HIERARCHICAL STRUCTURE OF DENTAL DATA IN THE RANDOM EFFECTS INCLUSION APPROACH

Tiago Peres da Silva SUGUIURA[1]
Omar Cléo Neves PEREIRA[1]
Waenya Fernandez CARVALHO[2]
Isolde Terezinha Santos PREVIDELLI[1]

- ABSTRACT: Data sets with complex structures is increasingly common in dental research. As consequences, statistical methods to analyze and interpret these data must be efficient and robust. Hierarchical structures is one of the most common kind of complex structures, and a proper approach is required. The multilevel modeling used to study hierarchical structures is a powerful tool which allows the collected data to be analyzes in several levels. This study has as objective to make a literature review on multilevel linear models and to illustrate a three level model through a matrix procedure, without the use of specific software to estimate the parameters. With this model, we analyzed the vertical gingival retraction when using the substances: Naphazoline Chloridrate, Aluminium Chloride and without any substance. The intraclass correlation coefficient on dental level within patients showed that the hierarchical structure was important to accommodate the dependence within clusters.

- KEYWORDS: Covariance matrix; Henderson's equation; mixed models; linear multilevel models; vertical gingival retraction.

1 Introduction

In 1976, Bennett published an important study about methods of teaching with elementary school students from England. The results suggested that formal methods of teaching were associated to a greater progress in basic students skills,

[1]Universidade Estadual de Maringá - UEM, Departamento de Estatística, CEP:87020-900, Maringá, PR, Brasil. E-mail: *tiago.suguiura@gmail.com; omarcnpereira@gmail.com; iprevidelli@uem.br*
[2]Universidade Estadual de Maringá - UEM, Centro de Ciências da Saúde, CEP:87020-900, Maringá, PR, Brasil. E-mail: *waenyacarvalho@gmail.com*

causing considerable controversy. The experimental data were analyzed using traditional multiple regression models which recognized the studied children as units of analysis and ignored their grouping with classes and teachers.

The results were statistically significant, however, Aitkin *et al.* (1981) showed that, considering the analysis with the children (units of analysis) grouped in clusters, the significant differences vanished, and the children whom received such formal teaching style, did not presented differences to those who did not received such treatment. According to Goldstein, 2011, this reanalysis is the first important example of a multilevel analysis in data form social sciences. In this case, any children, for being in the same class, might be similar in their performances. More information about the use of multilevel models regardind educational research can be found in (BOCK, 2014).

Statistical methods and algorithms were developed and, in 1986, the basis of multilevel analysis was established (LAIRD and WARE, 1982; MASON *et al*,1983, GOLDSTEIN, 1986).

In dental clinical trials, the use of these methods is essential due to its hierarchical structure (GILTHORPE, 2000; MDALA, 2012; MARTINS, 2014; KOLAWOLE, 2016; CHRCANOVIC, 2016). Analysis assuming independence of observations are inappropriate and, therefore, methodology which tends to group the data in patient level results in loss of valuable information and may not reflect the specific association. In such scenery, the estimates of standard errors are underestimated and, therefore, type I errors are inflated by all statistical tests using the supposition of independence (HANCOCK, 2010). The use of random effects in multilevel modeling is a ordinary and suitable manner of modeling such group structure.

2 Mixed models

Correlated data often appears in statistical analysis. Whether in the subjects grouping, or in repeated measures in the same experimental unit throughout time or space. Mixed models analysis provides a general and flexible approach in these situations, because allow a great variety of correlations structures to be modeled. A mixed model of data sets coming from repeated measures can be written as

$$\mathtt{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon_i} \tag{1}$$

where $\mathtt{y}_i = (y_{i1}, \cdots, y_{in_i})^T$ with dimension $(n_i \times 1)$ is the response profile of the i*th* experimental unit, $\boldsymbol{\beta}$ is a vector with dimension $(p \times 1)$ of unknown fixed effects parameters, $\mathbf{X}_i$ is a fixed effects design matrix with dimension $(n_i \times t)$, known and of full rank, $\mathbf{b}_i$ is a vector with dimension $(q_i \times 1)$ of random effects parameters $\mathbf{Z}_i$ is a random effects design matrix with dimension $(n_i \times q)$ known and of full rank and $\boldsymbol{\varepsilon}_i$ is a random errors vector with dimension $(n_i \times 1)$.

We assume that $\mathbf{b}_i \sim \mathbf{N}_q(\mathbf{0}, \mathbf{G})$ and $\boldsymbol{\varepsilon}_i \sim \mathbf{N}_{n_i}(\mathbf{0}, \mathbf{R}_i)$, with $\mathbf{G}$, with dimension $(q \times q)$ and $\mathbf{R}_i$ with dimension $(n_i \times n_i)$ positive defined symmetrical

matrices, and besides that, $\mathbf{b}_i$ and $\boldsymbol{\varepsilon}_i$ are independent random variables. Hence, $\mathtt{y_i} \sim \mathbf{N}(\mathbf{X}_i\beta, \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^T + \mathbf{R}_i)$.

The elements over the main diagonal of the matrix $\mathbf{G}$ represents each random effects variances in $\mathbf{b}_i$, and the other diagonals represents the covariances between two correspondent random effects. If the vector $\mathbf{b}_i$ has q random effects associated with the model, the matrix $\mathbf{G}$ is the following positive defined symmetrical matrix:

$$\mathbf{G} = Var(\mathbf{b}_i) = \begin{bmatrix} Var(b_{1i}) & cov(b_{1i}, b_{2i}) & \cdots & cov(b_{1i}, b_{qi}) \\ cov(b_{1i}, b_{2i}) & Var(b_{2i}) & \cdots & cov(b_{2i}, b_{qi}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(b_{1i}, b_{qi}) & cov(b_{2i}, b_{qi}) & \cdots & Var(b_{qi}) \end{bmatrix}$$

The same way we have that the n_i residuals in the vector $\boldsymbol{\varepsilon}_i$ are random variables which follows a normal multivariate distribution with mean $\mathbf{0}$, and covariance matrix positive defined and symmetric $\mathbf{R}_i$, defined as

$$\mathbf{R}_i = Var(\boldsymbol{\varepsilon}_i) = \begin{bmatrix} Var(\varepsilon_{1i}) & cov(\varepsilon_{1i}, \varepsilon_{2i}) & \cdots & cov(\varepsilon_{1i}, \varepsilon_{n_i i}) \\ cov(\varepsilon_{1i}, \varepsilon_{2i}) & Var(\varepsilon_{2i}) & \cdots & cov(\varepsilon_{2i}, \varepsilon_{n_i i}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\varepsilon_{1i}, \varepsilon_{n_i i}) & cov(\varepsilon_{2i}, \varepsilon_{n_i i}) & \cdots & Var(\varepsilon_{n_i i}) \end{bmatrix}$$

The matrix $\mathbf{X}_i$ of order $(n_i \times p)$ is the fixed effects specification matrix, that is, represents the known values of the p covariables, and its defined as follows:

$$\mathbf{X}_i = \begin{bmatrix} X_{1i}^{(1)} & X_{1i}^{(2)} & \cdots & X_{1i}^{(p)} \\ X_{2i}^{(1)} & X_{2i}^{(2)} & \cdots & X_{2i}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n_i i}^{(1)} & X_{n_i i}^{(2)} & \cdots & X_{n_i i}^{(p)} \end{bmatrix}$$

The matrix $\mathbf{Z}_i$ of order $(n_i \times q)$ is the random effects specification matrix, that is, represents the known values of the q covariables. The matrix $\mathbf{Z}_i$,

$$\mathbf{Z}_i = \begin{bmatrix} Z_{1i}^{(1)} & Z_{1i}^{(2)} & \cdots & Z_{1i}^{(q)} \\ Z_{2i}^{(1)} & Z_{2i}^{(2)} & \cdots & Z_{2i}^{(q)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n_i i}^{(1)} & Z_{n_i i}^{(2)} & \cdots & Z_{n_i i}^{(q)} \end{bmatrix},$$

will be structured according to the layout of the data

According to West (2014), in many cases, predictive variables with effects that vary randomly between individuals are represented in both matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$. For example, in a linear mixed model in which only the intercepts are random, the matrix $\mathbf{Z}_i$ will be simply composed of one column of 1s

To estimate $\boldsymbol{\beta}$ and predict $\mathbf{b}$, we used Henderson's equation (HENDERSON, 1950), which are given from the joint distribution of $\mathbf{b}$ and $\boldsymbol{\varepsilon}$.

Let g be the number of elements in $\mathbf{b}$ and n the dimension of $\mathtt{y}$, its joint distribution is given by

$$f(b,\varepsilon) = \frac{1}{(2\pi)^{(n+g)/2}} \left| \begin{array}{cc} \mathbf{G} & 0 \\ 0 & \mathbf{R} \end{array} \right|^{-\frac{1}{2}} \cdot \exp\left\{ -\frac{1}{2} \left[\begin{array}{c} \mathbf{b} \\ \mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{b} \end{array} \right]' \left[\begin{array}{cc} \mathbf{G}^{-1} & 0 \\ 0 & \mathbf{R}^{-1} \end{array} \right] \left[\begin{array}{c} \mathbf{b} \\ \mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{b} \end{array} \right] \right\} \quad (2)$$

To maximize $f(b.\varepsilon)$ regarding $\boldsymbol{\beta}$ and $\mathbf{b}$, means to minimize the exponential part of the equation(2)

$$\begin{aligned} Q &= \left[\begin{array}{c} \mathbf{b} \\ \mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{b} \end{array} \right]' \left[\begin{array}{cc} \mathbf{G}^{-1} & 0 \\ 0 & \mathbf{R}^{-1} \end{array} \right] \left[\begin{array}{c} \mathbf{b} \\ \mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{b} \end{array} \right] \\ &= \mathbf{b}'\mathbf{G}^{-1}\mathbf{b} + (\mathbf{y} - \mathbf{X}\beta\mathbf{Z}\mathbf{b})'\mathbf{R}^{-1}(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{b}) \end{aligned}$$

where is considered the independence of $\mathbf{b}$ and $\boldsymbol{\varepsilon}$. This leads to Henderson's mixed model equations

$$\begin{aligned} \frac{\partial Q}{\partial \beta} = 0 &\Leftrightarrow \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\widehat{\mathbf{b}} = \mathbf{X}'\mathbf{R}^{-1}\mathbf{Y} \\ \frac{\partial Q}{\partial b} = 0 &\Leftrightarrow \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\widehat{\boldsymbol{\beta}} + (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1})\widehat{\mathbf{b}} = \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Y} \end{aligned}$$

or in a matrix form

$$\left[\begin{array}{cc} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}) \end{array} \right] \left[\begin{array}{c} \widehat{\boldsymbol{\beta}} \\ \widehat{\mathbf{b}} \end{array} \right] = \left[\begin{array}{c} \mathbf{X}'\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Y} \end{array} \right]. \quad (3)$$

For the multilevel case, we can write $\mathtt{y_i} \sim \mathbf{N}(\mathbf{X}_i\beta, \mathbf{Z}_{1i}\mathbf{G}_1\mathbf{Z}_{1i}^T + \mathbf{Z}_{2i}\mathbf{G}_2\mathbf{Z}_{2i}^T + \mathbf{R}_i)$, in which $\mathbf{G}_1$ and $\mathbf{G}_2$ represents the covariance matrices for the random effects b_i's of the ith cluster, and the random effects $b_{j|i}$'s of the jth unit of analysis nested within the ith cluster, respectively, and $\mathbf{Z}_{1i}$ and $\mathbf{Z}_{2i}$ represents the specification matrices of the random effects for the clusters and the unit of analysis respectively.

To solve the equation (3), the matrix $\mathbf{Z}_i$ can be written as

$$\begin{aligned} \mathbf{Z}_i &= [\mathbf{Z}_{1i} \quad \mathbf{Z}_{2i}] \\ &= \left[\left(\begin{array}{cccc} Z_{1i}^{(1)} & Z_{1i}^{(2)} & \cdots & Z_{1i}^{(q)} \\ Z_{2i}^{(1)} & Z_{2i}^{(2)} & \cdots & Z_{2i}^{(q)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n_i i}^{(1)} & Z_{n_i i}^{(2)} & \cdots & Z_{n_i i}^{(q)} \end{array} \right) \quad \left(\begin{array}{cccc} Z_{1j|i}^{(1)} & Z_{1j|i}^{(2)} & \cdots & Z_{1j|i}^{(s)} \\ Z_{2j|i}^{(1)} & Z_{2j|i}^{(2)} & \cdots & Z_{2j|i}^{(s)} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n_i j|i}^{(1)} & Z_{n_i j|i}^{(2)} & \cdots & Z_{n_i j|i}^{(s)} \end{array} \right) \right], \end{aligned}$$

and the covariance matrix $\mathbf{G}$ of the random effects as

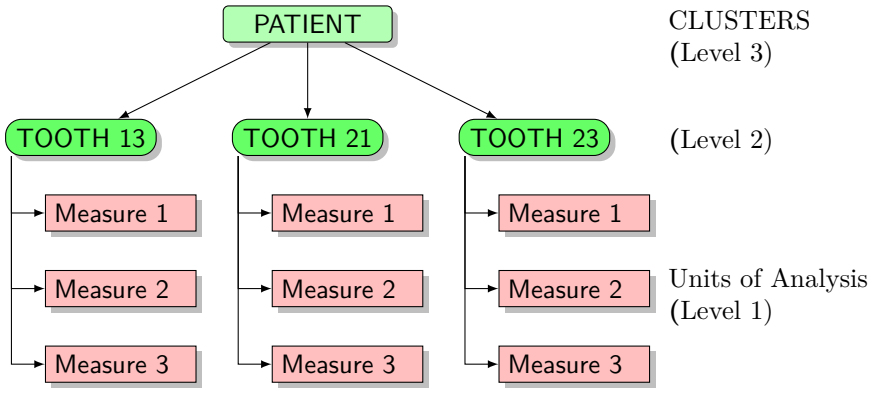
$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & 0 \\ 0 & \mathbf{G}_2 \end{bmatrix}.$$

3 Material and methods

This study were realized with the objective of evaluating the effects on the vertical gingival retraction regarding the application of a treatment with Naphazoline Chloridrate (C) against the usual treatment with Aluminium Chloride (H), in comparison with a placebo (P). 24 patients were selected and they had the same three teeth evaluated, teeth 13, 21 and 23 (Opinion 1.515.263 of the Ethics Committee from the Maringá State University).

All the 24 patients received the three treatments, one in each tooth, that is, each tooth in a patient received randomly a retractor wire with different chemical substances or no substance. Of all teeth from all the patients were taken three measures of the vertical gingival retraction.

The following diagram, represents a hierarchical structure with repeated measures, which represents the data structure regarding the study in question.



Initially the related isolation was performed with cotton rollers in the areas corresponding to the teeth to be evaluated. After cleaning with dental floss and cotton ball soaked in `chlorhexidine` at 2%, the teeth were rinsed an dried. A layer of photopolymerizable gum guard Top dam® was applied to the dental surface of the elements 13, 21 and 23 at the cervical margin of the gingival sulcus to record its initial position.

After the entire excess of the gingival contour was removed with the aid of a probe, the gel was polymerized. Then, using the double wire technique, the retraction wires were positioned. Initially, the retractor wire (ultrapak) nº 000 was placed within the gingival sulcus of the vestibular face of each teeth. Then, the (Ultrapak®) nº1 retractor wire was placed randomly over the first wire. After the placement of the wire 000, the retractor wire (Ultrapak®) nº 1 was installed at the first tooth soaked with Naphazoline Chloridrate (Legrand® - Group 1), over the second tooth, soaked in Aluminium Chloride (Hemostop® - Group 2), and over the third tooth, the wire was places without any substance (Placebo - Group 3).

The wires were soaked in their respective solutions for 7 minutes before being applied to the teeth. After a 4 minute period, the retractor wires were removed from within the gingival sulcus, the area was air dried and the molding was performed using addition silicone (polyvinyl siloxane)(3D - Angelus, Londrina - Brazil)

After the addition silicone prey was taken, the dental tray was removed from the mouth. After two hours of this molding, the cast was casted in a special type IV plaster, and then cut into small blocks, from which were taken 72 images (one image for each tooth with gingival retraction) by a camera coupled to a magnifying glass (Olympus SZ-STS). The images were analyzed with the Image Pro-Plus (version 4.5) program to measure the distances between the gingival protector Top dam® and the gingival level. For example, the three measurements taken from one tooth can be seen in Figure 1.

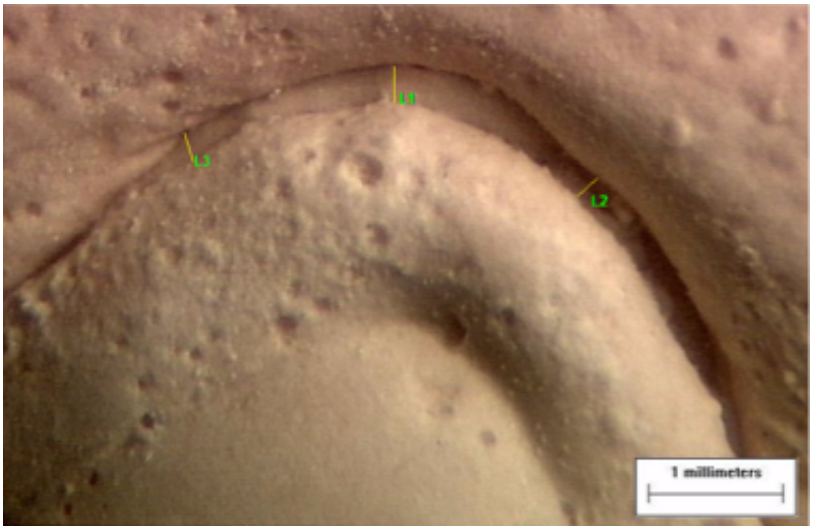


Figure 1 - The three measurements taken from one tooth casted in plaster.

4 Linear multilevel model

To model the response variable vertical gingival retraction, we will assume the treatments H, C and P as fixed effects, and the effects of the patients receiving treatment, as random. With this approach we mean that if we would repeat the same experiment with another patients, the expected values for the treatment would be very similar to those obtained in this study, however, the individual effects of the patients, may be different but with similar variance.

Considering the kth observation of the gingival retraction, $k = 1, \cdots, K$, with $K = 3$, at the jth tooth, which received randomly one of the treatments, $j = 1, \cdots, J$, and $J = 3$, of the ith patient, $i = 1, \cdots, I$, with $I = 24$, a linear model which seeks to relate the response variable y_{ijk}, with the treatments, taking into account the variability that may exist between patients and the variability within the teeth of each patient is given by

$$y_{ijk} = \mu + \alpha_{ij} + b_i + b_{j|i} + \varepsilon_{ijk} \tag{4}$$

where $b_i \sim N(0, \sigma_b^2)$, $b_{j|i} \sim N(0, \sigma_{b_{j|i}}^2)$ and $\varepsilon_{ijk} \sim N(0, \sigma^2)$, and are all mutually independent random variables. μ are the overall mean, α_{ij} represents the fixed effects regarding the treatment received by the jth tooth of the ith patient, b_i are the patients random effects, and $b_{j|i}$ are the random effects of the teeth within each patient.

We will test the null hypothesis $H_0 : \sigma_{b_{j|i}}^2 = 0$ against $H_1 : \sigma_{b_{j|i}}^2 \neq 0$, which is equivalent of testing $H_0 : b_{j|i} = 0 \ \forall \ ij$. From that, we have

$$y_{ijk} = \mu + \alpha_{ij} + b_i + \varepsilon_{ijk}. \tag{5}$$

We tested the null hypotheses by comparing (4) and (5) through an F-ratio test, which led to reject H_0, that is, there is strong evidence that, in many patients, the teeth influence the gingival retraction, disregarding the treatment received.

Generally, in mixed models, the random effects variance is more interesting than the random effects themselves. Therefore, we must estimate them (WOOD, 2006; FINCH, 2014).

For the dataset of vertical gingival retraction of this study, we estimated σ^2 using $\sigma^2 = \dfrac{RSS_\varepsilon}{(n - IJ)}$, where RSS_ε is the residual sum of squares of the model (4) and $n = IJK$. In order to estimate $\sigma_{b_{j|i}}^2$, we use the model which results from the mean of the K values of the teeth level, that is,

$$\bar{y}_{ij.} = \mu + \alpha_{ij} + b_i + b_{j|i} + \frac{1}{K}\sum_{k=1}^{K} \epsilon_{ijk}. \tag{6}$$

By defining $e_{ij} = b_{j|i} + \dfrac{1}{K}\displaystyle\sum_{k=1}^{K}\epsilon_{ijk}$, we have that $\text{var}(e_{ij}) = \sigma_{b_{j|i}}^2 + \dfrac{\sigma^2}{K}$, where e_{ij} are independents and identically distributed random variables such that $e_{ij} \sim N(0, \sigma_{b_{j|i}}^2 + \sigma^2/K)$. This way, we can rewrite the simplified model as

$$\bar{y}_{ij.} = \mu + \alpha_j + b_i + e_{ij}\,. \tag{7}$$

which is useful to estimate the residual variance $\hat{\sigma}^2_{b_{j|i}} = \dfrac{RSS_{b_{j|i}}}{IJ - I - J + 1} - \dfrac{\hat{\sigma}^2}{K}$, such that $RSS_{b_{j|i}}$ is the residual sum of squares of (7).

By taken the mean of the response variable for each patient we have $\bar{y}_{i..} = \mu + \dfrac{1}{J}\sum_{j=1}^{J}\alpha_{ij} + b_i + \dfrac{1}{J}\sum_{j=1}^{J} e_{ij}$. If $\mu'_i = \mu + \dfrac{1}{J}\sum_{j}\alpha_{ij}$ and $e_i = b_i + \dfrac{1}{J}\sum_{j} e_{ij}$, then we have

$$\bar{y}_{i..} = \mu'_i + e_i \ , \tag{8}$$

such that $e_i \sim N\left(0, \sigma^2_b + \dfrac{\sigma^2_{b_{j|i}}}{J} + \dfrac{\sigma^2}{JK}\right)$. Therefore, if RSS_b is the residual sum of squares of the model (8), an unbiased estimator of σ^2_b is given by

$$\hat{\sigma}^2_b = \frac{RSS_b}{I-1} - \frac{\hat{\sigma}^2_{b_{j|i}}}{J} - \frac{\hat{\sigma}^2}{JK}. \tag{9}$$

The intraclass correlation coefficient (ICC) is a measure that describes the similarity (or homogeneity) of observations within the same cluster (WEST, 2014; FINCH, 2014). For each level of grouping, an ICC value can be defined as functions of the variance components. This statistic takes values between zero and one, indicating that, the closer to zero, smaller the chance of grouping of the data, and the closer to one, greater is the chance of grouping. In other words, the greater the correlation between individuals, the greater is the inadequacy of the usual regression model. For a greater dependence between individuals of the same group, the greater the need for a regression method that respects the data aggregation structure.

In order to verify the values acquired when using the matrix with Henderson's equation, we will use the statistical software `R` (R CORE TEAM, 2017) with the package `nlme` (PINHEIRO *et al*, 2016), and its function `lme()`, which uses the maximum likelihood (ML) and restricted maximum likelihood (REML) methods for the estimation of parameters.

5 Results

The Table 1 presents part of the dataset of the vertical gingival retraction in its hierarchical structure.

Table 1 - Measures of the vertical gingival retraction by patients and teeth

Patients - Level 3	**Teeth - Level 2**	**Units of Analysis - Level 1**		
Patients	**Teeth**	**Measure 1**	**Measure 2**	**Measure 3**
Patient 1	Tooth 13	0,25778	0,187796	0,183607
Patient 1	Tooth 21	0,154878	0,136879	0,161837
Patient 1	Tooth 23	0,216886	0,203544	0,18065
Patient 2	Tooth 13	0,232762	0,279884	0,141471
Patient 2	Tooth 21	0,427033	0,383964	0,520655
Patient 2	Tooth 23	0,440261	0,27015	0,299254
Patient 3	Tooth 13	0,304009	0,203544	0,185695
Patient 3	Tooth 21	0,223543	0,248008	0,224964
Patient 3	Tooth 23	0,222686	0,204482	0,196372
⋮	⋮	⋮	⋮	⋮
Patient 24	Tooth 23	0,239643	0,248008	0,203231

To solve the system (3), we must define the matrices $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{R}$ and $\mathbf{G}$. According to the disposition of the values in the dataset, we define

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}_{216\times 3}, \mathbf{Y} = \begin{bmatrix} 0.257780 \\ 0.154878 \\ 0.216886 \\ 0.232762 \\ 0.427033 \\ 0.440261 \\ 0.304009 \\ 0.223543 \\ \vdots \\ 0.239643 \end{bmatrix}_{216\times 1}$$

The matrix $\mathbf{Z}$ must express both random effects used in the model, and according to the dataset, it can be written as

$\mathbf{Z} =$

P_1	P_2	P_3	$\cdots$	P_{24}	P_1/D_{13}	P_1/D_{21}	P_1/D_{23}	P_2/D_{13}	$\cdots$	P_{24}/D_{23}
1	0	0	$\cdots$	0	1	0	0	0	$\cdots$	0
1	0	0	$\cdots$	0	1	0	0	0	$\cdots$	0
1	0	0	$\cdots$	0	1	0	0	0	$\cdots$	0
0	1	0	$\cdots$	0	0	0	0	1	$\cdots$	0
0	1	0	$\cdots$	0	0	0	0	1	$\cdots$	0
0	1	0	$\cdots$	0	0	0	0	1	$\cdots$	0
⋮	⋮	⋮	$\ddots$	⋮	⋮	⋮	⋮	⋮	$\cdots$	⋮
0	0	0	$\cdots$	1	0	0	0	0	$\cdots$	1
0	0	0	$\cdots$	1	0	0	0	0	$\cdots$	1
0	0	0	$\cdots$	1	0	0	0	0	$\cdots$	1

$_{216\times 96}$

In this study, we have $\sigma = 0.03869$, $\sigma_b = 0.03924956$ e $\sigma_{b_{j|i}} = 0.06282674$. Therefore, the matrix $\mathbf{R}$ is defined as

$$\mathbf{R}_2 = \sigma^2 \mathrm{I} = 0.0014972 \times \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{216\times 216},$$

and the matrix $\mathbf{G}$, which regards the patients and the teeth within the patients random effects variances, which means

$$\mathbf{G}_1 = \sigma_b^2 \ \mathrm{I}_{24\times 24} = [0.001540523] \times \mathrm{I}_{24\times 24}$$
$$\mathbf{G}_2 = \sigma_{b_{j\beta}}^2 \ \mathrm{I}_{72\times 72} = [0.003947203] \times \mathrm{I}_{72\times 72}$$

can be written as

$$\mathbf{G} = \begin{bmatrix} \begin{bmatrix} 0.001540 & 0 & \cdots & 0 \\ 0 & 0.001540 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0.001540 \end{bmatrix} & & 0 & \\ & & & \\ 0 & & & \\ & & & \end{bmatrix} \begin{bmatrix} 0.003947 & 0 & \cdots & 0 \\ 0 & 0.003947 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0.003947 \end{bmatrix}_{96\times 96}$$

With these matrices implemented at the `R` software, the system (3) can be solved. To emphasize the matrix process results from the `R` software, we present part of its output:

```
[1]   0.191998  0.013502  0.054398    -0.014034  0.060240  0.004622
[7]  -0.039918 -0.021236 -0.036653     0.053653  0.018053  0.003654
[13]  0.027100 -0.031677 -0.034517    -0.022178  0.002574 -0.013885
[19] -0.035041  0.001480  0.040138    -0.004210 -0.004791 -0.003178
[25]  0.004162  0.026968  0.018673    -0.035748 -0.020093  0.019882
[31] -0.042346  0.026561  0.170135    -0.002020 -0.016731  ...
```

where the first three values (inside the box) represent the fixed effects related to the treatments, the next 24 represents the random effects regarding each patient, and the last 72 (presenting only a few terms) represents the random effects of each tooth nested within each patient.

In order to a statistical model be valid, we know that the randomness and unpredictability are crucial components. Therefore, we need to analyze the statistical errors of the model. For practical purposes, graphical displays of residues can be used to detect discrepancies in the model for the mean response or the presence of outliers observations which may require further investigations (FROST, J., 2012)

The Figure 2 presents three residual graphics of the chosen model. The first two show standardized residuals and the observed values regarding fitted values. With exception of a few values in the first one, these graphics do not indicate large deviations from the proposed linear model. The final graph, a Q–Q plot ("Q" stands for quantile), shows that the linearity of the points meet the assumption of normality.

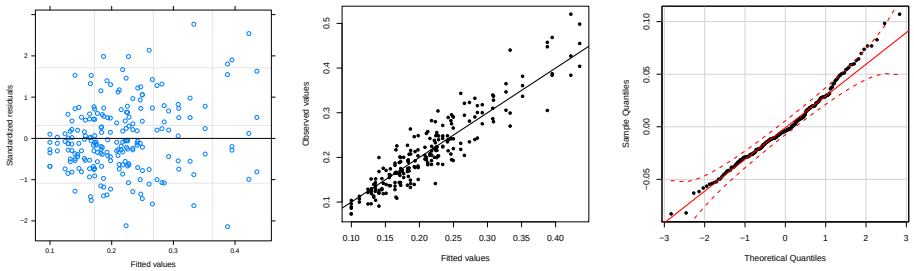


Figure 2 - Residual graphs for the linear multilevel model proposed. The left panel presents the stardardized residuals regarding the fitted values. The middle panel shows the observed values and the fitted values. The right panel is the Q-Q plot.

The estimated values, standard errors and the confidence intervals of 95% for the fixed effects are presented in Table 2.

Table 2 - Estimate, Standard Error (S.E.), lower and upper limits with 95% of confiance e p-value for the linear multilevel model

Parameter	**Estimate**	**S.E.**	**lower**	**upper**	**p-value**
Intercept	0.19199793	0.01579398	0.16077995	0.22321591	<0.0001
Colírio	0.01350160	0.01924896	-0.02524455	0.05224774	0.4866
Hemostop	0.05439754	0.01924896	0.01565139	0.09314369	0.0070

Using the Placebo as the intercept (basis for comparison), we have that the treatment with Naphazoline Chloridrate did not presented a significant p-value. That is, the null hypothesis of them being equal was not rejected. However, the treatment with Aluminium Chloride, got a significant p-value in comparison with the placebo, indicating that there might be differences between the treatments.

Therefore, according to the data, there are evidences that the Naphazoline Chloridrate (collyrium Legrand), when used as an agent of vertical gingival retraction, is not capable of increasing the vertical retraction when compared with the standard substance. However, the Hemostop treatment was significant, proving its already conventional use.

6 Discussion

The use of a multilevel model in order to adjust a hierarchical structure and its more complex dataset is presented in this work. The necessity of this kind of modeling can be measured with the ICC.

For the adopted model, the ICC for the patient level ICC_b, is given by

$$\begin{aligned} ICC_b &= \frac{\sigma_b^2}{\sigma_b^2 + \sigma_{b_{j|i}}^2 + \sigma^2} \\ &= 0.2205497 \ . \end{aligned}$$

In a similar way, the ICC for the teeth level, $ICC_{b_{j|i}}$, is given by

$$\begin{aligned} ICC_{b_{j|i}} &= \frac{\sigma_b^2 + \sigma_{b_{j|i}}^2}{\sigma_b^2 + \sigma_{b_{j|i}}^2 + \sigma^2} \\ &= 0.7856527 \end{aligned}$$

With this value, relatively high for the teeth within patients level ICC, we can say that the hierarchical structure adopted is important to model the dependency among individuals.

Beside that, a mixed model with only one random effect on the patient level was adjusted. This model was compared with the multilevel model through the restricted likelihood ratio test. We verified the necessity of maintaining the random effect regarding the teeth nested within the patient.

Throughout this study we verified that not using a model which contemplates the dependency among individuals of the same group increases the amount of Type I errors by all statistical tests which uses the supposition of independence.

With the results obtained, we verified the importance of multilevel models to accommodate dataset with hierarchical structure, common in dentistry. Beside that, the linear multilevel models are easy to manipulate, both algebraic and computationally.

Acknowledgments

We thank reviewers and editors for their suggestions.

SUGUIURA, T. P. S.; PEREIRA, O. C. N.; CARVALHO, W. F.; PREVIDELLI, I. T. S. Estrutura hierárquica de dados odontológicos na abordagem de inclusão de efeitos aleatórios. *Rev. Bras. Biom.,* Lavras, v.36, n.3, p.700-714, 2018.

- *RESUMO: Conjuntos de dados com estruturas complexas é cada vez mais comum em pesquisas odontológicas. Como consequência, os métodos estatísticos usados para análise e interpretação desses dados devem ser eficientes e robustos. Estruturas hierárquicas é um dos tipos mais comuns de estruturas complexas, e uma abordagem apropriada é necessária. A modelagem multinível utilizada para analisar estruturas hierárquicas é uma ferramente poderosa a qual permite analisar os dados coletados em mais de um nível. Este estudo tem como objetivo fazer uma breve revisão de literatura sobre modelos lineares multinível e ilustrar um modelo de três níveis através de um procedimento matricial, sem o uso de programas específicos para estimar os parâmetros. Com este modelo, avaliou-se fastamento gengival vertical em função da aplicação das substâncias Cloridrato de Nafazolina e Cloreto de Alumínio, e sem substância. O coeficiente de correlação intraclasse no nível dos dentes dentro dos pacientes mostrou que a estrutura hierárquica foi importante para acomodar a dependência dentro dos grupos.*

- *PALAVRAS-CHAVE: Afastamento gengival vertical; modelo misto; modelo multinível.*

References

AITKIN, M; ANDERSON, D; HINDE, J. Statistical modelling of data on teaching styles. *Journal of the Royal Statistics Society*, p.419–461, 1981.

BENNETT, H; JORDAN, J.; LONG, G; WADE, B., Teaching styles and pupil progress. JSTOR, 1976.

BOCK, R. D. *Multilevel analysis of educational data*. Berlin: Elsevier, 2014.

CHRCANOVIC, B. R. *et al*. Bruxism and dental implant failures: a multilevel mixed effects parametric survival analysis approach. *Journal of oral rehabilitation*, v.43, n.11, p.813-823, 2016.

FINCH, W. H.; BOLIN, J. E. ; KELLEY, K. M. *Multilevel Modeling using R*, London: CRC Press, 2014.

FROST, J. Why You Need to Check Your Residual Plots for Regression Analysis: Or, To Err is Human, To Err Randomly is Statistically Divine. 2012. available at: <http://blog.minitab.com/blog/adventures-in-statistics/ why-you-need-to-check-your-residual-plots-for-regression-analysis>.

GILTHORPE, M. S.; CUNNINGHAM, S. J. The application of multilevel, multivariate modelling to orthodontic research data. *Community dental health*, v.17, n.4, p.236–242, 2000.

GOLDSTEIN, H. Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, v.73, n.1, p.43–56, 1986.

GOLDSTEIN, H. *Multilevel statistical models*. New York: John Wiley & Sons, 2011. 922p.

HANCOCK, G. R.; MUELLER, R. O. *The reviewer's guide to quantitative methods in the social sciences*. Routledge, 2010.

HENDERSON, C. R. Estimation of genetic parameters. In: International Biometric Soc. 1441 I ST, NW, SUITE 700, WASHINGTON, DC. Biometrics, v.6, n.2, p.186–187, 1950.

KOLAWOLE, K. A.; FOLAYAN, M. O.; AGBAJE, H. O.; OYEDELE, T. A.; OZIEGBE, E. O.; ONYEJAKA, N. K.; CHUKWUMAH, N. M.; OSHOMOJI, O. V. Digit sucking habit and association with dental caries and oral hygiene status of children aged 6 months to 12 years resident in semi-urban Nigeria. *PloS one*, v.11, n.2, 2016.

LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. *Biometrics*, p.963–974, 1982.

MARTINS, M. T.; SARDENBERG, F.; ABREU, M. H.; VALE, M. P.; PAIVA, S. M. PORDEUS, I. A. Factors associated with dental caries in Brazilian children: a multilevel approach. *Community dentistry and oral epidemiology*, v.42, n.4, p.289-299, 2014.

MASON, W. M.; WONG, G. Y.; ENTWISLE, B. Contextual analysis through the multilevel linear model. *Sociological methodology*, v.1984, p.72–103, 1983.

MDALA, I.; HAFFAJEE, A. D.; SOCRANSKY, S. S.; BLASIO, B. F. D.; THORESEN, M.; OLSEN, I.; GOODSON, J. M. Multilevel analysis of clinical parameters in chronic periodontitis after root planing/scaling, surgery, and systemic and local antibiotics: 2-year results. *Journal of oral microbiology*, v.4, 2012.

PINHEIRO, J.; BATES, D.; DEBROY, S.; SARKAR, D.; nlme: Linear and Nonlinear Mixed Effects Models, 2016. R package version 3.1-124. Available at: <http://CRAN.R-project.org/package=nlme>.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2017. Available at: <https://www.R-project.org/>.

WEST, B. T.; WELCH, K. B.; GALECKI, A. T. *Linear mixed models: a practical guide using statistical software*. London: CRC Press, 2014.

WOOD, S. *Generalized additive models: an introduction with R*. London: CRC press, 2006.

Received on 22.03.2017.

Approved after revised on 01.09.2017.