



# **Comparação por simulação de metodologias para estimar risco relativo em dados correlacionados**

Autor: Daiane Rodrigues  
Orientador: Prof. Dra. Vanessa Bielefeldt Leotti  
Co-Orientadora: Prof. Dra. Suzi Alvez Camey

Porto Alegre, 09 de dezembro de 2016.

Universidade Federal do Rio Grande do Sul  
Instituto de Matemática e Estatística  
Departamento de Estatística

# **Comparação por simulação de metodologias para estimar risco relativo em dados correlacionados**

Autor: Daiane Rodrigues

Trabalho de Conclusão de Curso  
apresentado para obtenção  
do grau de Bacharel em Estatística.

Banca Examinadora:  
Professor Dra. Vanessa Bielefeldt Leotti (orientadora)  
Professor Dra. Suzi Alvez Camey (co-orientadora)  
Professor Dr. Álvaro Vigo

Porto Alegre, 09 de dezembro de 2016.



*Dedico este trabalho aos meus pais que sempre apoiaram as minhas escolhas.*

*“A alegria está na luta, na tentativa, no sofrimento envolvido e não na vitória propriamente dita.”*

Mahatma Gandhi

# Resumo

Com a crescente utilização de ensaios aleatorizados por *cluster*, onde há dependência entre os indivíduos do mesmo grupo, vem a necessidade de estudar a performance dos modelos que estimam o risco relativo (RR), no contexto de dados correlacionados. Sendo que há predominância destas análises de desempenho para dados independentes.

Neste artigo, serão comparados quatro métodos que permitem estimar RR em dados correlacionados. Os modelos analisados foram o log-binomial GEE, Poisson GEE, log-binomial de efeitos mistos frequentista e log-binomial de efeitos bayesiano. Ressaltamos que até o momento não se tem conhecimento sobre estudos que comparam os métodos mistos por simulação.

Desta maneira, os métodos foram comparados utilizando bancos de dados simulados, com as principais características controladas. A performance foi comparada com relação a convergência, vício relativo, erro quadrático médio, poder, cobertura e amplitude do intervalo de confiança/credibilidade. Percebe-se que os resultados mostraram uma superioridade dos modelos mistos, sendo que, em geral, o modelo bayesiano obteve desempenho superior ao frequentista. O modelo Bayesiano obteve menor EQM para 10 dos 12 cenários para a variável confundidora e em 11 dos 12 cenários para a variável de tratamento.

Palavras Chave: Risco relativo, Razão de prevalência, Dados dependentes, Modelo misto, GEE, Simulação.

# Sumário

|                                    |    |
|------------------------------------|----|
| 1. Introdução.....                 | 9  |
| 2. Metodologia.....                | 11 |
| 2.1 Modelos GEE.....               | 11 |
| 2.2 Modelos Misto.....             | 12 |
| 2.3 Simulação.....                 | 14 |
| 2.4 Propriedades.....              | 15 |
| 3. Resultados.....                 | 16 |
| 4. Discussão.....                  | 18 |
| 5. Referências bibliográficas..... | 20 |
| 6. Tabelas.....                    | 22 |
| Anexos.....                        | 27 |

Este artigo será submetido a “REVISTA CLINICAL TRIALS”



## 1. INTRODUÇÃO

O risco relativo (RR) é a medida de associação natural em delineamentos longitudinais, estudos em que os indivíduos são acompanhados ao longo do tempo e dada a exposição a determinado tratamento/fator, deseja-se estimar o risco associado ao desfecho.

Já a razão de chances (RC) é a medida de associação natural para estudos caso-controle, onde o objetivo é verificar se os casos diferem dos controles com relação à exposição a um fator. Torman e Camey (1) mostram que a razão de chances é apenas uma boa aproximação do risco relativo quando o desfecho é raro (<10%) entre os expostos e não expostos.

O risco relativo e a razão de chances podem ser confundidos. O RR é uma razão entre probabilidades, por sua vez a RC é uma razão entre chances. As probabilidades no cálculo do RR podem também ser denominadas como prevalências, no caso de estudos transversais, sendo assim a medida de associação chamada de razão de prevalências (RP).

Autores comentam como a RC é uma medida difícil de ser interpretada. Deeks (2) indica que a RC somente deve ser utilizada em estudos caso-controle, sendo que em ensaios clínicos e coortes o RR deve ser preferido, assim como a RP em estudos transversais, para que não haja comprometimento da interpretação da associação.

Há vários trabalhos sobre a estimação do RR em dados independentes, via modelos de regressão. Barros e Hirakata (3) compararam empiricamente o modelo log-binomial, Poisson com variância robusta, Cox com tempo constante e o método de Mantel-Haenszel e recomendaram utilizar a regressão de Poisson ou o modelo log-binomial. Zou (4) comparou por simulação o desempenho do modelo Poisson, com e sem variância robusta, o log-binomial e o método de Mantel-Haenszel e recomendou o uso do modelo Poisson com variância robusta ou log-binomial. Apesar de o modelo log-binomial apresentar boas propriedades, sendo indicado por vários autores, ele apresenta recorrentes problemas de convergência (3). Todavia há uma carência de propostas para quando os dados são correlacionados (5).

Em ensaios clínicos randomizados por *cluster*, onde grupos são aleatorizados para receber ou não o tratamento, é natural supor que haja alguma dependência entre indivíduos do mesmo *cluster*. Da mesma maneira, em estudos multiníveis a dependência está inerente aos dados, isto é, entre as unidades amostrais do mesmo nível. As alternativas para incorporar

essa correlação nas análises podem basear-se nas equações de estimação generalizadas (GEE – do inglês *Generalized Estimating Equations*) ou modelos mistos.

Yelland (5) e Zou e Donner (6) propuseram métodos GEE a fim de permitir a estimação do RR para dados correlacionados. Zou e Donner (6) comparam os modelos log-binomial GEE e Poisson GEE. Yelland (5) comparou os modelos log-binomial GEE, Poisson GEE, log-normal GEE e regressão logística expandida GEE. Ambos os trabalhos compararam os métodos por simulação, sendo que as matrizes de correlação de trabalho *exchangeable* e independente foram avaliadas. A recomendação de Zou e Donner (6) foi para a Poisson GEE, quando a log-binomial GEE não convergir; Yelland sugere a Poisson GEE ou a log-normal GEE como alternativa a esse problema de convergência. Os autores também concluíram que as duas estruturas de correlação pareceram ser igualmente eficientes.

Até o momento, não se tem conhecimento da comparação de performance, por simulação, do modelo log-binomial misto para estimar RR em dados correlacionados. Torman e Comey (7) abordaram o método frequentista e Bayesiano, com aplicação em um conjunto de dados provenientes de dados dependentes. Os autores comentam sobre a vantagem de o método Bayesiano superar os problemas comuns de convergência, além de estimar a variância do efeito aleatório e ICC de uma forma mais direta que os métodos GEE. O modelo misto frequentista não pode ser avaliado, pois não houve convergência. O método bayesiano já foi comparado por simulação, mas para estimação do ICC Yelland *et al* 2011 (14) e Zorzi *et al* (8).

Neste trabalho irá se avaliar quatro métodos que estimam RR ou RP para desfechos binários e dados dependentes. Compararemos a performance, por simulação, dos modelos log-binomial GEE, Poisson GEE, log-binomial misto Bayesiano e frequentista. Primeiramente será mostrado como cada modelo é definido e como a simulação foi conduzida, em seguida os resultados são apresentados, por final uma discussão sobre os resultados encontrados.

## 2. METODOLOGIA

Dados de um ensaio clínico randomizado em *cluster*, onde o objetivo é estimar o risco relativo de um tratamento e o risco relativo de uma variável confundidora binária foram simulados. Para estimação, quatro métodos serão abordados, dois modelos GEE, o log-binomial GEE e o Poisson GEE; e dois modelos mistos, o log-binomial misto frequentista e log-binomial misto bayesiano.

Para cada conjunto de dados simulado foi utilizado o software R versão 3.2.4 para ajustar os três métodos frequentistas, com o uso do pacote `geepack` (9) para os modelos GEE e o pacote `lme4` (10) para o modelo misto frequentista. A abordagem bayesiana para o modelo log-binomial misto, utilizando simulação Monte Carlo via Cadeias de Markov (MCMC), foi realizada por meio do programa JAGS 3.4.0 em conjunto com o `rjags` (11). A geração dos bancos de dados foi realizada através do SAS 9.0.

### 2.1 MODELOS GEE

O modelo log-binomial para dados independentes é um modelo linear generalizado com função de ligação logaritmo. A distribuição do desfecho é suposta binomial. O modelo log-binomial GEE, sugerido por Zou e Donner (6) e Yelland (5), é a extensão deste para dados correlacionados, modelando-se uma matriz de correlação de trabalho para as observações dentro do mesmo *cluster*.

No contexto do ensaio simulado, o modelo log-binomial GEE, com intercepto e dois preditores, pode ser escrito como:

$$\theta_i = P(Y_i = 1 | X_{1i}, X_{2i}) = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}) \quad 0 \leq \theta_i \leq 1 \quad (1)$$

Na equação 1,  $\theta_i$  denomina a probabilidade condicional de sucesso da  $i$ -ésima observação,  $X_{1i}$  é a variável que indica a presença do tratamento,  $X_{2i}$  é a variável confundidora binária. O RR do tratamento é estimado por  $\exp(\beta_1)$  e o RR da variável confundidora por  $\exp(\beta_2)$ . Uma propriedade importante que pode impactar na convergência do modelo é que os coeficientes estão condicionados aos valores em que  $0 < \theta_i \leq 1$ , para todos os indivíduos.

O modelo Poisson GEE assume distribuição Poisson para o desfecho. Assim como no modelo log-binomial, há também uma função de ligação logaritmo. Neste caso, se modela a esperança condicional do desfecho, dadas as variáveis preditoras, e esta será uma aproximação da probabilidade condicional de sucesso.

$$\theta_i = E(Y_i | X_{1i}, X_{2i}) = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}) \quad \theta_i \geq 0 \quad (2)$$

Ao se assumir distribuição Poisson para o desfecho, mesmo ele sendo de natureza dicotômica, o modelo apenas considera que  $\theta_i \geq 0$ . Por isso temos a desvantagem de poderem ser estimadas probabilidades maiores que 1. Assim como no modelo log-binomial, o risco relativo é estimado pela exponencial dos coeficientes do modelo.

Ambos os modelos GEE necessitam de especificação sobre a estrutura de correlação, para que a dependência nos dados possa ser modelada. Será assumida a estrutura de correlação do tipo *exchangeable*. Zou e Donner (6) e Yelland (5) testaram as matrizes *exchangeable* e independente, e os autores comentam que os resultados são robustos quanto a matriz de correlação.

## 2.2 MODELOS MISTO

O modelo log-binomial com efeitos mistos é também um modelo linear generalizado. O desfecho é binomial, com função de ligação logaritmo. Diferente dos modelos GEE, ele permite estimar a variância do efeito de *cluster* de uma forma direta.

O modelo misto pode ser definido como:

$$\theta_{ij} = P(Y_{ij} = 1 | X_{1ij}, X_{2ij}, a_j) = \exp(\beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + a_j) \quad 0 \leq \theta_i \leq 1 \quad (3)$$

Na equação 3, percebe-se que o modelo misto leva em consideração o *cluster* ao qual o indivíduo é pertencente dentro do preditor linear. Nesta equação,  $\theta_{ij}$  expressa a probabilidade de sucesso do i-ésimo indivíduo no j-ésimo *cluster*,  $a_j$  é o efeito devido ao j-ésimo *cluster*, onde se supõe que  $a_j \sim N(0, \sigma_a^2)$ . Os demais termos do modelo têm definições idênticas ao do

modelo log-binomial GEE, a principal diferença é que aqui há um índice indicativo para *cluster* nas variáveis preditoras.

Como todo modelo estatístico, o modelo log-binomial misto pode ser estimado via abordagem clássica ou Bayesiana. Na estatística clássica a única informação sobre o parâmetro de interesse vem da amostra coletada em estudos, as inferências são feitas com base na amostra e nos possíveis resultados do experimento. A metodologia Bayesiana, diferentemente da estatística clássica, considera o parâmetro de estudo como variável aleatória, assim ela possui uma distribuição de probabilidade associada, ignorando os possíveis resultados que poderiam ter sido gerados no experimento. Com esta abordagem, há a possibilidade de incorporar nos resultados informações anteriores sobre o parâmetro, através da distribuição a priori, combinando matematicamente os resultados encontrados na amostra com estas informações, assim obtendo como resultado a distribuição a posteriori.

A posteriori utilizada na metodologia bayesiana, principalmente em modelos multiparamétricos, pode ter forma analítica desconhecida. As simulações permitiram que este problema fosse superado, sendo que os métodos mais utilizados com este enfoque são simulações de Monte Carlo via cadeias de Markov (MCMC). Os resultados simulados podem ser utilizados somente se não há correlação entre os valores simulados.

Salmeron, Cano e Chirlaque (12) propuseram uma reparametrização do modelo log-binomial misto bayesiano baseado na matriz de variâncias e covariâncias da regressão de Poisson, utilizando a decomposição de Cholesky. Este método permite diminuir problemas de convergências caso utilizássemos a matriz de covariâncias do modelo log-binomial, fazendo com que sejam necessárias menos iterações.

O modelo log-binomial de efeitos mistos bayesiano reparametrizado é definido como:

$$\log(\theta_{ij}) = \theta_0 Z_{0ij} + \theta_1 Z_{1ij} + \theta_2 Z_{2ij} + a_j \quad (4)$$

Seja  $\hat{\Sigma}$  a matriz de covariâncias de máxima verossimilhança do estimador  $\hat{\beta}$  do modelo Poisson com variância robusta e seja  $L$  o fator triangular da decomposição de Choleski de  $\hat{\Sigma} = L^T L$ . A matriz de variáveis reparametrizadas  $Z_\epsilon$  é calculada como  $Z = X L^T$ . Assim, os termos  $\theta_0 Z_{0i}$ ,  $\theta_1 Z_{1i}$ ,  $\theta_2 Z_{2ij}$  são a reparametrizações para  $\beta_0$ ,  $\beta_1 X_{1ij}$ ,  $\beta_2 X_{2ij}$ . Para maiores detalhes

ver Salmeron, Cano e Chirlaque (12). A interpretação dos riscos relativos permanece a mesma, pois após a geração dos resultados simulados é possível recalcular os parâmetros originais.

As prioris atribuídas para o modelo foram todas não informativas, sendo que para  $\theta_0$ ,  $\theta_1$  e  $\theta_2$  foi suposta distribuição normal com média zero e variância 10000 e para  $\sigma_a$  distribuição uniforme com parâmetros 0,01 e 100. As prioris foram assim escolhidas para não se perder o poder de comparação com os demais modelos frequentistas. O número de iterações, o período de aquecimento (*burn-in*) e o *thin* das simulações foram definidos por meio de análise visual e a estatística de Gelman e Rubin, sendo que cada um dos parâmetros do modelo teve duas cadeias geradas utilizando um *burn-in* de 2000, com *thin* igual a cinco.

Para estimador bayesiano pontual, calculou-se a média, mediana e moda. Já para a estimativa bayesiana por intervalo, foi utilizado o intervalo de credibilidade (IC) de caudas iguais, com 0.95 de probabilidade. Em 2010 Torman e Camey (13) testaram estimadores bayesianos pontuais e por intervalo, onde a recomendação foi utilizar a moda como estimador pontual e o intervalo de credibilidade de caudas iguais por intervalo.

### 2.3 SIMULAÇÃO

Os bancos de dados utilizados nos diferentes cenários são todos simulados, concedidos pela professora Dra. Lisa N. Yelland. Em seus trabalhos, Zorzi *et al* (8) e Yelland *et al* (14) utilizaram os mesmos dados simulados para comparar metodologias de se calcular o coeficiente de correlação intraclasse.

Os dados foram gerados considerando 12 cenários, compostos por combinações do risco relativo da variável de tratamento, da variável confundidora e da variância do efeito aleatório. O risco basal (exponencial do intercepto) para todos os cenários é de 0,1.

Para cada cenário, mil conjuntos de dados foram gerados, sendo que cada conjunto possui 20 *clusters*, com 20 indivíduos, ou seja, há 400 observações em cada banco de dados. Os cenários considerados estão especificados na tabela 1.

A metade dos *clusters* em cada banco foi tratado/exposto, isto é, 10 indivíduos dos 20 para todos os *clusters*. Os desfechos binários foram gerados com probabilidade  $\theta_{ij} = \exp(\beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + a_j)$ , para o sujeito  $i$  no *cluster*  $j$ , sendo  $X_{1ij}$  a variável binária indicativa de tratamento/exposição,  $X_{2ij}$  a variável confundidora binária, gerada de uma

distribuição beta binomial com 0,5 de prevalência média por cluster e 0,05 de ICC,  $a_j$  é o termo de erro associado ao  $j$ -ésimo *cluster*, sendo que  $a_j \sim N(0, \sigma_a^2)$ .

O código utilizado nas simulações encontra-se em anexo.

## 2.4 PROPRIEDADES

A performance dos quatro métodos será verificada por meio da comparação das estimativas dos RR geradas por eles, com base nos verdadeiros valores dos parâmetros especificado para a geração dos dados. As propriedades calculadas para cada cenário simulado serão as seguintes:

- Convergência: calculada como o percentual de simulações onde o modelo convergiu, apenas para os métodos frequentistas, pois para o Bayesiano assumiu-se que as definições do MCMC eram suficientes para a convergência em todas as amostras.
- Vício relativo: definido como a média dos RR estimados menos o verdadeiro risco, dividido pelo verdadeiro risco.
- Erro quadrático médio (EQM): calculado pela soma da variância das estimativas e o quadrado do vício.
- Poder: calculado como o percentual de simulações em que a hipótese de nulidade de efeito de tratamento/covariável de confundimento foi rejeitada, já que tal hipótese nula era falsa para todos os cenários.
- Cobertura: calculada como o percentual dos intervalos de confiança/credibilidade que contiveram o verdadeiro RR.
- Amplitude: calculado através da média da diferença entre o limite superior e inferior dos intervalos de confiança/credibilidade.

### 3. RESULTADOS

Os resultados são apresentados para as estimativas do risco relativo de tratamento e para a covariável confundidora. Considerando o número de cenários e propriedades que foram avaliadas neste artigo, não é possível apresentar os resultados dos coeficientes do modelo, mas poderão ser consultados no material suplementar.

Além disso, no caso do modelo bayesiano, apenas a moda será mencionada, pois quando comparada com a média e a mediana ela apresentou menor EQM em geral.

Na tabela 2, têm-se os resultados das medidas de performance para as estimativas pontuais: EQM, vício relativo e convergência, calculados para os quatro modelos em cada um dos cenários. Em relação ao vício relativo, os modelos mistos (frequentista e Bayesiano) apresentaram melhor desempenho para estimação dos RR de tratamento e da variável confundidora em 11 dos 12 cenários avaliados, sendo o frequentista foi responsável por 6 dos 11. Os métodos GEE tiveram resultados bastante semelhantes. Para o EQM os modelos mistos também apresentaram melhores resultados, sendo que o Bayesiano obteve desempenho superior em 11 dos 12 cenários para a variável de tratamento e em 10 dos 12 cenários para a variável confundidora. Os métodos GEE não apresentaram problemas de convergência, enquanto o modelo misto frequentista sofreu com este problema em todos os cenários, observa-se que maiores problemas aconteceram a partir do cenário 6, conforme a variância entre *clusters* aumentava e os RR se distanciavam da unidade. O menor percentual de convergência aconteceu no cenário 12, onde o modelo convergiu em apenas 41,80% dos conjuntos de dados simulados.

A tabela 2 apresenta também os resultados para estimativas por intervalo, a cobertura, a amplitude e o poder. A cobertura foi maior que 95% apenas para os modelos mistos, sendo que para variável de tratamento, obteve este desempenho em 9 cenários para o Bayesiano e em 2 para o frequentista. Já para variável confundidora, a cobertura foi maior que 95% em 6 cenários no bayesiano e em 5 no frequentista. Para a amplitude do intervalo de confiança ou intervalo de credibilidade, o modelo misto Bayesiano obteve o pior desempenho em todos os cenários para ambas variáveis preditoras, possivelmente isso possa ter influenciado nos seus resultados de cobertura, pois este método gerou maiores intervalos. O poder foi menor que 25%, em todos os métodos, nos cenários em que os riscos relativos são iguais a 1,25, para



variável de tratamento ou para variável confundidora, independente da variância dos *clusters*. Em geral, os métodos GEE tiveram maior poder que os demais.

Os resultados, de uma maneira geral, mostram que os modelos mistos obtiveram um desempenho melhor que os modelos GEE nas propriedades avaliadas, sendo que o misto bayesiano teve vício relativo semelhante ao frequentista e desempenho superior com relação a EQM e cobertura para a maior parte dos cenários analisados.

## 4. DISCUSSÃO

Neste trabalho comparou-se a performance de quatro métodos para estimação de risco relativo no contexto de dados correlacionados, o log-binomial GEE, o Poisson GEE, o log-binomial de efeitos mistos frequentista e o log-binomial de efeitos mistos bayesiano. A maior inovação foi ter conseguido verificar o desempenho do modelo log-binomial de efeitos mistos, sendo a primeira vez que este método é comparado por simulação para estimar RR. Além disso, aqui se chega à conclusão de que, no contexto de dados correlacionados, os modelos de efeitos mistos são melhores que os GEE, especialmente o bayesiano. Em 2011, quando Yelland (5) comparou modelos para dados correlacionados, apenas métodos GEE foram abordados e a recomendação caiu sobre o log-binomial GEE.

Os quatro modelos foram analisados nos diferentes cenários que combinavam variadas configurações de riscos relativos e níveis de dependência. Assim tenta-se contribuir para a utilização de modelos que considerem a dependência existente nos dados, como os mistos e GEE, nos estudos onde essas abordagens são mais indicadas, como no caso de estudos randomizados por *cluster*.

Ilustraram-se os resultados pensando na estimativa de RR em delineamentos prospectivos. Entretanto, todos os resultados aqui encontrados, podem ser estendidos para delineamentos transversais, na estimativa da razão de prevalência para situações onde os dados são dependentes, o desfecho é dicotômico e as variáveis preditoras também são dicotômicas.

Como o poder foi baixo quando os riscos relativos são próximos de um, podemos inferir que um tamanho de amostra 400 não é suficiente para detectarmos essa associação, para isso deve-se aumentar o tamanho de amostra. Nesse sentido, ressalta-se a importância do cálculo do coeficiente de correlação intraclasse (ICC), pois a dependência entre as unidades amostrais irá consequentemente aumentar o tamanho de amostra, para que não haja perda de poder. Mais detalhes sobre esse tópico podem ser encontrados em Zorzi *et al* (8) e Yelland *et al* (14).

Em geral, os resultados das simulações mostram que os dois modelos GEE obtiveram desempenho semelhante para as propriedades analisadas e não apresentam problemas de convergência quando os preditores de interesse são dicotômicos. Entretanto, eles não obtiveram os resultados mais precisos quando comparado aos modelos de efeitos mistos que apresentou performance superior aos GEE na maior parte dos cenários. Sendo que modelo

Poisson GEE possui a desvantagem de poder estimar probabilidades maiores que um, pois restringe incorretamente o espaço paramétrico (15).

Uma limitação deste estudo é abrangência quanto as covariáveis, pois apenas preditores categóricos foram analisados. Planeja-se estender as simulações também para situações com covariáveis contínuas.

Destaca-se que uma das dificuldades para utilização do método MCMC é a necessidade de algum conhecimento sobre programação, pois sua implementação não existe em *softwares* estatísticos mais populares que utilizam menus. Porém, esses problemas vêm sendo superados por meio da utilização do aplicativo de web *Shiny*, para o software estatístico R. Essa ferramenta irá permitir que a aplicação de métodos bayesianos, como o modelo log-binomial de efeitos misto bayesiano, seja mais acessível a todos que quiserem utilizá-lo.

Conclui-se assim que o risco relativo é uma medida importante para medir associação em delineamentos longitudinais, como em estudos randomizados em *cluster*, onde há dependência entre os indivíduos do mesmo grupo. Nesses casos modelos GEE ou mistos podem ser uma boa alternativa para modelar esta correlação. Os modelos mistos apresentam desempenho superior aos demais, sendo que o modelo bayesiano, além de apresentar melhores propriedades, consegue contornar os problemas de convergência do método frequentista.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

1. Torman VBL, Camey SA. Modelos Bayesianos para Estimar Risco Relativo em Desfechos Binários e Politémicos [Doutorado]. Programa de Pós.
2. Deeks J. When can odds ratios mislead? Odds ratios should be used only in case-control studies and logistic regression analyses. *BMJ*. 24 de outubro de 1998;317(7166):1155–1156; author reply 1156–1157.
3. 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* [Internet]. 2003 Dec [cited 2016 Nov 17];3(1). Available from: <http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-3-21>
4. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am J Epidemiol*. 2004 Apr 1;159(7):702–6.
5. Yelland LN, Salter AB, Ryan P. Relative Risk Estimation in Cluster Randomized Trials: A Comparison of Generalized Estimating Equation Methods. *Int J Biostat*. 2011 Jan 21;7(1):1–26.
6. Zou G, Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat Methods Med Res*. 2013 Dec 1;22(6):661–70.
7. Torman VBL, Camey SA. Bayesian models as a unified approach to estimate relative risk (or prevalence ratio) in binary and polytomous outcomes. *Emerg Themes Epidemiol* [Internet]. 2015 Dec [cited 2016 Nov 17];12(1). Available from: <http://www.ete-online.com/content/12/1/8>.
8. Zorzi MD, Comparação de metodologias para estimar o coeficiente de correlação intraclasse no modelo log-binomial misto. 2015
9. Hojsgaard, S., Halekoh, U. & Yan J. (2006) The R Package geepack for Generalized Estimating Equations *Journal of Statistical Software*, 15, 2, pp1–11.  
Yan, J. & Fine, J.P. (2004) Estimating Equations for Association Structures *Statistics in Medicine*, 23, pp859–880.  
Yan, J (2002) geepack: Yet Another Package for Generalized Estimating Equations *R-News*, pp12-14.
10. Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software* 67, no. 1 (2015). doi:10.18637/jss.v067.i01.
11. Martyn Plummer (2016). *rjags: Bayesian Graphical Models using MCMC*. R package version 4-6. <https://CRAN.R-project.org/package=rjags>
12. Salmerón D, Cano JA, Chirlaque MD. Reducing Monte Carlo error in the Bayesian estimation of risk ratios using log-binomial regression models. *Stat Med*. 2015 Aug 30;34(19):2755–67.
13. Torman VBL, Camey SA. Bayesian Analysis of the Log-Binomial Model: a Comparison with the Frequentist Approach for the Estimation of Relative Risk. *Epidemiology*. 2013 [a ser submetido].
14. Yelland LN, Salter AB, Ryan P, Laurence CO. Adjusted intraclass correlation coefficients for binary data: methods and estimates from a cluster-randomized trial in primary care. *Clin Trials*. 1 de fevereiro de 2011;8(1):48–58.

15. Chu H, Cole SR. Estimation of risk ratios in cohort studies with common outcomes: a Bayesian approach. *Epidemiol.* novembro de 2010;21(6):855–62.

## 6. TABELAS

**Tabela 1:** Cenários de Simulação

| Cenário | RR Tratamento | RR do confundidor | $\sigma_a^2$ |
|---------|---------------|-------------------|--------------|
| 1       | 1,25          | 1,25              | 0,01         |
| 2       | 1,25          | 1,25              | 0,1          |
| 3       | 1,25          | 1,25              | 0,5          |
| 4       | 1,25          | 2,00              | 0,01         |
| 5       | 1,25          | 2,00              | 0,1          |
| 6       | 1,25          | 2,00              | 0,5          |
| 7       | 2,00          | 1,25              | 0,01         |
| 8       | 2,00          | 1,25              | 0,1          |
| 9       | 2,00          | 1,25              | 0,5          |
| 10      | 2,00          | 2,00              | 0,01         |
| 11      | 2,00          | 2,00              | 0,1          |
| 12      | 2,00          | 2,00              | 0,5          |

**Tabela 2:** Resultados das análises para EQM, vício relativo, convergência, cobertura, amplitude e poder.

| Cenário | Parâmetro | Modelo                       | EQM    | Vício relativo | Convergência | Cobertura | Amplitude | Poder |
|---------|-----------|------------------------------|--------|----------------|--------------|-----------|-----------|-------|
| 1       | RR1       | Log-binomial GEE             | 0,1483 | 5,42%          | 100,00       | 91,50     | 1,40      | 17,30 |
|         |           | Poisson GEE                  | 0,1484 | 5,45%          | 100,00       | 91,50     | 1,40      | 17,30 |
|         |           | Log-binomial misto           | 0,1489 | 5,43%          | 99,90        | 94,59     | 1,53      | 12,41 |
|         |           | Log-binomial misto bayesiano | 0,1202 | 3,18%          | -            | 96,40     | 1,82      | 10,30 |
|         | RR2       | Log-binomial GEE             | 0,1421 | 4,45%          | 100,00       | 93,10     | 1,40      | 16,20 |
|         |           | Poisson GEE                  | 0,1426 | 4,46%          | 100,00       | 93,20     | 1,40      | 16,10 |
|         |           | Log-binomial misto           | 0,1389 | 4,28%          | 99,90        | 94,99     | 1,45      | 12,71 |
|         |           | Log-binomial misto bayesiano | 0,1154 | 2,94%          | -            | 94,40     | 1,49      | 14,50 |
| 2       | RR1       | Log-binomial GEE             | 0,1857 | 6,22%          | 100,00       | 92,90     | 1,59      | 16,00 |
|         |           | Poisson GEE                  | 0,1853 | 6,23%          | 100,00       | 93,10     | 1,59      | 16,10 |
|         |           | Log-binomial misto           | 0,1851 | 6,08%          | 99,40        | 94,77     | 1,66      | 12,88 |
|         |           | Log-binomial misto bayesiano | 0,1475 | 3,68%          | -            | 96,40     | 2,05      | 11,10 |
|         | RR2       | Log-binomial GEE             | 0,1521 | 4,39%          | 100,00       | 92,90     | 1,40      | 16,80 |
|         |           | Poisson GEE                  | 0,1525 | 4,40%          | 100,00       | 93,00     | 1,40      | 16,80 |
|         |           | Log-binomial misto           | 0,1481 | 4,07%          | 99,40        | 94,97     | 1,46      | 12,07 |
|         |           | Log-binomial misto bayesiano | 0,1222 | 3,19%          | -            | 94,70     | 1,49      | 14,40 |
| 3       | RR1       | Log-binomial GEE             | 0,4646 | 10,25%         | 100,00       | 90,40     | 2,35      | 14,80 |
|         |           | Poisson GEE                  | 0,4597 | 10,18%         | 100,00       | 90,20     | 2,34      | 14,70 |
|         |           | Log-binomial misto           | 0,4201 | 8,67%          | 94,20        | 90,87     | 2,33      | 12,63 |
|         |           | Log-binomial misto bayesiano | 0,3357 | 7,39%          | -            | 94,30     | 3,50      | 8,20  |
|         | RR2       | Log-binomial GEE             | 0,1675 | 5,19%          | 100,00       | 91,00     | 1,41      | 17,20 |
|         |           | Poisson GEE                  | 0,1674 | 5,16%          | 100,00       | 91,30     | 1,42      | 17,40 |
|         |           | Log-binomial misto           | 0,1555 | 3,92%          | 94,20        | 93,63     | 1,46      | 13,48 |
|         |           | Log-binomial misto bayesiano | 0,1344 | 2,13%          | -            | 92,80     | 1,51      | 15,50 |
| 4       | RR1       | Log-binomial GEE             | 0,0930 | 2,97%          | 100,00       | 92,50     | 1,12      | 21,70 |
|         |           | Poisson GEE                  | 0,0935 | 3,04%          | 100,00       | 92,40     | 1,12      | 21,40 |
|         |           | Log-binomial misto           | 0,0916 | 2,41%          | 93,50        | 95,29     | 1,22      | 14,44 |
|         |           | Log-binomial misto bayesiano | 0,0808 | 3,07%          | -            | 96,70     | 1,41      | 12,20 |
|         | RR2       | Log-binomial GEE             | 0,2893 | 4,25%          | 100,00       | 93,20     | 2,00      | 88,00 |
|         |           | Poisson GEE                  | 0,2901 | 4,27%          | 100,00       | 93,20     | 2,00      | 87,70 |
|         |           | Log-binomial misto           | 0,2517 | 2,21%          | 93,50        | 96,68     | 2,05      | 85,67 |
|         |           | Log-binomial misto bayesiano | 0,2221 | 2,50%          | -            | 95,60     | 2,16      | 88,00 |
| 5       | RR1       | Log-binomial GEE             | 0,1335 | 3,89%          | 100,00       | 92,60     | 1,33      | 17,40 |
|         |           | Poisson GEE                  | 0,1329 | 3,92%          | 100,00       | 92,70     | 1,33      | 17,30 |
|         |           | Log-binomial misto           | 0,1232 | 2,77%          | 89,50        | 92,85     | 1,34      | 14,86 |
|         |           | Log-binomial misto bayesiano | 0,1125 | 3,19%          | -            | 95,50     | 1,65      | 11,20 |
|         | RR2       | Log-binomial GEE             | 0,3138 | 4,01%          | 100,00       | 93,70     | 2,00      | 86,80 |
|         |           | Poisson GEE                  | 0,3147 | 4,00%          | 100,00       | 93,80     | 2,00      | 86,60 |
|         |           | Log-binomial misto           | 0,2419 | 1,19%          | 89,50        | 95,64     | 2,01      | 84,25 |
|         |           | Log-binomial misto bayesiano | 0,2419 | 2,85%          | -            | 94,70     | 2,16      | 87,40 |
| 6       | RR1       | Log-binomial GEE             | 0,2718 | 2,97%          | 100,00       | 91,00     | 1,99      | 11,90 |
|         |           | Poisson GEE                  | 0,2658 | 2,87%          | 100,00       | 90,90     | 1,97      | 11,70 |
|         |           | Log-binomial misto           | 0,2399 | 0,33%          | 66,50        | 86,62     | 1,80      | 14,44 |
|         |           | Log-binomial misto bayesiano | 0,2339 | 10,54%         | -            | 94,50     | 2,82      | 7,30  |
|         | RR2       | Log-binomial GEE             | 0,4112 | 5,08%          | 100,00       | 92,90     | 2,06      | 84,60 |
|         |           | Poisson GEE                  | 0,4129 | 5,01%          | 100,00       | 93,20     | 2,06      | 84,40 |
|         |           | Log-binomial misto           | 0,4470 | 2,40%          | 66,50        | 89,17     | 2,04      | 79,40 |
|         |           | Log-binomial misto bayesiano | 0,2890 | 2,59%          | -            | 94,20     | 2,21      | 86,10 |

| Cenário | Parâmetro | Modelo                       | EQM    | Vício relativo | Convergência | Cobertura | Amplitude | Poder |
|---------|-----------|------------------------------|--------|----------------|--------------|-----------|-----------|-------|
| 7       | RR1       | Log-binomial GEE             | 0,3098 | 4,68%          | 100,00       | 92,60     | 1,99      | 86,50 |
|         |           | Poisson GEE                  | 0,3102 | 4,72%          | 100,00       | 92,50     | 1,99      | 86,30 |
|         |           | Log-binomial misto           | 0,3109 | 4,67%          | 99,20        | 95,16     | 2,17      | 82,66 |
|         |           | Log-binomial misto bayesiano | 0,2504 | 2,35%          | -            | 96,30     | 2,55      | 78,70 |
|         | RR2       | Log-binomial GEE             | 0,0919 | 3,17%          | 100,00       | 93,10     | 1,11      | 21,80 |
|         |           | Poisson GEE                  | 0,0931 | 3,20%          | 100,00       | 93,50     | 1,12      | 21,80 |
|         |           | Log-binomial misto           | 0,0856 | 2,63%          | 99,20        | 96,17     | 1,17      | 16,73 |
|         |           | Log-binomial misto bayesiano | 0,0774 | 2,12%          | -            | 95,30     | 1,20      | 18,30 |
| 8       | RR1       | Log-binomial GEE             | 0,3994 | 5,03%          | 100,00       | 92,10     | 2,31      | 75,20 |
|         |           | Poisson GEE                  | 0,3990 | 5,03%          | 100,00       | 92,00     | 2,31      | 75,60 |
|         |           | Log-binomial misto           | 0,3909 | 4,71%          | 96,00        | 93,54     | 2,38      | 73,44 |
|         |           | Log-binomial misto bayesiano | 0,3222 | 3,04%          | -            | 95,90     | 2,95      | 68,20 |
|         | RR2       | Log-binomial GEE             | 0,0942 | 3,14%          | 100,00       | 93,60     | 1,12      | 21,70 |
|         |           | Poisson GEE                  | 0,0954 | 3,16%          | 100,00       | 93,60     | 1,13      | 21,60 |
|         |           | Log-binomial misto           | 0,0823 | 1,77%          | 96,00        | 96,77     | 1,17      | 15,94 |
|         |           | Log-binomial misto bayesiano | 0,0786 | 2,17%          | -            | 95,60     | 1,19      | 19,10 |
| 9       | RR1       | Log-binomial GEE             | 1,0363 | 8,52%          | 100,00       | 90,40     | 3,53      | 45,90 |
|         |           | Poisson GEE                  | 1,0187 | 8,30%          | 100,00       | 90,30     | 3,51      | 45,80 |
|         |           | Log-binomial misto           | 0,8447 | 3,97%          | 80,90        | 90,23     | 3,30      | 43,88 |
|         |           | Log-binomial misto bayesiano | 0,9338 | 5,36%          | -            | 94,50     | 5,50      | 36,80 |
|         | RR2       | Log-binomial GEE             | 0,0930 | 2,96%          | 100,00       | 93,60     | 1,14      | 21,70 |
|         |           | Poisson GEE                  | 0,0931 | 2,92%          | 100,00       | 93,60     | 1,14      | 21,40 |
|         |           | Log-binomial misto           | 0,0849 | 0,85%          | 80,90        | 94,44     | 1,15      | 14,96 |
|         |           | Log-binomial misto bayesiano | 0,0781 | 2,22%          | -            | 95,20     | 1,17      | 20,00 |
| 10      | RR1       | Log-binomial GEE             | 0,1906 | 2,49%          | 100,00       | 92,70     | 1,58      | 95,50 |
|         |           | Poisson GEE                  | 0,1926 | 2,57%          | 100,00       | 92,90     | 1,59      | 95,60 |
|         |           | Log-binomial misto           | 0,1832 | 1,12%          | 73,30        | 95,77     | 1,69      | 93,86 |
|         |           | Log-binomial misto bayesiano | 0,1668 | 2,38%          | -            | 97,00     | 1,97      | 92,40 |
|         | RR2       | Log-binomial GEE             | 0,1831 | 3,04%          | 100,00       | 92,60     | 1,58      | 96,90 |
|         |           | Poisson GEE                  | 0,1871 | 3,13%          | 100,00       | 93,10     | 1,59      | 96,90 |
|         |           | Log-binomial misto           | 0,1468 | 1,25%          | 73,30        | 96,45     | 1,60      | 96,45 |
|         |           | Log-binomial misto bayesiano | 0,1516 | 1,62%          | -            | 95,00     | 1,71      | 97,90 |
| 11      | RR1       | Log-binomial GEE             | 0,2978 | 3,63%          | 100,00       | 91,00     | 1,94      | 85,90 |
|         |           | Poisson GEE                  | 0,2936 | 3,64%          | 100,00       | 91,40     | 1,94      | 86,60 |
|         |           | Log-binomial misto           | 0,2645 | 0,09%          | 57,60        | 89,93     | 1,83      | 84,55 |
|         |           | Log-binomial misto bayesiano | 0,2487 | 2,38%          | -            | 95,50     | 2,37      | 80,70 |
|         | RR2       | Log-binomial GEE             | 0,2044 | 2,78%          | 100,00       | 91,50     | 1,59      | 94,80 |
|         |           | Poisson GEE                  | 0,2059 | 2,83%          | 100,00       | 91,10     | 1,60      | 94,90 |
|         |           | Log-binomial misto           | 0,1785 | 1,90%          | 57,60        | 92,53     | 1,60      | 91,84 |
|         |           | Log-binomial misto bayesiano | 0,1680 | 2,22%          | -            | 94,00     | 1,71      | 95,70 |
| 12      | RR1       | Log-binomial GEE             | 0,6496 | 1,87%          | 100,00       | 89,00     | 2,81      | 43,80 |
|         |           | Poisson GEE                  | 0,6310 | 2,11%          | 100,00       | 89,80     | 2,77      | 44,80 |
|         |           | Log-binomial misto           | 0,6207 | 8,04%          | 41,80        | 81,10     | 2,37      | 43,78 |
|         |           | Log-binomial misto bayesiano | 0,5710 | 12,10%         | -            | 95,00     | 3,98      | 34,50 |
|         | RR2       | Log-binomial GEE             | 0,2178 | 2,75%          | 100,00       | 93,10     | 1,68      | 92,60 |
|         |           | Poisson GEE                  | 0,2193 | 2,70%          | 100,00       | 92,90     | 1,69      | 92,10 |
|         |           | Log-binomial misto           | 0,1853 | 2,72%          | 41,80        | 89,71     | 1,64      | 87,08 |
|         |           | Log-binomial misto bayesiano | 0,1786 | 3,43%          | -            | 95,10     | 1,77      | 93,90 |



# Anexos

## 1. Tabelas para os coeficientes

**\*Tabela 3:** Resultados para os coeficientes do modelo

| Cenário | Parâmetro | Modelo                       | Vício relativo | EQM  | Convergência | Cobertura | Poder | Amplitude |
|---------|-----------|------------------------------|----------------|------|--------------|-----------|-------|-----------|
| 1       | B1        | Log-binomial GEE             | 1,38%          | 0,08 | 100,00       | 91,50     | 81,30 | 1,00      |
|         |           | Poisson GEE                  | 1,40%          | 0,08 | 100,00       | 91,50     | 81,20 | 1,00      |
|         |           | Log-binomial misto           | 1,37%          | 0,08 | 99,90        | 94,59     | 76,68 | 1,10      |
|         |           | Log-binomial misto bayesiano | 8,26%          | 0,08 | -            | 96,40     | 10,30 | 1,25      |
|         | B2        | Log-binomial GEE             | 0,57%          | 0,07 | 100,00       | 93,10     | 80,60 | 1,02      |
|         |           | Poisson GEE                  | 0,57%          | 0,07 | 100,00       | 93,20     | 80,60 | 1,02      |
|         |           | Log-binomial misto           | 0,47%          | 0,07 | 99,90        | 94,99     | 80,28 | 1,06      |
|         |           | Log-binomial misto bayesiano | 3,10%          | 0,08 | -            | 94,40     | 14,50 | 1,06      |
| 2       | B1        | Log-binomial GEE             | 1,27%          | 0,09 | 100,00       | 92,90     | 75,70 | 1,12      |
|         |           | Poisson GEE                  | 1,29%          | 0,09 | 100,00       | 93,10     | 75,70 | 1,12      |
|         |           | Log-binomial misto           | 1,14%          | 0,09 | 99,40        | 94,77     | 72,94 | 1,17      |
|         |           | Log-binomial misto bayesiano | 8,25%          | 0,10 | -            | 96,40     | 11,10 | 1,37      |
|         | B2        | Log-binomial GEE             | 0,36%          | 0,08 | 100,00       | 92,90     | 81,80 | 1,02      |
|         |           | Poisson GEE                  | 0,36%          | 0,08 | 100,00       | 93,00     | 82,10 | 1,02      |
|         |           | Log-binomial misto           | 0,12%          | 0,08 | 99,40        | 94,97     | 80,99 | 1,06      |
|         |           | Log-binomial misto bayesiano | 1,83%          | 0,08 | -            | 94,70     | 14,40 | 1,06      |
| 3       | B1        | Log-binomial GEE             | 0,41%          | 0,20 | 100,00       | 90,40     | 52,60 | 1,52      |
|         |           | Poisson GEE                  | 0,43%          | 0,20 | 100,00       | 90,20     | 52,60 | 1,52      |
|         |           | Log-binomial misto           | 1,38%          | 0,20 | 94,20        | 90,87     | 53,18 | 1,53      |
|         |           | Log-binomial misto bayesiano | 1,81%          | 0,22 | -            | 94,30     | 8,20  | 1,96      |
|         | B2        | Log-binomial GEE             | 0,86%          | 0,08 | 100,00       | 91,00     | 81,10 | 1,01      |
|         |           | Poisson GEE                  | 0,83%          | 0,08 | 100,00       | 91,30     | 80,90 | 1,01      |
|         |           | Log-binomial misto           | 0,18%          | 0,08 | 94,20        | 93,63     | 81,74 | 1,05      |
|         |           | Log-binomial misto bayesiano | 3,62%          | 0,08 | -            | 92,80     | 15,50 | 1,06      |
| 4       | B1        | Log-binomial GEE             | 0,26%          | 0,05 | 100,00       | 92,50     | 91,80 | 0,83      |
|         |           | Poisson GEE                  | 0,32%          | 0,05 | 100,00       | 92,40     | 92,10 | 0,84      |
|         |           | Log-binomial misto           | 0,27%          | 0,05 | 93,50        | 95,29     | 90,37 | 0,92      |
|         |           | Log-binomial misto bayesiano | 1,95%          | 0,05 | -            | 96,70     | 12,20 | 1,03      |
|         | B2        | Log-binomial GEE             | 1,19%          | 0,06 | 100,00       | 93,20     | 30,80 | 0,91      |
|         |           | Poisson GEE                  | 1,20%          | 0,06 | 100,00       | 93,20     | 30,70 | 0,91      |
|         |           | Log-binomial misto           | 0,54%          | 0,05 | 93,50        | 96,68     | 26,74 | 0,95      |
|         |           | Log-binomial misto bayesiano | 2,10%          | 0,06 | -            | 95,60     | 88,00 | 0,95      |
| 5       | B1        | Log-binomial GEE             | 0,21%          | 0,07 | 100,00       | 92,60     | 84,10 | 0,97      |
|         |           | Poisson GEE                  | 0,26%          | 0,07 | 100,00       | 92,70     | 84,30 | 0,97      |
|         |           | Log-binomial misto           | 0,74%          | 0,07 | 89,50        | 92,85     | 84,58 | 0,99      |
|         |           | Log-binomial misto bayesiano | 2,31%          | 0,07 | -            | 95,50     | 11,20 | 1,16      |
|         | B2        | Log-binomial GEE             | 0,75%          | 0,06 | 100,00       | 93,70     | 31,50 | 0,91      |
|         |           | Poisson GEE                  | 0,73%          | 0,06 | 100,00       | 93,80     | 31,40 | 0,91      |
|         |           | Log-binomial misto           | 1,58%          | 0,05 | 89,50        | 95,64     | 30,50 | 0,94      |
|         |           | Log-binomial misto bayesiano | 1,33%          | 0,06 | -            | 94,70     | 87,40 | 0,95      |
| 6       | B1        | Log-binomial GEE             | 4,91%          | 0,16 | 100,00       | 91,00     | 61,00 | 1,41      |
|         |           | Poisson GEE                  | 4,88%          | 0,16 | 100,00       | 90,90     | 61,20 | 1,40      |
|         |           | Log-binomial misto           | 6,91%          | 0,15 | 66,50        | 86,62     | 67,37 | 1,32      |
|         |           | Log-binomial misto bayesiano | 16,42%         | 0,18 | -            | 94,50     | 7,30  | 1,78      |
|         | B2        | Log-binomial GEE             | 1,35%          | 0,07 | 100,00       | 92,90     | 30,70 | 0,92      |
|         |           | Poisson GEE                  | 1,27%          | 0,07 | 100,00       | 93,20     | 30,80 | 0,92      |
|         |           | Log-binomial misto           | 1,60%          | 0,07 | 66,50        | 89,17     | 35,94 | 0,93      |
|         |           | Log-binomial misto bayesiano | 1,51%          | 0,07 | -            | 94,20     | 86,10 | 0,95      |

| Cenário | Parâmetro | Modelo                       | Vício relativo | EQM   | Convergência | Cobertura | Poder | Amplitude |
|---------|-----------|------------------------------|----------------|-------|--------------|-----------|-------|-----------|
| 7       | B1        | Log-binomial GEE             | 1,36%          | 0,064 | 100,00       | 92,60     | 29,70 | 0,90      |
|         |           | Poisson GEE                  | 1,39%          | 0,064 | 100,00       | 92,50     | 29,30 | 0,90      |
|         |           | Log-binomial misto           | 1,34%          | 0,064 | 99,20        | 95,16     | 23,89 | 0,98      |
|         |           | Log-binomial misto bayesiano | 3,21%          | 0,066 | -            | 96,30     | 78,70 | 1,10      |
|         | B2        | Log-binomial GEE             | 0,55%          | 0,051 | 100,00       | 93,10     | 92,10 | 0,83      |
|         |           | Poisson GEE                  | 0,55%          | 0,051 | 100,00       | 93,50     | 91,90 | 0,84      |
|         |           | Log-binomial misto           | 0,14%          | 0,049 | 99,20        | 96,17     | 92,14 | 0,88      |
|         |           | Log-binomial misto bayesiano | 2,22%          | 0,050 | -            | 95,30     | 18,30 | 0,88      |
| 8       | B1        | Log-binomial GEE             | 0,82%          | 0,081 | 100,00       | 92,10     | 26,40 | 1,03      |
|         |           | Poisson GEE                  | 0,82%          | 0,081 | 100,00       | 92,00     | 26,50 | 1,03      |
|         |           | Log-binomial misto           | 0,52%          | 0,081 | 96,00        | 93,54     | 24,90 | 1,07      |
|         |           | Log-binomial misto bayesiano | 3,00%          | 0,085 | -            | 95,90     | 68,20 | 1,23      |
|         | B2        | Log-binomial GEE             | 0,45%          | 0,052 | 100,00       | 93,60     | 92,10 | 0,84      |
|         |           | Poisson GEE                  | 0,44%          | 0,053 | 100,00       | 93,60     | 92,00 | 0,84      |
|         |           | Log-binomial misto           | 0,69%          | 0,049 | 96,00        | 96,77     | 92,81 | 0,88      |
|         |           | Log-binomial misto bayesiano | 1,54%          | 0,052 | -            | 95,60     | 19,10 | 0,88      |
| 9       | B1        | Log-binomial GEE             | 0,92%          | 0,177 | 100,00       | 90,40     | 19,90 | 1,45      |
|         |           | Poisson GEE                  | 1,04%          | 0,176 | 100,00       | 90,30     | 19,90 | 1,44      |
|         |           | Log-binomial misto           | 4,46%          | 0,165 | 80,90        | 90,23     | 23,36 | 1,42      |
|         |           | Log-binomial misto bayesiano | 3,12%          | 0,191 | -            | 94,50     | 36,80 | 1,84      |
|         | B2        | Log-binomial GEE             | 0,26%          | 0,053 | 100,00       | 93,60     | 91,80 | 0,85      |
|         |           | Poisson GEE                  | 0,22%          | 0,053 | 100,00       | 93,60     | 91,80 | 0,85      |
|         |           | Log-binomial misto           | 1,66%          | 0,050 | 80,90        | 94,44     | 92,71 | 0,88      |
|         |           | Log-binomial misto bayesiano | 0,32%          | 0,049 | -            | 95,20     | 20,00 | 0,86      |
| 10      | B1        | Log-binomial GEE             | 0,34%          | 0,042 | 100,00       | 92,70     | 40,60 | 0,74      |
|         |           | Poisson GEE                  | 0,41%          | 0,042 | 100,00       | 92,90     | 40,40 | 0,75      |
|         |           | Log-binomial misto           | 1,00%          | 0,042 | 73,30        | 95,77     | 35,74 | 0,81      |
|         |           | Log-binomial misto bayesiano | 1,10%          | 0,042 | -            | 97,00     | 92,40 | 0,90      |
|         | B2        | Log-binomial GEE             | 1,04%          | 0,038 | 100,00       | 92,60     | 41,70 | 0,74      |
|         |           | Poisson GEE                  | 1,08%          | 0,039 | 100,00       | 93,10     | 40,70 | 0,75      |
|         |           | Log-binomial misto           | 2,99%          | 0,035 | 73,30        | 96,45     | 43,52 | 0,78      |
|         |           | Log-binomial misto bayesiano | 1,53%          | 0,038 | -            | 95,00     | 97,90 | 0,78      |
| 11      | B1        | Log-binomial GEE             | 0,39%          | 0,063 | 100,00       | 91,00     | 31,30 | 0,89      |
|         |           | Poisson GEE                  | 0,45%          | 0,062 | 100,00       | 91,40     | 31,10 | 0,89      |
|         |           | Log-binomial misto           | 3,13%          | 0,061 | 57,60        | 89,93     | 34,03 | 0,87      |
|         |           | Log-binomial misto bayesiano | 1,65%          | 0,063 | -            | 95,50     | 80,70 | 1,04      |
|         | B2        | Log-binomial GEE             | 0,49%          | 0,044 | 100,00       | 91,50     | 40,30 | 0,75      |
|         |           | Poisson GEE                  | 0,53%          | 0,044 | 100,00       | 91,10     | 39,70 | 0,75      |
|         |           | Log-binomial misto           | 4,08%          | 0,044 | 57,60        | 92,53     | 45,49 | 0,78      |
|         |           | Log-binomial misto bayesiano | 0,38%          | 0,043 | -            | 94,00     | 95,70 | 0,79      |
| 12      | B1        | Log-binomial GEE             | 9,22%          | 0,151 | 100,00       | 89,00     | 27,20 | 1,32      |
|         |           | Poisson GEE                  | 9,29%          | 0,148 | 100,00       | 89,80     | 26,60 | 1,30      |
|         |           | Log-binomial misto           | 15,58%         | 0,162 | 41,80        | 81,10     | 39,00 | 1,19      |
|         |           | Log-binomial misto bayesiano | 9,10%          | 0,153 | -            | 95,00     | 34,50 | 1,63      |
|         | B2        | Log-binomial GEE             | 0,33%          | 0,047 | 100,00       | 93,10     | 38,30 | 0,78      |
|         |           | Poisson GEE                  | 0,26%          | 0,047 | 100,00       | 92,90     | 38,00 | 0,79      |
|         |           | Log-binomial misto           | 5,07%          | 0,048 | 41,80        | 89,71     | 45,93 | 0,81      |
|         |           | Log-binomial misto bayesiano | 1,04%          | 0,046 | -            | 95,10     | 93,90 | 0,81      |

\*Para os coeficientes, a mediana foi utilizada como estimador bayesiano pontual, pois houve um erro no cálculo da moda.

## 2. Códigos

#GEE - Modelo log -binomial ou Poisson

```
fun_gee= function(banco,fam,des,trat,cov,dep,conf){
  mod1 <- geeglm(des~trat+cov,
    id=dep,family=fam(link=log),corstr = "exchangeable", data=banco)
  mod1=summary(mod1)
  a.1=mod1$coefficients$Estimate[1]
  a.2=mod1$coefficients$Estimate[2]
  a.3=mod1$coefficients$Estimate[3]

  b.1=mod1$coefficients$Std.err[1]
  b.2=mod1$coefficients$Std.err[2]
  b.3=mod1$coefficients$Std.err[3]

  c.1=mod1$coefficients$Pr[1]
  c.2=mod1$coefficients$Pr[2]
  c.3=mod1$coefficients$Pr[3]

  confint.geeglm = function(mod1, parm, level = sig, ...) {
    cc <- coef(summary(object))
    mult <- qnorm((1+conf)/2)
    citab <- with(as.data.frame(cc),
      cbind(lwr=Estimate-mult*Std.err,upr=Estimate+mult*Std.err))
    mult <- qnorm((1+conf)/2)
    d.1=mod1$coefficients$Estimate[1]-mult*mod1$coefficients$Std.err[1]
    d.2=mod1$coefficients$Estimate[1]+mult*mod1$coefficients$Std.err[1]
    d.3=mod1$coefficients$Estimate[2]-mult*mod1$coefficients$Std.err[2]
    d.4=mod1$coefficients$Estimate[2]+mult*mod1$coefficients$Std.err[2]
    d.5=mod1$coefficients$Estimate[3]-mult*mod1$coefficients$Std.err[3]
    d.6=mod1$coefficients$Estimate[3]+mult*mod1$coefficients$Std.err[3]
    e.1=exp(mod1$coefficients$Estimate[1])
    e.2=exp(mod1$coefficients$Estimate[2])
    e.3=exp(mod1$coefficients$Estimate[3])

    z=exp(as.vector(c(d.1,d.2,d.3,d.4,d.5,d.6)))
    correlacao = unlist(mod1$corr)

    resultados=matrix(c(a.1,a.2,a.3,b.1,b.2,b.3,c.1,c.2,c.3,d.3,d.5,d.2,d.4,d.6,z[1],z[3],z[5],z[2],z[4],z[6],e.1,e.2,e.3),
      ncol=8,nrow=3,dimnames =list(c("B0","B1","B2"),c("coef","erro","p-valor","B-lwr","B-upr","RR-lwr","RR-
      upr","RR")))
    resultados=list(resultados, correlacao)
  }

  #Log-binomial misto frequentista
  fun_misto= function(banco,fam,des,trat,cov,dep,conf){ mod1 = glmer(des ~ as.factor(trat)+ cov + (1 | dep),
    family = fam(link=log), data = banco, nAGQ=2)
  mod1=summary(mod1)
  a.1=mod1$coefficients[1,1]
  a.2=mod1$coefficients[2,1]
  a.3=mod1$coefficients[3,1]
  b.1=mod1$coefficients[1,2]
  b.2=mod1$coefficients[2,2]
  b.3=mod1$coefficients[3,2]
  c.1=mod1$coefficients[1,4]
  c.2=mod1$coefficients[2,4]
  c.3=mod1$coefficients[3,4]
```

```

#lmites IC
confint.geeglm = function(mod1, parm, level = sig, ...) {
  cc <- coef(summary(object))
  mult <- qnorm((1+conf)/2)
  citab <- with(as.data.frame(cc),
    cbind(lwr=Estimate-mult*Std.err,upr=Estimate+mult*Std.err))}
mult <- qnorm((1+conf)/2)
d.1=a.1-mult*b.1
d.2=a.1+mult*b.1

d.3=a.2-mult*b.2
d.4=a.2+mult*b.2

d.5=a.3-mult*b.3
d.6=a.3+mult*b.3

e.1=exp(a.1)
e.2=exp(a.2)
e.3=exp(a.3)

z=exp(as.vector(c(d.1,d.2,d.3,d.4,d.5,d.6)))
resultados=matrix(c(a.1,a.2,a.3,b.1,b.2,b.3,c.1,c.2,c.3,d.1,d.3,d.5,d.2,d.4,d.6,z[1],z[3],z[5],z[2],z[4],z[6],
e.1,e.2,e.3), ncol=8,nrow=3,dimnames =list(c("B0", "B1", "B2"),c("coef", "erro", "p-valor", "B-lwr", "B-upr", "RR-
lwr", "RR-upr", "RR")))
correl = c(unlist(mod1$varcor[1]),sqrt(unlist(mod1$varcor[1])));names(correl)=c("variancia", "desvio")
resultados = list(resultados,correl)
}

#Log-binomial misto Bayesiano
writeLines(
  "model{
  for (i in 1:N) {
  log(p[i]) <- theta0 * Z0[i] + theta1 * Z1[i] + theta2 * Z2[i] + u[cluster[i]]
  y[i] ~ dbern(p[i])}

  for(j in 1:ncl){
  u[j] ~ dnorm(0,tau)}

  theta0 ~ dnorm(0, 0.000001)
  theta1 ~ dnorm(0, 0.000001)
  theta2 ~ dnorm(0, 0.000001)

  sigma ~ dunif(0.01, 100)
  tau <- 1/(sigma*sigma)
  s2 <- sigma*sigma

  for (i in 1:N){
  ones[i] ~ dbern(C1[i])
  C1[i] <- step(1-p[i])}
  }", "model_yelland_reparam.txt")

library(rjags)
library(geepack)

poigee <- geeglm(formula = outcome_clus ~ as.factor(treatment) + as.factor(bincov_clus),
  data = dados,
  family = poisson(link = "log"),
  id = cluster_id,
  corstr = "independence")

```

```

S3 = poigee$geese$vbeta
X = model.matrix(poigee)
n = nrow(X)

Sigma <- S3
L <- chol(Sigma)
Z <- X%*%t(L)

dados.jags = list(N = n, ncl = 20, Z0 = Z[,1], Z1 = Z[,2], Z2 = Z[,3],
  y = dados$outcome_clus,
  cluster = dados$cluster_id)

th.ini1 = matrix(c(log(0.5),0,0),ncol=3)%*%solve(L)
th.ini2 = matrix(c(log(0.1),0,0),ncol=3)%*%solve(L)

inits.jags = list(
  list(theta0 = th.ini1[1], theta1 = th.ini1[2], theta2 = th.ini1[3], sigma=1, u=rep(-3,20)),
  list(theta0 = th.ini2[1], theta1 = th.ini2[2], theta2 = th.ini1[3], sigma=1, u=rep(-3,20)))

m <- jags.model("model_yelland_reparam.txt", dados.jags, inits.jags, n.chains=ncad, n.adapt=(nburn/2),
quiet=TRUE)
update(m, (nburn/2), progress.bar="none")
mcmc <- coda.samples(m, parameters, n.iter=niter,thin=nthin, progress.bar="none")

```