



MATHEUS HENRIQUE DAL MOLIN RIBEIRO

**MODELOS MISTOS EM ESTUDO DE ISQUEMIA
CEREBRAL**

Dissertação de Mestrado

Maringá - Paraná
Junho de 2015

MATHEUS HENRIQUE DAL MOLIN RIBEIRO

**MODELOS MISTOS EM ESTUDO DE ISQUEMIA
CEREBRAL**

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá como requisito parcial para obtenção do título de Mestre em Bioestatística.
Orientadora: Profa. Dra. Isolde Previdelli

Maringá - Paraná

Junho de 2015

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá – PR., Brasil)

R484m Ribeiro, Matheus Henrique Dal Molin
 Modelos mistos em estudo de isquemia cerebral /
Matheus Henrique Dal Molin Ribeiro. -- Maringá,
2015..
 39 f. : il., figs., tabs.

 Orientadora: Profa. Dra. Isolde Previdelli.
 Dissertação (mestrado) - Universidade Estadual de
Maringá, Centro de Ciências Exatas, Programa de Pós-
Graduação em Bioestatística, 2015.

 1. Dados Longitudinais. 2. Efeitos aleatórios 3.
Estrutura de covariância. 4. Latência. I.
Previdelli, Isolde, orient. II. Universidade
Estadual de Maringá. Centro de Ciências Exatas,
Programa de Pós-Graduação em Bioestatística. III.
Título.

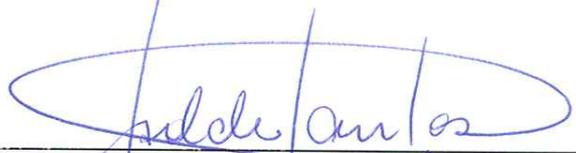
CDD 21.ed. 570.15195

MATHEUS HENRIQUE DAL MOLIN RIBEIRO

Modelos Mistos em Estudo de Isquemia Cerebral

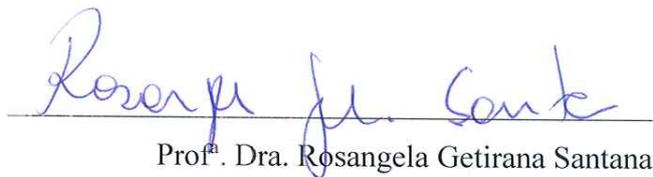
Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para a obtenção do título de Mestre em Bioestatística.

BANCA EXAMINADORA



Prof.^a Dra. Isolde Previdelli

Universidade Estadual de Maringá - UEM



Prof.^a Dra. Rosangela Getirana Santana

Universidade Estadual de Maringá - UEM



Prof. Dr. Silvado Cesar da Costa

Universidade Estadual de Londrina – UEL

Maringá, 26 de junho de 2015.

Dedico este trabalho a todos que fizeram parte da minha vida durante essa jornada.

Agradecimentos

Agradeço primeiramente a Deus, pois sem sua proteção tenho certeza que não teria conseguido chegar ao fim dessa jornada tão desafiadora.

A minha namorada, esposa, cúmplice Cristiane. Sem a sua paciência ao longo destes 16 meses tenho certeza que teria sido muito mais difícil do que foi. Agradeço por nunca ter me deixado desistir, pelas incansáveis noites que mesmo chegando do trabalho cansado, logo em seguida precisava viajar. Pelas motivações, pelas palavras de apoio. Pelos momentos de desabafo e reclamações. Sem seu apoio eu não teria vencido.

A minha família, em especial a minha mãe Sirlei por sempre ter me incentivado a estudar em busca de uma vida melhor. Por sempre ter me consolado e ajudado psicologicamente nos momentos difíceis. Por ter trabalhado durante sua vida para que eu pudesse seguir caminhos que nos conduzissem a uma vida melhor. Por ter ensinado valores que hoje fazem a diferença em minha personalidade, os quais são decisivos para que sejamos cidadãos melhores e para que tornemos o mundo que vivemos um lugar melhor. Ao meu pai Setembrino (*in memoriam*) por ter me protegido ao longo de todas as viagens realizadas.

Agradeço veementemente a minha orientadora Professora Isolde Previdelli por ter me acolhido como um filho, pelos seus ricos conselhos, por compartilhar seus conhecimentos e sobretudo, pela sua paciência. Por sempre ter me conduzido pelo melhor caminho e por nunca desistir de mim, mesmo nas horas mais difíceis. Por ter acreditado que era possível vencer todos os obstáculos que surgiram pelo caminho e por, muitos momentos, ter abdicado de suas atividades profissionais e familiares para que pudéssemos conversar. Por sempre ter valorizado todas as atividades e trabalhos desenvolvidos. Por ter se desdobrado em mil para que eu conseguisse alcançar os objetivos propostos. Enfim, agradeço pela maravilhosa pessoa que é.

Ao professor Humberto Milani do Programa de Ciências Farmacêuticas da UEM por todos os momentos de discussão e pela interdisciplinariedade construída entre as áreas da Estatística e da Saúde ao longo de todo o tempo de convivência. Pela disponibilização

dos dados utilizados neste trabalho e por nunca ter relatado em sanar dúvidas de cunho biológico. Certamente isso contribuiu muito para o enriquecimento do trabalho.

Agradeço a professora Rosângela por todos os momentos, desde sua primeira aula, por ter se dedicado para que aprendêssemos da melhor forma possível, sempre nos fazendo compreender os conceitos que nos cercavam. Obrigado por nessa última etapa ter contribuído para a melhora deste trabalho e pelas contribuições feitas durante o exame de qualificação, as quais foram cruciais para o bom andamento e conclusão do mesmo.

Agradeço aos demais professores do PBE, Eniuce, Vanderly, Eraldo, Edson pelos conhecimentos repassados.

Aos colegas de apartamento em Maringá, André, Eleomar e Djeison, por me acolherem.

Aos colegas do PBE, Marcos, Rafaela, Marina, Kelly, Viviane, Omar, Emerson, Sérgio, Oilson, Jardel, e tantos outros, os quais, considerando um teste de hipóteses com $p < 0,0000001$ foram estatisticamente significativos nessa jornada. Pelos momentos de estudo, pelas descontrações, pelas reclamações e sobretudo pelo companheirismo. Aos funcionários da secretaria do PBE, Rosângela e Daniel por sempre atenderem as solicitações de forma disposta.

Aos colegas de IFPR, Jean, Beth, Silvia, Marilei, Silvano, Augusto, Melania, Roberto, por sempre me incentivarem a seguir por esse caminho, mesmo cheio de barreiras, sempre lutando em busca do melhor.

A colega Cleonis, por sempre ter me aconselhado, pelas dúvidas sanadas e tantas outras conversas.

Agradeço aos meus professores da UTFPR - Pato Branco, por sempre incentivarem a seguir os estudos, em especial aos professores Carlos, Richard, Fredy, Samoara, Janecler, Janice, Gilberto e todos os demais do Departamento de Matemática.

De um modo geral, agradeço a todos que sempre me motivaram e incentivaram a seguir por este caminho que certamente será de sucesso.

*“Que os vossos esforços desafiem as impossibilidades,
lembrai-vos de que as grandes coisas do homem foram conquistadas do que parecia
impossível.”
(Charles Chaplin)*

Resumo

A modelagem de dados provenientes de estudos longitudinais destaca-se no cenário científico atual, principalmente nas áreas da Saúde e Ciências Biológicas, uma vez que é comum avaliar uma mesma unidade experimental em diferentes instantes, o que induz uma correlação entre as medidas. Faz-se necessária a modelagem da dependência intraunidade amostral por intermédio da escolha de uma estrutura de covariância que seja capaz de captar e acomodar a variabilidade dos dados. Contudo, o desconhecimento da metodologia para análise de dados correlacionados pode implicar na obtenção de resultados errôneos, os quais podem conduzir a conclusões equivocadas, bem como pode-se aumentar a ocorrência do erro tipo I e subestimar os erros-padrão das estimativas do modelo. Neste estudo, adotou-se um modelo misto gaussiano para variável resposta latência (tempo gasto para o animal encontrar o esconderijo verdadeiro) de um experimento que investiga o déficit de memória em animais submetidos a isquemia cerebral quando tratados com óleo de peixe, cuja estrutura de covariância para os efeitos aleatórios utilizada foi a não estruturada. Para a estimação dos parâmetros do modelo utilizou-se os métodos da máxima verossimilhança e máxima verossimilhança restrita. Com base no teste da razão de verossimilhança e critérios de informação foi adotada para os erros a matriz de covariância autorregressiva. As análises estatísticas residuais e de diagnóstico para o modelo foram satisfatórias, visto que os resultados obtidos corroboraram com as evidências biológicas, isto é, constatou-se a eficácia do tratamento para amenizar os efeitos cognitivos provocados pela isquemia cerebral.

Palavras-chaves: Dados Longitudinais; efeito aleatório; estrutura de covariância; latência.

Abstract

The modeling of data from longitudinal studies stands out in the current scientific scenario, especially in the areas of Health and Life Sciences, since it is common to evaluate the same experimental unit at different times, which induces a correlation between measurements. Thus, the modeling of the intra unit dependency sample is required through the choice of a covariance structure that is able to receive and accommodate the variability in data. However, the lack of methodology for correlated data analysis may result in getting erroneous results, which can lead to erroneous conclusions, and it can also increase the occurrence of Type I error and underestimate the standard errors of the model estimates. In this study, we adopted a Gaussian mixed model for variable response latency (the time it takes for the animal to find the real hideout) of an experiment investigating the memory deficits in animals subjected to cerebral ischemia when treated with fish oil, the structure of covariance for the random effects used was the unstructured one. To estimate the parameters of the model, we used the methods of maximum likelihood and restricted maximum likelihood. Based on the likelihood ratio test, information criteria and adjustments of models we adopted autoregressive covariance matrix for the errors. The residual statistics analysis and diagnosis for the model were satisfactory, since the obtained results corroborate with biological evidence, that is, it was found the effectiveness of treatment to alleviate the cognitive effects caused by cerebral ischemia.

Keywords: Longitudinal data; random effect; covariance structure; latency.

Lista de ilustrações

Figura 1 – Esquema do desenvolvimento do experimento.	14
Figura 2 – Labirinto Radial Aversivo	14
Figura 3 – Gráficos de perfis individuais e médios para variável resposta latência nos diferentes tratamentos.	26
Figura 4 – Gráficos de autocorrelação para modelos com estruturas VC e AR(1) para os erros.	28
Figura 5 – Resíduos condicionais padronizados e envelope simulado com 95% para o RCCM.	31
Figura 6 – Distância de Mahalanobis e Distância de Lesaffre-Verbeke Padronizada.	31
Figura 7 – Distância de Cook e 1º termo da decomposição da DCC para o modelo ajustado.	32
Figura 8 – Segundo e terceiro termos da Decomposição da DCC para o modelo ajustado.	32
Figura 9 – Razão de Covariância para β e θ	33

Sumário

1	Modelos Mistos em Estudo de Isquemia Cerebral	11
1.1	Introdução	11
1.2	Delineamento e Metodologia Estatística	12
1.2.1	Isquemia Cerebral	12
1.2.2	Delineamento Experimental	13
1.2.3	Descrição do Experimento	13
1.3	Modelo Misto Gaussiano	15
1.3.1	Estruturas de Covariância	18
1.3.2	Estimação dos Parâmetros do Modelo	18
1.3.3	Seleção de Modelos	19
1.3.3.1	Teste da Razão de Verossimilhança	20
1.3.3.2	Critérios de Informação	21
1.3.4	Análise de Diagnóstico	21
1.3.4.1	Análise de Resíduos	22
1.3.4.2	Análise de Influência	23
1.4	Resultados e Discussões	26
1.4.1	Análise Exploratória dos Dados	26
1.4.2	Ajuste e escolha do modelo	27
1.4.3	Análise de Diagnóstico para o modelo ajustado	30
1.5	Considerações Finais	35
1.6	Sugestões de Pesquisa	35
	Referências	36

Modelos Mistos em Estudo de Isquemia Cerebral

1.1 Introdução

Nos últimos anos, tem sido crescente o número de experimentos das áreas da Saúde e Ciências Biológicas em que variáveis respostas são analisadas ao longo do tempo para uma mesma unidade experimental com o objetivo de verificar, por exemplo, o desempenho de tratamentos alternativos que possam vir a ser utilizados na cura de pacientes com algum tipo de doença. Estudos farmacológicos com essas características são realizados por Ferreira et al. (2014), Godinho et al. (2015) e Bacarin et al. (2015). Esses autores avaliam longitudinalmente (testes de memória) o desempenho cognitivo de grupos de animais submetidos a tratamentos com medicamentos e tempo de duração diferentes. Estudos com essas características, em que uma ou mais variáveis respostas são avaliadas repetidamente em uma mesma unidade pertencem à classe das medidas repetidas, .

Um caso especial dos estudos com medidas repetidas são os estudos longitudinais. As medidas são realizadas sob diferentes condições de avaliação e estão dispostas ao longo do tempo, respeitando uma ordem de análise. Quando assim caracterizados, os dados apresentam uma estrutura hierárquica, ou seja, as observações para um mesmo indivíduo possuem uma estrutura de dependência. Isso exige a modelagem de uma estrutura de covariância para captar e acomodar a variabilidade inerente aos dados longitudinais.

Uma forma de levar em consideração essa dependência na modelagem dos dados é utilizar modelos lineares com a suposição de erros correlacionados em que utilizando uma matriz de correlação acomoda-se a variação individual.

Alternativamente, pode-se modelar esses dados recorrendo a inclusão de um efeito aleatório no modelo. Nesse caso, tem-se um modelo com efeitos fixos e aleatórios, o qual

denomina-se modelo misto, cujo foco é acomodar a correlação entre as medidas repetidas e avaliar o comportamento individual e coletivo (de grupos) ao longo do tempo. Essa é uma classe de modelos bastante flexível, visto que proporcionam uma maior versatilidade na modelagem da estrutura de covariância dos dados para que se acomode as variações intra e entre indivíduos. De um modo geral, eles investigam e acomodam simultaneamente o efeito aleatório e a estrutura de correlação das medidas repetidas.

Para esses modelos é comum considerar a variável resposta seguindo uma distribuição normal, uma vez que tal abordagem está consolidada teoricamente desde a formulação, métodos de ajustes, avaliação dos mesmos e análise de diagnóstico (análise dos resíduos e influência), conforme observa-se em Diggle (1988), Diggle et al. (2002), Fitzmaurice, Laird e Ware (2004), Nobre (2004), Nobre e Singer (2007), Nobre e Singer (2011) com disponibilidade para o uso nos aplicativos como R e SAS como verifica-se em Pinheiro e Bates (2000) e Littel et al. (2006).

O escopo central deste trabalho é utilizar a metodologia dos modelos mistos gaussianos para modelar a variável resposta latência (lat - tempo para o animal encontrar o esconderijo verdadeiro) de um experimento farmacológico que tem por objetivo avaliar o efeito do tratamento com óleo de peixe em animais submetidos a isquemia cerebral, isto é, verificar se animais tratados apresentam melhora em seu quadro clínico (amnésia) em relação aos animais que não foram tratados.

1.2 Delineamento e Metodologia Estatística

1.2.1 Isquemia Cerebral

A expectativa de vida do ser humano cresceu consideravelmente nos últimos anos. Assim, existe a necessidade de desenvolvimento de estratégias farmacológicas para o tratamento e/ou prevenção de doenças neurodegenerativas agudas¹ ou crônicas². Nesse sentido, o interesse de pesquisadores da área farmacêutica em descobrir tratamentos que minimizem os efeitos de uma série de doenças, cresce constantemente, com objetivo de melhorar a saúde da população como um todo.

A Isquemia cerebral ocorre quando o fluxo sanguíneo cerebral é interrompido em uma área do cérebro (isquemia focal) ou no cérebro como um todo (isquemia global), e representa uma das principais causas de morte ou invalidez permanente. Isquemia cerebral pode ocorrer por diferentes causas, tais como: (i) oclusão embólica/trombótica de um ou mais vasos, (ii) parada cardíaca reversível, (iii) fibrilação atrial, (iv) hipotensão grave, ou ainda (v) durante ou após procedimentos diagnósticos e/ou cirúrgicos. A doença cerebral isquêmica é altamente prevalente, sendo que inúmeros fatores de risco, como obe-

¹ Lesão cerebral de origem isquêmica, infecciosa, traumática.

² Doença de Parkinson, Doença de Alzheimer entre outras.

sidade, hipertensão, diabetes, dislipidemias, sedentarismo, tabagismo, entre outros, desempenham papel importante na sua epidemiologia. Nos indivíduos que sobrevivem a um evento isquêmico cerebral, a gravidade das sequelas depende da extensão e duração da isquemia, e na maioria dos casos o indivíduo terá deficiências cognitivas/executivas (aprendizagem, memória, raciocínio, atenção, tomada de decisão e planejamento), sensoriais e motoras (paralisia) (LIM et al., 2004; ANDERSON; ARCINIEGAS, 2010).

De acordo com o Ministério da Saúde acidentes vasculares cerebrais (AVC), de origem isquêmica ou hemorrágica, são os primeiros causadores de óbitos no Brasil e terceira maior causa no mundo, perfazendo um total de 6 milhões de vítimas anualmente. Devido políticas de prevenção, o número de óbitos por conta do AVC foi reduzido em 32% entre 2000 e 2010 (BRASIL, 2014). Ressalta-se que, além de mortes, o AVC deixa sequelas graves em aproximadamente 50% dos sobreviventes, com ônus socioeconômico considerável por conta da incapacitação profissional permanente.

É importante salientar, que muitas pesquisas tem sido desenvolvidas para nortear novas propostas de tratamento nos ensaios clínicos com intuito de aliviar os efeitos de lesões isquêmicas no cérebro humano, como por exemplo, no Programa de Ciências Farmacêuticas da UEM, com os estudos do grupo de pesquisa em Isquemia Cerebral e Farmaconeuroproteção, sob a liderança do Professor Doutor Humberto Milani, em conjunto com professores do Programa de Mestrado em Bioestatística.

Os dados utilizados neste estudo são resultados de um experimento realizado no Laboratório de Isquemia Cerebral e Neuroproteção da UEM e concluído no segundo semestre de 2014.

1.2.2 Delineamento Experimental

O experimento foi realizado segundo um delineamento inteiramente casualizado com sete tratamentos e cinco repetições (Dias de Testes de Memória - TM). Foram avaliados os tratamentos "Sham"(15 animais), OP 4 horas (16 animais), OP 8 horas (13 animais), OP 12 horas (13 animais), Veículo 4 horas - V4 (10 animais), Veículo 8 horas - V8 (10 animais), Veículo 12 horas - V12 (10 animais).

1.2.3 Descrição do Experimento

Em todo o experimento foram utilizados 85 ratos da espécie Wistar com idade de 3 meses e peso variando entre 280 e 300 gramas. A Figura 1 mostra um esquema do desenvolvimento do experimento para avaliar como varia a eficácia antiamnésica do óleo de peixe (OP, 300 mg/kg) em função do momento em que o tratamento foi iniciado após a isquemia.

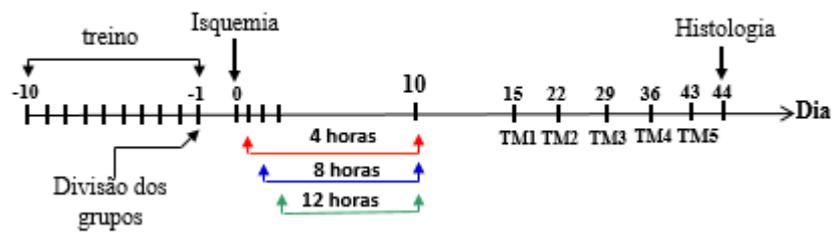


Figura 1 – Esquema do desenvolvimento do experimento.

Inicialmente, todos os animais foram treinados no teste do labirinto radial aversivo (PAGANELLI et al., 2004) (LRA, Figura 2) por 10 dias consecutivos.

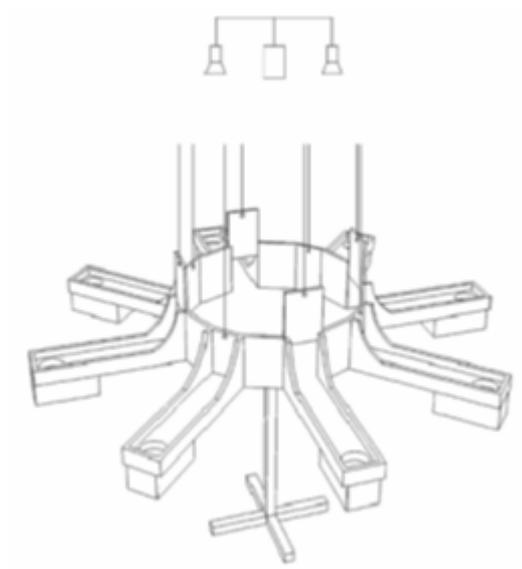


Figura 2 – Labirinto Radial Aversivo

A tarefa no LRA consiste em que o animal aprenderá em qual dos oito braços encontra-se a área de segurança localizada na extremidade de um dos oito braços radiais (esconderijo). Duas lâmpadas fluorescentes encontram-se cerca de 1 metro acima da superfície do LRA, cuja luminosidade torna-se aversiva para o rato. Para fugir deste ambiente, o animal procura avidamente por algum lugar "seguro".

Um ou dois dias pós-treino, os animais foram submetidos à isquemia cerebral e tratados com: "Sham"(Grupo falso isquêmico), Óleo de Peixe (OP 4, 8 e 12 horas pós-isquemia - Grupo isquêmico tratado com Óleo de Peixe), Veículo (V 4, 8 e 12 horas pós-isquemia - Grupo isquêmico tratado com Veículo). O que diferencia um rato isquêmico de um falso isquêmico é o fato de que ratos isquêmicos são submetidos a lesão cerebral e tratados com um placebo ou OP, enquanto os falsos isquêmicos são submetidos apenas aos procedimentos cirúrgicos, um grupo controle.

Após o início do tratamento, as doses de medicamento foram veiculadas uma vez ao dia durante 10 dias. Passados 15 dias da isquemia, os animais foram avaliados, quanto

sua capacidade de lembrar da tarefa aprendida durante o treino, ou seja, memória retrógrada. Os testes de memória (TM) foram aplicados 1 vez/semana, durante 5 semanas.

A capacidade de aprendizagem e memória no LRA foi avaliada em cinco dias, com intervalo de uma semana entre uma avaliação e outra (3 tentativas por dia), e é medida por três parâmetros (variáveis resposta): (i) Latência; (ii) Número de erros de memória de referência (memória a longo prazo) e (iii) Número de erros de memória de trabalho (memória a curto prazo).

A latência expressa o tempo gasto (segundos) pelo animal para encontrar o esconderijo durante cada tentativa. Enquanto explora o labirinto, o animal comete erros ao entrar em braços contendo os esconderijos falsos (7 no total). Em cada tentativa, o animal cometerá um erro de memória de referência (7 no máximo) ao entrar, pela primeira vez, em um braço falso. Caso ele retorne ao braço previamente visitado, registra-se, então, um erro de memória de trabalho. A memória de referência é de longo prazo, e se consolida ao longo das várias tentativas e sessões (dias) de teste. Já a memória de trabalho é de curto prazo e contexto-dependente, relevante somente para um dado momento. Para registrar a latência, erro de referência diário e erro de trabalho diário foi computada a média aritmética das tentativas diárias.

Os procedimentos operacionais para indução da isquemia seguiram os "Princípios básicos para a utilização animal", conforme aprovado pelo Comitê de Ética em Experimentação Animal da Universidade Estadual de Maringá.

1.3 Modelo Misto Gaussiano

Ao avaliar estudos biológicos longitudinais, a modelagem dos dados utilizando um modelo de regressão com suposição de independência entre observações de uma mesma unidade experimental não é adequada, uma vez que se deve considerar a correlação existente entre as mesmas. Uma forma de acomodar a correlação serial é realizando a modelagem da dependência incluindo uma estrutura de covariância no modelo, do mesmo modo que efeitos aleatórios para que seja possível avaliar o comportamento intrínseco a cada unidade experimental.

O uso dos modelos mistos está inserido no contexto da saúde quando se deseja avaliar/modelar o comportamento de unidades experimentais e grupos de tratamento observados ao longo do tempo.

Nas pesquisas encontradas na literatura, cujas análises estatísticas voltam-se para essa metodologia, é possível citar o trabalho de Dzialowski et al. (2007) que utilizaram-na para comparação, entre grupos, do grau de edema cerebral em ratos avaliados longitudinalmente (ao longo do tempo) após reperfusão arterial precoce (ou imediata). Wang et al.

(2009) e Sussman et al. (2012) empregaram modelos mistos para avaliação longitudinal do grau de lesão da massa branca cerebral medida por ressonância magnética e das alterações hemodinâmicas cerebrais³ globais e regionais medidas ao longo do tempo por meio de absorção ótica de tomografia fotoacústica comparadas entre o cortex esquerdo e direito após hipóxia-isquemia⁴, respectivamente. Da mesma forma, Dhungana et al. (2013) adotaram modelos mistos para análise dos dados em um experimento que avaliou se o envelhecimento agrava o dano isquêmico cerebral induzido por acidente vascular cerebral em ratos com infecção periférica crônica. Um modelo misto com efeito aleatório na unidade experimental foi utilizado por Heikkinen et al. (2014) para avaliar a perda neural entre os grupos de trabalho, em um estudo que objetivou-se identificar a suscetibilidade da isquemia encefálica global e local em ratos com Alzheimer. Holliday et al. (2015) utilizaram modelos mistos para estimar correlações genéticas em um estudo que avaliou a sobreposição genética entre subtipos de acidente vascular cerebral isquêmico. De um modo geral, percebe-se a importância do uso dos modelos mistos para avaliação de dados provenientes de ensaios clínicos associados a isquemia cerebral.

Segundo West, Welch e Galecki (2014), um modelo misto é um modelo linear paramétrico para dados em clusters, medidas repetidas e dados longitudinais, os quais quantificam a relação entre uma variável resposta dependente e vários preditores lineares, e que contém efeitos fixos e aleatórios. Com o uso desses modelos corrige-se a distorção que surge ao se utilizar modelos de regressão para respostas independentes para analisar dados correlacionados.

A forma geral desses modelos é dada por

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \quad (1.1)$$

em que \mathbf{Y}_i é um vetor de dimensão $n_i \times 1$, composto por variáveis respostas para a i -ésima unidade experimental $i = 1, \dots, n$ e n_i sendo o número de medidas longitudinais para a i -ésima unidade experimental; \mathbf{X}_i ($n_i \times p$) é uma matriz de delineamento dos efeitos fixos, ou seja, que representa valores conhecidos de p covariáveis; $\boldsymbol{\beta}$ ($p \times 1$) é um vetor de parâmetros, desconhecidos, associado aos efeitos fixos; \mathbf{Z}_i ($n_i \times q$) é uma matriz de delineamento dos efeitos aleatórios, ou seja, que representa valores conhecidos de q covariáveis, \mathbf{b}_i ($q \times 1$) é um vetor de efeitos aleatórios e $\boldsymbol{\epsilon}_i$ ($n_i \times 1$) é um vetor dos erros aleatórios assumidos independentes entre si. Assume-se que $\mathbf{b}_i \sim N_q(0, \mathbf{G})$ e $\boldsymbol{\epsilon}_i \sim N_{n_i}(0, \mathbf{R}_i)$, em que seguem uma distribuição normal multivariada com vetor de médias 0 e matrizes (simétricas) de variância e covariância \mathbf{G} e \mathbf{R}_i respectivamente, com $\boldsymbol{\epsilon}_i$ e \mathbf{b}_i independentes.

Segundo Laird e Ware (1982), marginalmente os \mathbf{Y}_i são independentes e seguem

³ Alterações do fluxo sanguíneo.

⁴ Lesão seletiva da massa branca cerebral.

uma distribuição normal multivariada com média $X_i\beta$ e matriz de variâncias e covariâncias,

$$V_i = Z_i G Z_i' + R_i, \quad (1.2)$$

sendo V_i o i -ésimo bloco da matriz bloco diagonal V , cujos elementos são chamados de componentes de variância, os quais compõem o vetor θ . Alternativamente escreve-se

$$Y_i \sim N(X_i\beta, V_i). \quad (1.3)$$

De acordo com Perri e lemma (1999), a análise de um modelo misto consiste de duas partes: uma análise para a parte aleatória e outra para a parte fixa. A análise da parte aleatória baseia-se na predição dos efeitos aleatórios, na presença de efeitos fixos e estimação dos componentes de variância. Já a análise da parte fixa consiste de estimação e testes de hipóteses sobre funções estimáveis dos efeitos fixos. Em geral, tanto a predição dos efeitos aleatórios quanto estimação dos efeitos fixos dependem da estimação dos componentes de variância.

É comum que os modelos mistos, para dados longitudinais, apresentem efeitos aleatórios no intercepto e na inclinação individual, os quais podem ser identificados por intermédio dos gráficos de perfis individuais. De acordo com Fausto et al. (2008), os efeitos aleatórios no intercepto representam a heterogeneidade natural entre os indivíduos decorrentes de características não mensuradas. Todavia, parte da variabilidade dos dados pode ser influenciada por medidas não avaliadas em relação as covariáveis, entre os indivíduos e dentro dos indivíduos. Quando isso ocorre, necessita-se da inclusão de um efeito nas covariáveis de interesse. Assim, adicionar um efeito aleatório na variável tempo é justificado pela hipótese de se ter uma variabilidade do desfecho dependendo do tempo.

Para análise dos dados do experimento apresentado na seção 1.2.2 é especificado um modelo misto gaussiano com efeito aleatório no intercepto e na inclinação, sob a suposição de que a distribuição de probabilidade da variável resposta latência seja aproximadamente normal. Considerando que animais diferentes respondem inicialmente de maneira distinta, isto é, a latência para cada animal no primeiro momento de avaliação é diferente dos demais, implica na inclusão de um efeito aleatório no intercepto. Da mesma forma, a resposta para cada rato para os diferentes TM difere dos demais, ou seja, existem inclinações diferentes. Logo, há necessidade de atribuir um efeito aleatório a variável tempo para cada animal.

Diante disso, o modelo misto proposto inicialmente é dado por

$$y_{ijk} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t_{ik} + \alpha_j + \gamma(\alpha_j t_{ik}) + \epsilon_{ijk}, \quad (1.4)$$

sendo y_{ijk} a latência do i -ésimo rato, $i=1, \dots, 85$, observado no j -ésimo tratamento, $j=1, \dots, 7$, durante o k -ésimo dia (TM), $k = 1, \dots, 5$ a qual segue uma distribuição normal multivariada com média $X_i\beta$ e matriz de variância e covariância V_i ; β_0 é o parâmetro relativo ao

intercepto (média geral); α_j é o efeito relativo ao j -ésimo tratamento; $t_{ik} = 1, 2, 3, 4, 5$ é o tempo (dia de TM) avaliado como uma variável contínua; γ é o parâmetro para interação entre o j -ésimo tratamento com o k -ésimo dia de TM; b_{i0} e $b_{i1}t_{ik}$ são os efeitos aleatórios no intercepto e inclinação individual e ϵ_{ijk} é o erro aleatório.

Para o modelo dado por 1.4 assume-se que $\mathbf{b}_i = (b_{i0}, b_{i1})' \sim N(0, \mathbf{G})$ e o erro ϵ_{ijk} associado a resposta y_{ijk} segue uma distribuição normal multivariada com média 0 e matriz de variância e covariância \mathbf{R}_i , isto é, $\epsilon_i \sim N(0, \mathbf{R}_i)$, sendo que

$$\mathbf{G} = \begin{bmatrix} \sigma_{b_{i0}}^2 & \sigma_{b_{i0}, b_{i1}} \\ \sigma_{b_{i0}, b_{i1}} & \sigma_{b_{i1}}^2 \end{bmatrix} \quad e \quad \mathbf{R}_i = \begin{bmatrix} var(\epsilon_{1i}) & cov(\epsilon_{1i}, \epsilon_{2i}) & \dots & cov(\epsilon_{1i}, \epsilon_{5i}) \\ cov(\epsilon_{1i}, \epsilon_{2i}) & var(\epsilon_{2i}) & \dots & cov(\epsilon_{2i}, \epsilon_{5i}) \\ \vdots & \dots & \ddots & \vdots \\ cov(\epsilon_{1i}, \epsilon_{5i}) & cov(\epsilon_{2i}, \epsilon_{5i}) & \dots & var(\epsilon_{5i}) \end{bmatrix}$$

1.3.1 Estruturas de Covariância

Segundo Singer, Nobre e Rocha (2015), um dos principais pontos da modelagem de dados longitudinais é a escolha da estrutura de covariância. Isso possibilita fazer inferências adequadas para que se obtenha resultados válidos para o experimento. Diversas são as estruturas encontradas na literatura, entretanto, podemos destacar a Componente de Variância (VC), Composta Simétrica (CS), Autorregressiva de Ordem 1 (AR(1)) e a Não-Estruturada (UN). Essa é uma das versatilidades de se utilizar modelos mistos, sendo que existem 10 estruturas disponíveis na biblioteca nlme do *software* R 3.1.3 e 23 estruturas disponíveis na *procedure* MIXED do *software* SAS 9.4.

Na estrutura VC assume-se que as variâncias são constantes em n_i ocasiões distintas. Na estrutura CS tem-se que variâncias e covariâncias são homogêneas. Já para estrutura AR(1), as variâncias homogêneas e covariâncias reduzindo-se exponencialmente à medida que se aumenta o intervalo de tempo. Para a UN todas as variâncias e covariâncias são desiguais. Outras estruturas de covariância são apresentadas por Fitzmaurice, Laird e Ware (2004); Littell et al. (2006) entre outros.

1.3.2 Estimação dos Parâmetros do Modelo

Em geral, os métodos utilizados na estimação dos parâmetros relativos aos efeitos fixos e componentes de variância do modelo misto baseiam-se na teoria da Máxima Verossimilhança, isto é, nos métodos da Máxima Verossimilhança (MV) e Máxima Verossimilhança Restrita (MVR) os quais serão utilizados neste estudo. Todavia, ajustes alternativos podem ser realizados utilizando metodologias Bayesianas conforme observa-se em Resende et al. (2001), Molenberghs e Verbeke (2005), Carlin e Louis (2009), Rizzato (2011) entre outros.

Tendo em conta que para o modelo 1.1 $Y \sim N(X\beta, V)$, os métodos da MV e MVR consistem em maximizar o logaritmo das funções de verossimilhança marginal e verossimilhança marginal restrita da distribuição normal multivariada, em relação ao vetor de parâmetros relativos aos efeitos fixos (β), tal como em relação ao vetor dos componentes de variância (θ). A estimativa dos parâmetros do vetor de efeitos fixos depende dos componentes de variância, os quais nem sempre são conhecidos. Assim, deve-se estimá-los.

Para estimar os componentes de variância, em ambos os casos, maximiza-se as respectivas funções de verossimilhança em relação a θ , considerando uma estimativa de β obtida substituindo-se V por \tilde{V} , sendo \tilde{V} uma matriz auxiliar. A equação obtida para maximização é não-linear, isto é, não possui solução analítica. Logo, métodos iterativos como algoritmo de *Newthon-Raphson*, *Score de Fisher* ou *EM* podem ser utilizados.

Problemas surgem quando obtêm-se as estimativas de θ , $\hat{\theta}$, pelo método da MV, isto é, as mesmas tendem a ser viesadas. Uma maneira de corrigir tal problema é o uso do método da MVR. O mesmo fornece estimadores não viesados para os componentes de variância, uma vez que esse considera a perda dos graus de liberdade devido a estimação dos p parâmetros de β (PERRI; IEMMA, 1999).

De forma geral, tanto a estimativa de β quanto o preditor dos efeitos aleatórios podem ser obtidos conjuntamente. Nesse caso, considerando a função de densidade de probabilidade conjunta da variável resposta e efeitos aleatórios, obtêm-se as equações dos modelos mistos cuja solução fornecem as estimativas $\hat{\beta}$ e o preditor de b , \hat{b} . Utilizando uma extensão do Teorema de Gauss-Markov, mostra-se que o estimador de β é o melhor estimador não viesado (*Best linear unbiased estimator - BLUE*) e o preditor dos efeitos aleatórios, é o melhor preditor linear não viesado (*Best Linear Unbiased Predictor - BLUP*). Eles são melhores, pois minimizam a variância amostral, lineares, pois são uma função linear de y e não viesados, pois seu valor esperado é igual a estimativa (HARVILLE, 1976).

1.3.3 Seleção de Modelos

Selecionar o modelo adequado para o conjunto de dados requer cuidados, pois deseja-se obter um modelo parcimonioso. Contudo, a partir de uma escolha inadequada os resultados obtidos podem não ser válidos, assim como conduzirem a resultados viesados.

Alguns autores têm direcionado esforços no que tange a seleção de modelos mistos utilizando diferentes técnicas, como é o caso de Pu e Niu (2006), que propõem o uso de uma generalização do critério de informação generalizado (GIC) desenvolvido por Rao e Wu (1989). Já Srivastava e Kubokawa (2010) sugerem o uso de extensões do critério de informação AIC condicional quando as observações tenham sido obtidas de clusters. Da mesma forma, Ibrahim et al. (2011) selecionam efeitos fixos e aleatórios usando máxima verossimilhança penalizada em conjunto com outras técnicas, Muller, Scealy e Welsh

(2013) estudam a seleção de modelos seguindo técnicas Bayesianas e métodos de encolhimento. Fan, Qin e Zhu (2014) propõem o uso de um estimador robusto penalizado para seleção das variáveis do modelo.

Em suma, observa-se a gama de opções existentes para seleção de modelos dessa classe. Não obstante, as formas mais usuais de seleção de modelos recaem sobre o Teste da Razão de Verossimilhança (TRV), da mesma forma que aos critérios de Informação de Akaike (AIC) e Bayesiano Schwarz (BIC), os quais serão utilizados neste estudo, pois estão consolidadas teoricamente e computacionalmente.

1.3.3.1 Teste da Razão de Verossimilhança

Testes da razão de verossimilhança são baseados na comparação dos valores das funções de log-verossimilhança para dois modelos encaixados. Neste caso, tem-se um modelo geral chamado de irrestrito (M_0) e um modelo restrito (M_1), sendo M_1 um caso particular de M_0 .

O TRV pode ser utilizado para testar hipóteses sobre os efeitos fixos assim como sobre os efeitos aleatórios. A estatística do teste é dada por

$$LR = -2\{l(\hat{\beta}_1, \hat{\theta}_1) - l(\hat{\beta}_0, \hat{\theta}_0)\} \quad (1.5)$$

em que $l(\hat{\beta}_1, \hat{\theta}_1)$ é a função de log-verossimilhança sob o modelo restrito e $l(\hat{\beta}_0, \hat{\theta}_0)$ é a função de log-verossimilhança sob o modelo irrestrito. Sob a hipótese nula, a estatística LR é assintoticamente distribuída como uma χ_{gl}^2 , com graus de liberdade igual a diferença entre o número de parâmetros do modelo irrestrito com o número de parâmetros do modelo restrito, a qual pode ser utilizada para verificar a significância do teste.

Ao testar-se hipóteses sobre modelos encaixados com diferentes estruturas de covariância, cujo método da MVR é utilizado na estimação dos parâmetros, na estatística do TRV deve-se substituir o logaritmo da função de verossimilhança pelo logaritmo da função de verossimilhança restrita. Nesse caso, tem-se o Teste da Razão de Verossimilhança Restrita (TRVR).

Quando deseja-se verificar a inclusão de efeitos aleatórios no modelo, segundo Verbeke e Molenberghs (2000), deve-se tomar cuidado com a distribuição assintótica do teste, uma vez que os parâmetros de covariância (elementos da matriz e variância e covariância dos efeitos aleatórios) podem situar-se na fronteira do espaço paramétrico. Nesse caso, de acordo com Self e Liang (1987), a distribuição assintótica do teste, será comparada a uma mistura de distribuições χ_q^2 e χ_{q+1}^2 , ambas com pesos 0,5 em que q e $q+1$ representam o número de efeitos aleatórios de cada modelo. Nesse contexto, deve-se testar se a variância associada ao efeito aleatório é igual a zero ou não (WEST; WELCH; GALECKI, 2014).

1.3.3.2 Critérios de Informação

Os critérios de informação AIC e BIC baseiam-se no valor máximo da função de log-verossimilhança ou log-verossimilhança restrita, bem como no número de parâmetros do modelo e observação do conjunto de dados. Ambos os critérios podem ser usados na comparação de modelos encaixados ou não. Quando os ajustes dos modelos são realizados utilizando o método da MV, os critérios podem ser comparados entre quaisquer modelos ajustados para o mesmo conjunto de dados. No entanto, quando o ajuste for realizado usando o método da MVR, seus valores, e incluindo log-verossimilhança, só podem ser comparados entre modelos que possuem a mesma estrutura de efeitos fixos.

Cada critério possui suas particularidades, sendo que o critério de Akaike tende a indicar modelos mais complexos como adequados. O mesmo torna-se adequado à medida que o tamanho amostral cresce. Já o critério BIC penaliza modelos com um número maior de parâmetros, e de acordo com Gurka (2006), quando se desejar selecionar a estrutura de covariância para o modelo, ele deve ser preferido. Todavia, a decisão não deve ser baseada estritamente neles. De forma geral, modelos que apresentarem menores valores de AIC e/ou BIC devem ser preferidos.

1.3.4 Análise de Diagnóstico

Com a análise de diagnóstico busca-se verificar a adequabilidade do modelo aos dados. É uma etapa importante na modelagem, uma vez que auxilia na verificação do afastamento dos pressupostos e permite detectar observações extremas que venham a interferir nos resultados. É possível dividir essa fase entre a análise dos resíduos e influência.

Por meio da análise dos resíduos, verifica-se se o modelo ajustado é adequado, ou seja, se suposições de independência, homocedasticidade e normalidade são satisfeitas. Em determinadas ocasiões esses pressupostos são violados. Consequentemente, pontos com características atípicas podem ser identificados. Assim, deve-se conhecer a natureza destes valores para verificar se os mesmos realmente possuem comportamentos atípicos ou se isso pode ocorrer dentro do grupo ao qual pertencem.

Como parte do diagnóstico, é necessário realizar uma análise detalhada para o modelo ajustado, isto é, verificar a possível existência de pontos com características influentes. Isso posto, investiga-se a existência de observações que estejam causando algum tipo de perturbação nos dados, tal qual nas hipóteses do modelo, as quais possam estar influenciando nas estimativas dos parâmetros.

1.3.4.1 Análise de Resíduos

Segundo Nobre e Singer (2007), para os modelos mistos, os resíduos mais usuais são

- Resíduos Marginais: $\hat{r} = Y - X\hat{\beta}$, preditores dos erros marginais $r = Y - X\beta$;
- Resíduos Condicionais: $\hat{\epsilon} = Y - X\hat{\beta} - Z\hat{b}$, preditores dos erros condicionais $\epsilon = Y - X\beta - Zb$;
- Resíduos dos Efeitos Aleatórios: $Z\hat{b}$, preditores dos efeitos aleatórios Zb .

Os resíduos marginais são utilizados para aferir a suposição de linearidade entre $E[Y]$ e as covariáveis. Em vista disso, constrói-se um gráfico dos resíduos \hat{r} *versus* o valor das covariáveis. Para que a suposição seja atendida espera-se que os elementos dos gráficos variem aleatoriamente.

Tais resíduos podem ser utilizados para verificar quais unidades amostrais a estrutura de covariância utilizada é adequada. Para isso, deve-se construir um gráfico das unidades observadas *versus* uma padronização na distância de Lesaffre-Verbeke, sugerida por Singer, Nobre e Rocha (2015). Observações fora dos limites estabelecidos indicam que a estrutura de covariância adotada para os erros é inadequada à elas.

Para avaliar a suposição de homocedasticidade dos resíduos, Nobre e Singer (2007) sugerem o uso do gráfico dos valores preditos *versus* resíduos condicionais padronizados, uma vez que é possível que os $\hat{\epsilon}$ tenham variâncias desiguais. A padronização é feita segundo a expressão a seguir:

$$\hat{\epsilon}_i^* = \frac{\hat{\epsilon}_i}{\sigma\sqrt{q_{ii}}} \quad (1.6)$$

em que $\hat{\epsilon}_i$ é o i -ésimo elemento de $\hat{\epsilon}$, q_{ii} é o i -ésimo elemento da diagonal de $Q = W - WX(X'WX)^{-1}X'W$ em que $W = \sigma^2V^{-1}$. Para que a condição de homocedasticidade seja satisfeita um comportamento aleatório em torno do zero é necessário.

Resíduos que dependem de mais de uma fonte de variação tendem a ser confundidos, como é o caso dos resíduos condicionais. Isso pode dificultar a verificação da suposição de normalidade dos resíduos. Para correção desse impedimento, Hilden-Milton (1995 apud NOBRE; SINGER, 2007) definem uma razão de confundimento e nesta aplicam uma linearização de modo a minimizar o confundimento dos resíduos condicionais, obtendo assim, o que nomeia-se resíduo condicional com confundimento mínimo (RCCM). Desse modo, para avaliar o pressuposto de normalidade por meio de uma padronização dos RCCM, esse autor sugere o uso de um gráfico QQ-plot com envelope simulado, sendo

que se os pontos mantêm-se no interior do envelope e a condição de normalidade é satisfeita.

Os resíduos dos efeitos aleatórios são utilizados para verificar a existência de unidades experimentais *outliers*, em que $Z\hat{b}$ reflete a diferença entre a resposta predita para o i -ésimo indivíduo e a média geral (NOBRE, 2004). No entanto, ao invés de utilizar \hat{b}_i ou $Z_i\hat{b}_i$ para aferir tal situação, Waternaux, Laird e Ware (1999) sugerem o uso da distância de Mahalanobis (DM) (M_i), a qual é dada pela expressão

$$M_i = \hat{b}'_i \hat{V} [\hat{b}_i - \mathbf{b}_i]^{-1} \hat{b}_i, \quad (1.7)$$

em que $\hat{V}[\hat{b}_i - \mathbf{b}_i]$ é a variância do preditor dos efeitos aleatórios e sob a validade do modelo, tem-se que $M_i \sim \chi^2_{n_i}$ para n_i suficientemente grande. Nesse caso, um gráfico de M_i em função dos índices de unidades amostrais podem ser empregados para identificar a presença daquelas que podem ser consideradas discrepantes, as quais estarão fora dos limites estabelecidos (NOBRE, 2004; NOBRE; SINGER, 2007).

1.3.4.2 Análise de Influência

Dentre as formas mais comuns para se avaliar a influência de observações que estejam exercendo influência desproporcional sobre as estimativas dos parâmetros do modelo, destacam-se a omissão/eliminação de observações e identificação de pontos de alavanca.

Pontos influentes são interpretados como aqueles que, devido a sua presença ou não no conjunto de dados, possam alterar aspectos importantes no ajuste dos modelos. A base da análise de influência está associada ao diagnóstico para uma determinada observação (ou subconjunto de observações) e quantifica-se o efeito da omissão dessas observações sobre os resultados da análise de todo o conjunto de dados.

Recorrendo a eliminação de uma observação ou de um grupo de observações, do conjunto de dados, avalia-se os efeitos provocados, positivos ou negativos, por sua ausência. Essa técnica foi inicialmente desenvolvida por Cook (1977) e explorada por vários autores.

De acordo com Nobre e Singer (2007), para modelos mistos, a distância de Cook é dada por

$$D_I = \frac{(\hat{\beta} - \hat{\beta}_I)(X'V^{-1}X)(\hat{\beta} - \hat{\beta}_I)}{c} = \frac{(\hat{Y} - \hat{Y}_I)'V^{-1}(\hat{Y} - \hat{Y}_I)}{c}, \quad (1.8)$$

em que c é o parâmetro de escala e I é o conjunto de observações retiradas do conjunto de dados. É evidente que I pode ser um conjunto formado pelas observações de cada uni-

dade experimental com característica influente, bem como ser composta por uma unidade experimental com todas as suas observações.

No contexto dos modelos de efeitos aleatórios, isto é, dos modelos mistos, Tan, Ouwens e Berger (2001) propõe o uso da Distância de Cook Condicional (DCC), visto que a distância de Cook usual possui determinadas limitações no que tange a identificação de observações que estejam exercendo influência desproporcional nas estimativas dos efeitos aleatórios.

A DCC é expressa por

$$D_i^{Cond} = \sum_{j=1}^c \frac{P'_{j(i)} P_{j(i)}}{k}, \quad (1.9)$$

em que $P_{j(i)} = (X_j \hat{\beta} + Z_j \hat{b}_j) - (X_j \hat{\beta}_{(i)} + Z_j \hat{b}_{j(i)})$ com $\hat{\beta}_i$ e $\hat{b}_{j(i)}$ representando, respectivamente, $\hat{\beta}$ e \hat{b}_j quando elimina-se da amostra a i -ésima observação e $k = \sigma^2([n - 1]c + p)$.

A DCC é decomposta em três componentes, em que por meio de cada uma, é possível avaliar as influências em partes distintas do modelo.

A primeira componente ($DCi1$) permite avaliar a influência da i -ésima observação em relação aos efeitos fixos. Já a segunda componente ($DCi2$) permite verificar quais observações estão exercendo influências nas estimativas dos efeitos aleatórios.

$$D_{i1}^{Cond} = \frac{(\hat{Y} - \hat{Y}_i)' (\hat{Y} - \hat{Y}_i)}{k}. \quad e \quad D_{i2}^{Cond} = \frac{(\hat{b} - \hat{b}_i)' Z' Z (\hat{b} - \hat{b}_i)}{k}. \quad (1.10)$$

A terceira componente ($DCi3$) permite avaliar a influência sobre os efeitos fixos e aleatórios simultaneamente.

$$D_{i3}^{Cond} = \frac{2(\hat{\beta} - \hat{\beta}_i)' X' Z (\hat{b} - \hat{b}_i)}{k}. \quad (1.11)$$

Graficamente, para identificar um ponto influente constrói-se um gráfico da DCC *versus* o valor observado, ou das suas partições, sendo que um ponto será candidato a influente se apresentar alto valor no gráfico quando comparado aos demais.

Mediante a omissão das observações, é possível aferir a precisão das estimativas dos parâmetros relativos aos efeitos fixos, assim como dos componentes de variância modelo utilizando a Razão de Covariância (*Covariance Ratio - CovRatio*). Nesse caso são comparados os determinantes das matrizes de covariância dos dados reduzidos e completos.

As expressões para o cálculo desta medida para β e para θ são dadas respectivamente por:

$$CovRatio(\boldsymbol{\beta}) = \frac{\det(Var[\hat{\boldsymbol{\beta}}_I])}{\det(Var[\hat{\boldsymbol{\beta}}])} \quad e \quad CovRatio(\boldsymbol{\theta}) = \frac{\det(Var[\hat{\boldsymbol{\theta}}_I])}{\det(Var[\hat{\boldsymbol{\theta}}])} \quad (1.12)$$

sendo que $Var(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ e para os componentes de variância, a matriz de variância utilizada nesse caso é a inversa da matriz das segundas derivadas, isto é, da matriz Hessiana.

Uma razão maior que 1 indica maior precisão nas estimativas com os dados completos, enquanto uma *CovRatio* menor que 1 implica em uma maior precisão considerando os dados reduzidos. Para investigar a influência, visualmente constrói-se um gráfico das unidades experimentais *versus* suas respectivas *CovRatio* (SCHABENBERGER, 2004).

Um ponto de alavanca é definido como aquele que possui forte influência no seu valor predito, isto é, um ponto que reflete forte variação em relação ao seu valor predito (NOBRE; SINGER, 2011). Conceitos de alavancagem surgem para avaliação de pontos de alavanca no que tange os modelos mistos.

Da mesma maneira que na DCC, deve-se verificar pontos que estejam alavancando as estimativas dos efeitos fixos e aleatórios, e nesse caso, deve-se modificar a matriz de alavancagem. Em geral, um ponto será de alavanca quando estiver diferenciando-se, evidentemente, em relação as demais observações.

Em síntese, para que sejam identificados os pontos de alavanca, gráficos de alavancagem *versus* unidades ou índices de observação podem ser utilizados, em que pontos fora dos limites estabelecidos são candidatos a alavanca.

Mais detalhes sobre análise de diagnóstico para modelos lineares mistos são abordados por Hilden-Milton (1995), Nobre (2004), Nobre e Singer (2007) e Nobre e Singer (2011) entre outros autores.

1.4 Resultados e Discussões

Para a realização dos ajustes e análises do modelo adotado foram utilizados os softwares R versão 3.1.3 e SAS versão 9.4, sendo que todas as hipóteses foram testadas com nível de significância de 5%. Os gráficos utilizados para análise de diagnóstico foram feitos com base nos algoritmos disponibilizados em <http://www.ime.usp.br/~jmsinger>.

1.4.1 Análise Exploratória dos Dados

A Figura 3 representa os perfis individuais e médios dos grupos Sham, OP 4, 8 e 12h e Veículo 4, 8 e 12h.

Pela Figura 3a, identifica-se que na sua maioria, animais possuem latências iniciais diferentes, bem como inclinações diferentes ao longo dos TM. Considerando a Figura 3b, observa-se que a medida que os TM são aplicados, os animais do grupo OP, independente do tempo de início do tratamento, levam um tempo inferior, quando comparado com o grupo Veículo, para encontrar o esconderijo verdadeiro, indicando ocorrer um efeito de tempo (dias de teste) para os grupos OP.

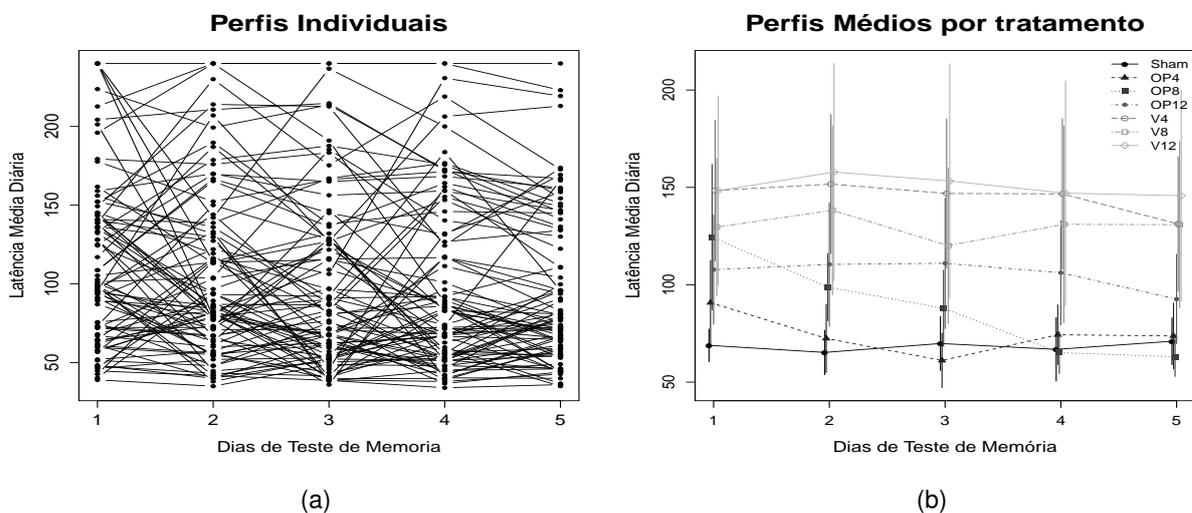


Figura 3 – Gráficos de perfis individuais e médios para variável resposta latência nos diferentes tratamentos.

A matriz abaixo contém as correlações observada dos dados. Quando mensurada a correlação entre os dias 1 e 2, 1 e 3 os valores são próximos, sendo que o mesmo ocorre para os dias finais. O mesmo pode-se observar para os demais dias. Em geral, tem-se valores decrescentes.

$$\begin{bmatrix} 1,0000 & 0,5893 & 0,5929 & 0,4525 & 0,4522 \\ & 1,0000 & 0,7803 & 0,6338 & 0,6921 \\ & & 1,0000 & 0,6720 & 0,6402 \\ & & & 1,0000 & 0,8768 \\ & & & & 1,0000 \end{bmatrix}$$

A julgar pela análise descritiva dos dados (Tabela 1) a latência do grupos OP4 foi inferior a do grupo V4 independente do TM, ou seja, animais tratados levaram um tempo menor para executar a tarefa. O mesmo cenário se reproduziu para os demais tratamentos. Isso indica que animais do grupo veículo possuem uma dificuldade maior para realizar a tarefa quando comparados com os animais tratados com óleo de peixe.

Tabela 1 – Média (\bar{x}) e desvio-padrão (σ) das latências médias para os tratamentos OP4 e V4.

Teste de Memória	OP4		V4	
	\bar{x}	σ	\bar{x}	σ
1	90,92 s	72,44 s	148,36 s	50,40 s
2	72,44 s	33,90 s	151,70 s	49,95 s
3	61,11 s	26,29 s	146,89 s	53,55 s
4	74,32 s	28,80 s	146,67 s	54,03 s
5	73,73 s	31,77 s	131,23 s	48,26 s

1.4.2 Ajuste e escolha do modelo

Considerando o modelo geral dado por 1.4, foram avaliados outros dois modelos, modelo 2 o qual tem efeito aleatório apenas no intercepto e modelo 3 com efeito aleatório apenas na inclinação. Para verificar a necessidade de manter o efeito aleatório no intercepto e posteriormente o efeito aleatório na inclinação, testou-se as seguintes hipóteses:

$$T_1 : \begin{cases} H_0 : g_{11} = 0 \\ H_1 : g_{11} > 0 \end{cases} \quad e \quad T_2 : \begin{cases} H_0 : g_{22} = 0 \\ H_1 : g_{22} > 0 \end{cases}$$

A partir dos resultados obtidos por meio do TRVR ($p < 0,05$), com 5% de significância, existem evidências amostrais para rejeitar a hipótese nula em ambos os casos, isto é, deve-se manter o efeito aleatório no intercepto e na inclinação individual.

Um modelo sem interação entre os efeitos do tratamento e o tempo foi comparado (M_4) com o modelo 1.4 (M_1). Os resultados das comparações usando o critério BIC foram favoráveis a não mantê-la no mesmo ($BIC_{M_4} = 4338,48$ e $BIC_{M_1} = 4360,18$). Pelos valores do critério AIC ($AIC_{M_1} = 4287,25$ e $AIC_{M_4} = 4289,85$) o resultado é favorável ao modelo com a interação. Ao observar os ajustes de ambos, biologicamente os resultados mais coerentes foram obtidos com o modelo sem a interação e assim, esse foi adotado.

Para escolha da estrutura de covariância R_i , dado o modelo misto com as restrições já obtidas, foram ajustados modelos com as estruturas de covariância VC ($M_{4.1}$), CS ($M_{4.2}$), AR(1) ($M_{4.3}$) e UN ($M_{4.4}$) para avaliação da qual captaria e acomodaria adequadamente a variabilidade inerente aos dados. Considerando os conjuntos de ajustes dos modelos propostos com as diferentes estruturas avaliadas, modelos com estruturas CS ($BIC_{M_{4.2}} = 4300,42$) e UN ($BIC_{M_{4.4}} = 4303,00$) não acomodaram a variabilidade adequada, visto que obteve-se estimativas de parâmetros com erros-padrão elevados. Avaliados os critérios de informação, o valor do BIC foi favorável a escolha da estrutura AR(1) ($BIC_{M_{4.3}} = 4290,90$), embora o valor do BIC para estrutura VC tenha sido muito próximo ($BIC_{M_{4.1}} = 4291,06$). Para decidir qual das duas estruturas adotar, utilizou-se o TRVR, em que testou-se hipóteses apresentadas em T_3 para comparação dos modelos $M_{4.1}$ e $M_{4.3}$.

$$T_3 : \begin{cases} H_0 : M_{4.1} = M_{4.3} \\ H_1 : M_{4.1} \neq M_{4.3} \end{cases}$$

Com base no teste de hipótese conduzido, com 5% de significância, existem evidências amostrais para rejeitar a hipótese nula, isto é, adotar a estrutura AR(1) para os erros - $\chi_1^2 = 6,19$ e $p\text{-valor} = 0,0128$.

Como informação adicional a escolha da matriz de covariância, os gráficos de autocorrelação (Figuras 4a e 4b) reforçam a escolha da estrutura AR(1), pois todas as autocorrelações estão contidas dentro do intervalo de confiança. Corroborando com o que é proposto por Pinheiro e Bates (2000).

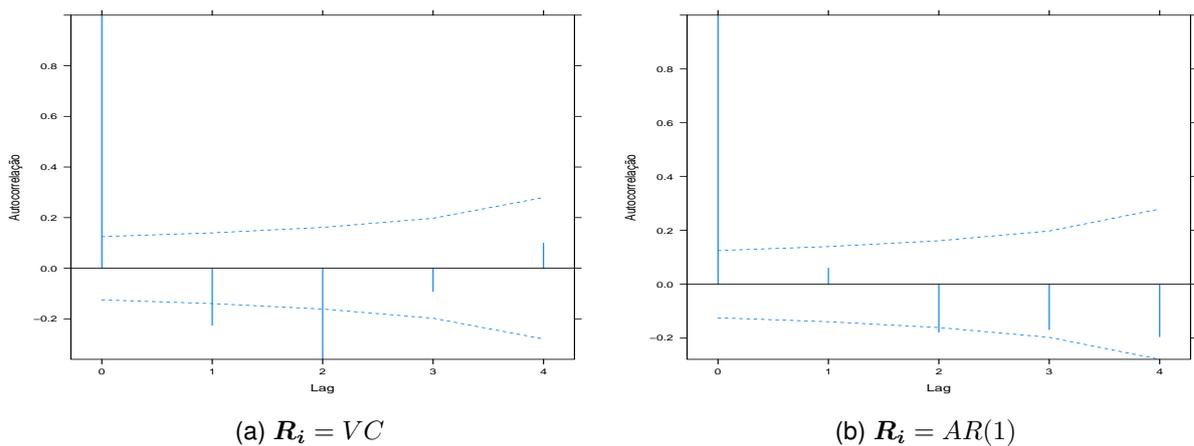


Figura 4 – Gráficos de autocorrelação para modelos com estruturas VC e AR(1) para os erros.

Baseando-se nos resultados até aqui obtidos, conclui-se que o modelo sem a interação entre os efeitos principais, com efeito aleatório no intercepto e inclinação ($G=UN$),

matriz de variância e covariância AR(1) para os erros é adequado para modelagem dos dados, em que as expressões para $Var(y_{ijk})$ e $Cov(y_{ijk}, y_{ijn})$, com $k \neq n$ são dadas por

$$Var(y_{ijk}) = \sigma_{b_{0i}}^2 + 2t_{ik}\sigma_{b_{0i}b_{1i}} + t_{ik}^2\sigma_{b_{1i}}^2 + \sigma^2, \quad (1.13)$$

$$Cov(y_{ijk}, y_{ijn}) = \sigma_{b_{0i}}^2 + t_{ik}t_{in}\sigma_{b_{1i}}^2 + (t_{ik} + t_{in})\sigma_{b_{0i}b_{1i}} + \sigma^2\rho^{|k-n|}, \quad (1.14)$$

sendo que ambas as expressões dependem do tempo, ou seja, do momento de avaliação.

O modelo final é dado por

$$y_{ijk} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t_{ik} + \alpha_j + \epsilon_{ijk}. \quad (1.15)$$

Pelo teste dos efeitos fixos verificou-se que existe diferença entre os tratamentos ($F_{6,78} = 10,84, p < 0,001$), da mesma maneira que entre o desempenho dos ratos ao longo dos testes de memória ($F_{1,339} = 8,57, p = 0,0036$).

A partir das estimativas de MVR dos parâmetros do modelo final (Tabela 2), dado que o grupo Sham foi adotado como referência, verifica-se que à medida que os dias passam, a latência tende a diminuir em 4,11 unidades em relação ao valor observado inicialmente. A julgar pelas estimativas dos parâmetros, a estimativa relativa ao tratamento com OP 4 horas possui valor inferior ao seu erro-padrão. Tal ocorrência pode estar associada a presença de pontos influentes, isto é, unidades experimentais que possam estar exercendo influência desproporcional nas estimativas e erros-padrão dos seus respectivos parâmetros, assim como pela presença de *outliers*.

Tabela 2 – Estimativas de MVR, erro-padrão, *p*-valor e intervalo de 95% de confiança para os parâmetros do modelo

Efeito	Parâmetro	Est	Ep	p-valor	Ic (95%)	
					Li	Ls
Intercepto	β_0	80,84	9,67	<0,001	61,81	99,87
OP4	α_2	7,12	12,13	0,5590	-17,04	31,28
OP8	α_3	20,47	12,79	0,1136	-4,99	45,95
OP12	α_4	36,32	12,79	0,0058	10,84	61,80
V4	α_5	75,73	13,78	<0,001	48,28	103,18
V8	α_6	61,42	13,78	<0,001	33,97	88,87
V12	α_7	81,39	14,78	<0,001	51,96	110,83
Dia (TM)	β_1	-4,11	1,40	0,0036	-6,88	-1,35
Componentes de Variância						
	$\sigma_{b_{0i}}^2$	1165,20			1098,26	1232
	$\sigma_{b_{1i}}^2$	42,85			29,83	55,46
	$\sigma_{b_{0i}b_{1i}}$	-124,82			-209,72	-39,74
	σ^2	1071,63			1007,49	1135,78
	ρ	0,2980			0,0199	0,5334

Os modelos marginais ajustados para o tratamento com OP são dados por:

$$\begin{aligned} E(y_{i2k}) &= 87,96 - 4,11t_{ik} \quad \text{se } OP_4; \\ E(y_{i3k}) &= 101,31 - 4,11t_{ik} \quad \text{se } OP_8; \\ E(y_{i3k}) &= 117,16 - 4,11t_{ik} \quad \text{se } OP_{12}. \end{aligned}$$

O modelo condicional para o primeiro animal tratado com OP_4 , 16º animal do experimento, é dado por

$$E(y_{12k}) = (87,96 - 11,28) - (3,28 + 4,11)t_{1k}, \quad (1.16)$$

sendo que o mesmo pode ser feito para os demais 84 animais.

O primeiro bloco da matriz diagonal V , V_1 , cujos valores são obtidos por meio das expressões 1.13 e 1.14, é escrito como

$$V_1 = \begin{bmatrix} 2030,0 & 1195,8 & 889,6 & 740,8 & 638,9 \\ 1195,8 & 1908,9 & 1117,6 & 854,3 & 748,3 \\ 889,6 & 1117,6 & 1873,5 & 1125,1 & 904,6 \\ 740,8 & 854,3 & 1125,1 & 1923,8 & 1218,2 \\ 638,9 & 748,3 & 904,6 & 1218,2 & 2059,9 \end{bmatrix}$$

1.4.3 Análise de Diagnóstico para o modelo ajustado

Pela Figura 5a é possível perceber que há observações localizadas na região esquerda do gráfico concentradas em torno de zero e outras mais dispersas a direita, bem como observações que estão fora dos limites, umas mais próximas, outras mais distantes. Em especial, destacam-se as observações das unidades experimentais 39, 43, 51, 57, 80. Todavia, não existem evidências para constatar-se a presença de heterocedasticidade para os resíduos condicionais padronizados. Por intermédio da Figura 5b observa-se que a suposição de normalidade dos mesmos é atendida.

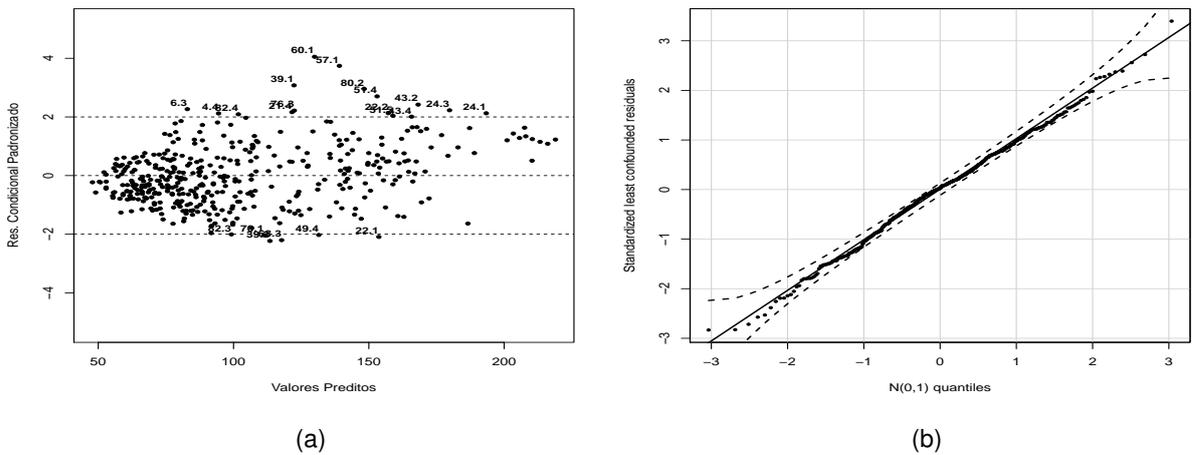


Figura 5 – Resíduos condicionais padronizados e envelope simulado com 95% para o RCCM.

Analisando a Figura 6a, é possível identificar observações discrepantes, que em sua maioria pertencem ao grupo veículo. Pela Figura 6b, para as observações fora dos limites a estrutura AR(1) não captou adequadamente a variabilidade, sendo que algumas também foram identificadas na Figura 6a. Tais unidades são pontos discrepantes e a estrutura de covariância intraunidade amostral adotada não acomodou satisfatoriamente a variação individual destas observações.

Considerando todas as observações identificadas nas Figuras 6a e 6b 70% pertencem aos grupos Veículo, 25% pertencem aos grupos Óleo de Peixe e apenas 5% ao grupo Sham. Os comportamentos atípicos para o tratamento com veículo são esperados, uma vez que animais deste grupo não recebem tratamento. Contudo, para ratos tratados com Óleo de Peixe isso indica que os mesmos encontraram dificuldade para realizar a tarefa.

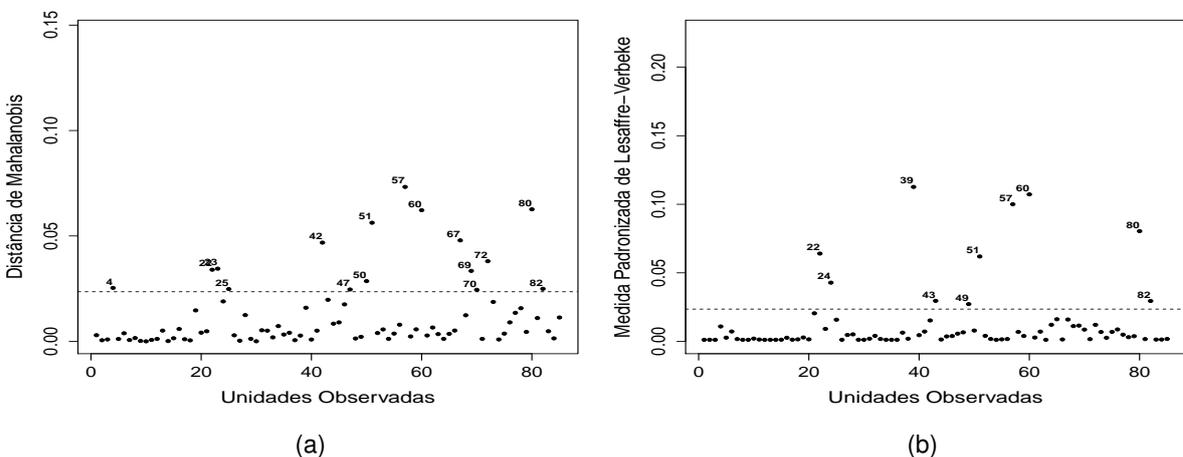


Figura 6 – Distância de Mahalanobis e Distância de Lesaffre-Verbeke Padronizada.

Avaliando-se os gráficos da DCC (Figuras 7a e 7b) identifica-se que há um grande número de observações em torno de zero, no entanto, algumas fora, entre as quais quatro distanciam-se significativamente do limite estabelecido, que foram consideradas com comportamento atípico na Figura 5a.

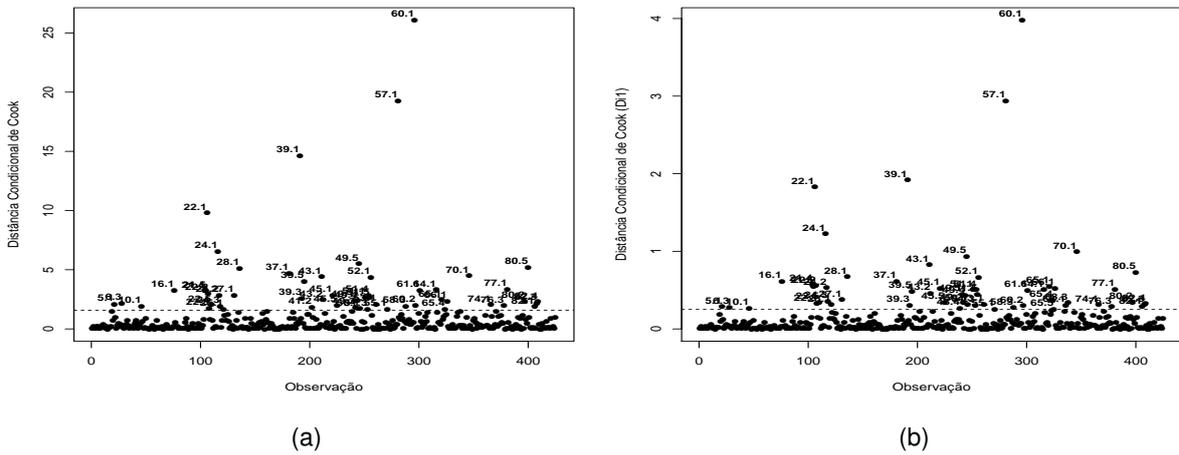


Figura 7 – Distância de Cook e 1º termo da decomposição da DCC para o modelo ajustado.

As unidades observadas nas componentes da decomposição da DCC (Figuras 8a e 8b) são as mesmas, em geral, identificadas nas Figuras 7a e 7b. Assim, tais observações são candidatas a pontos influentes, isto é, podem estar influenciando de maneira desproporcional nas estimativas dos efeitos fixos e aleatórios do modelo. Todavia, essas unidades que destoam das demais pertencem aos grupos Veículo, em sua maioria, o que justifica o comportamento, conforme descrito na seção 1.2.3.

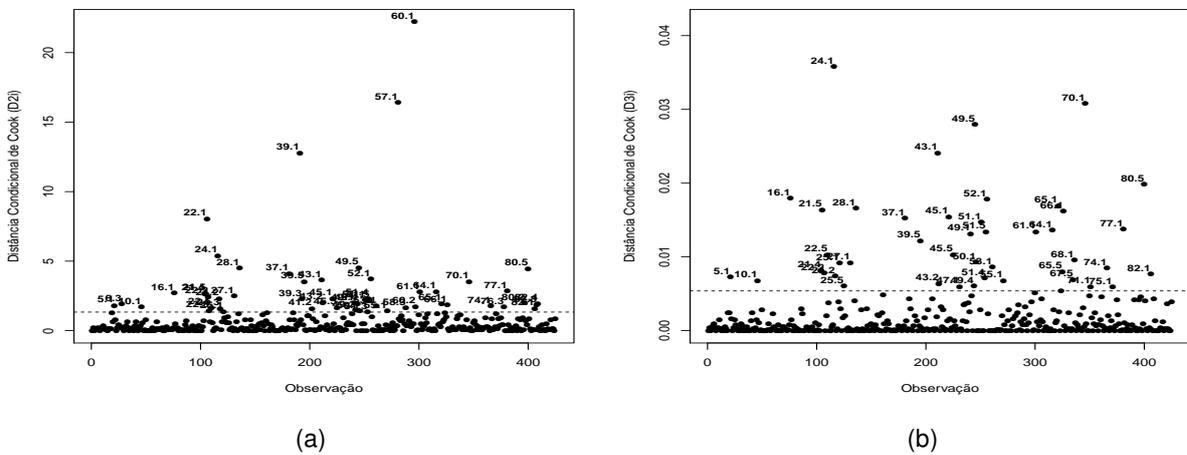


Figura 8 – Segundo e terceiro termos da Decomposição da DCC para o modelo ajustado.

Pela análise da Razão de Covariância ($CovRatio(\beta)$, $CovRatio(\theta)$) constatou-se que na maioria os pontos identificados nas Figuras 7a, 7b, 8a e 8b são os mesmos detectados nas Figuras 9a e 9b. Em suma, estas observações podem estar exercendo influência e alterando a precisão das estimativas dos parâmetros do modelo.

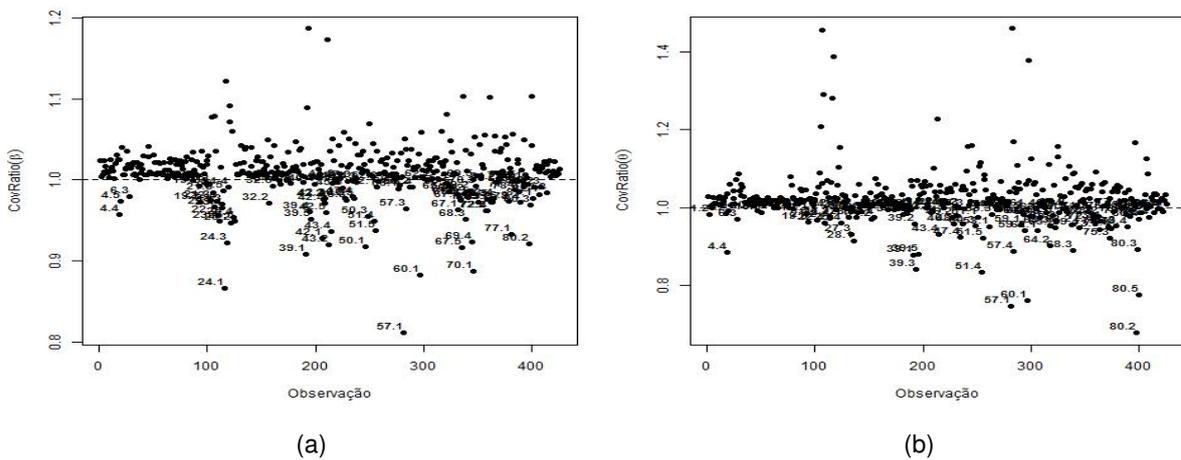


Figura 9 – Razão de Covariância para β e θ .

Das observações com características influentes identificadas, em geral, aproximadamente 46,87%, 10,9%, 12,5%, 9,4%, 20,3% são observações do 1º, 2º, 3º, 4º e 5º dias de TM respectivamente. O fato da maior presença de observações ser do primeiro TM justifica a possível influência, uma vez que essa é a primeira avaliação pós-isquemia. Pontos de alavanca não foram identificados no conjunto de dados.

Quando essas observações foram excluídas coletivamente, as maiores influências foram exercidas sobre as estimativas dos parâmetros relativos aos tratamentos OP4 e OP8. Essa exclusão implicou na redução de aproximadamente 34% cada uma das estimativas. Já seus erros-padrão, proporcionalmente em relação as estimativas iniciais não foram alterados significativamente. Para as demais estimativas, intercepto, dia, OP12, V4, V8 e V12, foram reduzidas em 0,01%, 5,78%, 6,84% e 11,3% respectivamente. Para influência sobre os componentes de variância, a maior redução foi em relação a estimativa de $\sigma_{b_{0i}}^2$, aproximadamente 34,71%. Com base nesses resultados, evidencia-se a influência das observações detectadas.

Tendo em vista que as observações detectadas são pontos influentes, após o reajuste do modelo com a omissão delas, as decisões obtidas com os dados completos não foram alteradas. Quando analisado o grupo o qual pertenciam (V4, V8, V12, OP8 e OP12), o comportamento atípico e influente é considerado normal, visto que o desempenho de cada animal é intrínseco a sua natureza biológica. Assim, as mesmas devem permanecer no conjunto de dados.

De um modo geral, pode-se dizer que o modelo misto gaussiano adotado mostrou-se adequado para responder os objetivos do experimento estudado, uma vez que os resultados estatísticos corroboram com as evidências biológicas. E ainda, verificou-se que quando confrontados com o grupo 'Sham' os tratamentos com Veículo 4, 8 e 12 horas pós-isquemia apresentaram uma latência superior, indicando que os animais desse grupo esqueceram da tarefa realizada pré-isquemia, conforme os resultados da Tabela 2.

No entanto, a defasagem cognitiva foi reduzida significativamente pelos tratamentos com OP 4, 8 e 12 horas pós-isquemia implicando parcialmente que a memória foi restabelecida, de acordo com suas limitações ($OP4$ vs $V4$ - $t=-5,04$; $OP8$ vs $V8$ $t=-2,88$; $OP12$ vs $V12$ $t=-2,97$; $p < 0,05$). Da mesma forma, animais tratados com OP 4 e 8 horas pós-isquemia foram capazes de lembrar das tarefas realizadas antes da intervenção ($OP4$ vs 'Sham'; $OP8$ vs 'Sham' ; $p > 0,05$). Embora o desempenho na tarefa do LRA dos animais do tratamento com OP12 tenha sido superior em relação aos animais do tratamento V12, o déficit de memória não foi aliviado de modo que seu desempenho se equiparasse ao dos animais do grupo Sham ($OP12$ vs 'Sham'; $p < 0,05$).

1.5 Considerações Finais

O propósito desse estudo foi utilizar a metodologia dos modelos mistos gaussianos para análise de dados longitudinais de um estudo farmacológico, cujo objetivo foi de aferir a eficácia do tratamento com óleo de peixe para amenizar os efeitos da isquemia cerebral em ratos, uma vez que, modelos animais que reproduzem tanto as doenças quanto as sequelas das mesmas nos seres humanos representam o ponto de partida para a investigação destes tratamentos.

Em suma, pressupostos de homocedasticidade e normalidade para os resíduos não foram violados. Mesmo que indivíduos com comportamentos atípicos e influentes tenham sido identificados, as análises estatísticas residuais e de influência para o modelo foram satisfatórias, uma vez que pressupostos básicos foram verificados. Constatou-se que as estruturas de covariância para os efeitos aleatórios e erros (UN e AR(1)) acomodaram de maneira adequada a variabilidade e a dependência entre e intra-indivíduos.

Os modelos mistos empregados neste estudo evidenciaram que o tratamento com óleo de peixe é eficaz quando iniciado 4 ou 8 horas pós-isquemia. Essa eficácia foi reduzida, contudo, quando o início do tratamento foi 12 horas pós-isquemia. Uma melhora clínica é observada com maior intensidade quando o tratamento foi iniciado 4 horas pós-isquemia, indicando que quanto mais rápido se inicia o tratamento, melhor será a eficácia anti-ansiolítica do OP.

1.6 Sugestões de Pesquisa

Com o intuito de melhorar o ajuste do modelo, recomenda-se:

- Avaliar a variável resposta latência considerando outras distribuições como Gama e Inversa Gaussiana utilizando a metodologia dos Modelos Mistos Generalizados;
- Fazer análise de diagnóstico para o modelo proposto no item anterior com base no trabalho de Pinho, Nobre e Singer (2015);
- Modelar as demais variáveis resposta (erro de referência diário e erro de trabalho diário) com possíveis modelos longitudinais binomial, Poisson e Poisson Inflacionado de Zero.

Referências

- ANDERSON, C. A.; ARCINIEGAS, D. B. **Cognitive sequelae of hypoxic-ischemic brain injury**: a review. *NeuroRehabilitation*, v. 26, p. 47–63, 2010.
- BACARIN, C.; Sá-NAKANISHI, A. B.; BRACHT, A.; MATSUSHITA, M.; PREVIDELLI, IS; MORI, M.; MARIA, R.; OLIVEIRA, W.; H, M. **Fish Oil Prevents Oxidative Stress and Exerts Sustained Antiamnesic Effect After Global Cerebral Ischemia**. *CNS & Neurological Disorders*, v. 14, p. 400 – 410, 2015.
- BRASIL, P. **População deve ficar atenta aos riscos do AVC**. 2014. Disponível em: <<http://www.brasil.gov.br/saude/2012/10/no-dia-mundial-do-avc-pais-alerta-populacao-contr-a-doenca>>. Acesso em: 20 abr. 2015.
- CARLIN, B. P.; LOUIS, T. A. **Bayesian Methods for Data Analysis**. Chapman & Hall / CRC Press, 2009.
- COOK, R. D. **Detection of influential observation in linear regression**. *Technometrics*, v. 39, p. 15–18, 1977.
- DHUNGANA, H.; MALM, T.; DENES, A.; VALONEN, P.; WOJCIECHOWSKI, S.; MAGGA, J.; SAVCHENKO, E.; HUMPHREYS, N.; GRENCIS, R.; ROTHWELL, N.; KOISTINAHO, J. **Aging aggravates ischemic stroke-induced brain damage in mice with chronic peripheral infection**. *Aging Cell*, v. 12, n. 5, p. 842–850, 2013.
- DIGGLE, P. J. **An approach to the analysis of repeated measurements**. *Biometrics*, v. 44, p. 959 – 971, 1988.
- DIGGLE, P. J.; HEAGERTY, P.; LIANG, K.-Y.; ZEGER, S. **Analysis of longitudinal data**. Oxford University Press, p. 396, 2002.
- DZIALOWSKI, I.; KLOTZ, E.; GOERICKE, S.; DOERFLER, A.; FORSTING, M.; KUMMER, R. von. **Ischemic Brain Tissue Water Content**: Ct monitoring during middle cerebral artery occlusion and reperfusion in rats. *Radiology*, v. 243, n. 3, p. 720–726, 2007.
- FAN, Y.; QIN, G.; ZHU, Z. Y. **Robust Variable Selection in Linear Mixed Models**. *Communications in Statistics - Theory and Methods*, v. 43, n. 21, p. 4566–4581, 2014.
- FAUSTO, M. A.; CARNEIRO, M.; ANTUNES, C. M. d. F.; PINTO, J. A.; COLOSIMO, E. A. **O modelo de Regressão Linear misto**: Uma aplicação a dados antropométricos

desbalanceados. *Caderno de Saúde Pública, Rio de Janeiro*, v. 24, n. 3, p. 513 – 524, 2008.

FERREIRA, E. D.; ROMANINI, C. V.; OLIVEIRA, J. N.; PREVIDELLI, I. T.; MELO, S. R.; OLIVEIRA, R. M.; MILANI, H. **Fish oil provides a sustained anti-amnesic effect after acute, transient forebrain ischemia but not after chronic cerebral hypoperfusion in middle-aged rats**. *Behavioural Brain Research*, v. 26, p. 101 – 110, 2014.

FITZMAURICE, G.; LAIRD, N. M.; WARE, J. H. **Applied Longitudinal Analysis**. Boston - Massachusetts, 2004.

GODINHO, J. et al. **Cilostazol but not sildenafil prevents memory impairment after chronic cerebral hypoperfusion in middle-aged rats**. *Behavioural Brain Research*, p. 61–68, 2015.

GURKA, M. J. **Selecting the Best Linear Mixed Model under REML**. *The American Statistician*, Taylor & Francis, Ltd. on behalf of the American Statistical Association, v. 60, n. 1, p. 19–26, 2006.

HARVILLE, D. **Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects**. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 4, n. 2, p. 384–395, 1976.

HEIKKINEN, R.; MALM, T.; HEIKKIL, J.; MUONA, A.; HEIKKI, T.; KOISTINAHO, M.; KOISTINAHO, J. **Susceptibility to Focal and Global Brain Ischemia of Alzheimer Mice Displaying A β Deposits: effect of immunoglobulin**. *Aging and Disease*, p. 76–87, 2014.

HILDEN-MILTON, J. A. **Multilevel diagnostics for mixed and hierarchical linear models**. 125 p. Tese (Thesis (Ph.D.) — University of California, 1995.

HOLLIDAY, E. G.; TRAYLOR, M.; MALIK, R.; BEVAN, S.; FALCONE, G.; HOPEWELL, J. C.; CHENG, Y.-C.; COTLARCIUC, I.; BIS, J. C.; BOERWINKLE, E.; BONCORAGLIO, G. B.; CLARKE, R.; COLE, J. W.; FORNAGE, M.; FURIE, K. L.; IKRAM, M. A.; JANNES, J.; KITTNER, S. J.; LINCZ, L. F.; MAGUIRE, J. M.; MESCHIA, J. F.; MOSLEY, T. H.; NALLS, M. A.; OLDMEADOW, C.; PARATI, E. A.; PSATY, B. M.; ROTHWELL, P. M.; SESHADRI, S.; SCOTT, R. J.; SHARMA, P.; SUDLOW, C.; WIGGINS, K. L.; WORRALL, B. B.; ROSAND, J.; MITCHELL, B. D.; DICHGANS, M.; MARKUS, H. S.; LEVI, C.; ATTIA, J.; WRAY, N. R. **Genetic Overlap Between Diagnostic Subtypes of Ischemic Stroke**. *Stroke*, v. 46, n. 3, p. 615–619, 2015.

IBRAHIM, J. G.; ZHU, H.; GARCIA, R. I.; GUO, R. **Fixed and Random Effects Selection in Mixed Effects Models**. *Biometrics*, Blackwell Publishing Inc, v. 67, n. 2, p. 495–503, 2011.

LAIRD, N. M.; WAIRE, J. H. **Random-Effects Models for Longitudinal Data**. *Biometrics*, v. 38, n. 4, p. 513 – 524, 1982.

LIM, C.; ALEXANDER, M. P.; LAFLECHE, G.; SCHNYER; VERFAELLIE, M. **The neurological and cognitive sequelae of cardiac arrest**. *Neurology*, 2004.

LITTEL, R. C.; STROUP, W. W.; WOLFINGER, R. D.; SCHABENBERGER, O. **SAS for Mixed Models**. Cary: Statistics Analysis System Institute, p. 834, 2006.

- MOLENBERGHS, G.; VERBEKE, G. **Models for Discrete Longitudinal Data**. New York:Springer, p. 683, 2005.
- MULLER, S.; SCEALY, J. L.; WELSH, A. H. **Model Selection in Linear Mixed Models**. *Statist. Sci.*, The Institute of Mathematical Statistics, v. 28, n. 2, p. 135–167, 2013.
- NOBRE, J. S. **Métodos de Diagnóstico para Modelos Lineares Mistos**. Dissertação (Mestrado em Estatística) — Universidade de São Paulo, São Paulo, 2004.
- NOBRE, J. S.; SINGER, J. M. **Residual Analysis for Linear Mixed Models**. *Biometrical Journal*, v. 49, p. 863 – 875, 2007.
- NOBRE, J. S.; SINGER, J. M. **Leverage analysis for linear mixed models**. *Journal of Applied Statistics*, v. 38, n. 5, p. 1063–1072, 2011.
- PAGANELLI, R. A.; BENETOLLI, A.; LIMA, K. C. M.; CESTARI-JUNIOR, L. A.; FILHO, L. A. F.; MILANI, H. A novel version of the 8-arm radial maze: effects of cerebral ischemia on learning and memory. *Journal of Neuroscience Methods*, v. 132, n. 1, p. 9 – 18, 2004.
- PERRI, S. H. V.; IEMMA, A. F. **Procedimento "MIXED" do SAS para análise de modelos mistos**. *Scientia Agrícola*, scielo, v. 56, p. 959 – 967, 1999.
- PINHEIRO, J. C.; BATES, D. M. **Mixed Effects in S and S-plus**. [S.l.]: New York, Springer, 2000.
- PINHO, L. G. B.; NOBRE, J. S.; SINGER, J. M. **Cook's distance for generalized linear mixed models**. *Computational Statistics & Data Analysis*, v. 82, p. 126 – 136, 2015.
- PU, W.; NIU, X.-F. **Selecting mixed-effects models based on a generalized information criterion**. *Journal of Multivariate Analysis*, v. 97, n. 3, p. 733 – 758, 2006.
- RAO, R.; WU, Y. **A strongly consistent procedure for model selection in a regression problem**. *Biometrika*, v. 76, n. 2, p. 369–374, 1989.
- RESENDE, M. D. V. et al. **Análise de Modelos Lineares Mistos Via Inferência Bayesiana**. *Rev. Mat. Estat.*, p. 41–70, 2001.
- RIZZATO, F. B. **Modelos para análise de dados discretos longitudinais com superdispersão**. 145 p. Tese (Doutorado em Estatística e Experimentação Agrícola) — Escola Superior de Agricultura "Luiz Queiroz", Universidade de São Paulo, São Paulo, 2011.
- SCHABENBERGER, O. **Mixed Model Influence Diagnostics**. SAS Institute Inc, Cary, NC, 2004. Disponível em: <<http://www2.sas.com/proceedings/sugi29/189-29.pdf>>. Acesso em: 23 mai. 2015.
- SELF, S. G.; LIANG, K.-Y. **Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions**. *Journal of the American Statistical Association*, v. 82, n. 398, 1987.
- SINGER, J. M.; NOBRE, J. S.; ROCHA, F. M. M. **Análise de Dados Longitudinais: Versão preliminar**. São Paulo: IME - USP, 2015. Disponível em: <<http://www.ime.usp.br/~jmsinger/MAE0610/Singer&Rocha&Nobre2015fev.pdf>>. Acesso em: 20 fev. 2015.

- SRIVASTAVA, M. S.; KUBOKAWA, T. **Conditional information criteria for selecting variables in linear mixed models**. *Journal of Multivariate Analysis*, v. 101, n. 9, p. 1970 – 1980, 2010.
- SUSSMAN, C. B.; ROSSIGNOL, C.; ZHANG, Q.; JIANG, H.; ZHENG, T.; STEINDLER, D.; YOUNG, L.; WEISS, M. **Photoacoustic tomography can detect cerebral hemodynamic alterations in a neonatal rodent model of hypoxia-ischemia**. *Acta Neurobiologiae Experimentalis*, p. 253 – 263, 2012.
- TAN, F. E. S.; OUWENS, M. J. N.; BERGER, M. P. F. **Detection of influential observations in longitudinal mixed effects, regression models**. *The Statistician*, v. 50, p. 271–284, 2001.
- VERBEKE, G.; MOLENBERGHS, G. **Linear Mixed Models for Longitudinal Data**. New York:Springer, p. 569, 2000.
- WANG, S.; WU, E.; CAI, K.; LAU, H.; CHEUNG, P. L. K. **Mild Hypoxic-Ischemic Injury in the Neonatal Rat Brain**:longitudinal evaluation of white matter using diffusion tensor mr imaging. *American Of Neurology*, p. 1907 – 1913, 2009.
- WATERNAUX, C.; LAIRD, N. M.; WARE, J. H. **Methods for analysis of longitudinal data: Blood-lead concentrations and cognitive development**. *J. Am. Stat. Assoc.*, v. 84, p. 33–41, 1999.
- WEST, B. T.; WELCH, K. B.; GALECKI, A. T. **Linear mixed models: a practical guide using statistical software**. 2. ed. [S.l.]: Chapman & Hall/CRC, Boca Raton, 2014. 337 p.