

---

Um estudo de estresse através dos níveis de cortisol  
em crianças

*Karine Zanuto Mendes*

---



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS  
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA  
UFSCar-USP

KARINE ZANUTO MENDES

**UM ESTUDO DE ESTRESSE ATRAVÉS DOS NÍVEIS DE  
CORTISOL EM CRIANÇAS**

Dissertação apresentada ao Departamento de Estatística – DEs-UFSCar e ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestra em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.

Orientador: Prof. Dr. Adriano Polpo de Campos

**São Carlos  
Junho de 2017**



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS  
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA  
UFSCar-USP

KARINE ZANUTO MENDES

**A STUDY OF STRESS THROUGH CORTISOL LEVELS IN  
CHILDREN**

Master dissertation submitted to the Departamento de Estatística – DEs-UFSCar and to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master joint Graduate Program in Statistics.

Advisor: Prof. Dr. Adriano Polpo de Campos

**São Carlos  
June 2017**





UNIVERSIDADE FEDERAL DE SÃO CARLOS  
Centro de Ciências Exatas e de Tecnologia  
Programa Interinstitucional de Pós-Graduação em Estatística

---

Folha de Aprovação

---

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a defesa de dissertação de mestrado da candidata Karine Zanuto Mendes realizada em 26/05/2017:

---

Prof. Dr. Adriano Polpo de Campos  
UFSCar

---

Prof. Dr. Flávio Bambirra Gonçalves  
UFMG

---

Prof. Dr. Victor Fossaluzza  
USP

Certifico que a sessão de defesa foi realizada com a participação à distância do membro Prof. Dr. Flávio Bambirra Gonçalves e, depois das arguições e deliberações realizadas, o participante à distância está de acordo com o conteúdo do parecer da comissão examinadora redigido no relatório de defesa do(a) aluno(a) Karine Zanuto Mendes.

---

Prof. Dr. Adriano Polpo de Campos  
Presidente da Comissão Examinadora  
UFSCar





# Agradecimentos

Agradeço primeiramente aos meus pais, Carlos e Silmara por todo o apoio durante todos esses anos, pela oportunidade de estudar durante tanto tempo e pelo carinho e amor que sempre me deram.

Agradeço a minha irmã Aline, que é, e sempre será o meu maior exemplo de vida. Que sempre me apoiou e torceu pelo meu sucesso.

A minha amiga Estela, que foi a minha segunda família em São Carlos, e que com certeza, é alguém que eu quero levar por toda a minha vida.

A minha amiga Nádia e família, por sempre me convidar para os almoços em família aos domingos e por me tratarem como se eu realmente fizesse parte dela.

Aos meus amigos da turma de 2010 da Estatística UFSCar, em especial, aos meus amigos Nicholas, Bárbara, Natália e Marcela, por toda a paciência e pelos estudos realizados em conjunto que sempre me ajudaram muito e foi de imensa importância para mim, além da amizade construída.

As minhas amigas Taís e Gretta, que são pessoas que eu tanto admiro e que me proporcionaram momentos de tantas risadas.

As minhas amigas do vôleitão, Lígia, Ju, Ice, Ti, Nize e Carol, por estarem comigo nos melhores e piores momentos de toda a graduação e mestrado.

Ao meu orientador e grande amigo Adriano Polpo de Campos, que nunca duvidou da minha capacidade e por ser esse ser humano tão especial que sempre fez de tudo pra me ajudar nos momentos mais difíceis.

Ao Programa Interinstitucional de Pós-Graduação em Estatística e a cada um de seus docentes, por colaborarem na minha formação.

E por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro.



# Resumo

O nível de cortisol é considerado uma forma de medir o estresse de pessoas. Um estudo foi realizado a fim de verificar se crianças que trabalhavam nas ruas durante o dia tem estresse mais alto do que crianças que não trabalhavam. O nível de cortisol de uma pessoa pode ser considerado uma função crescente até atingir um máximo e depois decrescente (função quasicôncava). O cortisol das crianças foram coletados 4 vezes ao dia, sendo considerado dois grupos de crianças: aquelas que trabalham na rua e aquelas que ficavam em casa. Para a análise dos dados, foi considerada uma metanálise de um modelo de dados funcionais sob enfoque Bayesiano. Cada indivíduo é analisado por um modelo de dados funcionais e a metanálise foi usada para fazer uma inferência para cada grupo. A geração de uma amostra da distribuição a posteriori foi obtida pelo o método de Gibbs com Metrópolis-Hasting. Na comparação das curvas calculamos a probabilidade a posteriori ponto-a-ponto da função do cortisol de um grupo ser maior do que a do outro.

**Palavras-chave:** Eixo HPA, cortisol, dados funcionais, estatística Bayesiana, metanálise.



# Abstract

The level of cortisol is considered as a measure of people's stress. We perform an statistical analysis of the data from a study conducted to evaluate if children that work on the streets during the day have higher stress than children who does not work. The cortisol level of a person can be considered as an increasing function until reaching a maximum level and then decreasing to almost zero (quasi-concave function). Children's cortisol were collected 4 times in one day, where two groups of children were considered: those who work in the street and those who stay at home. To analyse the data we considered a meta-analysis of a functional data model under Bayesian approach. Each individual is analysed by a functional data model, and then, a meta-analysis was used to have inference for each group. We used the Gibbs Metropolis-Hastings method to sample from the posteriori distribution. Also, we calculated the pointwise posterior probability of the cortisol function of one group being greater than the cortisol function of other group to compare the groups.

**Key words** : HPA axis, cortisol, functional data, Bayesian statistics, meta-analysis.



## Lista de ilustrações

Figura 1	–	Boxplot para os tempos de observações do cortisol das crianças, em que <i>hora0</i> , <i>hora1</i> , <i>hora2</i> e <i>hora3</i> representam o horário das 4 medições. . . . .	13
Figura 2	–	Boxplot para os tempos de observações do cortisol das crianças, em que <i>t1</i> , <i>t2</i> e <i>t3</i> representam o tempo da primeira, segunda e terceira medição após o acordar. . . . .	14
Figura 3	–	Dados observados referentes a 4 indivíduos, sendo (a) e (b) crianças na rua e (c) e (d) em casa. . . . .	15
Figura 4	–	Média a posteriori das funções de cortisol e intervalo de credibilidade ponto-a-ponto de 95%, referente aos 4 indivíduos da Figura 3. . . . .	22
Figura 5	–	Metanálise, em que as linhas contínuas representam a média da medida metanalítica de cada grupo e as linhas tracejadas representam o intervalo de credibilidade central ponto-a-ponto, com 95% de probabilidade (curvas vermelhas - indivíduos que trabalham nas ruas e curvas verdes - indivíduos que não trabalham nas ruas). . . . .	23
Figura 6	–	Probabilidade ponto-a-ponto a posteriori dos níveis de cortisol do grupo <b>rua</b> serem maiores do que o do grupo <b>não rua</b> . . . . .	24
Figura 7	–	Média metanalítica marginal a posteriori dos parâmetros da função base de cada grupo (curvas vermelhas - indivíduos que trabalham nas ruas e curvas verdes - indivíduos que não trabalham nas ruas). . . . .	25

## Lista de tabelas

Tabela 1	–	Número de famílias considerando a quantidade de filhos na rua e filhos em casa. . . . .	12
----------	---	---	----

# Sumário

	<b>Lista de ilustrações . . . . .</b>	<b>7</b>
	<b>Lista de tabelas . . . . .</b>	<b>7</b>
<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>9</b>
<b>2</b>	<b>DESCRIÇÃO DO ESTUDO . . . . .</b>	<b>11</b>
<b>3</b>	<b>METODOLOGIA E ANÁLISE DOS DADOS . . . . .</b>	<b>17</b>
<b>3.1</b>	<b>Dados funcionais . . . . .</b>	<b>17</b>
<b>3.2</b>	<b>Comparação entre os níveis de cortisol de indivíduos que trabalhavam nas ruas e os que não trabalhavam. . . . .</b>	<b>20</b>
<b>4</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>27</b>
	<b>Referências . . . . .</b>	<b>31</b>
<b>5</b>	<b>APÊNDICES . . . . .</b>	<b>34</b>
<b>5.1</b>	<b>Funções densidade a posteriori dos indivíduos que trabalhavam nas ruas. . . . .</b>	<b>34</b>
<b>5.2</b>	<b>Funções densidade a posteriori dos indivíduos que não trabalhavam. . . . .</b>	<b>49</b>



# 1 Introdução

Estudos mostram que crianças com traumas na infância possuem uma atividade aumentada no eixo hipotálamo-pituitária-adrenal (HPA). “*Este eixo exerce um papel fundamental na resposta aos estímulos externos e internos, incluindo os estressores psicológicos*” (Mario F Juruena, 2004). O cortisol, que é o produto final do eixo HPA, está diretamente envolvido na resposta ao estresse e é responsável pelas variações nas reações comportamentais, emocionais e cognitivas dos indivíduos. Considerado como o hormônio do estresse, ele estimula uma resposta do corpo antes de situações emergenciais, aumentando a pressão arterial e o açúcar no sangue, para propiciar energia muscular. Para a obtenção de energia, todas as funções anabólicas de recuperação, renovação e criação de tecidos são paralisadas. Uma vez que o estresse é pontual, os níveis hormonais e o processo fisiológico volta à normalidade, mas quando este se prolonga, a atividade do eixo HPA aumenta e, conseqüentemente os níveis de cortisol no organismo disparam e causam um aumento no risco de doenças psicóticas na fase adulta, como por exemplo a esquizofrenia (Mello et al., 2003). Já segundo Juruena et al. (2004), o aumento da atividade do eixo HPA ocorre em pessoas depressivas.

O estudo que motivou este trabalho foi realizado pelo PROVE - grupo de pesquisa em psiquiatria, ligado à Universidade Federal de São Paulo (UNIFESP), que atende crianças, jovens e adultos que desenvolvem transtornos mentais após uma experiência de violência. As famílias foram incluídas no estudo após o cuidador principal assinar um termo de consentimento livre e esclarecido, aprovado pelo comitê de ética institucional da UNIFESP (Mello et al., 2013).

O estudo selecionou crianças que trabalhavam nas ruas (em geral, vendendo balas no farol, pedindo dinheiro, etc) e que não trabalhavam (ficavam em casa). Os cuidadores destas crianças foram treinados para a coleta da saliva das crianças (a saliva é utilizada para medir o nível de cortisol). A saliva foi coletada em 4 instantes de tempo durante o dia, pois os especialistas entendem que o cortisol tem uma característica funcional, sendo baixo ao acordar, crescendo rapidamente até atingir

seu máximo e diminuindo gradativamente após o máximo atingido.

Considerando todos estes fatos, foi possível utilizar técnicas de dados funcionais que levam em conta todas as características do comportamento do cortisol ao longo do dia para termos uma comparação mais realista.

A sequência deste trabalho está organizada da seguinte forma: No Capítulo 2, dedicamo-nos a detalhar o problema e descrever as características dos dados observados. No Capítulo 3, apresentamos o método de dados funcionais, discutindo o seu uso no nosso problema específico, a análise exploratória dos dados funcionais, as análises inferenciais, a metanálise e a probabilidade a posteriori para a comparação dos dois grupos (**rua** e **não rua**). Por fim, no Capítulo 4, são apresentados os comentários finais, destacando os resultados obtidos, as dificuldades encontradas e as soluções utilizadas.

## 2 Descrição do estudo

Para esse estudo, foram selecionadas 211 crianças que tinham idades entre 7 e 14 anos. Apenas 191 foram incluídas no projeto, sendo 126 crianças que trabalhavam nas ruas durante o dia (grupo **rua**) e 65 que não trabalhavam nas ruas (grupo **não rua**). Além disso, algumas dessas crianças pertenciam a mesma família. Foram excluídas do estudo as crianças que possuíam retardo mental, distúrbios mentais e neurológicos que não estavam sob controle ou doenças que podem alterar o eixo HPA (Síndrome de Cushing, por exemplo, ou deficiência de glicocorticóides)(Mello et al., 2013).

Os pesquisadores do PROVE acreditam que as quantidades de cortisol medidas nas salivas das crianças que trabalhavam nas ruas é maior que as medidas das crianças que não trabalhavam, em função do estresse gerado por tal situação.

As medidas de cortisol foram idealmente tomadas em 4 tempos distintos: ao acordar, meia hora após acordar, às 12h e às 16h.

Durante a coleta dos dados ocorreram alguns problemas com os horários de medição (não foram respeitados os horários previstos) e os horários que as crianças acordavam não eram homogêneos (crianças que acordavam às 5 horas e outras que acordavam às 10 horas). Outro possível problema foi o fato dos cuidadores das crianças terem feito a coleta, apesar de terem passado por treinamento para tal. A saliva foi coletada em suas casas e refrigeradas por uma noite em uma geladeira doméstica, antes de ser refrigerada a  $-20^{\circ}\text{C}$  até a análise. A unidade de medida utilizada foi  $\mu\text{g}/\text{dL}$ .

A Tabela 1 apresenta o número de famílias, considerando a quantidade de filhos na rua e filhos em casa, sendo que a maioria das famílias possuem até 3 filhos. Notamos que das 26 famílias que possuem 1 filho, 23 crianças trabalhavam nas ruas. Para as famílias com 2 filhos percebemos que somente 6 destas têm os dois filhos que não trabalhavam, resultado ainda mais evidente para as famílias com 3 filhos em que apenas uma delas tem os 3 filhos sem trabalho. Nas famílias com 4 a 7 crianças

temos que em nenhum dos casos todos os filhos não trabalhavam.

Tabela 1 – Número de famílias considerando a quantidade de filhos na rua e filhos em casa.

Número de filhos em casa	Número de filhos na rua					
	0	1	2	3	4	5
0	0	23	10	10	1	1
1	3	7	8	0	1	0
2	6	7	2	0	0	0
3	1	0	1	0	0	0
4	0	0	1	0	0	0
5	0	0	0	0	0	0
6	0	1	0	0	0	0

Sabemos que, segundo especialistas, o nível de cortisol cresce até 30 minutos após o despertar, decaindo após atingir o máximo. Como ressaltado anteriormente, os horários das medições foram irregulares.

Na Figura 1, verificamos que a variabilidade das medições para “hora0” e “hora1” foi grande. Notamos que, “hora0” é a medição ao acordar e “hora1” seria 30 minutos após ao acordar, logo, há grande variação nos horários em que as crianças acordam. Já para as medições que eram para serem feitas às 12h e às 16h, temos que as medianas para esses horários estão bem próximas dos valores propostos pelo estudo, porém verificamos que muitos indivíduos tem suas medições após os horários previstos para a coleta. Desta forma, trabalhamos com o horário ao acordar como o tempo 0 (zero) e os outros instantes de tempo foram considerados como o tempo decorrido após a criança acordar. Na Figura 2 apresentamos os resultados para os três instantes de tempo após acordar, visto que no primeiro instante todos os indivíduos tem o horário de medição considerado 0 (zero). Notamos que a mediana do “t1” é próxima de 30 minutos, porém existem alguns indivíduos com medições superior a este horário.

Na Figura 3 apresentamos quatro indivíduos. Consideramos o método de dados funcionais para realizar a análise dos dados. Para cada indivíduo, observamos a função do cortisol ao longo do dia (após acordar). Entretanto, como características de dados funcionais, de fato, o que observamos são alguns pontos da função e não a

função completamente. Além disso, neste caso específico, apenas 4 pontos da função foram coletados e isso dificulta bastante a análise estatística.

Com isso, temos uma importante pergunta: O cortisol das crianças que trabalhavam nas ruas é maior do que o das crianças que não trabalhavam?

Um fato importante a ser destacado é que consideramos que todos os indivíduos observados são independentes. Sabemos que vários indivíduos possuem irmãos no conjunto de dados (mesma família) o que indica uma possível dependência entre os indivíduos. Entretanto, a complexidade do problema nos levou a considerar a independência entre os indivíduos.

No Capítulo 3 apresentamos como os quatro pontos observados de cada criança foram tratados para obtermos as funções de cortisol dos indivíduos e descrevemos a técnica utilizada para a análise destes dados funcionais.

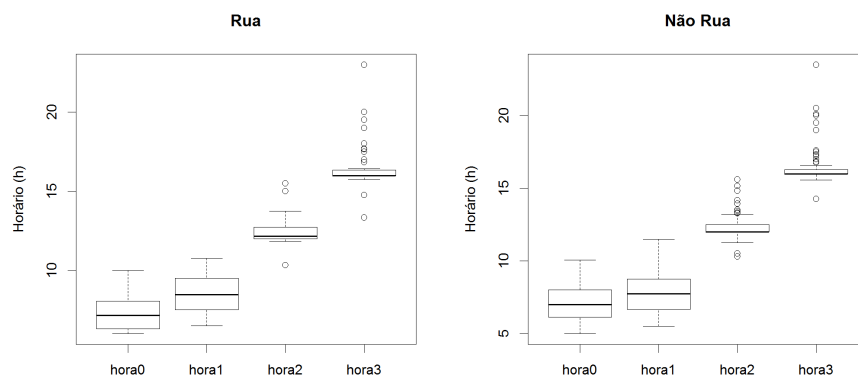


Figura 1 – Boxplot para os tempos de observações do cortisol das crianças, em que *hora0*, *hora1*, *hora2* e *hora3* representam o horário das 4 medições.

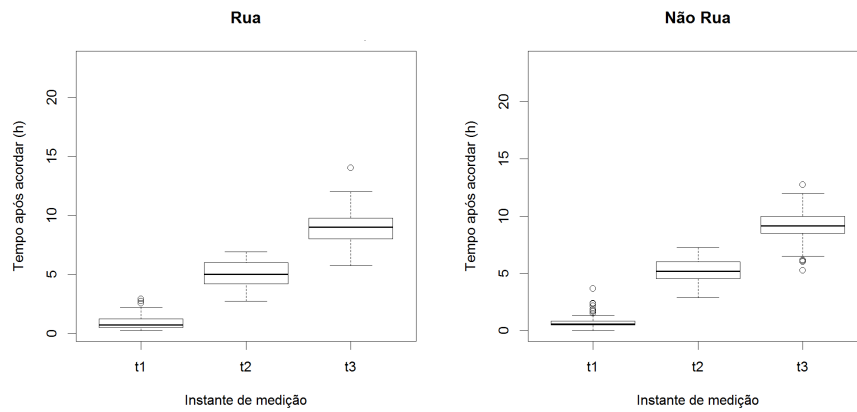


Figura 2 – Boxplot para os tempos de observações do cortisol das crianças, em que  $t_1$ ,  $t_2$  e  $t_3$  representam o tempo da primeira, segunda e terceira medição após o acordar.

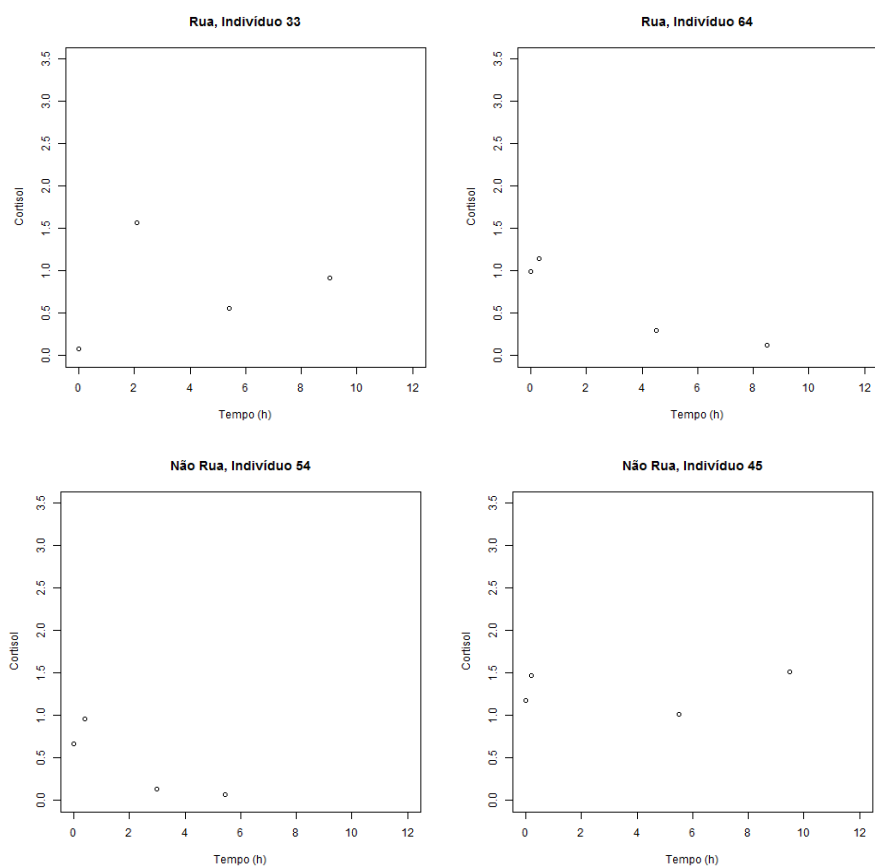


Figura 3 – Dados observados referentes a 4 indivíduos, sendo (a) e (b) crianças na rua e (c) e (d) em casa.





## 3 Metodologia e Análise dos dados

Como já discutido, a comparação entre os níveis de cortisol dos indivíduos que trabalham nas ruas e os que não trabalham foi baseada em uma análise de dados funcionais. Desta forma, apresentamos na Seção 3.1 a construção das funções de cada indivíduo e o método para análise inferencial.

### 3.1 Dados funcionais

Primeiramente precisamos definir o que é um dado funcional. Em geral, uma análise de dados funcionais (ADF) é composta por dados em que a  $i$ -ésima observação é um conjunto de pontos provenientes de uma função real, ou seja,  $x_i(t)$ ,  $i = 1, \dots, n$ ,  $t \in T$ , em que  $T$  é um intervalo real, portanto, cada  $x_i$  é um ponto em algum espaço de funções  $H$  (Ramsay & Dalzell, 2012).

Embora algumas idéias sejam de muito tempo atrás, o termo *análise de dados funcionais* (ADF) foi criado em 1991 por Ramsay e Dalzell. A principal novidade é a noção de ADF como uma maneira de pensar e não mais como um conjunto de métodos e técnicas.

Alguns objetivos da ADF são:

- representar os dados de forma a auxiliar futuras análises;
- estudar fontes importantes de padrão e variação entre os dados;
- explicar a variação em um resultado ou variável dependente, usando informações da variável independente.
- comparar dois ou mais conjuntos de dados com respeito a certos tipos de variação, onde dois conjuntos podem conter diferentes conjuntos de replicações das mesmas funções ou diferentes funções para um conjunto comum de replicações.

Neste caso, queremos uma função que seja simples, mas ao mesmo tempo flexível para acomodar as características que os especialistas descrevem sobre a natureza do cortisol (nível baixo ao acordar, cresce rapidamente atingindo um pico pouco tempo após acordar e decai ao longo do dia). Com isto em mente, construímos uma função base para cada indivíduo da seguinte forma: consideramos uma função que tenha as características de interesse (crescimento, pico, decaimento) e, para cada observação do cortisol, tomamos uma função com o ponto de máximo no valor de tempo observado, obtendo uma função para cada ponto observado; a função base a ser considerada é a média das funções de cada ponto.

Este processo é descrito por:

1. Consideramos o  $i$ -ésimo indivíduo, cujos valores observados são  $(c_{i1}, t_{i1}), \dots, (c_{in_i}, t_{in_i})$ .
2. A função a ser utilizada para cada ponto observado é

$$h(t \mid k, \lambda) = \frac{kt^{k-1}}{\lambda^k} \exp \left\{ - \left( \frac{t}{\lambda} \right)^k \right\}, \quad (3.1)$$

que possui as características desejadas (crescimento, pico, decaimento).

3. Obtemos os valores de  $k_j$  e  $\lambda_j$  para cada par observado  $(c_{ij}, t_{ij})$ ,  $j = 1, \dots, n_i$ , tal que para cada par  $(c_{ij}, t_{ij})$ , obtemos um par  $(k_j, \lambda_j)$  tal que o ponto de máximo de  $h(\cdot)$  no ponto  $t_{ij}$  seja igual a  $c_{ij}$ . O par  $(k_j, \lambda_j)$  é determinístico e especifica uma função  $h(\cdot \mid k_j, \lambda_j)$ .
4. A função base, do  $i$ -ésimo indivíduo, é dada por

$$\tilde{h}_i(t) = \sum_{j=1}^{n_i} \frac{h(t \mid k_j, \lambda_j)}{n_i} \quad (3.2)$$

Uma vez que temos a função base construída, passamos para a etapa de estimação dos níveis de cortisol de cada indivíduo, que consiste na estimação de uma função não negativa  $g(t \mid a, b) = e^{a+bt}$  que é multiplicada a função base. A função  $g(\cdot)$  tem o papel de “adequar” a função base aos dados observados, uma vez que a função base é uma média de funções que garantem a característica de interesse mas não

necessariamente próxima aos valores observados. Desta forma, multiplicar a função em cada  $t$  por diferentes valores, obtemos uma função estimada do cortisol pelo tempo para cada indivíduo.

O modelo considerado, para o  $i$ -ésimo indivíduo é

$$C_{ij} \mid t_{ij}, a_i, b_i, \sigma_i \sim N(\tilde{h}_i(t_{ij})g(t_{ij} \mid a_i, b_i), \sigma_i),$$

sendo os parâmetros a serem estimados  $a_i, b_i \in \mathbb{R}$  e  $\sigma_i > 0$ ,  $C_{ij}$  são variáveis aleatórias condicionalmente independentes,  $j = 1, \dots, n_i$  e  $c_{ij}$  o seu valor observado em  $t_{ij}$ .

Sendo  $\text{dados}_i = \{c_{i1}, \dots, c_{in_i}, t_{i1}, \dots, t_{in_i}\}$ , a função de verossimilhança é

$$L_i(a_i, b_i, \sigma_i \mid \text{dados}_i) = \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left\{ - \frac{[c_{ij} - \tilde{h}_i(t_{ij})g(t_{ij} \mid a_i, b_i)]^2}{2\sigma_i^2} \right\}. \quad (3.3)$$

Consideramos o paradigma Bayesiano neste trabalho e os resultados inferencias são baseados nele. Como o modelo é construído a partir de uma função base multiplicada pela função estimada, consideramos prioris informativas para os parâmetros do modelo, de tal forma que a priori tenha maior densidade próxima a função multiplicativa igual a 1 ( $g(t_{ij} \mid a_i, b_i) = 1$ ), assim, consideramos a priori que  $a_i \sim N(0, 1)$ ,  $b_i \sim N(0, 1)$  e  $\sigma \sim \text{Gama}(1, 1)$ , todas independentes. Neste caso, a distribuição a posteriori é dada por,

$$\pi_i(a_i, b_i, \sigma_i \mid \text{dados}_i) \propto L_i(a_i, b_i, \sigma_i \mid \text{dados}_i) \pi(a_i) \pi(b_i) \pi(\sigma_i).$$

Assim, o logarítmo da posteriori para o  $i$ -ésimo indivíduo pode ser escrito como

$$\log \pi_i(a_i, b_i, \sigma_i \mid \text{dados}_i) \propto -n_i \log(\sigma_i) - \sum_{j=1}^{n_i} \left\{ \frac{[c_{ij} - \tilde{h}_i(t_{ij})g(t_{ij} \mid a_i, b_i)]^2}{2\sigma_i^2} \right\} - \frac{a_i^2}{2} - \frac{b_i^2}{2} - \sigma_i. \quad (3.4)$$

As log-densidades condicionais são

$$\begin{aligned}\log \pi_i(a_i | b_i, \sigma_i, \text{dados}_i) &\propto -\sum_{j=1}^{n_i} \left\{ \frac{[c_{ij} - \tilde{h}_i(t_{ij})g(t_{ij} | a_i, b_i)]^2}{2\sigma_i^2} \right\} - \frac{a_i^2}{2}, \\ \log \pi_i(b_i | a_i, \sigma_i, \text{dados}_i) &\propto -\sum_{j=1}^{n_i} \left\{ \frac{[c_{ij} - \tilde{h}_i(t_{ij})g(t_{ij} | a_i, b_i)]^2}{2\sigma_i^2} \right\} - \frac{b_i^2}{2}, \\ \log \pi_i(\sigma_i | a_i, b_i, \text{dados}_i) &\propto -n_i \log(\sigma_i) - \sum_{j=1}^{n_i} \left\{ \frac{[c_{ij} - \tilde{h}_i(t_{ij})g(t_{ij} | a_i, b_i)]^2}{2\sigma_i^2} \right\} - \sigma_i.\end{aligned}$$

Como as funções de densidade condicional não têm uma forma conhecida, geramos uma amostra da distribuição a posteriori dos parâmetros para cada indivíduo usando o método iterativo de Gibbs com Metrópolis-Hastings (Gelman et al., 2004).

Para obter a função a posteriori de um grupo de indivíduos, consideramos o método de metanálise proposto por Martins (2013). Desta forma, a função de densidade metanalítica a posteriori é dada por

$$\pi(a, b, \sigma | \text{dados}) = \sum_{i=1}^n \frac{\pi_i(a, b, \sigma | \text{dados}_i)}{n},$$

em que  $\text{dados} = \{\text{dados}_1, \dots, \text{dados}_n\}$ ,  $n$  é o total de indivíduos no grupo e  $a, b \in \mathbb{R}$ ,  $\sigma > 0$  são os parâmetros de interesse. Note que a função de densidade metanalítica a posteriori tem apenas 3 parâmetros referentes a cada grupo. Agora não temos mais os  $a_i, b_i, \sigma_i$  de cada um dos indivíduos.

### 3.2 Comparação entre os níveis de cortisol de indivíduos que trabalhavam nas ruas e os que não trabalhavam.

Para esse estudo, as crianças que possuíam menos do que 3 medições de cortisol foram excluídas das análises. Assim, ficamos com um total de 178 crianças, sendo 116 crianças que trabalhavam nas ruas durante o dia e 62 que não trabalhavam.

Considerando o modelo proposto na seção anterior, geramos uma amostra da distribuição a posteriori tamanho 10000, um “burn-in” (aquecimento) de tamanho

10000 e um “jump” (salto) de tamanho 100. Para  $a_i$  e  $b_i$  utilizamos a priori uma distribuição Normal(0,1). Para o parâmetro  $\sigma_i$  utilizamos a distribuição Gama(1,1).

O resultado obtido é exemplificado pelos quatro indivíduos na Figura 3. Na Figura 4 apresentamos as médias a posteriori e os percentis (ponto-a-ponto) 2, 5% e 97, 5% para os 4 indivíduos e verificamos que, em geral, a função representa bem o cortisol. O resultado da comparação entre os grupos é apresentado na Figura 5. Verificamos que, no geral, as curvas são próximas. Para os indivíduos que não trabalham nas ruas durante o dia, o cortisol é um pouco maior ao acordar e depois decaem mais rapidamente (curva verde contínua). Já para os indivíduos que trabalham, o cortisol é um pouco mais baixo ao acordar e decaem mais lentamente durante o dia (curva vermelha contínua). Vale ressaltar que, entre 2 e 4 horas após acordar, as curvas dos dois grupos estão praticamente sobrepostas, o que indica que não temos diferenças entre os níveis de cortisol de ambos os grupos durante esse intervalo de tempo.

Como medida de evidência para a diferença entre as curvas dos grupos, calculamos a probabilidade a posteriori de a curva de um grupo ser superior a de outro dado  $t$ , isto é,  $\Pr(C_{rua}(t) > C_{nr}(t) \mid T = t, \theta)$ , em que  $C_{rua}(\cdot)$  e  $C_{nr}(\cdot)$  são as curvas referentes aos grupos **rua** e **não rua** e  $\theta$  os parâmetros considerados, que, neste caso, são  $a$ ,  $b$  e  $\sigma$ . Para este cálculo, utilizamos um procedimento de Monte Carlo dada as amostras a posteriori da medida metanalítica de cada grupo.

Na Figura 6 verificamos que, no geral, a probabilidade ficou em torno de 50%, sendo que até aproximadamente 2h30min após acordar, o cortisol dos indivíduos que trabalham nas ruas é menor do que o cortisol dos indivíduos que não trabalham. A partir de 2h30min após acordar, os indivíduos que trabalham passam a ter uma probabilidade maior do que os indivíduos que não trabalham de possuir um cortisol mais alto. Ainda assim, visualizamos que, de forma geral, essa probabilidade está bem próxima de 0,5, o que nos leva a conclusão que não há diferença dos níveis de cortisol dos dois grupos. Além disso, as densidades metanalíticas marginais de ambos os grupos são muito similares, evidenciando a não diferença entre os grupos (ver Figura 7).

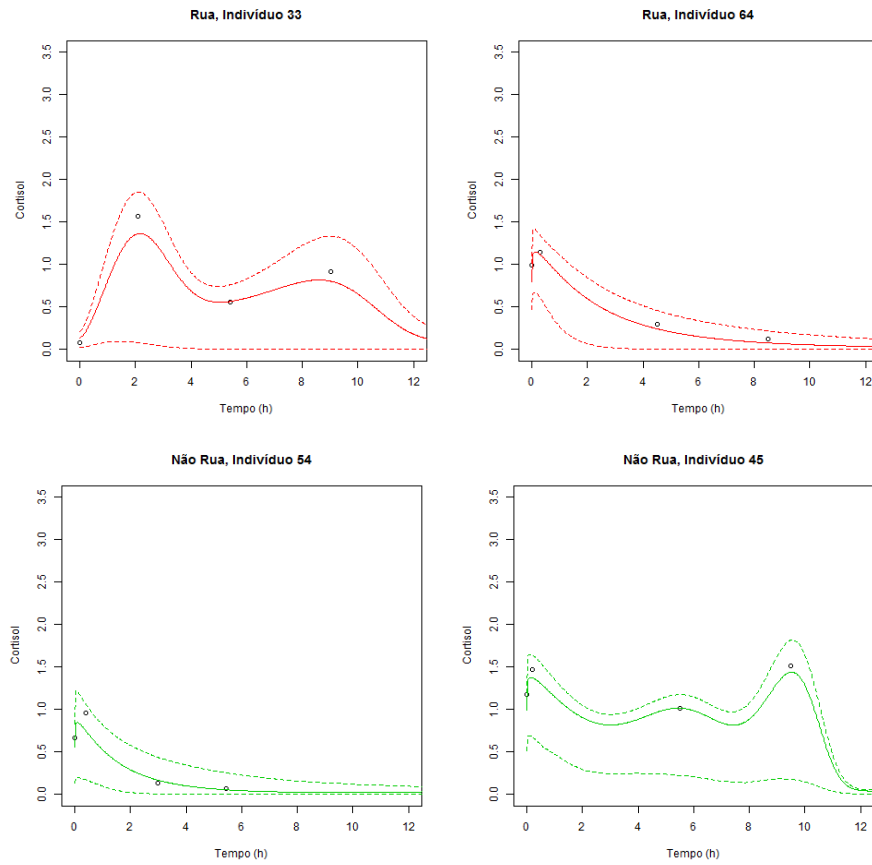


Figura 4 – Média a posteriori das funções de cortisol e intervalo de credibilidade ponto-a-ponto de 95%, referente aos 4 indivíduos da Figura 3.

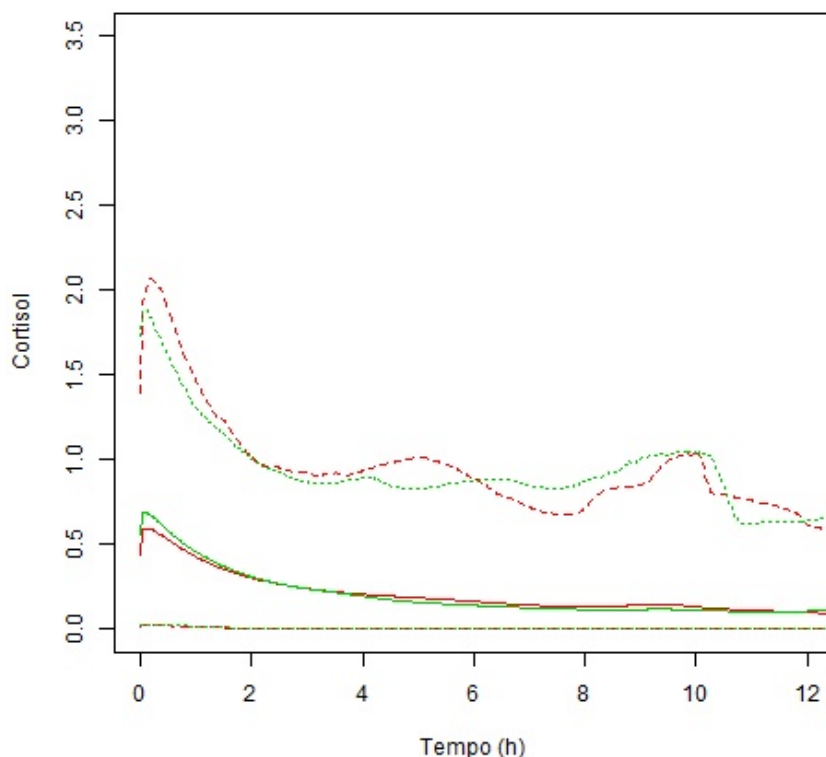


Figura 5 – Metanálise, em que as linhas contínuas representam a média da medida metanalítica de cada grupo e as linhas tracejadas representam o intervalo de credibilidade central ponto-a-ponto, com 95% de probabilidade (curvas vermelhas - indivíduos que trabalham nas ruas e curvas verdes - indivíduos que não trabalham nas ruas).

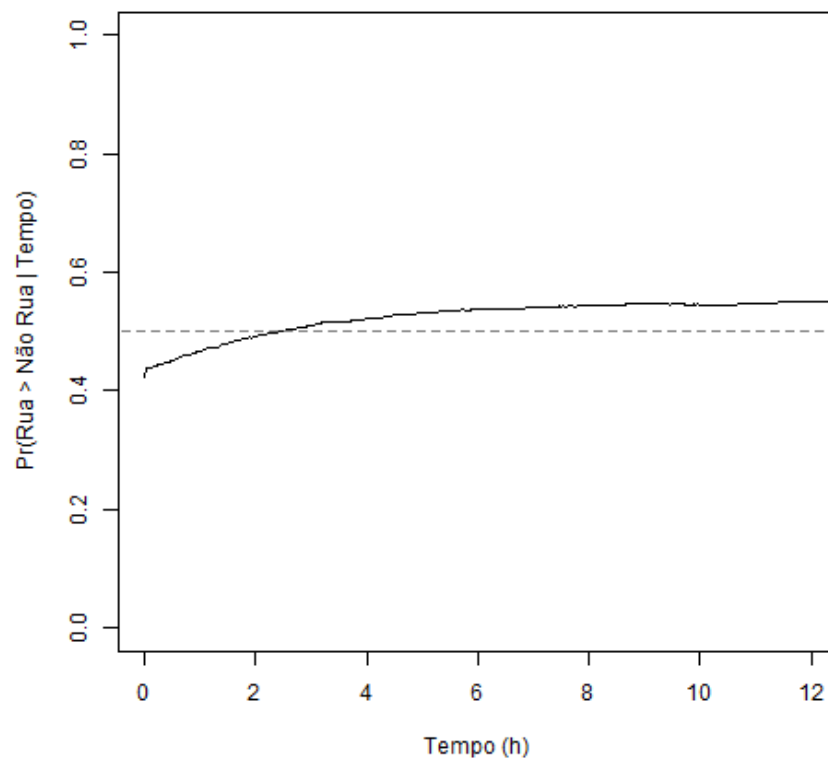


Figura 6 – Probabilidade ponto-a-ponto a posteriori dos níveis de cortisol do grupo **rua** serem maiores do que o do grupo **não rua**.



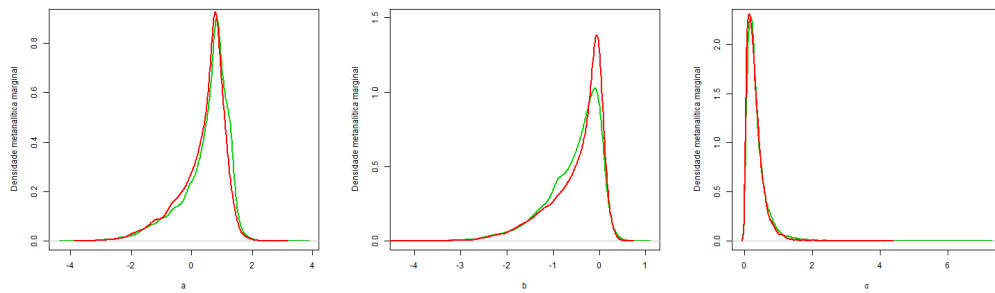


Figura 7 – Média metanalítica marginal a posteriori dos parâmetros da função base de cada grupo (curvas vermelhas - indivíduos que trabalham nas ruas e curvas verdes - indivíduos que não trabalham nas ruas).



## 4 Considerações finais

A descrição do problema pelos médicos nos levou a utilizar uma técnica de análise para dados funcionais e, na parte inferencial, fizemos o uso do método Bayesiano. O grande desafio estava em trabalhar com apenas quatro (em alguns casos três) observações da função do cortisol de cada indivíduo. Este desafio foi abordado fixando uma função base que foi construída com as características que os médicos entendiam ser o padrão de cortisol ao longo do dia para humanos. Uma vez a função base construída, partimos para a parte inferencial, dando maior flexibilidade para ter uma melhor representação da função do cortisol de cada indivíduo. Note que tal função base pode ser considerada como extremamente conservadora ou subjetiva, pois leva fortemente em conta a opinião dos médicos. Entretanto, ressaltamos que o objetivo principal do estudo era o de comparar os dois grupos de crianças e, considerando isso, entendemos que o processo inferencial de comparação entre os grupos não fica prejudicado, uma vez que partimos, como base, da informação que os médicos dispunham.

Uma vez que obtivemos a distribuição a posteriori para a função do cortisol de cada indivíduo, precisávamos de um procedimento para fazer a inferência para os grupos. Assim, propusemos o uso da metanálise Bayesiana (Martins, 2013) que nos permitiu encontrar a distribuição metanalítica de cada um dos dois grupos de interesse. Outro fato importante foi a construção da banda de credibilidade. Em geral, os estudos, mesmo os que tratam de dados funcionais, utilizam regiões de credibilidade/confiança ponto-a-ponto. Olshen et al. (1989) abordou o problema da região de confiança utilizando um método de bootstrap e construção de limities para banda contendo a confiança desejada da função estar inteiramente dentro da banda. No nosso caso, por tratar-se de uma abordagem Bayesiana, desenvolvemos uma banda de credibilidade que tivesse a menor área. O método proposto é aproximado para a obtenção de tal banda.

Para a comparação entre os grupos, propusemos utilizar a sobreposição de ban-

das, graficamente. Como testes de hipótese para funções não são simples de ser realizados, nós propusemos utilizar a probabilidade a posteriori ponto-a-ponto da função do cortisol de um grupo ser maior que a do outro. Uma vez com esta probabilidade calculada para os diversos instantes de tempo, a desenhamos em um gráfico que, se a probabilidade a posteriori estiver próxima de meio (uma linha horizontal), entendemos não haver diferença entre os grupos. Mas se esta probabilidade se afastar de meio, não só podemos evidenciar a diferença entre as funções, como também saber se a diferença é na função inteira, em apenas um pedaço, se é maior ou menor, tendo assim um processo inferencial completo para a comparação entre os grupos.

Como visto na Seção 3.2, conseguimos uma função que se ajusta muito bem aos dados e que, ao mesmo tempo, mantém as premissas dos especialistas. Na metanálise verificamos que as curvas para ambos os grupos são bem próximas, sendo razoavelmente maior a curva do grupo **não rua** (curva verde) até 2h após acordar. De 2 até 4 horas após acordar as curvas parecem serem iguais e, após isto, a curva dos indivíduos do grupo **rua** (curva vermelha) se torna minimamente maior. Portanto, entendemos que o nível de cortisol, ao longo do dia, tem um comportamento similar nos dois grupos, não sendo possível diferenciá-los.

Verificamos que os valores das probabilidades a posteriori ponto-a-ponto da função de um grupo ser maior que a função do outro grupo eram próximos de 0,5, o que indica que não existe diferença nos níveis de cortisol de ambos os grupos, ou seja, o fato da criança trabalhar ou não nas ruas, não interfere no nível de cortisol delas e, conseqüentemente, não podemos concluir que crianças que trabalham nas ruas são mais estressadas que crianças que não trabalham.

Um detalhe importante a ser considerado que nos foi reportado pelos médicos, é que as crianças deste estudo eram, em geral, de situação muito carente. As crianças que ficavam em casa muitas vezes sofriam maus-tratos, e por isso, trabalhar nas ruas (vender bala no farol, entre outros) poderia ser menos problemático do que ficar em casa. Além disso, temos um problema na aquisição dos dados. Primeiramente que, com apenas quatro observações de uma função, fica difícil fazer inferências, e por fim, a qualidade destas informações coletadas pelos cuidadores, aparentemente ocorreram sem um grande rigor.

Entendemos que nas condições do problema, apesar de considerarmos independência entre os indivíduos, os resultados obtidos foram satisfatórios. A conclusão pela não diferença dos grupos é condizente com o relatado pelos médicos. Do ponto de vista estatístico, como propostas futuras, sugerimos investigar um método para a construção da banda de credibilidade e, principalmente, para comparação de funções.

Destacamos que, as contribuições do ponto de vista estatístico foram: a construção da função base e uma função multiplicativa para a análise de dados funcionais. Esta proposta mostrou-se útil quando temos poucas observações da função de interesse e temos disponível a informação de especialistas para a construção da função base. Para a estimação da função do cortisol de cada grupo, propusemos o uso da metanálise Bayesiana. Não encontramos na literatura, métodos para a comparação de grupos em análise funcional. Então, desenvolvemos um método para a comparação das funções dos dois grupos, baseado na probabilidade a posteriori ponto a ponto.



## Referências

- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Chapman & Hall/CRC, 2a. edition.
- Juruena, M., Cleare, A., & Pariante, C. (2004). The hypothalamic pituitary adrenal axis, glucocorticoid receptor function and relevance to depression. *Rev Bras Psiquiatr*, 26, 189–201.
- Mario F Juruena, a. A. J. C. e. C. M. P. (2004). O eixo hipotálamo-pituitária-adrenal, a função dos receptores de glicocorticóides e sua importância na depressão. *Rev Bras Psiquiatr*.
- Martins, C. B. (2013). *Metanálise caso a caso sob a perspectiva Bayesiana*. PhD thesis, IME-USP.
- Mello, A., Juruena, M., Maciel, M., Cavalcante-Nobrega, L., Cividanes, G., Fossaluza, V., Medeiros, M., Mello, M., Cleare, A., & Mari, J. (2013). The impact of an intervention program to reduce stress for at risk children: the cortisol awakening response before and after two years of intervention with children working on the streets. *Submetido*.
- Mello, A., Mello, M., & Carpenter, L.L. Price, L. (2003). Update on stress and depression: the role of the hypothalamic-pituitary-adrenal (hpa) axis. *LW Rev Bras Psiquiatr*, 25, 231–238.
- Olshen, R. A., Biden, E. N., Wyatt, M. P., & Sutherland, D. H. (1989). Gait analysis and the bootstrap. *The Annals of Statistics*, 17(4), 1419–1440.
- Ramsay, J. O. & Dalzell, C. J. (27/04/2012). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol.53, No. 3 (1991), pp. 539-572.

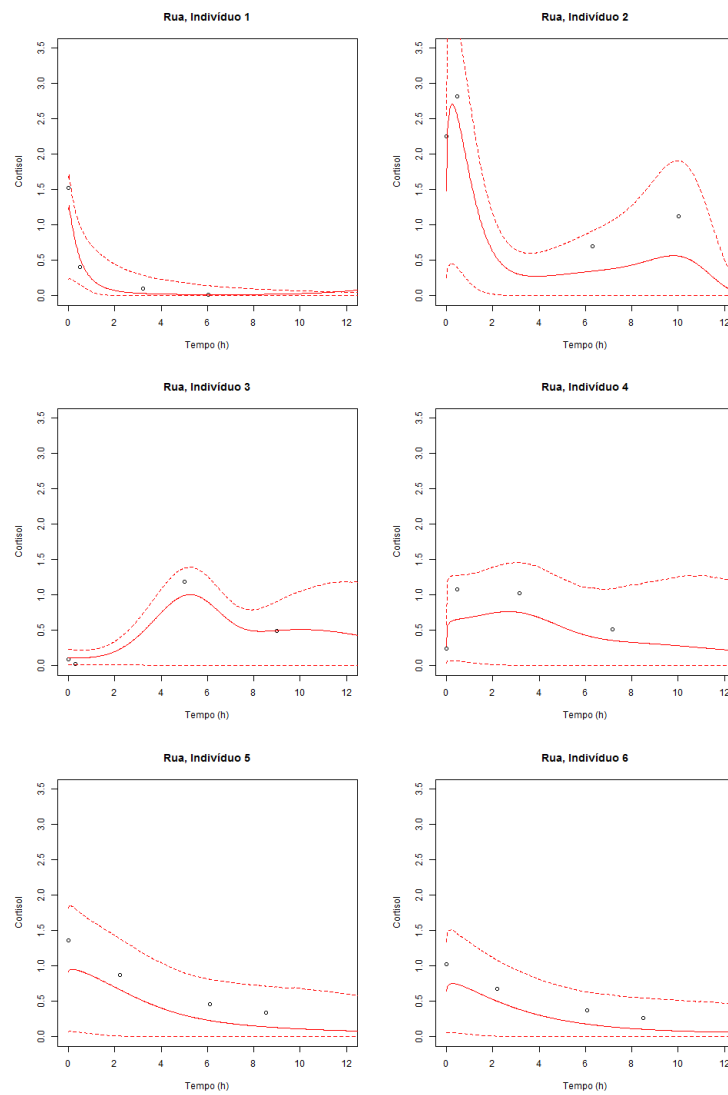


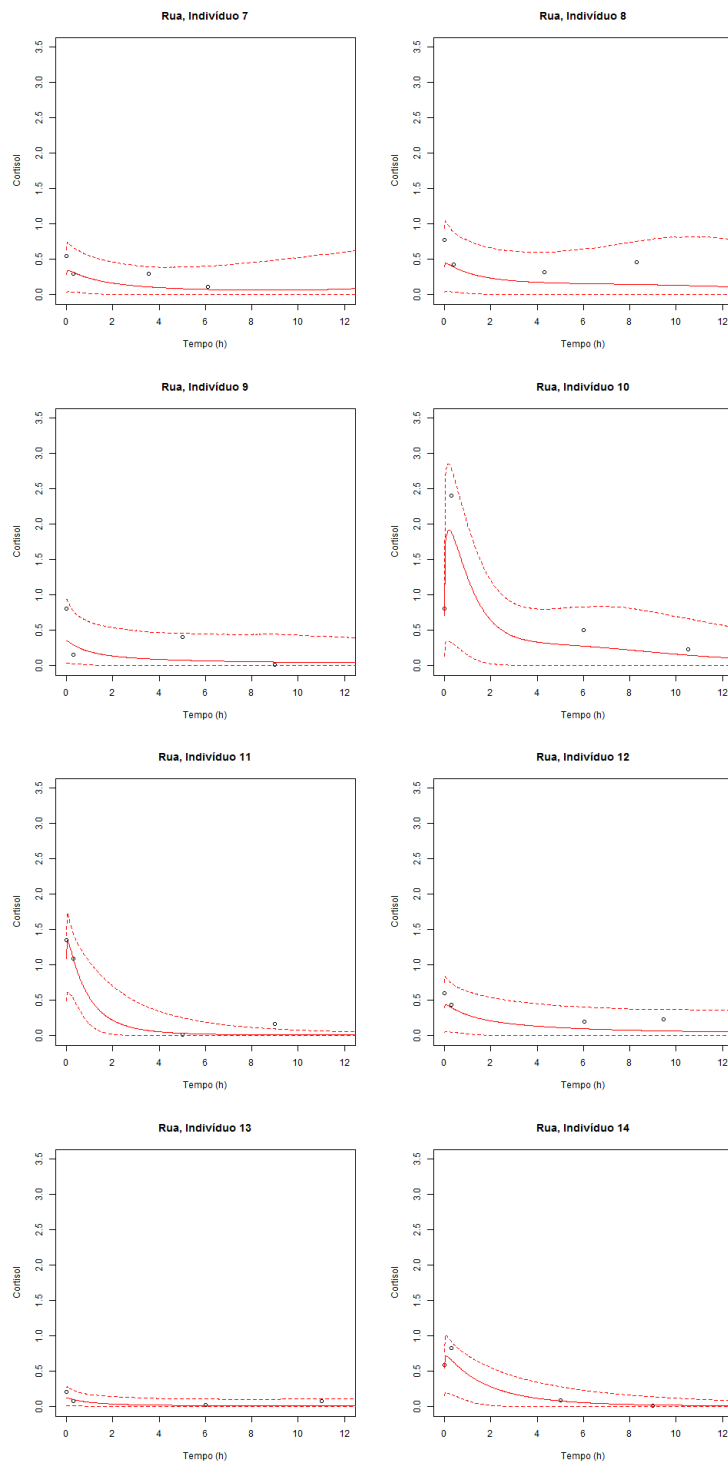


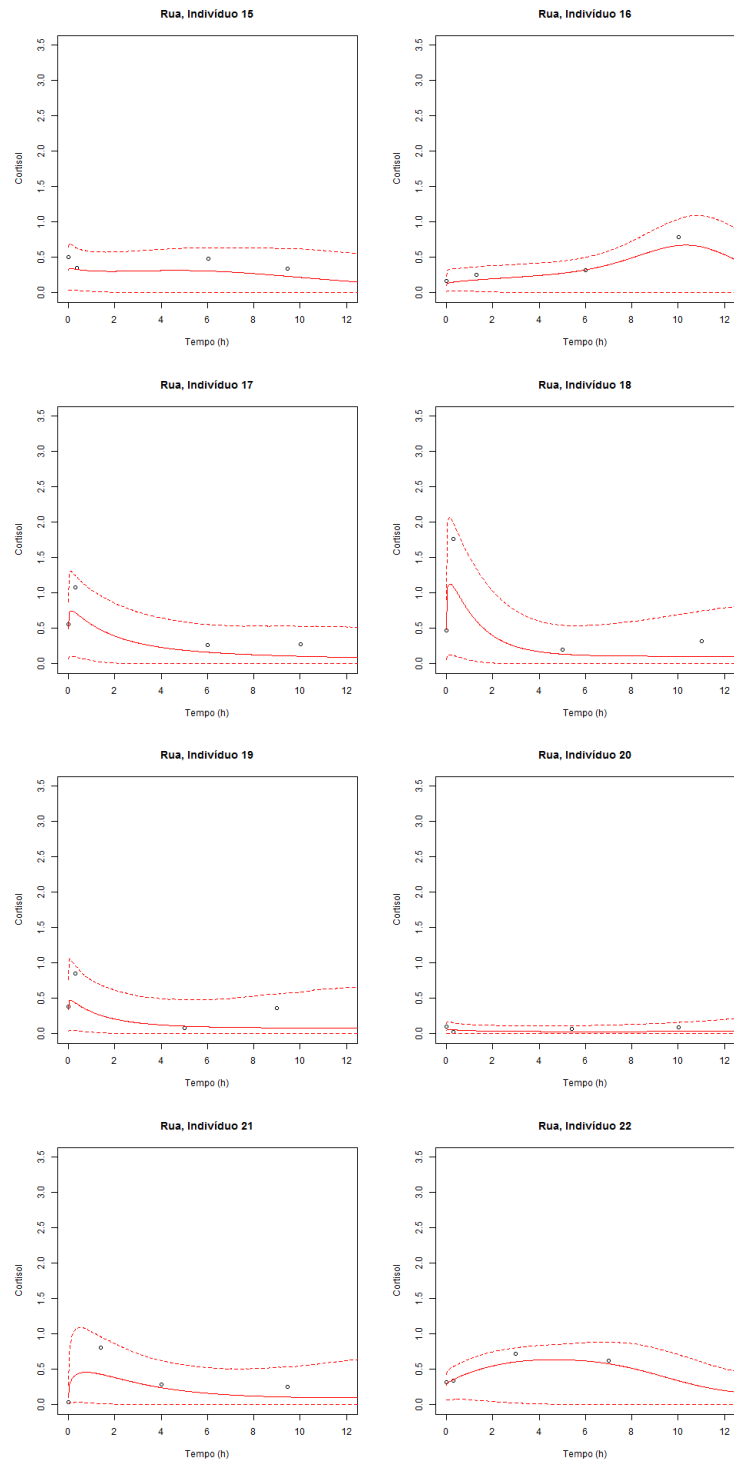


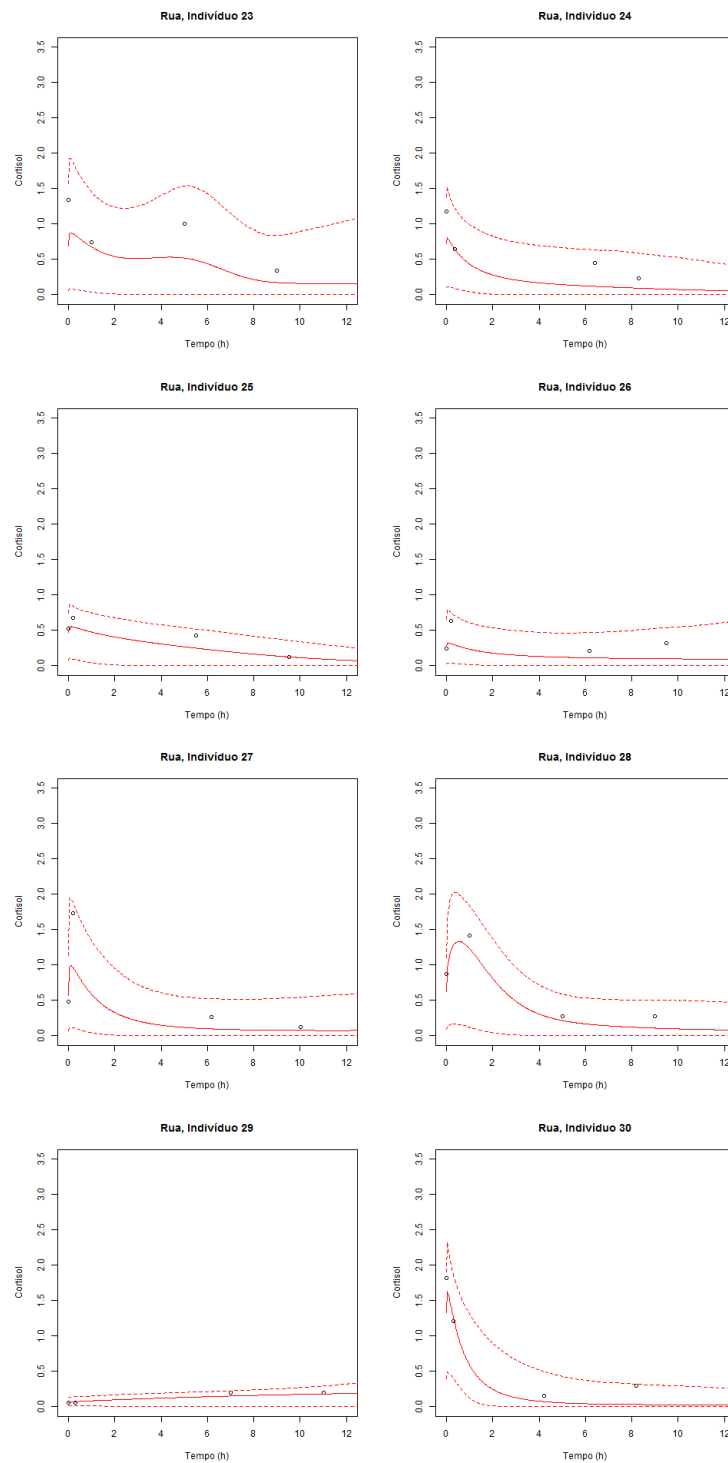
## 5 Apêndices

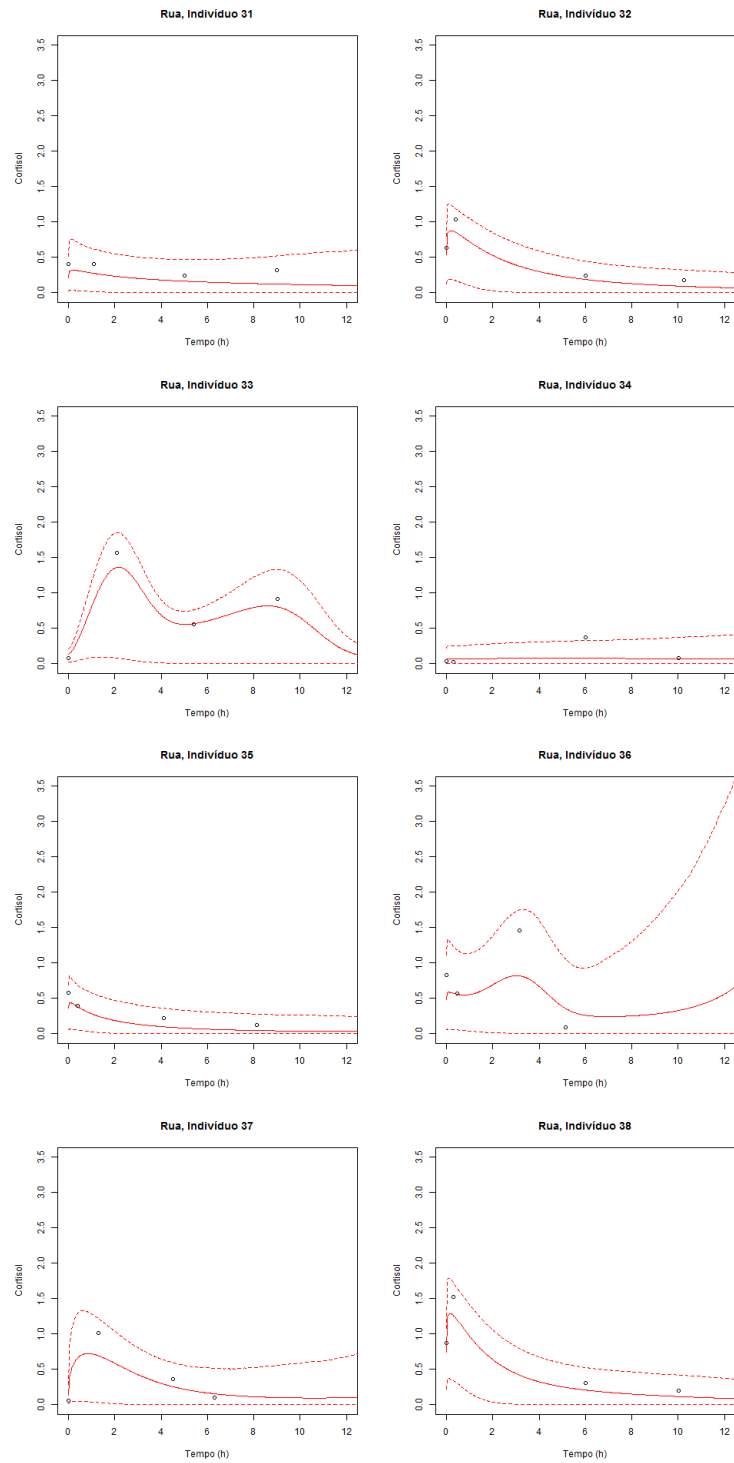
### 5.1 Funções densidade a posteriori dos indivíduos que trabalhavam nas ruas.

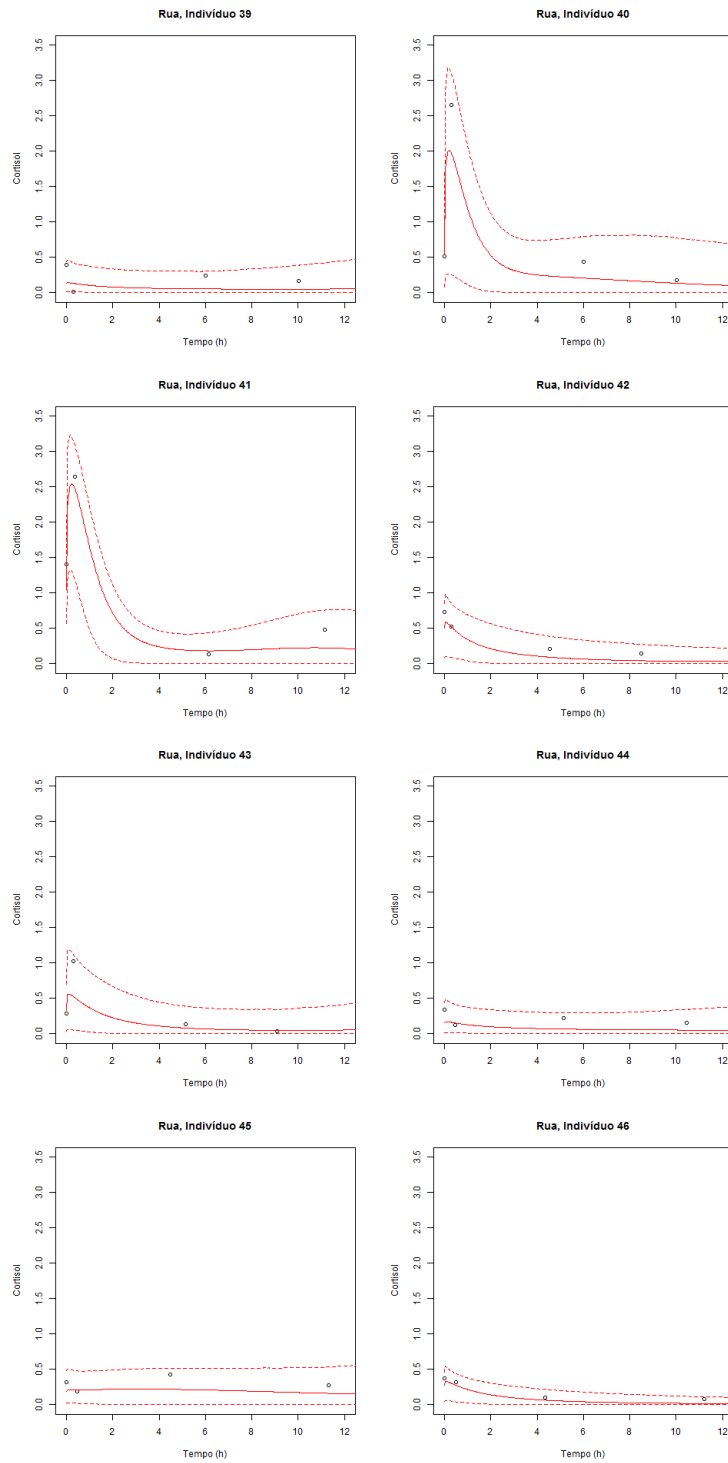


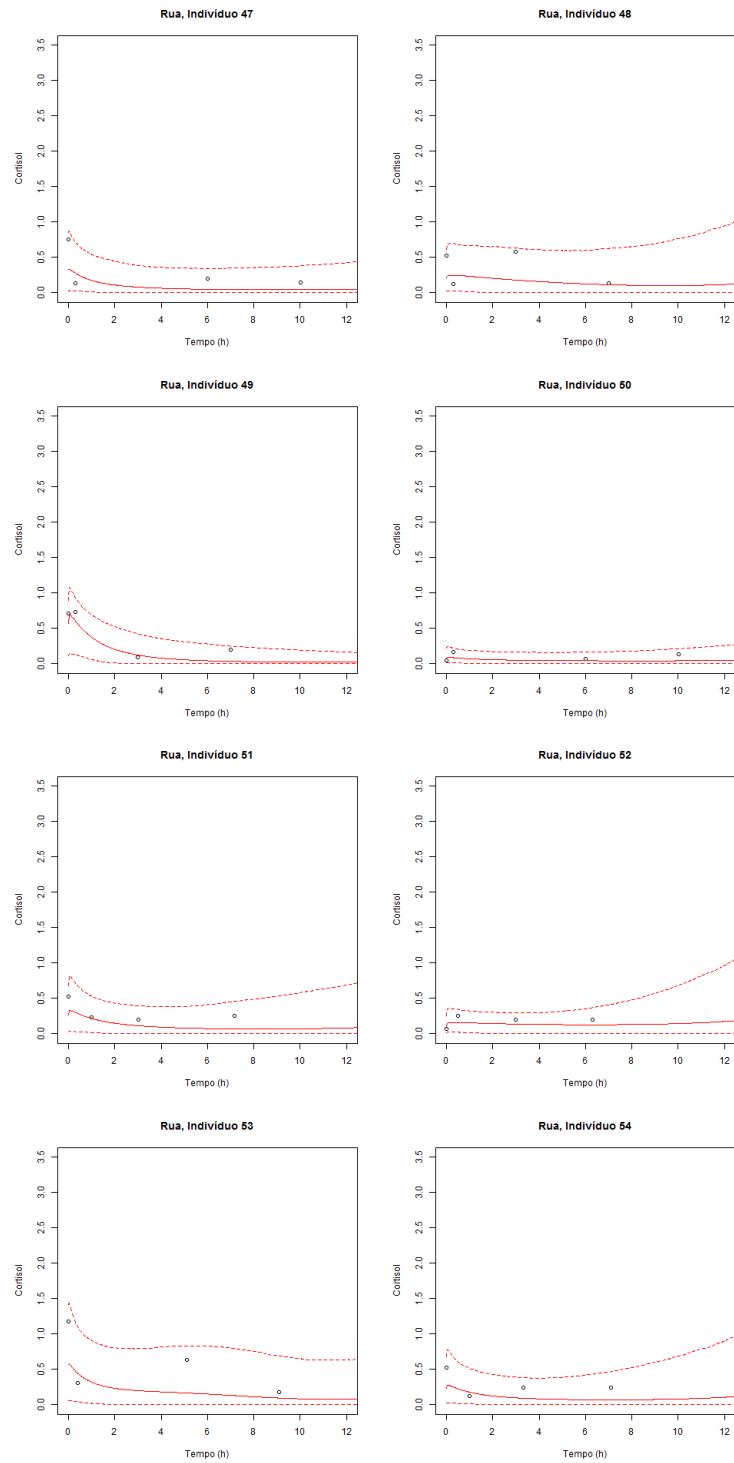




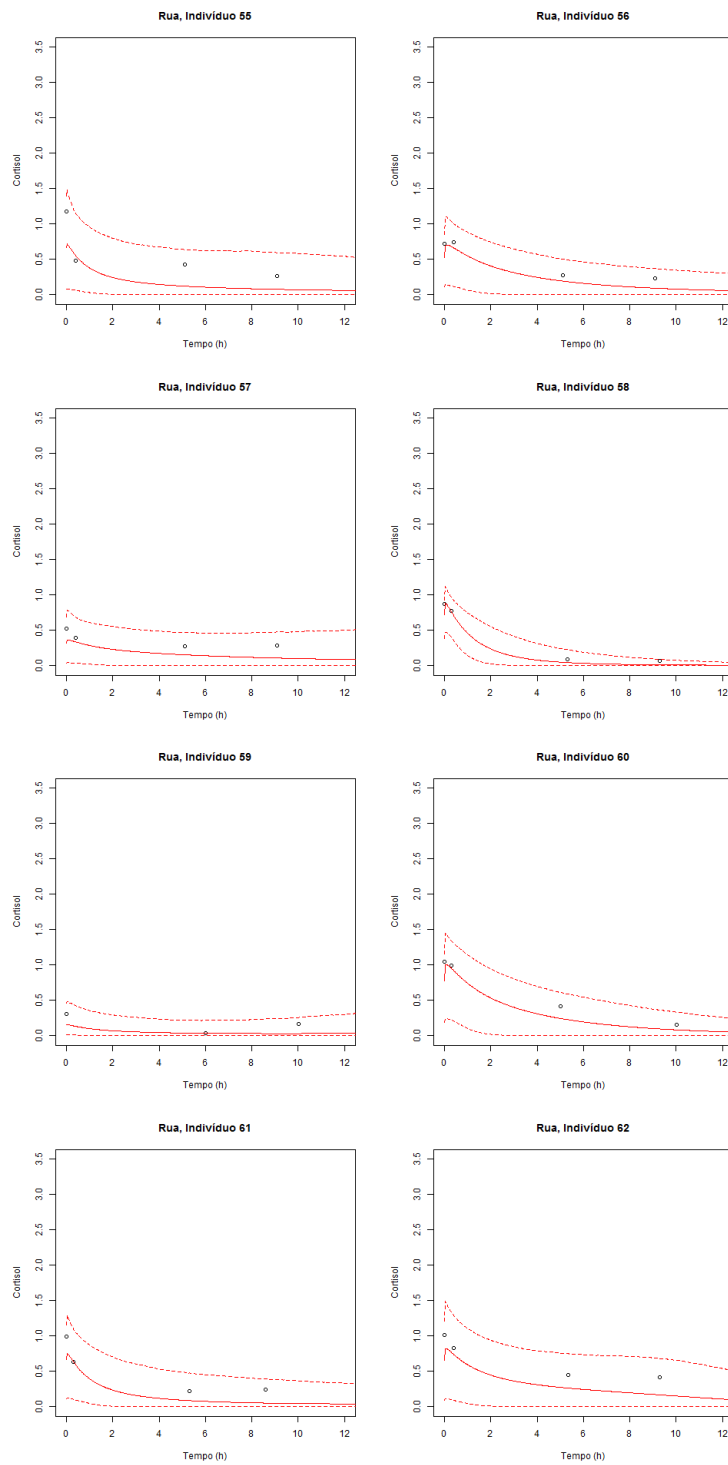


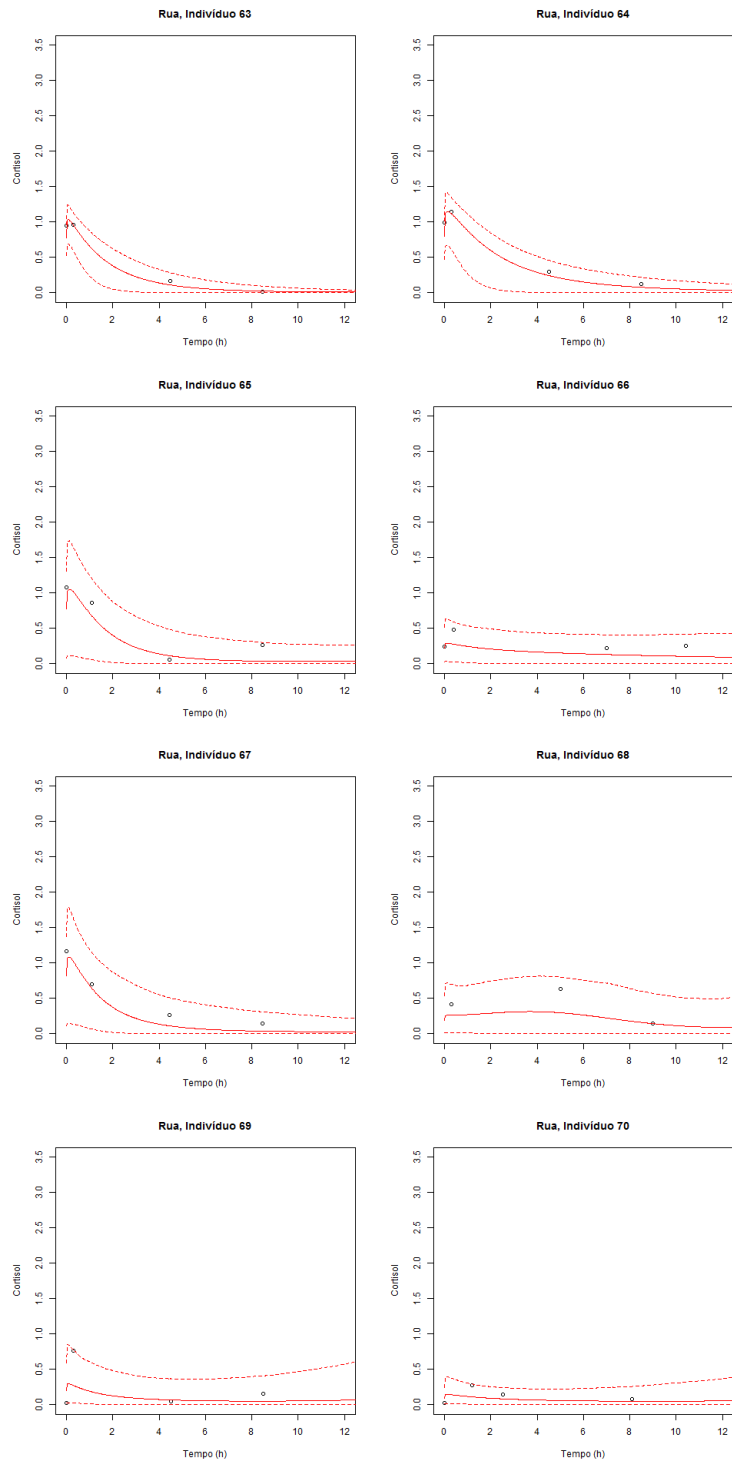


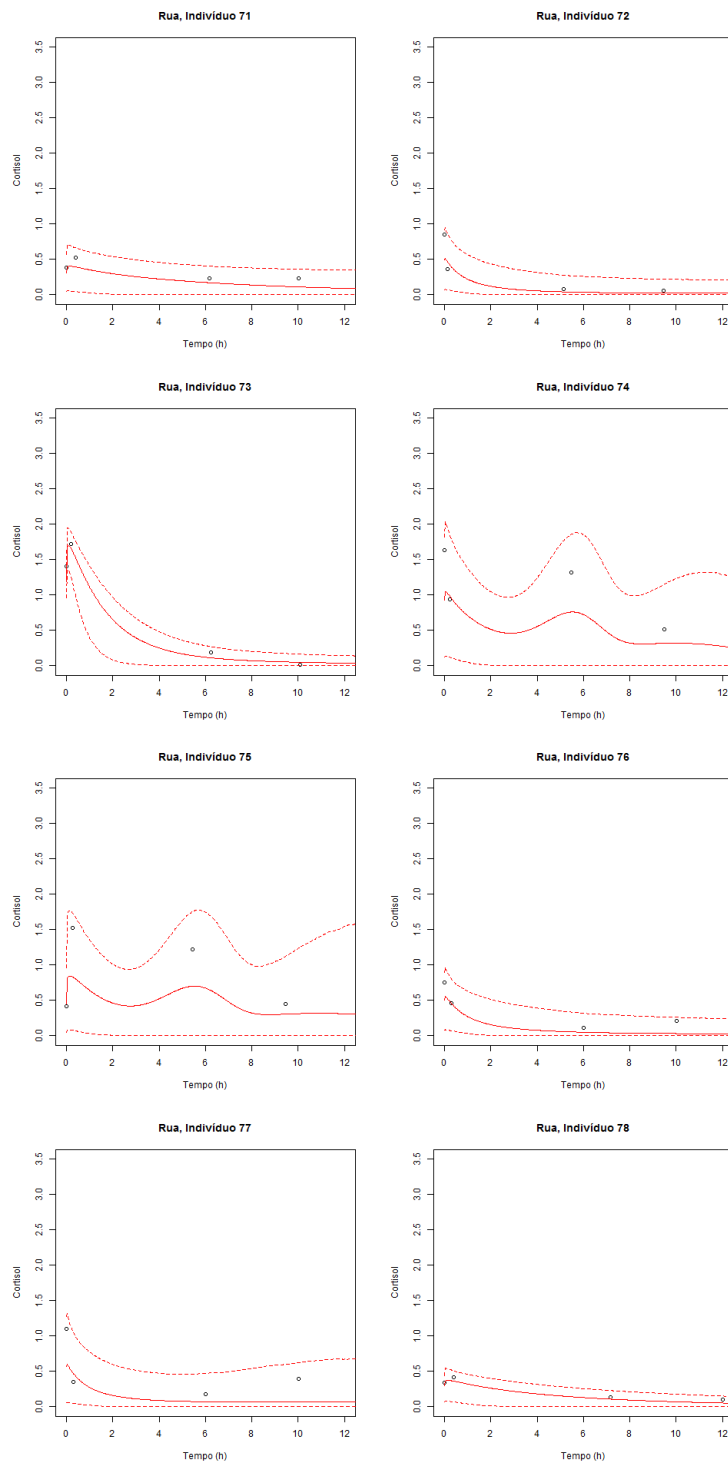


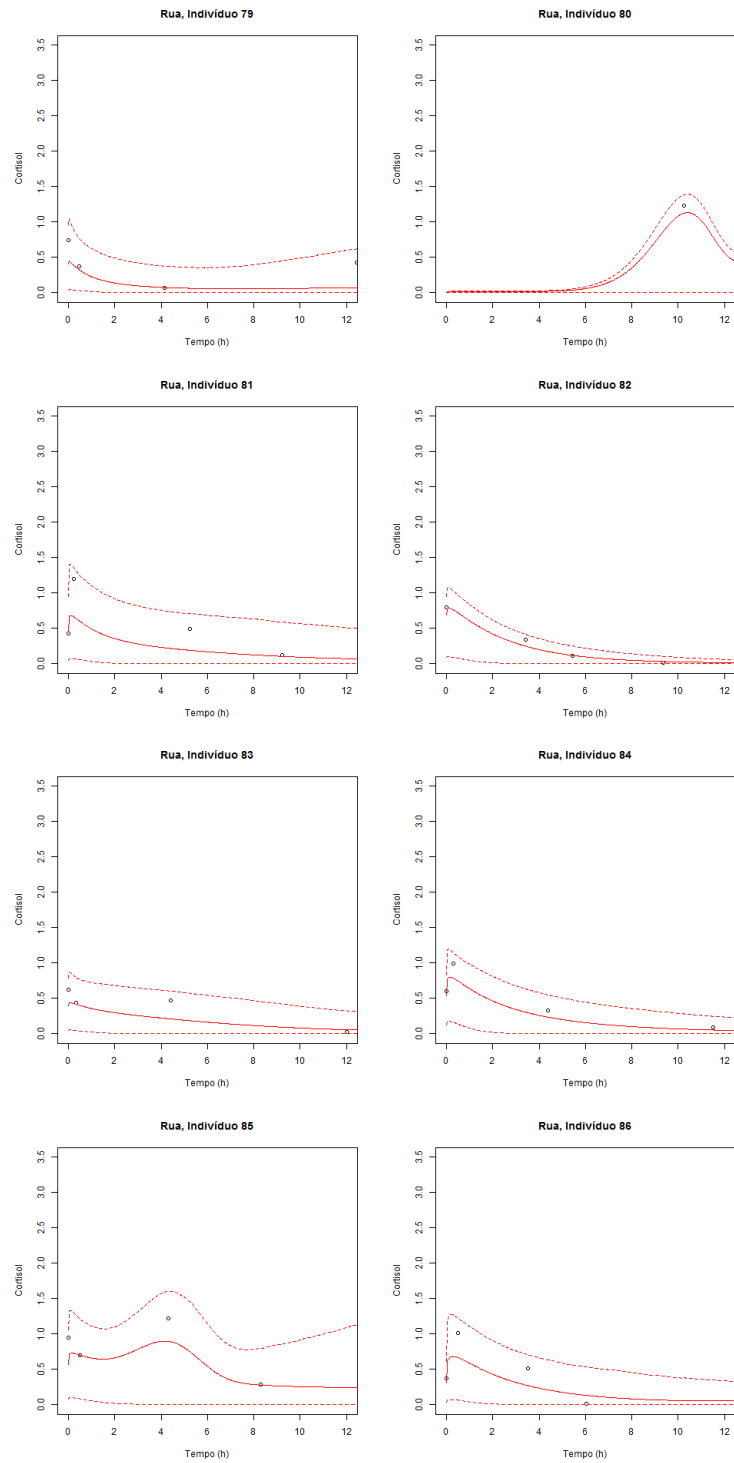


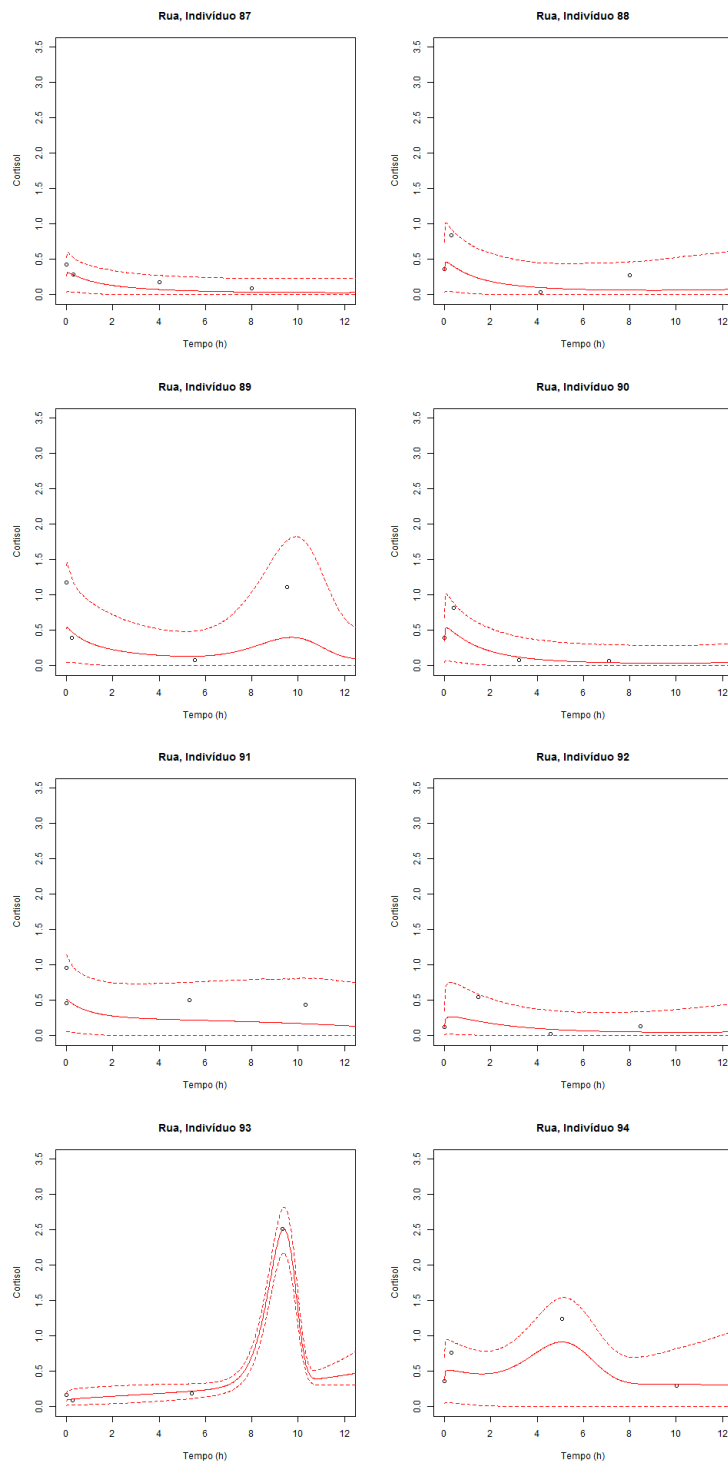


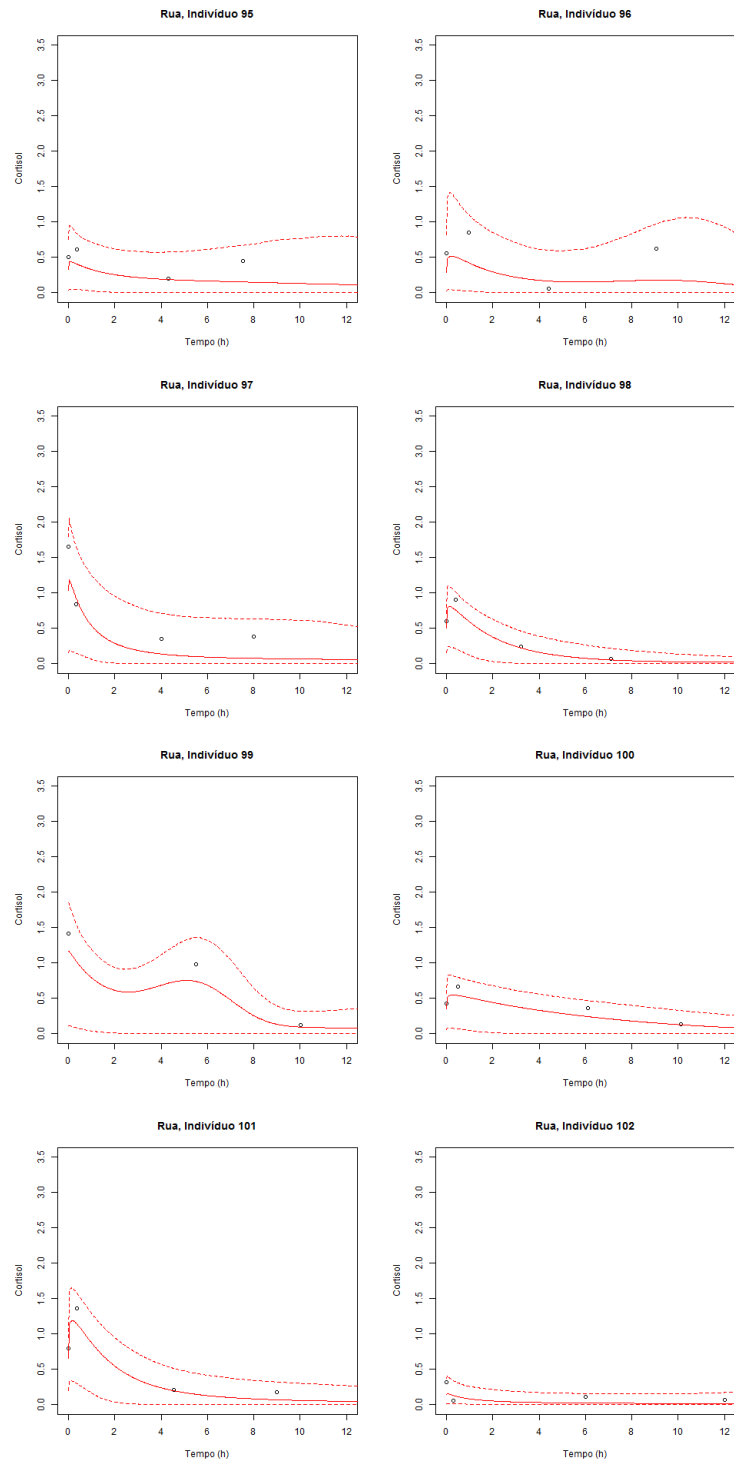


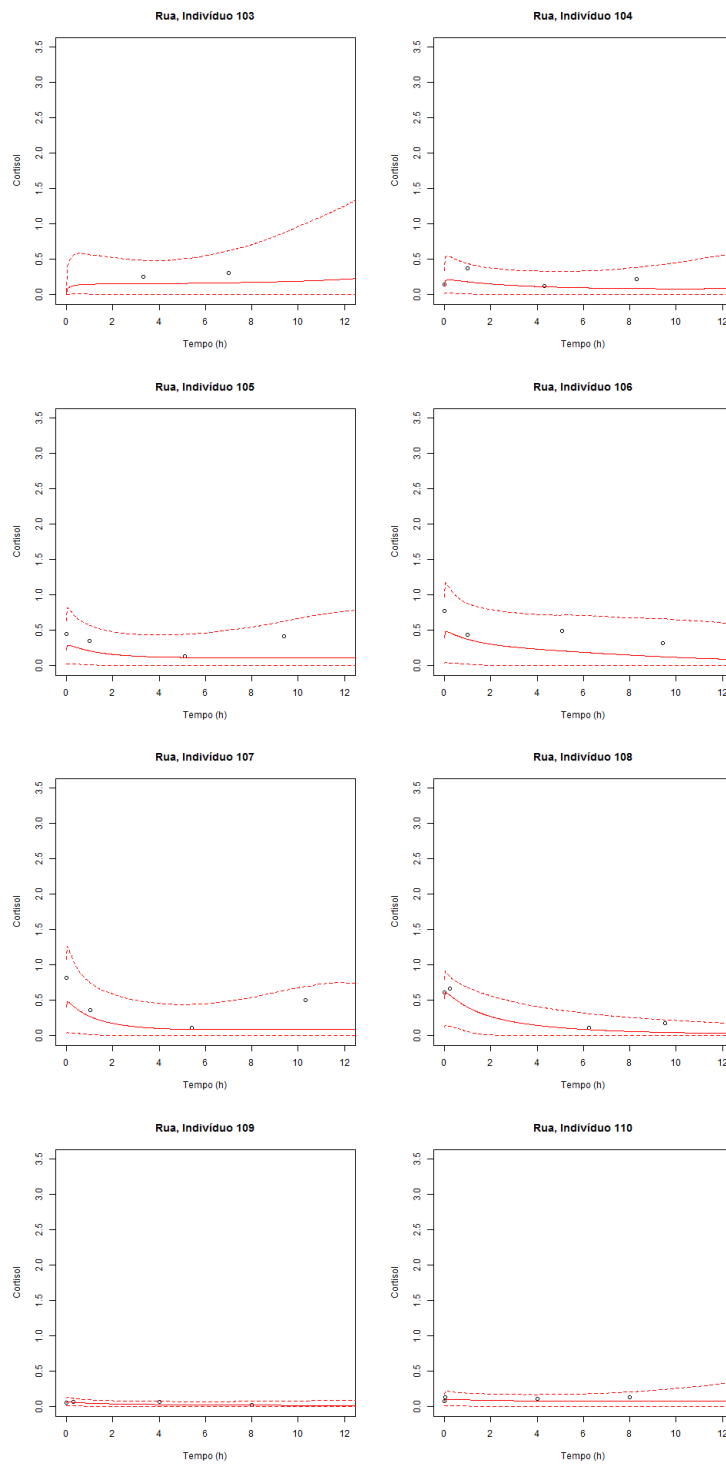


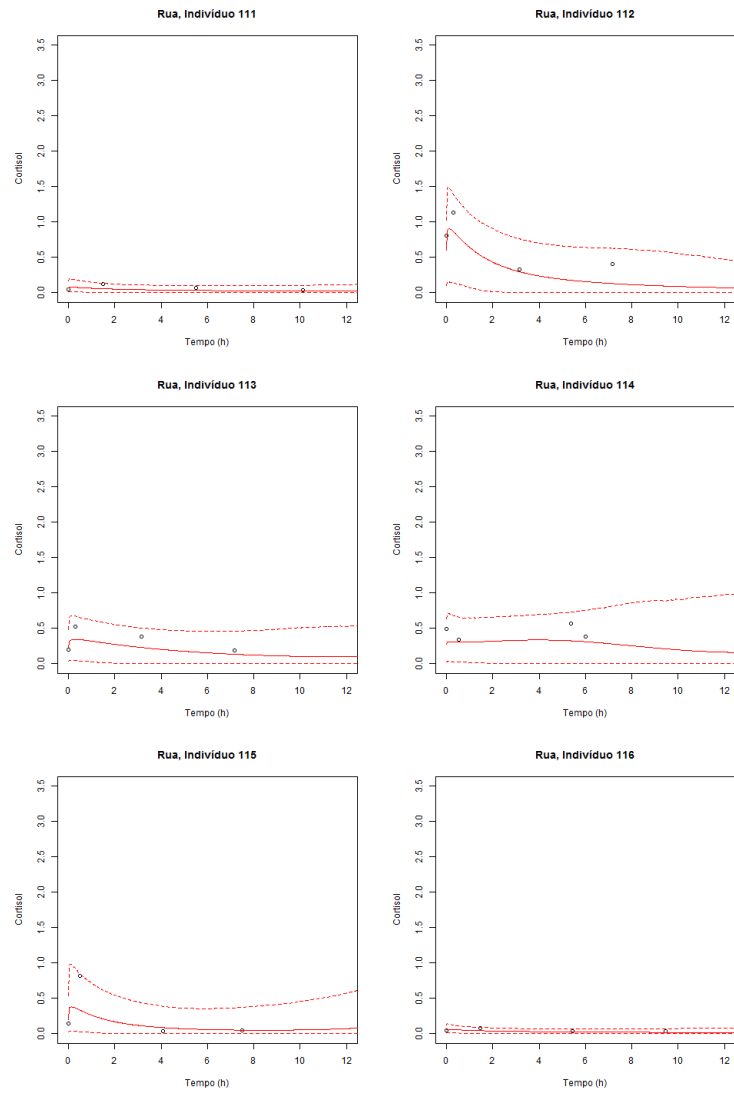














## 5.2 Funções densidade a posteriori dos indivíduos que não trabalhavam.

