

Leticia M S Coutinho<sup>1,II</sup>

Marcia Scazufca<sup>II,III</sup>

Paulo R Menezes<sup>I,II</sup>

# Métodos para estimar razão de prevalência em estudos de corte transversal

## Methods for estimating prevalence ratios in cross-sectional studies

---

### RESUMO

**OBJETIVO:** Comparar empiricamente as regressões de Cox, log-binomial, Poisson e logística para estimar razões de prevalência em estudos de corte transversal.

**MÉTODOS:** Foram utilizados dados de um estudo epidemiológico transversal (n=2.072), de base populacional, realizado com idosos na cidade de São Paulo (SP), entre maio de 2003 e abril de 2005. Diagnósticos de demência, possíveis casos de transtorno mental comum e autopercepção de saúde ruim foram escolhidos como desfechos com prevalência baixa, intermediária e alta, respectivamente. Foram utilizadas variáveis de confusão com duas ou mais categorias ou valores contínuos. Valores de referência para estimativas por ponto e por intervalo para as razões de prevalência (RP) foram obtidos pelo método de estratificação de Mantel-Haenszel. Estimativas ajustadas foram calculadas utilizando regressões de Cox e Poisson com variância robusta, e regressão log-binomial. *Odds ratios* (OR) brutos e ajustados foram obtidos pela regressão logística.

**RESULTADOS:** As estimativas por ponto e por intervalo obtidas pelas regressões de Cox e Poisson foram semelhantes à obtida pela estratificação de Mantel-Haenszel, independentemente da prevalência do desfecho e das covariáveis do modelo. O modelo log-binomial apresentou dificuldade de convergência quando o desfecho tinha prevalência alta e havia covariável contínua no modelo. A regressão logística produziu estimativas por ponto e por intervalo maiores do que as obtidas pelos outros métodos, principalmente para os desfechos com maiores prevalências iniciais. Se interpretados como estimativas de RP, os OR superestimariam as associações para os desfechos com prevalência inicial baixa, intermediária e alta em 13%, quase 100% e quatro vezes mais, respectivamente.

**CONCLUSÕES:** Em análise de dados de estudos de corte transversal, os modelos de Cox e Poisson com variância robusta são melhores alternativas que a regressão logística. O modelo de regressão log-binomial produz estimativas não enviesadas da RP, mas pode apresentar dificuldade de convergência quando o desfecho é muito freqüente e a variável de confusão é contínua.

**DESCRITORES:** Estudos Transversais. Técnicas de Estimativa. Razão de Prevalências. Modelos Logísticos. Estudo Comparativo.

<sup>I</sup> Departamento de Medicina Preventiva. Faculdade de Medicina Universidade de São Paulo. São Paulo, SP, Brasil

<sup>II</sup> Núcleo de Epidemiologia. Hospital Universitário. Universidade de São Paulo. São Paulo, SP, Brasil

<sup>III</sup> Departamento de Psiquiatria. Faculdade de Medicina Universidade de São Paulo. São Paulo, SP, Brasil

#### Correspondência | Correspondence:

Paulo Rossi Menezes  
Departamento de Medicina Preventiva  
Faculdade de Medicina da Universidade de São Paulo  
Av. Dr. Arnaldo 455  
01246-903 São Paulo, SP, Brasil  
E-mail: pmenezes@usp.br

Recebido: 27/11/2007

Revisado: 13/5/2008

Aprovado: 4/6/2008

---

## ABSTRACT

**OBJECTIVE:** To empirically compare the Cox, log-binomial, Poisson and logistic regressions to obtain estimates of prevalence ratios (PR) in cross-sectional studies.

**METHODS:** Data from a population-based cross-sectional epidemiological study (n = 2072) on elderly people in Sao Paulo (Southeastern Brazil), conducted between May 2003 and April 2005, were used. Diagnoses of dementia, possible cases of common mental disorders and self-rated poor health were chosen as outcomes with low, intermediate and high prevalence, respectively. Confounding variables with two or more categories or continuous values were used. Reference values for point and interval estimates of prevalence ratio (PR) were obtained by means of the Mantel-Haenszel stratification method. Adjusted PR estimates were calculated using Cox and Poisson regressions with robust variance, and using log-binomial regression. Crude and adjusted odds ratios (ORs) were obtained using logistic regression.

**RESULTS:** The point and interval estimates obtained using Cox and Poisson regressions were very similar to those obtained using Mantel-Haenszel stratification, independent of the outcome prevalence and the covariates in the model. The log-binomial model presented convergence difficulties when the outcome had high prevalence and there was a continuous covariate in the model. Logistic regression produced point and interval estimates that were higher than those obtained using the other methods, particularly when for outcomes with high initial prevalence. If interpreted as PR estimates, the ORs would overestimate the associations for outcomes with low, intermediate and high prevalence by 13%, almost by 100% and fourfold, respectively.

**CONCLUSIONS:** In analyses of data from cross-sectional studies, the Cox and Poisson models with robust variance are better alternatives than logistic regression is. The log-binomial regression model produces unbiased PR estimates, but may present convergence difficulties when the outcome is very prevalent and the confounding variable is continuous.

**DESCRIPTORS:** Cross-Sectional Studies. Estimation Techniques. Prevalence Ratio. Logistic Models. Comparative Study.

---

## INTRODUÇÃO

Em estudos de corte transversal com desfechos binários, a associação entre exposição e desfecho é estimada pela razão de prevalência (RP). Quando é necessário ajustar para potenciais variáveis de confusão, normalmente são usados modelos de regressão logística. Este tipo de modelo produz estimativas de *odds ratios* (OR), frequentemente interpretado como uma estimativa da RP. Entretanto, o OR não se aproxima muito bem da RP quando o risco inicial é alto, e nessas situações, interpretar o OR como se fosse a RP é inadequado.<sup>1,2,9,12</sup>

Alguns modelos estatísticos alternativos que podem estimar diretamente a RP e seu intervalo de confiança vêm sendo discutidos na literatura.<sup>1,4,6,10,12,14</sup> Modelos de regressão de Cox, log-binomial e Poisson têm sido sugeridos como boas alternativas para obter estimativas da RP ajustadas para variáveis de confusão. Usando

dados adaptados de um estudo de corte transversal, Barros & Hirakata<sup>1</sup> (2003) mostraram que estes modelos produzem estimativas ajustadas da RP muito similares àquelas obtidas pelo método de Mantel-Haenszel (MH).

O objetivo do presente estudo foi comparar empiricamente os modelos de regressão de Cox, log-binomial, Poisson e logístico para estimar a RP ajustada, cotejando seus resultados com os obtidos pelo método de MH.

## MÉTODOS

Os dados utilizados foram obtidos de um estudo de corte transversal, de base populacional, cujo objetivo foi estimar a prevalência de demência e outros problemas de saúde mental em idosos (65 anos ou mais),

residentes em uma área economicamente precária do distrito do Butantã, na cidade de São Paulo (SP), entre maio de 2003 e abril de 2005.<sup>8</sup> Procedimentos padronizados foram utilizados para avaliar funcionamento cognitivo e sintomas psiquiátricos. Foram coletadas informações de características sociodemográficas e socioeconômicas. Um total de 2.072 participantes foi incluído no estudo.

Três desfechos foram escolhidos: diagnóstico de demência, possíveis casos de transtorno mental comum (TMC), e autopercepção de saúde ruim. O diagnóstico de demência foi obtido por um procedimento desenvolvido pelo “Grupo de Pesquisa em Demência 10/66”, para ser usado em estudos de base populacional em países em desenvolvimento, com uma avaliação detalhada do início e curso da demência.<sup>7</sup> A classificação dos participantes como possíveis casos de TMC foi obtida com o *Self Report Questionnaire* (SRQ-20), um questionário desenvolvido pela Organização Mundial da Saúde para estudos em países em desenvolvimento.<sup>11</sup> O ponto de corte utilizado foi 4/5, de acordo com a validação da versão brasileira do SRQ-20.<sup>9</sup> Autopercepção de saúde foi avaliada com uma simples questão (“No geral, como você classificaria sua saúde nos últimos 30 dias?”), com as opções de respostas (“muito boa”, “boa”, “regular”, “ruim”, e “muito ruim”) agrupadas para classificar os participantes como tendo autopercepção de saúde boa (“muito boa” e “boa”) ou ruim (“regular”, “ruim” e “muito ruim”). A escolha dos três desfechos foi baseada em suas prevalências (baixa para demência, intermediária para TMC e alta para autopercepção de saúde ruim). Cada desfecho foi associado a uma exposição principal e dois potenciais fatores de confusão. Tanto para o desfecho demência como para TMC, a exposição principal foi nível educacional e as variáveis de confusão foram idade e sexo; para autopercepção de saúde ruim, a exposição principal foi presença de episódios depressivos, diagnosticados de acordo com critério para depressão da CID-10, e as variáveis de confusão foram renda e sexo.

Em relação a estudos anteriores, ampliamos a aplicação destes métodos para situações com duas variáveis de confusão, sendo algumas com dois ou mais níveis de exposição ou com valores contínuos, para verificar as estimativas por ponto e por intervalo da RP geradas por cada modelo multivariado. Desfechos de diferentes frequências foram analisados, para examinar como os modelos de Cox, log-binomial, Poisson e logístico comportam-se para estimar a RP, à medida que a prevalência do desfecho aumenta.

Valores de referência para as estimativas ajustadas das RP e respectivos intervalos de 95% de confiança (IC 95%), para as associações entre cada desfecho e respectiva exposição principal, foram obtidos pela estratificação de Mantel-Haenszel, controlando para o efeito das potenciais variáveis confundidoras. Estimativas da

RP com respectivos IC 95% foram calculadas usando modelos de regressão de Cox, log-binomial e Poisson, e os OR brutos e ajustados (IC 95%) também foram calculados pela regressão logística. Em seguida, para cada desfecho de interesse, uma variável de confusão foi testada com valores contínuos. As regressões de Cox e Poisson foram realizadas fixando tempo de seguimento igual à unidade para todos os participantes e utilizando estimadores de variância robusta. O *software* estatístico usado para este estudo foi o Stata versão 9.0.

O modelo de regressão de Poisson é geralmente usado em epidemiologia para analisar estudos longitudinais onde a resposta é o número de episódios de um evento ocorridos em um determinado período de tempo. Para estudos de coorte onde todos os indivíduos têm tempo de seguimento igual, a regressão de Poisson pode ser usada com o valor de tempo sob risco igual a 1 para cada indivíduo. Se o modelo ajustar corretamente os dados, esta aproximação fornece uma estimativa correta do risco relativo ajustado.<sup>4</sup> Em estudos de corte transversal é possível atribuir o valor unitário ao tempo de seguimento de cada participante, como estratégia para obtenção da estimativa por ponto da razão de prevalência, pois não há seguimento real dos participantes nesse tipo de estudo epidemiológico. No entanto, quando a regressão de Poisson é aplicada a dados binomiais, o erro para o risco relativo é superestimado, pois a variância da distribuição de Poisson aumenta progressivamente, enquanto a variância da distribuição binomial tem seu valor máximo quando a prevalência é 0,5. Este problema pode ser corrigido usando o procedimento de variância robusta, conforme proposto por Lin & Wei (1989).<sup>3</sup> A regressão de Poisson com variância robusta não tem dificuldade para convergir e produz resultados semelhantes àqueles obtidos pelo uso do procedimento de Mantel-Haenszel quando a covariável de interesse é categórica.<sup>6,14</sup>

O modelo de regressão de Cox geralmente é usado para analisar o tempo até um evento. Em estudos de corte transversal nenhum período de tempo é observado, mas se um período de risco constante for atribuído a todos os indivíduos no estudo, a razão de risco estimada pela regressão de Cox é igual à RP, similarmente à regressão de Poisson. Entretanto, o uso da regressão de Cox sem qualquer ajuste para análise de estudos de corte transversal pode também levar a erros na estimativa do intervalo de confiança, que pode ser maior do que deveria. O método de variância robusta também pode ser utilizado nessas situações.<sup>3</sup>

O modelo de regressão log-binomial é um modelo linear generalizado onde a função de ligação é o logaritmo da proporção em estudo e a distribuição do erro é binomial. Ele estima diretamente a razão de prevalência para variáveis dicotômicas. No entanto, pode ocorrer falta de convergência ao tentar fornecer estimativas do parâmetro. Normalmente este problema é devido ao

método de Newton, usado para encontrar um mínimo ou um máximo dessa função, que pode não ser capaz de encontrar a estimativa de máxima verossimilhança quando a solução está nas margens do restrito intervalo do parâmetro. Petersen & Deddens<sup>6</sup> (2003) sugeriram o método COPY, um macro para o *software* SAS, que pode gerar uma estimativa e erros padrões aproximados quando o comando Proc Genmod – geralmente é usado no SAS para distribuição binomial com função de ligação logarítmica – não consegue convergir.

A regressão logística tem sido usada amplamente em estudos epidemiológicos com desfechos binários para obter estimativas não enviesadas do OR ajustado para uma ou mais variáveis de confusão. É possível calcular a RP a partir da estimativa do OR, com IC 95%, mas o cálculo é complexo e requer programas computacionais para calcular estimativas de variância usando módulos de matriz.<sup>5</sup>

## RESULTADOS

O desfecho demência (baixa prevalência: 5,1%) mostrou associações estaticamente significativas com nível de escolaridade e faixa etária, mas não se associou com sexo (Tabela 1). O fator de risco nível de escolaridade também mostrou associação estatisticamente significativa com faixa etária ( $p < 0,01$ ) e sexo ( $p < 0,01$ ). Houve confusão relativa à faixa etária na estimativa da associação entre nível educacional e prevalência de demência, como mostra a estratificação de MH (Tabela 2). Comparando os resultados dos diferentes modelos com a exposição principal e uma variável de confusão com quatro níveis de exposição, as estimativas do ponto e respectivos IC 95% obtidos pelos modelos de Poisson, Cox e log-binomial aproximaram-se da estimativa obtida pela estratificação de MH (Tabela 2), com diferença de uma ou duas unidades no segundo dígito decimal. Os resultados observados ao adicionar mais uma potencial variável de confusão (sexo) aos modelos

**Tabela 1.** Prevalência de autopercepção de saúde ruim, demência e transtorno mental comum segundo exposição principal e variáveis de confusão. São Paulo, SP, maio/2003 a abril/2005.

Característica	n (%)	Autopercepção de saúde ruim	Desfecho	
			Demência	Transtorno mental comum
Escolaridade (anos)				
Nenhuma	793 (38,3)		59 (7,4)	334 (42,1)
1 ou mais anos	1279 (61,7)		46 (3,6)	450 (35,2)
p-valor			<0,01*	<0,01*
Faixa etária (anos)				
65 a 69	889 (42,9)		20 (2,3)	311 (35,0)
70 a 74	556 (26,8)		11 (2,0)	211 (38,0)
75 a 79	346 (16,7)		27 (7,8)	141 (40,8)
80 anos ou mais	281 (13,6)		47 (16,7)	121 (43,1)
p-valor			<0,01*	<0,01*
Sexo				
Feminino	1255 (60,6)	726 (57,9)	68 (5,4)	580 (46,2)
Masculino	817 (39,4)	387 (47,4)	37 (4,5)	204 (25,0)
p-valor		<0,01*	0,37	<0,01*
Episódio depressivo				
Não caso	1925 (92,9)	993 (51,6)		
Caso	99 (4,8)	87 (87,9)		
p-valor		<0,01*		
Renda				
Até R\$240 **	643 (31,0)	409 (63,6)		
R\$241 a R\$360	402 (19,4)	230 (57,2)		
R\$361 a R\$700	515 (24,9)	257 (49,9)		
R\$701 ou mais	512 (24,7)	217 (42,4)		
p-valor		<0,01*		

\* Diferença entre os grupos testada pelo qui-quadrado, significativa se  $p < 0,05$

\*\* 1 USD = R\$2,86

**Tabela 2.** Estimativas da razão de prevalência e intervalos de 95% de confiança (IC 95%) para a associação entre nível educacional e demência, controlando para faixa etária, faixa etária e sexo, e idade e sexo, usando estratificação de Mantel-Haenszel e modelos de regressão de Cox, Poisson, log-binomial e logística. São Paulo, SP, maio/2003 a abril/2005.

Modelo	(Demência, escolaridade e faixa etária)		(Demência, escolaridade e faixa etária) + sexo		(Demência, escolaridade e idade (contínua)) + sexo	
	Estimativas por ponto	IC 95%	Estimativas por ponto	IC 95%	Estimativas por ponto	IC 95%
RP Mantel-Haenszel	1,47	1,01;2,15	1,50	1,02;2,19	-	-
RP Cox robusto	1,47	1,01;2,13	1,46	1,01;2,12	1,43	0,98;2,09
RP Poisson robusto	1,47	1,01;2,13	1,46	1,01;2,12	1,43	0,98;2,09
RP log-binomial	1,48	1,02;2,15	1,47	1,01;2,13	1,43	0,97;2,10
OR regressão logística	1,53	1,01;2,30	1,52	1,00;2,29	1,49	0,98;2,25

Estimativa bruta: 2,07 (IC 95%: 1,42;3,01).

RP: Razão de prevalência

de Cox, Poisson e log-binomial produziram estimativas por ponto com 2 ou 3 unidades na segunda casa decimal menores que o ponto estimado por MH, e o IC 95% foi mais estreito que o intervalo de confiança de MH. Ao colocar a variável idade como contínua, a confusão foi mais bem ajustada pelos modelos, e a estimativa para a associação entre nível educacional e demência deixou de ser significativa. A regressão logística produziu uma estimativa por ponto aproximadamente 13% maior e IC 95% mais amplo que o obtido pelos outros modelos de regressão em todas as situações.

O desfecho TMC (prevalência intermediária: 37,8%) mostrou associações estatisticamente significativas com nível educacional, sexo e faixa etária (Tabela 1). Houve confusão devido a sexo e faixa etária na estimativa da associação entre nível educacional e risco de TMC, como mostra a estratificação de MH (Tabela 3). Comparando os resultados dos diferentes modelos, ambos na situação com a exposição principal e uma variável de confusão (sexo) com dois níveis de exposição e quando adicionada uma potencial variável de confusão extra (faixa etária) com quatro níveis de exposição, as estimativas do ponto e respectivos IC 95% obtidos pelos modelos de Poisson, Cox e log-binomial foram idênticos aos obtidos pela estratificação de MH

(Tabela 3). Quando a variável idade foi usada na forma contínua, os modelos de Cox, Poisson e log-binomial produziram estimativas por ponto e respectivos IC 95% quase idênticos. A regressão logística produziu estimativas por ponto quase 100% maiores que as obtidas pelos outros modelos de regressão, com intervalos de 95% de confiança maiores.

O desfecho autopercepção de saúde ruim (alta prevalência: 53,8%) mostrou associações estatisticamente significativas com episódio depressivo, sexo e renda (Tabela 1). A variável de exposição principal “episódio depressivo” também esteve associada com renda ( $p=0,04$ ). Quase não houve confusão devido à renda ou sexo na estimativa da associação entre episódio depressivo e autopercepção de saúde ruim, como mostrou a estratificação de MH (Tabela 4). Quando comparados os resultados dos diferentes modelos com a exposição principal e uma variável de confusão (renda) com quatro níveis de exposição, ou quando adicionada uma outra potencial variável de confusão (sexo) a cada modelo, as estimativas por ponto obtidas pelos modelos de Poisson e Cox e respectivos IC 95% foram idênticas às obtidas pela estratificação de MH. As estimativas por ponto obtidas pelo modelo log-binomial foram próximas da unidade do que aquelas obtidas pelos outros dois modelos.

**Tabela 3.** Estimativas da razão de prevalência e intervalos de 95% de confiança (IC95%) para a associação entre nível educacional e transtorno mental comum, controlando para sexo, sexo e faixa etária, e sexo e idade, usando estratificação de Mantel-Haenszel e modelos de regressão de Cox, Poisson, log-binomial e logística. São Paulo, SP, maio/2003 a abril/2005.

Modelo	(TMC, escolaridade e sexo)		(TMC, escolaridade e sexo) + faixa etária		(TMC, escolaridade e sexo) + idade (contínua)	
	Estimativas por ponto	IC 95%	Estimativas por ponto	IC 95%	Estimativas por ponto	IC 95%
RP Mantel-Haenszel	1,16	1,04;1,29	1,12	1,01;1,25	-	-
RP Cox robusto	1,16	1,04;1,29	1,12	1,01;1,25	1,12	1,01;1,25
RP Poisson robusto	1,16	1,04;1,29	1,12	1,01;1,25	1,12	1,01;1,25
RP log-binomial	1,16	1,04;1,29	1,12	1,01;1,25	1,12	1,00;1,24
OR regressão logística	1,30	1,07;1,57	1,23	1,01;1,49	1,22	1,00;1,48

Estimativa bruta: 1,22 (IC 95%: 1,10;1,36).

TMC: Transtorno mental comum

**Tabela 4.** Estimativas da razão de prevalência e intervalos de 95% de confiança (IC95%) para a associação entre episódio depressivo e autopercepção de saúde ruim, controlando para faixa de renda, faixa de renda e sexo, e renda e sexo, usando estratificação de MH, modelos de regressão de Cox, Poisson e log-binomial e regressão logística. São Paulo, SP, maio/2003 a abril/2005.

Modelo	(Autopercepção da saúde ruim, episódios depressivos, renda)		(Autopercepção da saúde ruim, episódios depressivos, renda) + sexo		(Autopercepção da saúde ruim, episódios depressivos, renda [contínua]) + sexo	
	Estimativas por ponto	IC 95%	Estimativas por ponto	IC 95%	Estimativas por ponto	IC 95%
RP Mantel-Haenszel	1,65	1,52;1,79	1,67	1,53;1,81	-	-
RP Cox robusto	1,65	1,52;1,79	1,66	1,52;1,81	1,64	1,52;1,80
RP Poisson robusto	1,65	1,52;1,79	1,66	1,52;1,81	1,66	1,52;1,80
RP log-binomial	1,57	1,46;1,69	1,56	1,45;1,67	1,61	1,49;1,74
OR regressão logística	6,99	3,69;13,21	7,10	3,8;13,4	7,03	3,7;13,3

Estimativa bruta: 1,72 (IC 95%: 1,58;1,87)

Quando a variável renda foi usada na forma contínua, os resultados dos modelos de Cox e Poisson foram semelhantes. No entanto, foi difícil atingir convergência usando o modelo log-binomial. A regressão logística produziu estimativas por ponto que, se interpretadas como estimativas de RP, seriam mais de quatro vezes maiores que as estimativas obtidas pelos outros modelos de regressão, e seus IC 95% foram maiores.

## DISCUSSÃO

Estudo prévio<sup>1</sup> mostrou que em estudos de corte transversal os modelos de regressão de Cox e Poisson, com variância robusta, e o modelo de regressão log-binomial geram estimativas adequadas da razão de prevalência e seus intervalos de confiança, independentemente da prevalência de base. Em estudo recente sobre essa questão, Petersen & Deddens<sup>6</sup> (2008) defendem, com base em dados reais e simulados, que a regressão de Poisson estima melhor a RP para desfechos muito frequentes, quando comparada ao modelo de regressão log-binomial. Entretanto, os autores sugerem a regressão log-binomial para prevalências intermediárias como o melhor método.

Nós exploramos o desempenho desses métodos para diferentes prevalências de desfechos de interesse, para mais de uma variável de confusão, e para covariáveis contínuas, mostrando que os três métodos geraram estimativas corretas por ponto e por intervalo em todas as situações, embora os modelos log-binomial tivessem dificuldade de convergir na situação de um desfecho muito prevalente e covariável contínua. Para os três desfechos investigados, os modelos de regressão de Cox e Poisson apresentaram estimativas da RP e IC 95% idênticos e muito similares àquelas obtidas pela referência (estratificação de MH). O uso de métodos robustos para a estimativa da variância nos modelos de Cox e de Poisson corrigiu para a superestimação da variância, e produziu intervalos de confiança adequados. Os modelos de Cox e Poisson também se comportaram bem com a presença de covariáveis contínuas.

Os modelos de regressão log-binomial também se comportaram bem na maioria das situações testadas, gerando estimativas por ponto e por intervalo próximas às obtidas pela estratificação de MH. Entretanto, quando a prevalência do desfecho foi alta, o modelo log-binomial produziu estimativas mais próximas da unidade que aquelas obtidas pela estratificação de MH e pelas regressões de Cox e Poisson. Também, quando uma das covariáveis foi contínua, o modelo log-binomial apresentou dificuldade de convergência, como descrito previamente.<sup>1,6</sup>

As estimativas de OR obtidas pelos modelos de regressão logística foram próximas às estimativas da RP quando a prevalência do desfecho era baixa (demência), embora já fosse observada uma tendência do OR ser maior que a RP. Na situação da prevalência intermediária (TMC), o OR foi quase duas vezes a RP. Em outras palavras, se interpretarmos o OR como a RP, pareceria que o aumento relativo no risco de TMC para aqueles com menor nível educacional seria 23% maior que o risco para aqueles com melhor nível educacional, no lugar de 12% de aumento, como mostrado pela RP. O OR obtido quando a prevalência era alta (autopercepção de saúde ruim) foi quatro vezes maior que as estimativas de RP obtidas pela estratificação de MH, pelos modelos de Cox, Poisson e log-binomial, mostrando a inadequação de interpretar as estimativas de OR como se fossem estimativas da RP nestas situações.

O presente estudo tem algumas limitações. As associações nas três situações examinadas não eram fortes, fazendo com que as estimativas obtidas pelos vários métodos testados ficassem próximas. Também, os efeitos de confusão não foram acentuados, a não ser para o efeito da idade no risco de demência. No entanto, o presente estudo dá suporte adicional para o uso das técnicas dos modelos testados como alternativas à regressão logística, disponíveis na maioria dos pacotes estatísticos utilizados para análise de estudos epidemiológicos.

## REFERÊNCIAS

1. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol.* 2003;3:21. DOI: 10.1186/1471-2288-3-21
2. Davies HT, Crombie IK, Tavakoli M. When can odds ratios mislead? *BMJ.* 1998;316(7136):989-991.
3. Lin DY, Wei LJ. The robust Inference for the Cox Proportional Hazards Model. *J Am Stat Assoc.* 1989;84(408):1074-8. DOI: 10.2307/2290085
4. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol.* 2003;157(10):940-3. DOI: 10.1093/aje/kwg074
5. Oliveira NF, Santana VS, Lopes AA. Razões de proporções e uso do método delta para intervalos de confiança em regressão logística. *Rev Saude Publica.* 1997;31(1):90-9.
6. Petersen MR, Deddens JA. A comparison of two methods for estimating prevalence ratios. *BMC Med Res Methodol.* 2008;8:9. DOI: 10.1186/1471-2288-8-9
7. Prince M, Acosta D, Chiu H, Scazufca M, Varghese M, Dementia diagnosis in developing countries: a cross-cultural validation study. *Lancet.* 2003;361(9361):909-17. DOI: 10.1016/S0140-6736(03)12772-9
8. Scazufca M, Menezes PR, Vallada HP, Crepaldi AL, Pastor-Valero M, Coutinho LMS, et al. High prevalence of dementia among older adults from poor socioeconomic backgrounds in São Paulo, Brazil. *Int Psychogeriatrics.* 2008;20(2):394-405. DOI: 10.1017/S1041610207005625
9. Scazufca M, Menezes PR, Vallada H, Araya R. Validity of the self reporting questionnaire-20 in epidemiological studies with older adults. *Soc Psychiatry Psychiatr Epidemiol.* 2008; Sep 8. [Epub ahead of print].
10. Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: what is to be done? *Occup Environ Med.* 1998;55(4):272-7.
11. World Health Organization. A user's guide to the Self-Reporting Questionnaire (SRQ). Geneva; 1994.
12. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA.* 1998;280(19):1690-1. DOI: 10.1001/jama.280.19.1690
13. Zocchetti C, Consonni D, Bertazzi PA. Relationship between prevalence rate ratios and odds ratios in cross-sectional studies. *Int J Epidemiol.* 1997;26(1):220-3. DOI: 10.1093/ije/26.1.220
14. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol.* 2004;159(7):702-6. DOI: 10.1093/aje/kwh090

---

Pesquisa financiada pela Wellcome Trust, UK.

PRMenezes e MSCaufca são parcialmente financiados pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq –bolsa produtividade em pesquisa).

Artigo baseado na dissertação de mestrado de LMS Coutinho, apresentada ao Programa de Pós-Graduação em Ciências da Faculdade de Medicina da Universidade de São Paulo, em 2007.