RESEARCH ARTICLE

Repeated Measures Models Applied to Cancer Patients Treated with Exergames

Isabela Pagani Heringer de Miranda[1], Denismar Alves Nogueira[1,2]*, Natalia da Silva Martins[1,2], Ricardo da Silva Alves[3], Leonardo Cesar Carvalho[3,4]

Abstract

Objective: The objective of this study was to define an appropriate linear model to analyse data on muscular fatigue in cancer patients over time through repeated measures techniques. **Methods:** Using the split plot in time system and linear mixed models, three groups of individuals were compared as to the methods used to reduce muscle fatigue. Group Cancer consisted of individuals who had already been treated; group Control consisted of healthy individuals and group Chemo / radio-therapy consisted of individuals diagnosed with cancer undergoing chemo and radiation therapy. Sessions were tested with exergames. A series of muscle strength data for each of the six muscles studied, in the pre-treatment, mid-treatment and final sessions. **Result:** The structure that best fit the covariance matrix was ARMA (1,1), according to AIC and BIC. There were significant differences and tendencies in the data series, especially for the left tibial muscle, in which the interactions between group and session and between group and time were significant, showing that exergames treatment increased muscle strength in debilitated patients and, with 20 sessions, the groups equalled in muscle strength. **Conclusion:** The linear mixed model proved to be efficient in modelling plots subdivided in time. Identifying the best structure of the covariance matrix allowed us to better estimate the effects, using tests appropriately to verify differences between factors that were not detected when using the median frequency of strength.

Keywords: Linear mixed model- split plot- analysis of variance

Introduction

In studies with repeated measures the difficulty is in modelling longitudinal data, since it is necessary to evaluate the changes that occur over time. It is expected that non-zero correlation occurs between measurements in time and that there is heterogeneity of variances, so it is possible to consider that the responses of closer times are more correlated than those of more distant times (Littell et al., 2006).

In order to model longitudinal data, different covariance matrix structures of the experimental error can be verified by searching, through some selection criterion, the one that best represents the data (Faraway, 2016).

The use of linear mixed models is a useful statistical procedure when the interest is in finding a longitudinal model that represents the data correctly, avoiding mistaken inferences.

Studies involving debilitated individuals, as is the case of individuals diagnosed with cancer, are usually studies of repeated measures (Hsiao, 2014), as these individuals are followed over a period of time and more than one measurement is performed during the study period.

Individuals diagnosed with cancer suffer from the side effects caused by the disease and the treatments applied. The most commonly reported side effect of cancer survivors has been muscle fatigue. Measuring fatigue in cancer patients is not an easy task, since the physical and psychological conditions of these patients are restricted (Andersen, et al., 2009; Prinsen et al, 2015; Da Silva Alves et al., 2017).

In recent years, the use of video games has been commercially used in the rehabilitation process (Parry et al., 2014). Some studies show that healthy individuals who underwent therapy with exergames, compared to the traditional method, had better physical fitness (Battaglini et al., 2006). Thus, it is believed that virtual reality therapy can present an efficient, pleasant and complementary means for the rehabilitation of physically unconditioned patients, as is the case of patients with cancer.

There are, in the literature, longitudinal studies involving individuals that, for data analysis purposes, consider the median or average frequency of the series of longitudinal measures. It is believed that fatigued patients generate series with trends for failing to maintain the strength in the evaluation period and that the use of median

[1]Applied Statistics and Biometrics Program, [2]Department of Statistic, [3]Physical Therapy Course, [4]Bioscience Program, Federal University of Alfenas, Alfenas, Brazil. *For Correspondence: denismar.nogueira@unifal-mg.edu.br

or average frequency does not portray this information.

The objective of this study was to define the appropriate linear model to analyse data on muscular fatigue in an experiment with cancer patients over time through repeated measures techniques.

Materials and Methods

The data used are from the study by Alves et al., (2017). This study follows the norms of Resolution 466/12 of the National Health Council and was approved on December, 03 2014, under the CAAE nº 38628314.3.0000.5142 by the Research Ethics Committee (CEP) of the Federal University of Alfenas-MG. It is registered in the Brazilian Registry of Clinical Trials under the number RBR-9t48g5.

Initially the study consisted of 135 individuals with ages ranging from 18 to 80, with 105 of them diagnosed with some type of cancer and 30 healthy volunteers.

Volunteers and patients with cognitive disorders that restrict the explanation of the handling of virtual environments and methods of evaluation, patients with severe infectious diseases, incapacitating diseases that restrict the movement of upper and lower limbs, patients with myopathies and diseases with a recognized alteration of collagen, with neurological abnormalities, those who, for various reasons, were unable to complete the three sessions of treatment, those who for personal reasons did not want to participate and those who refused to sign the Free and Informed Consent Form were excluded. For more details see Alves et al., (2017).

After exclusion of patients and volunteers, the study consisted of 19 individuals forming three groups: Group Cancer (G1): group composed of 7 patients diagnosed with stages 0, I, II and III, treated at Santa Casa de Alfenas-MG and at health centres in the city of Alfenas-MG, being monitored after chemo and radiation therapy; Group Control (G2): group consisting of 7 volunteers without diagnosis of cancer, living in the city of Alfenas - MG and Group Chemo and radiotherapy (G3): group composed of 5 patients with diagnosis of cancer in stages 0, I, II and III, who are undergoing chemotherapy at Santa Casa de Alfenas, MG, and radiotherapy in Poços de Caldas, MG.

The strength in the Gastrocnemius and Tibial muscles was evaluated in dynamometer and surface electromyography. Using the Trigno 8 Channel Wireless device (EMGworks, Delsys Inc., Boston) equipped with the software EMGworks 4.0 Acquisition® for the collection of electromyographic signals and the EMGworks 4.0 Analysis® for analysis, it was evaluated the force (in kgf) with that each participant can hold the heel lifted for about one minute (60 seconds) and the behaviour of that force, whether or not it is fatiguing in the evaluated period of time, generating a series of data in time.

Participants were verbally stimulated to maintain maximum contraction and after each attempt, five minutes of rest were allowed. For all the participants, three evaluations were carried out by the same examiner in three times, referring to the following moments: Pre-treatment (0 - no video game contact); half of the treatment (10th - after 10 sessions with video game) and post-treatment (20th - after 20 sessions with video game).

The muscle strength data sets of each individual in each session consisted of about 350 data. It was found that the initial and final 12 seconds were for rest and therefore excluded and 15 data were selected for analysis. This selection was made using the systematic sampling method, which consists of randomly selecting the first data to be considered and then selecting for the sample every umpteenth data in the sample mark.

The experimental model used was the split plot in time with factorial on the plot assuming a completely randomized design (Montgomery, 2008). The response variables were obtained by muscle strength in each muscle measured in the study - Right Medial Gastrocnemius (RMG), Left Medial Gastrocnemius (LMG), Right Lateral Gastrocnemius (RLG), Left Lateral Gastrocnemius (LLG), Right Tibial (RT) and Left Tibial (LT). The experiment consists of 57 plots (7 individuals measured in 3 sessions in G1 and G2 and 5 individuals measured in 3 sessions in G3), 855 subplots (15 times evaluated in each individual - series) and 2 factors (Group - 3 levels and Session - 3 levels) being individuals assumed as random effect. The linear mixed model is given by:

$$y_{ijkl} = \mu + G_i + S_k + GS_{ik} + I(GS)_{j(ik)} + T_l + GT_{il} + ST_{kl} + GST_{ikl} + \varepsilon_{ijkl}$$

in which, y_{ijkl} is the observed value of the muscular strength of the jth individual in the experimental plot that received the Ith level of factor Group, the kth level of factor Session and the lth level of factor Time; μ is a constant; G_i is the effect of the Ith level of factor Group, with $i = 1,...,3$; S_k is the effect of the kth level of factor Session, $k=1,...,3$; GS_{ik} is the effect of the interaction between the ith level of factor Group and the kth level of factor Session; $I(GS)_{j(ik)}$ is the effect of the jth individual within the interaction between Group and Session, with $j=1,...7$; T_l is the effect of the lth level of factor Time, with $l=1,...,15$; GT_{il} is the effect of the interaction between the ith level of factor Group and the lth level of factor Time; ST_{kl} is the effect of the interaction between the kth level of factor Session and the lth level of factor Time; GST_{ikl} is the effect of triple interaction between the ith level of factor Group, kth level of factor Session and the lth level of factor Time and ε_{ijkl} is the general error of the model. The nested term of the Individual within the Group x Session interaction consists of the Error relating to the plot factors. The groups and sessions were considered independent, so the covariance structure was the identity.

Based on split plot design over time and linear mixed models (Henderson, 1975), it can be written in the matrix form $y = X\beta + Zu + \varepsilon$ where y is the vector of observations; X is the matrix of experimental planning for fixed effects (group, session, time, double interactions GS, GT, ST and triple interaction GST); β is the vector of fixed, unknown parameters; Z is the experimental planning matrix for random effects (individual within GS interaction); u is the vector of random, unknown effects and ε is the vector of unobservable random errors.

The random error vectors u and ε are assumed to be distributed as N (0, G) and N (0, R), respectively. In fact,

component G is a scalar that represents the variance of the random effects. On the other hand, R is the residual covariance matrix.

For the results obtained of an ANOVA (split-plot in time) to be valid, the covariance matrix of the experimental error (R) must satisfy the condition of Huynh and Feldt (1970), that is, it must be spherical. The Mauchly sphericity test (Mauchly, 1940) was used to verify that matrix R meets this condition. If not, other structures are tested for the covariance matrix using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). If AIC and BIC do not agree, evidence-strength measures should be used to facilitate choice (Burnham and Anderson, 2004; Silva et al., 2015). The strength of evidence of a model, calculated in proportion to the complete set of candidate models, receives the name of model weight. The interpretation of the weights is direct: indicate the probability that the model is the best among the set of models tested. The weights of those models in which AIC and BIC did not agree were calculated and the model chosen was the one that presented the highest weight, that is, more likely to be the best model among the candidates.

Considering the adequate structure of matrix R, the Analysis of Variance was performed according to the linear mixed model, in which the individuals (repetition) were considered random (LIU et al., 2007).

The software used was R (R Core Team, 2016) and the procedures PROC GLM, PROC REG and PROC MIXED of SAS (System Analysis Statistical Institute - SAS Institute, 2004).

All structures are present in the SAS PROC MIXED and can be verified in Jennrich and Schluchter (1986) and Wolfinger (1996).

The graphs were constructed in the R software, plotting the observed values versus the wastes and the muscular force versus the measurement of the time. The sphericity tests were performed using the command REPEATED of the PROC GLM procedure of the SAS software. This command is also used to make corrections of degrees of freedom.

The normality and Tukey tests and the selection of the best structure for the covariance matrix were performed using the PROC MIXED of SAS. This procedure allows the separation, in the model, of fixed effects and random effects. Regression analysis was performed using the SAS PROC REG procedure.

Results

Some socio demographic and clinical characteristics of the study participants can be seen in Table 1.

Analysing the results of the Mauchly sphericity test for the muscles of the study, it is verified that in all six cases, the sphericity hypothesis of the covariance matrix was rejected with p-value less than 1%. The structure that best fit the data was the Autoregressive and 1st Order Moving Averages (ARMA (1,1)). The structures tested and the AIC and BIC estimates for muscle data are presented in Table 2.

The results presented in Table 2 show that for the RLG muscle only the Group was significant, that is, at a 5% significance level, muscle strength does not change over time, nor with the increase in the number of treatment sessions, but there is evidence that at least one group differs from the others.

For the RMG muscle, the significant factors were Session and Time at the 5% level of significance, indicating that, on average, there are differences between muscle strength with increasing number of sessions and over time. The factor Group was not significant, so even if there are differences between sessions and time, the groups are statistically the same, i.e., the muscular strength in the

Table 1. Socio Demographic and Clinical Characteristics of the Study Participants

Characteristics (Means ± SD)	G1	G2	G3
Age (years)	63.29±7.34	56.73±11.94	57.13±16.74
BMI (kg/m²)	24.72±4.57	26.24±4.67	24.81±4.63
Practice of physical activity (week)	1.21±1.76	1.80±2.11	1.00±1.85
Gender (n)			
F	2	3	1
M	5	4	4
Diagnosis of Cancer (%)			
Gastrointestinal tract	6.66		7.14
Breast	20.00	-	42.86
Abdominopelvic	20.00	-	35.71
Oropharynx	26.67	-	0.00
Others	26.67	-	14.29
Stage (%)			
0	8.33	-	14.29
I	33.33	-	21.43
II	50.00	-	35.71
III	8.33	-	28.57

BMI, Body mass index; F, Female; M, Male; Other, Leukaemia, Lymphoma; Bone Cancer, Brain and Chronic Myelogenous Lymphoma.

Table 2. ANOVA Results for the Six Muscles of the Study and Structures Tested and Estimates of AIC and BIC

Source of Variation	DF	RLG p-value	RMG p-value	LLG p-value	LMG p-value	RT p-value	LT p-value
Group (G)	2	0.0305*	0.1669	0.2179	0.7193	0.0480*	0.0105*
Session (S)	2	0.1461	0.0183*	0.4299	0.0431*	0.0520	0.2190
GxS	4	0.7748	0.2948	0.2202	0.0734	0.2621	0.0303
Time (T)	14	0.9456	0.0426*	0.0918	0.0544	0.0002**	0.3181
GxT	28	0.5155	0.6708	0.3149	0.5747	0.3678	0.0107*
SxT	28	0.6705	0.7245	0.9259	0.6568	0.2025	0.3352
GxSxT	56	0.9042	0.2919	0.2230	0.1124	0.6451	0.2619
Criterion Structures							
	TOEP	7020.5	7150.3	7067.1	7150.3	6610.2	6563.0
AIC	ARMA(1,1)	7025.6	7136.3	7075.5	7060.4	6603.2	6554.0
	FA1(2)	7028.3	7185.7	7069.3	7065.7	6600.5	6600.1
BIC	ARMA(1,1)	7031.7	7142.4	7081.6	7066.5	6609.3	6560.1

*, Significant at 5% probability; **, significant at 1% probability

individuals of the three groups does not differ.

The results show that none of the factors under study was significant for the LLG muscle, that is, there is no evidence that the treatment significantly alter muscle strength of the individuals for this muscle.

For the LMG muscle, the Session was significant at the 5% level, indicating that there were differences in muscular strength of the individuals with the increase in the number of treatment sessions.

For the RT muscle, the factors Group and Time were significant with p values of 0.0480 and 0.0002, respectively. There is no evidence that muscle strength changes with increasing treatment sessions. You can study the differences between groups without taking into account the session or the time, because the interactions between these factors are not significant, indicating that the effects of factor Group occur independently of the effects of factors Session and Time. In the same way, one can study time without taking into account group or session.

Finally, the results for the LT muscle show significant interactions. At 5% level of significance, the interactions between Group and Session (GS) and Group and Time (GT) are significant, indicating that the effects of group occur in a way dependent on the effects of session and time, so we cannot study one factor without considering the other. With this, it is necessary to perform the unfolding of one factor within the other.

The Tukey - Kramer tests comparing the significant simple factors appropriate for each muscle are presented in Table 3.

The Tukey - Kramer test for the groups, using the averages of the RLG muscle data, defines that the group composed of individuals diagnosed with cancer (G1) who have already been treated with chemo and radiotherapy have, on average, muscle strength equal to those individuals of group Control (G2), who are the healthy individuals. The group consisting of individuals diagnosed with cancer undergoing chemo and radiotherapy (G3) is statistically different from G1, but does not differ from G2.

The muscle strength of individuals in the RMG and LMG muscles is statistically the same in sessions 0 and 20 and in sessions 10 and 20. This strength is statistically

Table 3. Averages of Muscle Strength for Each Factor Followed by the Tukey Test

Muscle	Factors	Means	Test Result [1]
	Cancer	161.04	a
RLG	Control	149.96	ab
	Chemo/Radio	142.52	b
	Session 0	129.47	a
RMG	Session 10	82.84	b
	Session 20	101.03	ab
	Session 0	161.36	a
LMG	Session 10	136.27	b
	Session 20	155.46	ab
	Cancer	114.48	ab
RT	Control	117.60	a
	Chemo/Radio	100.42	b

[1]Means followed by the same letters do not differ significantly to the level of 5% probability based on Tukey-Kramer test.

Table 4. Group Split Analysis within the Three Sessions and Time within the Three Groups for Variable Left Tibia

Source of Variation	DF	F test	p-value
Group: S0	2	5.44	0.0074**
Group: S10	2	5.21	0.0089**
Group: S20	2	0.22	0.8003
Residual	48		
Time: G1	14	2.59	0.0012**
Time: G2	14	1.57	0.0837
Time: G3	14	0.62	0.8471
Residual	672		

**, significant at 1% probability

different between sessions 0 and 10. It is found that individuals lose strength from session 0 to session 10 but continuing the treatment applied, muscle strength significantly increases from session 10 to session 20.

The results for the RT muscle indicate that the mean muscle strength of individuals in G1 is statistically equal to the average strength of G2 individuals. G3 individuals, affected by cancer-fighting treatments, have mean muscle strength significantly lower than the mean strength of individuals in group Control. The adjusted regression equations of strength (Y) for factor Time (X) in the RMG and RT muscles were $y = 94.77+1.21\ x$ ($R^2 = 74.95$) and $y = 117.20 - 0.80\ x$ ($R^2 = 65.35$), respectively.

Analysing the equation for the RMG muscle, it is verified that the muscular strength is equal to 94.77 Kgf at time 0 (beginning) and this strength increases by 1.21 Kgf at each unit of time, that is, the individuals tend on average to increase muscle strength over time, this is a sign of improvement in fatigue. The model explains 74.95% of the variability of the data. However, for the RT muscle, muscle strength is equal to 117.20 Kgf at time 0 and this strength decreases by 0.80 Kgf at each unit of time, that is, individuals tend on average to fade over time. The model explains 63.35% of the variability of the data.

The analyses of the split between Group vs. Session and Time vs. Group for the LT muscle are presented in Table 4.

Analysing the results of the group split within the sessions, presented in Table 5, it is verified that for the strength data of the left tibial muscle, at least one group differs from the others in sessions 0 and 10, but in session 20 the groups are statistically the same. The

Tukey - Kramer test was used to compare groups within the sessions. Table 5 presents the test results.

The results of the Tukey - Kramer test for groups within session 0 show that there is no difference between G1 (group consisting of individuals with cancer) and G3 (group composed of individuals with cancer under treatment), but these groups differ from G2 (group composed of healthy individuals). The group composed of healthy individuals (G2), has on average greater muscle strength, and this strength is statistically superior to the muscular strength of the groups composed of individuals with cancer already treated or under treatment.

With 10 sessions of the treatment, there is no difference between G1 and G2, that is, the individuals in the group consisting of cancer patients who have already been treated with chemo and radiotherapy (G1), with 10 treatment sessions on average equal to the group composed of healthy individuals, recovering muscle strength. G3 remains different from G2 and now becomes different from G1 as well. The group composed of individuals with cancer under treatment (G3) continues to have the lowest average muscle strength. The results show that with 20 treatment sessions, the three groups can be considered statistically equal, that is, with 20 sessions of treatment, muscle strength, on average, in the three groups can be considered the same. The results from Table 5 for the time split within the groups show that factor time within G1 was significant at the 1% level of significance. The time within G2 and within G3 was not significant. The regression model of the force (Y) with respect to the time (X) within G1 was built, which

Table 5. Averages of the LT Muscle Strength for Each Group within Each Session Followed by the Tukey Test

Session	Groups	Means	Test Result[1]
	Cancer	94.69	a
S0	Control	123.50	b
	Chemo/Radio	82.95	a
	Cancer	111.97	a
S10	Control	99.09	a
	Chemo/Radio	70.12	b
	Cancer	111.21	a
S20	Control	103.44	a
	Chemo/Radio	105.52	a

[1]Means followed by the same letters do not differ significantly to the level of 5% probability based on Tukey-Kramer test.

is represented by the equation:

According to the simple regression model, for time within G1, muscle strength is equal to 104.00 Kgf at time 0 and this strength increases on average 0.25 Kgf at each unit of time. The model for time within G1, explains 3.13% of the variability of the data. This value may be small due to the amplitude of the data. Since the gradient of the regression equation in G1 is positive, we have a gain in muscle strength in individuals with cancer.

Discussion

Rejecting the sphericity hypothesis of the covariance matrix implies that the structure of the covariance matrix of the experimental error is different from the Component Variance (CV), Composite Symmetry (CS) or Huynh-Feld (HF) structures. If one of these structures were considered, the results of the inference of the factors associated with the subplots might not be reliable because they do not depict the true structure of the covariance matrix. In this way, it required a more accurate modelling for hypothesis testing.

Considering that the structure of matrix R that best represents the data is the one that presents the lowest value of AIC and BIC, it is noticed that in some situations the AIC and BIC criteria do not present the lowest value related to the same structure. In these cases, the strength of evidence of a model was calculated by measuring its distance to the best model, i.e., to the model with the lowest AIC or BIC value.

In the view of the calculation and interpretation of the weights, the structure of matrix R was defined to analyse the data from muscular strength in the six muscles of the study.

The structure that best fit the data was the ARMA (1,1), which are time series data with autoregressive parameter, moving averages component and residual variance. Thus, the best model is one that takes into account the existing dependence in time, in which the measure of the current muscular strength depends on the strength measured previously.

In conclusion, the methodology of the linear mixed models proved to be efficient in modelling plots subdivided in time, for being able to take into account different structures of the covariance matrix. The structure of the covariance matrix that best fit the data was the ARMA (1,1), evidencing the existing dependence in time. In the work of Alves et al. (2017), the significant results compared G1 and G3 with G2, but the results were not conclusive when compared to G1 and G3. Identifying the best structure of the covariance matrix allowed us to better estimate the effects, using the tests appropriately to verify differences between factors that were not detected when using the median frequency of strength, including differences between groups G1 and G2.

References

Adamsen L, Andersen C, Midtgaard J, et al (2009). Struggling with cancer and treatment: young athletes recapture body control and identity through exercise: qualitative findings from a supervised group exercise program in cancer patients of mixed gender undergoing chemotherapy. *Scand J Med Sci Sports*, **19**, 55–66.

Battaglini C, Bottaro M, Dennehy C, et al (2006). Efeitos do treinamento de resistência na força muscular e níveis de fadiga em pacientes com câncer de mama. *Rev Bras Med Esporte*, **12**, 153–8.

Burnham KP, Anderson DR (2004). Multimodel inference understanding AIC and BIC in model selection. *Sociol Methods Res*, **33**, 261–304.

Da Silva Alves R, Iunes DH, Pereira IC, et al (2017). Influence of exergaming on the perception of cancer-related fatigue. *Games Health J*, **6**, 119-26.

Faraway JJ (2006). Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. London: Chapman and Hall/CRC, **124**, p 331.

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, **31**, 423–47.

Hsiao C (2014). Analysis of panel data. [S.l.]: Cambridge university press, p 526.

Huynh H, Feldt LS (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *J Am Stat Assoc*, **65**, 1582–9.

Jennrich RI, Schluchter MD (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, **42**, 805–20.

Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006). SAS for mixed models. [S.l.]: SAS institute.

Liu C, Cao D, Chen P, Zagar T (2007). Random and repeated statements-how to use them to model the covariance structure in proc mixed. In: SAS Conference Proceedings: Midwest SAS User Group. [S.l.: s.n.].

Mauchly JW (1940). Significance test for sphericity of a normal n-variate distribution. *Ann Math Stat*, **11**, 204–9.

Montgomery DC (2008). Design and analysis of experiments. [S.l.]: John Wiley and Sons, p 752.

Parry I, Carbullido C, Kawada J, et al (2014). Keeping up with video game technology: Objective analysis of xbox kinectTM and playstation 3 moveTM for use in burn rehabilitation. Burns, Elsevier, **40**, pp 852–9.

Prinsen H, Van Dijk JPV, Zwarts MJ, et al (2015). The role of central and peripheral muscle fatigue in postcancer fatigue: a randomized controlled trial. J Pain Symptom Manage, Elsevier, **49**, pp 173–82.

R Core Team (2016). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna. https://www.R-project.org/.

Silva EM, Duarte JB, Reis AJS (2015). Seleção da matriz de variância-covariância residual na análise de ensaios varietais com medidas repetidas em cana-de-açúcar. *Ciência Rural SciELO Brasil*, **45**, 993–9.

Wolfinger RD (1996). Heterogeneous variance: Covariance structures for repeated measures. *J Agric Biol Environ Stat*, **1**, 205–30.