Massachusetts Institute of Technology Press; 1980.
35. McCullagh P. Regression models for ordinal data (with discussion). *J R Stat Soc Series B Stat Methodol* 1980; 42:109-42.
36. Engel J. Polytomous logistic regression. *Stat Neerl* 1988; 42:233-52.
37. Mukherjee B, Liu I, Sinha S. Analysis of matched case-control data with multiple ordered disease states: possible choices and comparisons. *Stat Med* 2007; 26:3240-57.
38. Andrade VR. Distribuição espacial do risco de dengue em região do Município de Campinas [Tese de Doutorado]. Campinas: Universidade Estadual de Campinas; 2009.
39. Greenland S. Alternative models for ordinal logistic regression. *Stat Med* 1994; 13:1665-77.

Recebido em 08/Jan/2009

Versão final reapresentada em 30/Set/2009

Aprovado em 16/Out/2009