Chilean Journal of Statistics
Vol. 7, No. 2, September 2016, 51-68

Research Paper

Modeling ordinal longitudinal outcomes: an applied perspective of marginal and conditional approaches

Nívea B. da Silva[1,*], Leila D. A. F. Amorim[2], Enrico A. Colosimo[1] and Leo Heller[3]

[1]Department of Statistics, Federal University of Minas Gerais, Minas Gerais, Brazil
[2]Department of Statistics, Federal University of Bahia, Bahia, Brazil
[3]Department of Sanitary Engineering, Federal University of Minas Gerais, Minas Gerais, Brazil

(Received: May 16, 2016 · Accepted in final form: September 25, 2016)

Abstract

Polytomous responses modeling, particularly ordinal ones, has been the subject of increasing interest in recent years. It has been gaining ground in research on quality of life and health status indicators, among other topics. Its use ranges from cross-sectional studies, which assume independence between observations, to longitudinal studies, in which two or more responses from the same individual are observed over time. There are several methods proposed in the literature to model data of this nature in cross-sectional studies, and in practice the most common is the proportional odds model. There is also the partial proportional odds model, which generalizes the proportional odds model and allows for non-proportionality in a subset of covariates. In longitudinal studies, the usual models - marginal and generalized linear mixed (conditional) models - for analysis of correlated data can also be used to model polytomous responses. In this paper, models for ordinal polytomous responses are discussed from an applied perspective of conditional and marginal structures. The specification and interpretation of the models are illustrated and discussed by analyzing two real data sets.

Keywords: conditional models · marginal models · ordinal responses · partial proportional odds model · proportional odds model.

Mathematics Subject Classification: Primary 62J12 · Secondary 62P10.

1. Introduction

In recent years, many methods have been proposed for modeling polytomous responses, and the interest in this type of modeling has been growing for both cross-sectional and longitudinal studies. In epidemiology, for example, there is often interest in estimating the risk of adverse events, and researchers may choose to classify the response of interest into two or more categories to estimate the relative risk or odds ratio, depending on the design of the study (Ananth and Kleinbaum, 1997). The use of ordinal polytomous responses has gained significant ground in studies on quality of life, health status indicators, and even in evaluating the severity of certain diseases. In clinical trials, responses on an ordinal scale are often used to quantify the symptoms or condition of a patient, and they may be used

*Corresponding author. Email: nivea.bispo@gmail.com

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

to evaluate the effectiveness of post-operative procedures (Parsons et al., 2009).

Typical approaches for analyzing longitudinal data can be used to model polytomous responses. For instance, using the marginal approach, some authors proposed different procedures for modeling polytomous responses in longitudinal studies. In most studies, the authors adopted the generalized estimating equations (GEE1) proposed by Liang and Zeger (1986) and Zeger and Liang (1986), and made use of the proportional odds model. Clayton (1992), for instance, considered GEE1 to model ordinal responses, but rewrote it based on binary variables and used the proportional odds model to fit the models. Heagerty and Zeger (1996) proposed a new set of estimating equations to analyze correlated ordinal responses. These authors have extended the proposed method of Carey et al. (1993) to accommodate correlated ordinal responses, simultaneously modeling the parameters of marginal mean and association.

In addition to the class of marginal models, numerous studies have suggested modeling ordinal responses by using conditional models. Hedeker and Gibbons (1994) proposed a random effects model for analyzing ordinal responses in longitudinal studies that uses probit and logistic link functions, using multilevel terminology. Hedeker and Mermelstein (1998, 2000) described an extension of the proportional odds model for longitudinal data allowing non-proportionality of odds in a subset of the predictor variables. Hedeker and Gibbons (1994, 2006) also presented a model within a multilevel representation that accommodates multiple random effects for analyzing longitudinal ordinal responses, in which variables can be included to explain inter and intra-individual variations.

Even though, there is available literature proposing a variety of methods for analysis of polytomous responses, they are presented in a fragmented way. Thus, in this paper, the literature on marginal and conditional models are systematized and summarized to model ordinal polytomous responses in longitudinal studies. Model's specification is fully presented to provide a comprehensive overview of these methodologies for the reader. Differences in the interpretation of the results from distinct methodologies are highlighted and illustrated by analyzing two real data sets. In the first one, data from an epidemiological study conducted with 635 children up to 5 years of age, for which the main objective was to assess the impact of rainwater harvesting systems on children's health, is used. The second application relates to a study of 49 pregnant women who were followed until delivery, which aimed to compare two methods of analgesia for pain during labor. The two data sets differ in sample size, in balance, in the number of repeated measurements per individual, and also in the equidistance of the time measurements. It is important to note that the choice of modeling strategy in a longitudinal study depends essentially on the question of interest to the researcher. In the following sections, marginal and conditional classes of models are presented for ordinal responses.

This paper is organized as follows. Section 2 presents in full details the two real data examples involving ordinal longitudinal responses. Detailed specifications for the models for analyzing longitudinal ordinal responses from the perspective of marginal and conditional approaches are presented in Section 3. Results from the models fitted for both examples are presented and discussed in Section 4. Finally, the paper ends with some final remarks in Section 5.

2. Motivating datasets

Two real data examples involving ordinal longitudinal responses are described in this section. The first one, in Section 2.1, is related to parasitic level measured three times in a year period and the second one, in Section 2.2, is related to pain intensity during labor measured repeatedly in an ordinal scale since application of anesthesia until delivery.

2.1 Study on rainwater harvesting systems and children's health

A prospective cohort epidemiological study was conducted with 635 children up to 5 years old, who were selected and monitored for a period of one year (2009-2010). The main objective of the study was to evaluate the impact of the implementation of rainwater harvesting systems, built mostly by the One Million Cisterns Program (Programa 1 Milhão de Cisternas - P1MC), on health of children in rural families living in Chapada do Norte and Berilo, in Médio Vale do Jequitinhonha-Minas Gerais-Brazil. In the first stage of the study, half of the children had access to cisterns for storing rainwater (Group 1), and the other half did not have (Group 2) (Fonseca, 2012). In the study, two health indicators were analyzed: the occurrence of diarrhea and the presence of intestinal parasites (protozoa diners: Endolimax nana, Entamoeba coli and Iodamoeba butschlii; pathogenic protozoa: Entamoeba histolytica/ dispar and Giardia lamblia; helminths: hookworm, Ascaris lumbricoides, Hymenolepis nana, Enterobius vermicularis, Strongyloides stercoralis and Trichuris trichiura) in three longitudinal study phases. The response of interest is the parasitic level (1: poly-infected child; 2: monoinfected child; 3: uninfected child). Each child was evaluated at three phases of the study: baseline, 6 months after the beginning of the study and, finally, one year from the beginning. The predictor variables are group (1: with cistern; 0: without cistern) and age of the children at baseline, in months. The proportion of older children ($\geq$24 months) at baseline are slightly higher in the cistern group (63.8%) compared to those without cistern (55.8%). Furthermore, 13.2% of children under 12 months were infected (mono- or poly-infected) in the baseline, while this percentage was 59.8% for older children ($\geq$ 24 months). Analyses presented here include information for 635 children who had parasitic level measured in at least one of the phases of the study.

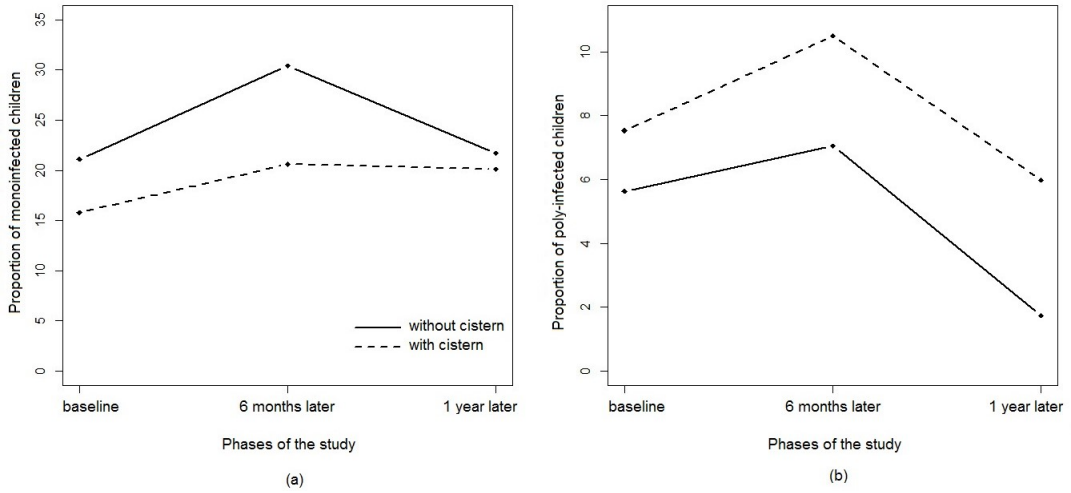

Figure 1. Trend plots for: (a) Proportion of monoinfected children and (b) Proportion of poly-infected children in the three phases of the study by cistern group

Figure 1 shows the proportion of monoinfected and poly-infected children by cistern group at three phases of the study. It is observed that the proportion of poly-infected children is relatively higher in the group with cistern for the three phases of the studies, while the opposite behavior was observed for mono-infection.

2.2 Study on analgesia in childbirth

This study was conducted in Minas Gerais state, Brazil, in order to compare two techniques of analgesia for labor pain. There were 49 patients who were monitored throughout

labor until birth of the child. Normal delivery was expected and indicated by the medical doctors to all patients included in the study. Pain intensity was subjectively assessed by the patient, and measurements of blood pressure, maternal heart rate, consumption of oxytocin, sedation level, signs of respiratory depression, apnea, and other variables were recorded every 5 minutes during the first 30 minutes after the onset of anesthesia, and thereafter every 30 minutes until delivery. One of the used techniques was epidural analgesia (the gold standard), which is a local anesthetic. The other one, whose efficiency was to be compared to the gold standard, involved continuous intravenous infusion of remifentanil, an opioid that has an onset of action within 1 to 3 minutes. Among the 49 pregnant women, 33 received epidural anesthesia and 16 received remifentanil. The intensity of the pain normally depends on the degree of uterine cervix dilatation, and it is generally mild. It is of a colic-type in the initial period, when the dilatation of the cervix is less than 3 cm, and the pain becomes more intense with the progression of labor (Soares, 2013).

The response of interest is the intensity of pain as measured by a Visual Analog Scale (VAS) (1: tolerable and mild pain; 2: moderate pain that causes discomfort; 3: intense and unbearable pain), and it was measured every 5 minutes (5, 10, 15, 20, 25, 30 minutes) for the first half hour after anesthesia, and thereafter every 30 minutes until delivery (60, 90, 120, ..., minutes). Only measurements from the first 90 minutes were used for data analysis. Thus, $0, 5, 10, 15, 20, 25, 30, 60, 90$ minutes denotes the time points at which the intensity of pain was assessed. Time 0 refers to the time immediately before to anesthesia. Predictor variables considered were the treatment group (0: peridural; 1: remifentanil) and patient age.

Figure 2 shows the trajectories, by treatment group, for probability of pain during the follow-up of 90 minutes. According to descriptive analysis, there is a difference between the groups regarding pain intensity, and the probability of intense pain, for example, is higher among pregnant women in the group receiving remifentanil. We observed that the women receiving epidural anesthesia are slight older than those receiving remifentanil. However, due to small sample sizes, we were not able to describe the data by age groups.

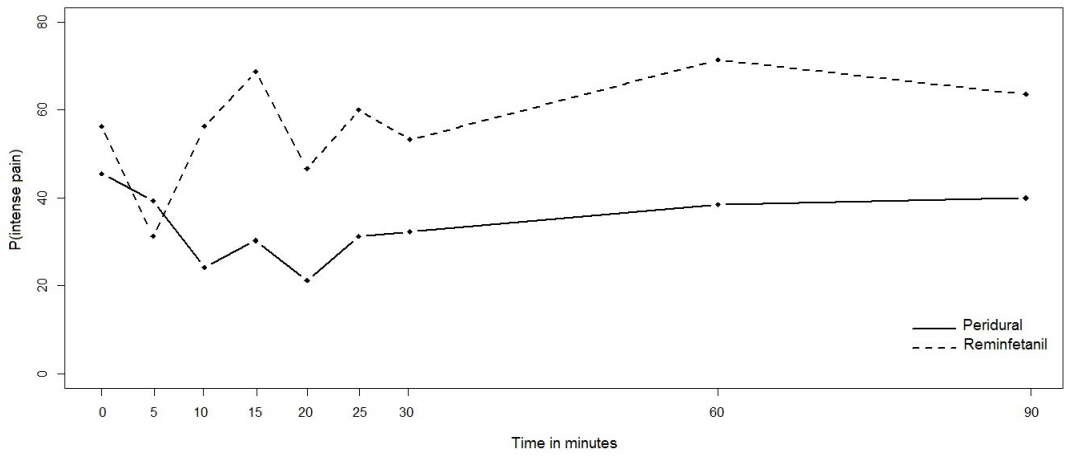

Figure 2. Probability of pain intensity in the two treatment groups for the first 90 min of follow-up

3. Longitudinal data set modeling

In a longitudinal study, participants are referred to as individuals or subjects. Accordingly, the response of the i-th individual, $i = 1, 2, \ldots, n$, taken repeatedly over time, is defined as $\mathbf{Y}_{ij}$, and may be grouped into a $m_i \times 1$ vector represented by $\mathbf{Y}_i = (\mathbf{Y}_{i1}, \ldots, \mathbf{Y}_{im_i})^T$,

with $j = 1, \ldots, m_i$, where m_i represents the number of measures over time. Here $\mathbf{Y}_{ij}$ is an ordinal categorical response assuming values from 1 to K. There is a $p \times 1$ vector of predictor variables associated with each $\mathbf{Y}_{ij}$ that may or may not change over time, i.e., $\mathbf{X}_{ij} = (\mathbf{X}_{ij1}, \ldots, \mathbf{X}_{ijp})^T$.

Marginal and conditional cumulative logit models for ordinal longitudinal data are presented in this section. GEE (Generalized Estimating Equations) is presented in Section 3.1 and Conditional model in Section 3.2. In section 3.3 the proportionality assumption of odds will be discussed.

3.1 Marginal model

Marginal model for ordinal response $\mathbf{Y}_{ij}$ can be generally defined as a partial proportional odds model (Peterson and Harrell, 1990). A natural extension of this model for longitudinal data is given by:

$$\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij} \leq k|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}^T)}{1 - \mathrm{P}(\mathbf{Y}_{ij} \leq k|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}^T)}\right] = \boldsymbol{\gamma}_k + \mathbf{X}_{ij}^T\boldsymbol{\beta} + \widetilde{\mathbf{X}}_{ij}^T\boldsymbol{\varrho}_k, \quad k = 1, 2, \ldots, K-1. \tag{1}$$

Model (1) is an extension of the proportional odds model (McCullagh, 1980; Walker and Duncan, 1967), in which the effect $\boldsymbol{\varrho}_k$ varies according to the k cutoff responses, with $\widetilde{\mathbf{X}}_{ij}$ being a subset of $\mathbf{X}_{ij}$ for which the effect varies with k. If $\boldsymbol{\varrho}_k = 0$, then model (1) reduces to the proportional odds model.

In the proportional odds model, changes in the $K-1$ cumulative logits over time are related to the predictor variables. Although the model includes $K-1$ intercepts $\boldsymbol{\gamma}_k$, it assumes that the effects of predictor variables are the same across the $K-1$ logits, which is equivalent to assuming that the effects of predictor variables on the cumulative odds are proportional (Fitzmaurice et al., 2011). One advantage of the proportional odds model is that regardless of the number of categories, the interpretation of the slope parameter is the same. Another desirable feature is that the slope parameter exponentiated is interpreted as an odds ratio.

Thus, in a study whose response has, for example, 3 ordered categories, we have 2 cumulative logits based on cumulative probabilities of the model (1):

$$\mathrm{logit}_1 = \log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij} \leq 1|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}^T)}{1 - \mathrm{P}(\mathbf{Y}_{ij} \leq 1|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}^T)}\right] = \gamma_1 + \mathbf{X}_{ij}^T\boldsymbol{\beta} + \widetilde{\mathbf{X}}_{ij}^T\boldsymbol{\varrho}_1$$

and

$$\mathrm{logit}_2 = \log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij} \leq 2|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}^T)}{1 - \mathrm{P}(\mathbf{Y}_{ij} \leq 2|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}^T)}\right] = \gamma_2 + \mathbf{X}_{ij}^T\boldsymbol{\beta} + \widetilde{\mathbf{X}}_{ij}^T\boldsymbol{\varrho}_2.$$

The logits 1 and 2, respectively, represent the log of the odds for the most favorable category compared to other categories, and the log of odds for the 2 most favorable categories compared to the last category.

The cumulative probabilities might be modelled treating the ordinal response as a set of $K-1$ variables of the form:

$$\mathbf{U}_{ijk} = \begin{cases} 1, \text{ if } & \mathbf{Y}_{ij} \leq k \\ 0, \text{ if } & \mathbf{Y}_{ij} > k. \end{cases}$$

A general specification of the marginal model for ordinal responses follows by:

(1) $\text{logit}\ (F_{ijk}) = \boldsymbol{\gamma}_k + \mathbf{X}_{ij}^T\boldsymbol{\beta} + \widetilde{\mathbf{X}}_{ij}^T\boldsymbol{\varrho}_k$, in which $F_{ijk} = \text{E}(U_{ijk}) = \text{P}(\mathbf{Y}_{ij} \leq k)$;
(2) $\text{Var}(U_{ijk}|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}) = F_{ijk}(1 - F_{ijk})$;
(3) when specifying the intra-individual association, the correlation between the components of $(\mathbf{U}_{ij,1}, \ldots, \mathbf{U}_{ij,K-1})$ in the j-th time point is a known function of F_{ijk}.

The components of $(\mathbf{U}_{ij,1}, \ldots, \mathbf{U}_{ij,K-1})$ in part 3 above are correlated, and this correlation follows from the fact that the probabilities of K multinomial responses must necessarily add up to 1 for a given point in time (Hedeker and Gibbons, 1994). Thus, for instance, the correlation between $\mathbf{U}_{ijk_1}$ and $\mathbf{U}_{isk_2}$ can be expressed by:

$$\text{Corr}(\mathbf{U}_{ijk_1}, \mathbf{U}_{isk_2}) = \frac{F_{ijk_1} - F_{ijk_1}F_{isk_2}}{\sqrt{F_{ijk_1}F_{isk_2}(1 - F_{ijk_1})(1 - F_{isk_2})}}.$$

As an alternative to using models for correlation, it is possible to specify the association between pairs of ordinal responses by making use of the overall odds ratio as a measure of association. Lipsitz et al. (1991) used the odds ratio as a measure of association when the response of interest is binary, in contrast to the model for correlation proposed by Prentice (1988), while Heagerty and Zeger (1996) considered the overall odds ratio as a measure of association when defining the marginal model for a pair of ordinal repeated responses. Thus, the overall odds ratio for the i-th individual, with responses $\mathbf{Y}_{is} = k_1$ and $\mathbf{Y}_{ij} = k_2$, is defined as:

$$\alpha_{isj}(k_1, k_2) = \frac{\overline{F}_{isj}(k_1, k_2)\left[1 - F_{isk_1} - F_{ijk_2} + \overline{F}_{isj}(k_1, k_2)\right]}{\left[F_{isk_1} - \overline{F}_{isj}(k_1, k_2)\right]\left[F_{ijk_2} - \overline{F}_{isj}(k_1, k_2)\right]}, \tag{2}$$

in which $\overline{F}_{isj}(k_1, k_2) = \text{P}(\mathbf{Y}_{is} \leq k_1, \mathbf{Y}_{ij} \leq k_2)$.

The parameter $\alpha_{isj}(k_1, k_2)$ in expression (2) can be seen as a marginal odds ratio, which seeks to capture the association between pairs of responses $\mathbf{Y}_{is}$ and $\mathbf{Y}_{ij}$ in categories k_1 and k_2, respectively.

The estimation of the parameters in the marginal model for ordinal responses is done using the GEE extension proposed by Heagerty and Zeger (1996). According to these authors, if the primary interest is to estimate the parameter $\boldsymbol{\beta}$, the 1^{st} order estimating equations are given by:

$$\mathcal{U}_1^\star(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^{n}\left[\frac{\partial\boldsymbol{\mu}_i}{\partial\boldsymbol{\beta}}\right]^T\mathbf{V}_{i11}^{-1}(\mathbf{U}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta}))$$

and

$$\mathcal{U}_2^\star(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^{n}\left[\frac{\partial\boldsymbol{\sigma}_i}{\partial\boldsymbol{\alpha}}\right]^T\mathbf{V}_{i22}^{-1}(\mathbf{S}_i - \boldsymbol{\sigma}_i(\boldsymbol{\beta}, \boldsymbol{\alpha})),$$

$\boldsymbol{\mu}_i = g^{-1}(\text{logit}(F_{ijk}))$; $\mathbf{V}_{i11} = \mathbf{A}_i^{1/2}\mathbf{R}_i(\boldsymbol{\alpha})\mathbf{A}_i^{1/2}$, where $\mathbf{A}_i$ is a diagonal matrix such that $\mathbf{A}_i = \text{diag}\{\text{Var}(U_{ijk}|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij})\}$; $\mathbf{R}_i(\boldsymbol{\alpha})$ is a $m_i \times m_i$ matrix, known as the "working correlation matrix"; $\mathbf{U}_i = (\mathbf{U}_{i1}^T, \ldots, \mathbf{U}_{im_i}^T)^T$; and $\mathbf{V}_{i22}$ denotes a covariance matrix corresponding to the following Kronecker product: $\mathbf{S}_{i(s,j)} = (\mathbf{U}_{is} - \boldsymbol{\mu}_{is}) \otimes (\mathbf{U}_{ij} - \boldsymbol{\mu}_{ij})$. Furthermore, $\sigma_i = \text{E}(\mathbf{S_i})$.

Estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ in the marginal model for ordinal responses is analogous to the Lipsitz et al. (1991) proposition. The difference for ordinal responses is that the covariance matrix for the vector of responses has a block-diagonal structure, with the covariance

determined by $\boldsymbol{\mu}_i$. Thus, the following algorithm summarizes the estimation process for GEE1 in the case of ordinal responses:

(1) Obtain initial estimates for $(\widehat{\boldsymbol{\beta}}^{(0)}, \widehat{\boldsymbol{\alpha}}^{(0)})$. In general, it is assumed that $\boldsymbol{\alpha}^{(0)} = 0$, and $\widehat{\boldsymbol{\beta}}^{(0)}$ is estimated from a proportional odds model for independent observations;
(2) Using the iterative Gauss-Seidel process, $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are updated, and the process continues until convergence is achieved, such that:

$$\widehat{\boldsymbol{\beta}}^{(m+1)} = \widehat{\boldsymbol{\beta}}^{(m)} + \left(\sum_{i=1}^{n}\left[\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right]^T \mathbf{V}_{i11}^{-1}\left[\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right]\right)^{-1}\left(\sum_{i=1}^{n}\mathcal{U}_1^{\star}(\boldsymbol{\beta}^{(m)}, \boldsymbol{\alpha}^{(m)})\right),$$

$$\widehat{\boldsymbol{\alpha}}^{(m+1)} = \widehat{\boldsymbol{\alpha}}^{(m)} + \left(\sum_{i=1}^{n}\left[\frac{\partial \boldsymbol{\sigma}_i}{\partial \boldsymbol{\alpha}}\right]^T \mathbf{V}_{i22}^{-1}\left[\frac{\partial \boldsymbol{\sigma}_i}{\partial \boldsymbol{\alpha}}\right]\right)^{-1}\left(\sum_{i=1}^{n}\mathcal{U}_2^{\star}(\boldsymbol{\beta}^{(m)}, \boldsymbol{\alpha}^{(m)})\right).$$

Parameter $\boldsymbol{\alpha}$ is treated as a perturbation in the 1^{st} order GEEs. This set of parameters was initially estimated by the method of moments (Liang and Zeger, 1986; Zeger and Liang, 1986). Carey et al. (1993) proposed the use of ALR (Alternating Logistic Regression) and the parameter $\boldsymbol{\alpha}$ was estimated by using the odds ratio as a measure of association. Heagerty and Zeger (1996) extended the method of Carey et al. (1993) to accommodate ordinal responses, and used the global odds ratio as a measure of association. In this case, a second estimation equation for parameter $\boldsymbol{\alpha}$ is given by:

$$\mathcal{U}_2^{\star\star}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{i=1}^{n}\left[\frac{\partial \xi_i}{\partial \boldsymbol{\alpha}}\right]^T \mathbf{M}_i^{-1}(\mathbf{U}_i^{\star} - \boldsymbol{\xi}_i) = 0, \tag{3}$$

where $\boldsymbol{\xi}_{i(s,j)(k_1,k_2)} = \mathrm{E}(\mathbf{U}_{is,k_1}|\mathbf{U}_{ij,k_2})$; $\mathbf{M}_i = \mathrm{diag}[\xi_i(1-\xi_i)]$; and $\mathbf{U}_i^{\star}$ is defined as:

$$\mathbf{U}_i^{\star} = ((\mathbf{U}_{i1}\otimes \mathbf{1}_K)^T, (\mathbf{U}_{i2}\otimes \mathbf{1}_K)^T, \ldots, (\mathbf{U}_{i(m_i-1)}\otimes \mathbf{1}_K)^T)^T.$$

In the Kronecker product, a vector of $\mathbf{1}$'s is used to represent the fact that the elements of vector $\mathbf{U}_{ij}$ are repeated K times. In addition, $(\mathbf{U}_i^{\star} - \xi_i)$ represents the conditional residual formed by all different pairs of ordinal responses, where $\xi_i = \mathrm{E}(\mathbf{U}_i^{\star}|\mathbf{U}_i^{\star\star})$, and $\mathbf{U}_i^{\star\star} = ((\mathbf{1}_K\otimes \mathbf{U}_{i2})^T, (\mathbf{1}_K\otimes \mathbf{U}_{i3})^T, \ldots, (\mathbf{1}_K\otimes \mathbf{U}_{im_i})^T)^T$.

Estimating equation (3) is nothing more than a regression of $\mathbf{U}_i^{\star}$ on $\mathbf{U}_i^{\star\star}$, where the conditional expectations are given by ξ_i. The estimation algorithm for ALR is similar to the one previously described for 1^{st} order GEEs, alternating between steps to update $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Estimation of $\boldsymbol{\beta}$ via ALR has the same robustness and consistency as those obtained by GEE1, the difference lies in the efficiency of parameter $\boldsymbol{\alpha}$ (Heagerty and Zeger, 1996). The asymptotic normality of $(\boldsymbol{\beta}, \boldsymbol{\alpha})$ shown by Liang and Zeger (1986) applies to the estimation for ordinal response models. Thus, the asymptotic variance of the ALR estimators is given by:

$$\mathbf{V} = \left[\sum_{i=1}^{n}\mathrm{E}\left(\mathbf{D}_{1i}^T B_{1i}\mathbf{D}_{1i}\right)\right]^{-1} \times \left\{\sum_{i=1}^{n}\mathrm{E}\left[\begin{pmatrix}\mathbf{U}_{1i}^{\star}\\ \mathbf{U}_{2i}^{\star\star}\end{pmatrix}\begin{pmatrix}\mathbf{U}_{1i}^{\star}\\ \mathbf{U}_{2i}^{\star\star}\end{pmatrix}^T\right]\right\} \times \left[\sum_{i=1}^{n}\mathrm{E}\left(\mathbf{D}_{1i}^T\mathbf{B}_{1i}\mathbf{D}_{1i}\right)\right]^{-1}, \tag{4}$$

where $\mathbf{D}_{1i} = \begin{bmatrix}\frac{\partial \mu_i}{\partial \boldsymbol{\beta}} & 0\\ 0 & \frac{\partial \xi_i}{\partial \boldsymbol{\alpha}}\end{bmatrix}$; and $\mathbf{B}_{1i} = \begin{bmatrix}\mathbf{V}_{i11} & 0\\ 0 & \mathbf{V}_{i22}\end{bmatrix}^{-1}$.

It is worth noting that the marginal model has focused on the inferences about the population mean, and it is one of the most popular methods in longitudinal modeling being solely based on assumptions about the average response.

3.2 Conditional or Mixed models

The conditional model (in the sense of a random effect) for cumulative probabilities can be described as in Hedeker and Mermelstein (2000), in terms of the following logit:

$$\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}{1-\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}\right]=\boldsymbol{\gamma}_k+\mathbf{X}_{ij}^T\boldsymbol{\beta}+\widetilde{\mathbf{X}}_{ij}\boldsymbol{\varrho}_k+\mathbf{Z}_{ij}^T\mathbf{b}_i,\quad k=1,2,\ldots,K-1,\ (5)$$

$\mathbf{Z}_{ij}$ is a q-vector of predictor variables that is a subset of $\mathbf{X}_{ij}$, usually those ones that change over time, and $\mathbf{b}_i\sim N_q(0,\mathbf{G})$. Again, if the effect $\boldsymbol{\varrho}_k$ in (5) is zero, then the expression reduces to the proportional odds mixed effects.

A general specification for the model of mixed effects in ordinal responses considers:

(1) $\text{logit}\left(\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})\right)=\boldsymbol{\gamma}_k+\mathbf{X}_{ij}^T\boldsymbol{\beta}+\widetilde{\mathbf{X}}_{ij}\boldsymbol{\varrho}_k+\mathbf{Z}_{ij}^T\mathbf{b}_i$;
(2) conditional to the vector of random effects $\mathbf{b}_i$, the values of $\mathbf{Y}_{ij}$ are independent and have a multinomial distribution for $j=1,\ldots,m_i$;
(3) the effects $\mathbf{b}_i$ are assumed to follow a multivariate normal distribution with mean 0 and a $q\times q$ covariance matrix $\mathbf{G}$ (for a model with random slopes and intercept).

Thus, the accumulated logits for a response $\mathbf{Y}_{ij}$ that has, for example, three categories can be expressed as:

$$\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}\leq 1|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}{1-\mathrm{P}(\mathbf{Y}_{ij}\leq 1|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}\right]=\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}=1|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}{\mathrm{P}(\mathbf{Y}_{ij}=2\quad\text{or}\quad 3|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}\right]=\gamma_1+\mathbf{X}_{ij}^T\boldsymbol{\beta}+\widetilde{\mathbf{X}}_{ij}^T\boldsymbol{\varrho}_1+Z_{ij}^T\mathbf{b}_i$$

and

$$\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}\leq 2|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}{1-\mathrm{P}(\mathbf{Y}_{ij}\leq 2|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}\right]=\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}=1\quad\text{or}\quad 2|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}{\mathrm{P}(\mathbf{Y}_{ij}=3|\mathbf{b}_i,\mathbf{X}_{ij},\widetilde{\mathbf{X}}_{ij})}\right]=\gamma_2+\mathbf{X}_{ij}^T\boldsymbol{\beta}+\widetilde{\mathbf{X}}_{ij}\boldsymbol{\varrho}_2+\mathbf{Z}_{ij}^T\mathbf{b}_i.$$

Therefore, for an ordinal response containing three categories, the model simultaneously describes the effects of predictor variables under all $K-1$ comparisons between the probabilities (Hedeker and Mermelstein, 2000). The inclusion of random effects in the model has implications for the interpretation of the parameters. Conditional at the random effect, a model as in (5) has the same slope, but different $\mathbf{b}_i$ for all subjects. Therefore, the interpretation of the regression coefficients in this class of models is done at individual level, and conditional to the random effect.

After the three specifications of the mixed model for ordinal responses, it follows that the joint probability for $\mathbf{Y}_i$ and $\mathbf{b}_i$ can be expressed as: $f(\mathbf{Y}_{ij},\mathbf{b}_i)=f(\mathbf{Y}_{ij}|\mathbf{b}_i)f(\mathbf{b}_i)$ and likelihood function is expressed by:

$$
\begin{aligned}
L(\mathbf{Y}_i|\boldsymbol{\beta}, \boldsymbol{\varrho}_k, \mathbf{G}) &= \prod_{i=1}^{n} \int_{\mathbf{b}_i} \prod_{j=1}^{m_i} f(\mathbf{Y}_i|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}, \mathbf{b}_i) f(\mathbf{b}_i) d\mathbf{b}_i \\
&= \prod_{i=1}^{n} \int_{\mathbf{b}_i} \prod_{j=1}^{m_i} \left[\prod_{k=1}^{K} \left(\mathrm{P}(\mathbf{Y}_i \leq k|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}, \mathbf{b}_i) - \mathrm{P}(\mathbf{Y}_i \leq k-1|\mathbf{X}_{ij}, \widetilde{\mathbf{X}}_{ij}, \mathbf{b}_i) \right)^{\mathrm{I}_{i,k}} \right] f(b_i) d\mathbf{b}_i \\
&= \prod_{i=1}^{n} \int_{\mathbf{b}_i} \prod_{j=1}^{m_i} \left[\prod_{k=1}^{K} \left(\frac{e^{\gamma_k + \mathbf{X}_{ij}^T \boldsymbol{\beta}^T + \widetilde{\mathbf{X}}_{ij}^T \boldsymbol{\varrho}_k^T}}{1 + e^{\gamma_k + \mathbf{X}_{ij}^T \boldsymbol{\beta}^T + \widetilde{\mathbf{X}}_{ij}^T \boldsymbol{\varrho}_k^T}} - \frac{e^{\gamma_{k-1} + \mathbf{X}_{ij}^T \boldsymbol{\beta}^T + \widetilde{\mathbf{X}}_{ij}^T \boldsymbol{\varrho}_{k-1}^T}}{1 + e^{\gamma_{k-1} + \mathbf{X}_{ij}^T \boldsymbol{\beta}^T + \widetilde{\mathbf{X}}_{ij}^T \boldsymbol{\varrho}_{k-1}^T}} \right)^{\mathrm{I}_{i,k}} \right] f(\mathbf{b_i}) d\mathbf{b}_i,
\end{aligned} \tag{6}
$$

where $\mathrm{I}_{i,k} = \begin{cases} 1, \text{ if } \quad \mathbf{Y} = k \\ 0, \text{ otherwise} \end{cases}$.

Parameters of the mixed model are estimated by maximizing the function in (6). It is necessary to use numerical approximation methods, such as Laplace or Gaussian quadrature, because this expression has no closed form. It is also possible to use the method of penalized quasi-likelihood (PQL). In some situations the PQL method is preferable to approximate numerical methods due to the computational time being much greater in this last method.

3.3 Proportionality Assumption

The cumulative logit model, commonly known as proportional odds model (McCullagh, 1980; Walker and Duncan, 1967), is one of the most used in the modeling of ordinal responses, and considers that each cumulative logit has its own intercept. The proportional odds model assumes that all the odds ratio are identical between the $K-1$ intercepts, that is, the model assumes that all observations have a common variance, implying that there is an approximately linear growth of the odds ratio. According to McCullagh (1980), the parameters $\boldsymbol{\gamma}_k$ are of little interest in the interpretation of the model, and are usually referred to as cut-off points. The parameter $\boldsymbol{\beta}$ describes the increase in the log of odds for any category of response. In the expression described in (1), if $\boldsymbol{\varrho}_k = 0$, we have the proportional odds model.

The assumption that $\boldsymbol{\beta}_k = \boldsymbol{\beta}$, $\forall k$, that is, the odds are proportional, can be verified individually or globally via likelihood ratio test (LRT) or a test based on score statistic. The global test evaluates only if the model violated the global assumption of proportionality of the odds. From the global test is not possible to identify which covariates violated the assumption. To assess each covariate individually, we also use a likelihood ratio test or the test based on score statistic whose asymptotic distribution in this case will be chi-square with $K-1$ degrees of freedom.

In marginal modeling the global assumption of proportionality of odds can be verified through a score test (Stiger et al., 1999), given by:

$$
S = \mathcal{U}_1^{\star T} \mathbf{V}^{-1} \mathcal{U}^{\star}{}_1, \tag{7}
$$

where $\mathcal{U}_1^{\star}$ is the 1^{st} order estimating equation; and $\mathbf{V}$ is the asymptotic variance, known as sandwich estimator. Under null hypothesis, the score statistic in (7) has an asymptotic chi-square distribution with $(K-2) \times p$ d.f., where p is the number of predictor variables in the model. When the assumption of proportionality of odds applies to some but not all of the covariates, the model (1) can be used to obtain the parameter estimates.

As in the marginal modeling, the global assumption of proportionality of odds can also be checked in the conditional modeling. In this case a LRT can be used, where the deviance of

the model assuming proportional odds (Dev_{PO}) is compared to the deviance of the partial proportional odds model (Dev_{PPO}). The test statistic is given by:

$$\begin{aligned} LRT &= -2\ln\left[L(\mathbf{Y}_i|\beta,\mathbf{G}) - L(\mathbf{Y}_i|\beta,\varrho_k,\widetilde{\mathbf{X}}_{ij},\mathbf{G})\right] \\ &= \text{Dev}_{PO} - \text{Dev}_{PPO}. \end{aligned} \tag{8}$$

Under the null hypothesis, the asymptotic distribution of statistics in (8) is chi-square with degrees of freedom given by $l = (d.f_{PPO} - d.f_{PO})$ (Hedeker and Gibbons, 1994).

4. Data applications and results

In this section the marginal and conditional models for ordinal response presented previously are used in the two real data applications, which were described in subsections 2.1 and 2.2.

Adjustments were made using the R-3.2.0 (R Core Team, 2015) and SAS 9.0 (SAS Institute Inc., 2002) softwares. For fitting the marginal proportional odds model, the functions `ordgee`$(\cdot)$, from the `geepack` library, and `repolr`$(\cdot)$, from the `repolr` library, were used in R. Such functions use an extended method from those proposed by Heagerty and Zeger (1996). The `repolr`$(\cdot)$ function also tests, through score statistics, the assumption of proportional odds. Different structures (exchangeable, unstructured, and independent) were used for the working matrix $R(\alpha)$. For fitting the marginal partial proportional odds model in SAS, the Proc GENMOD was used. The functions `clmm`$(\cdot)$ and `clmm2`$(\cdot)$ from the `ordinal` library in R software were used for fitting the conditional models. Adaptive Gaussian quadrature, with 50 points, was considered. Syntax for fitting the models described in this paper using R and SAS are available in the supplementary online material.

4.1 Rainwater harvesting systems and child health

In this application the response variable is the parasitic level (1: poly-infected child; 2: monoinfected child; 3: uninfected child) of the child i adjusted at each phase $j = 1, 2, 3$ of the study. The predictors variables considered were the treatment group, phase of the study and age of child.

The global assumption of proportional odds was evaluated for the marginal model by the score test, and the test indicated that the assumption was violated (p-value=0.014). For the conditional model the global assumption of proportionality of odds was tested by LRT test, and it was also rejected (p-value<0.001). In both models all variables violated the assumption of proportionality of odds (p-values< 0.05), so $X_{ij}^T\boldsymbol{\beta} = 0$. For this reason, it was decided to consider the partial proportional odds model, which ensures that the effects of all variables vary according to the category of response. The partial proportional odds models considered in the final adjustment were:

– **Marginal model:**

$$\log\left[\frac{\text{P}(\mathbf{Y}_{ij} \leq k|\widetilde{\mathbf{X}}_{ij})}{1-\text{P}(\mathbf{Y}_{ij} \leq k|\widetilde{\mathbf{X}}_{ij})}\right] = \boldsymbol{\gamma}_k + \varrho_{1k}\text{group}_{ij} + \varrho_{2k}\text{phase}_j + \varrho_{3k}\text{age}_i.$$

– **Conditional models:**

$$\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_i,\widetilde{\mathbf{X}}_{ij})}{1-\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_i,\widetilde{\mathbf{X}}_{ij})}\right]=\boldsymbol{\gamma}_k+\varrho_{1k}\mathrm{group}_{ij}+\varrho_{2k}\mathrm{phase}_j+\varrho_{3k}\mathrm{age}_i+b_i,$$

$$\log\left[\frac{\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_{1i},\mathbf{b}_{2i},\widetilde{\mathbf{X}}_{ij})}{1-\mathrm{P}(\mathbf{Y}_{ij}\leq k|\mathbf{b}_{1i},\mathbf{b}_{2i},\widetilde{\mathbf{X}}_{ij})}\right]=\boldsymbol{\gamma}_k+\varrho_{1k}\mathrm{group}_{ij}+\varrho_{2k}\mathrm{phase}_j+\varrho_{3k}\mathrm{age}_i+b_{1i}+b_{2i}\mathrm{phase}_j,$$

for $k=1,2,3$, $j=1,2,3$ and $i=1,\ldots,635$.

The marginal partial proportional odds model was adjusted considering the structures independent and exchangeable for the working matrix $\mathrm{R}(\alpha)$. We were unable to adjust the unstructured form for $\mathrm{R}(\alpha)$ since this structure was not available in the proc GENMOD (SAS 9.0). Due to the significance of the parameter α in exchangeable structure, this was the chosen structure for $\mathrm{R}(\alpha)$. Table 1 presents partial proportional odds estimates. The interpretation of the estimated parameters in this model considers that the $\boldsymbol{\varrho}_k$ effect differs for each cumulative logit. For example, it can be observed that the odds of poly-infection for children in house with cistern is $1.95(e^{0.667})$ times the odds of those who did not have one. Calculating the cumulative effect of mono or poly infection, it is observed that the odds decreases ($e^{-0.103}=0.902$) for children who have cistern at home compared to children who did not have one. There is an increased odds of poly-infection for each one-month increase in the child's age ($e^{0.014}=1.014$), the same was true for the odds of mono or poly-infection ($e^{0.020}=1.02$). Across different phases, there is a reduction in the odds of mono or poly-infection ($e^{-0.036}=0.96$). Exchangeable association structure indicates that the odds of a child with mono or poly-infection in one phase remains in such condition in the next phase is 2 ($e^{0.692}$). Because of the association structure, there is an equal odds that a non-infected child in a particular phase remains in such condition in the next phase.

Table 1. Estimates for the marginal partial proportional odds model in the study on rainwater harvesting systems and child health

	$R(\alpha)$ independent			$R(\alpha)$ exchangeable		
Variable	$\widehat{\beta}$	SE	p-value	$\widehat{\beta}$	SE	p-value
γ_1	-2.312	0.311	-	-2.342	0.311	-
γ_2	-1.478	0.152	-	-1.506	0.152	-
Group_1 (with cistern)	0.667	0.203	0.001	0.667	0.203	0.001
Group_2	-0.127	0.117	0.278	-0.103	0.117	0.377
Age_1	0.013	0.008	0.080	0.014	0.007	0.065
Age_2	0.020	0.004	<0.001	0.020	0.004	<0.001
Phase_1	-0.037	0.185	0.044	-0.036	0.018	0.049
Phase_2	0.005	0.011	0.633	0.001	0.011	0.607
$\widehat{\alpha}^*$	-			0.692	0.137	<0.001

* intra-individual association based on the log of odds ratio

The conditional partial proportional odds model considers random effects only in the intercept (lowest AIC when compared to the model with random effect in the intercept and slope) (Table 2).

It can be observed from the conditional partial proportional odds model results that given two children with the same random effects, the odds of poly-infection are greater

Table 2. Parameter estimates for the conditional partial proportional odds model in the study on rainwater harvesting systems and child health.

Variable	Conditional random intercept		Conditional random intercept and slope	
	$\widehat{\beta}$	SE	$\widehat{\beta}$	SE
γ_1	-3.623	0.432	-3.614	0.416
γ_2	-1.526	0.214	-1.522	0.215
Group_1(with cistern)	0.557	0.214*	0.511	0.215*
Group_2(with cistern)	-0.124	0.118	-0.123	0.118
Age_1	0.033	0.008*	0.033	0.008*
Age_2	0.019	0.004*	0.019	0.004*
Phase_1	-0.191	0.135	-0.177	0.134
Phase_2	0.032	0.070	0.029	0.071
$g_{11} = \text{Var}(b_{1i})$	0.046		0.046	
$g_{22} = \text{Var}(b_{2i})$			0.042	
AIC	2198.98		2200.44	

* p-value significant

among those in the group with cistern ($e^{0.557} = 1.75$) as compared to those in the group without cistern. Moreover, conditioned on the random effect, the odds of poly-infection increases with each additional month of child's age ($e^{0.033} = 1.04$). The same occurs with the chance of mono or poli infection ($e^{0.019} = 1.02$). It is important to note that although the phase of the study was not statistically significant, there is a reduction ($e^{-0.191} = 0.826$) in the odds of poly-infection over time among children who have the same random effect.

4.2 Analgesia in childbirth

For this application the response variable is pain intensity (1: tolerable and mild pain; 2: moderate pain; 3: intense and unbearable pain) for the parturient i fitted at each time $j = 0, 5, 10, 15, 20, 25, 30, 60, 90$, in minutes. The predictors were the treatment group, age of parturient and the time until the child-birth.

For the marginal model, the global assumption of odds proportionality was tested using the score test, and it was not rejected (p-value=0.07). The same occurred with the conditional model using the LTR test (p-value=0.08). Thus, in this application $\widetilde{\mathbf{X}}_{ij}\boldsymbol{\varrho}_k = 0$, and the following proportional odds models are assumed to be the final models:

– **Marginal model**

$$\log\left[\frac{\text{P}(\mathbf{Y}_{ij} \leq k|\mathbf{X}_{ij})}{1-\text{P}(\mathbf{Y}_{ij} \leq k|\mathbf{X}_{ij})}\right] = \gamma_k + \beta_1\text{group}_i + \beta_2\text{age}_i + \beta_3\text{time}_j.$$

– **Conditional models**

$$\log\left[\frac{\text{P}(\mathbf{Y}_{ij} \leq k|\mathbf{b}_i, \mathbf{X}_{ij})}{1-\text{P}(\mathbf{Y}_{ij} \leq k|\mathbf{b}_i, \mathbf{X}_{ij})}\right] = \gamma_k + \beta_1\text{group}_i + \beta_2\text{age}_i + \beta_3\text{time}_j + b_i,$$

$$\log\left[\frac{\text{P}(\mathbf{Y}_{ij} \leq k|\mathbf{b}_{1i}, \mathbf{b}_{2i}, \mathbf{X}_{ij})}{1-\text{P}(\mathbf{Y}_{ij} \leq k|\mathbf{b}_{1i}, \mathbf{b}_{2i}, \mathbf{X}_{ij})}\right] = \gamma_k + \beta_1\text{group}_i + \beta_2\text{age}_i + \beta_3\text{time}_j + b_{1i} + b_{2i}\text{time}_{ij},$$

for $k = 1, 2, 3$ e $i = 1, \ldots, 49$.

Three different structures for the matrix $R(\boldsymbol{\alpha})$ were considered for the marginal model. In the third model, 36 combinations of the 9 measures of time are considered, and thus $\boldsymbol{\alpha}$ is formed by 36 correlation parameters. Results in Table 3 were very similar for the three structures. The second model (exchangeable structure) was chose for a matter of interpretation and parsimonious, since the single component of $\boldsymbol{\alpha}$ was statistically significant.

Effect of group was the only one statistically significant, indicating that the odds of feeling slight pain will be less ($e^{-1.485} = 0.23$) among women who receive reminfentanil compared to those who receive peridural analgesia. Because the odds are proportional, this also expresses the odds that a woman in labor who receives reminfentanil experience mild to moderate pain compared to those who receive a peridural analgesia. Since age is negatively associated to pain intensity, we believe that the unadjusted results presented in Figure 2 may be spurious, because we observed that the women receiving epidural anesthesia are slight older than those receiving remifentanil. The intra-individual association (α) was statistically significant (p-value<0.001), revealing a change in the odds of a patient to remain in a certain response category from one follow-up time to another. In such a case, if there were not so many follow-up times to be estimated, it is advisable to consider the unstructured association structure.

Table 3. Parameter estimates for the marginal proportional odds model in the study of analgesia in childbirth.

	$R(\alpha)$ independence			$R(\alpha)$ exchangeable			$R(\alpha)$ unstructured		
Variable	$\widehat{\beta}$	SE	p-value	$\widehat{\beta}$	SE	p-value	$\widehat{\beta}$	SE	p-value
γ_1	0.151	0.737	-	0.232	0.776	-	0.253	0.776	-
γ_2	1.797	0.757	-	1.864	0.794	-	1.885	0.792	-
Group (Remifentanil)	-1.398	0.392	0.001	-1.485	0.438	0.001	-1.523	0.491	0.002
Age	-0.045	0.036	0.208	-0.047	0.038	0.207	-0.051	0.037	0.178
Time	-0.004	0.004	0.922	-0.002	0.004	0.738	-0.002	0.004	0.641
$\widehat{\alpha}^*$	-			1.023	0.219	<0.001	$\widehat{\alpha}_{0,5}$=1.900	0.764	0.013
							$\widehat{\alpha}_{0,10}$=1.076	0.627	0.086
							$\vdots$		
							$\widehat{\alpha}_{60,90}$=0.921	0.711	0.195

*intra-individual association based on the log of odds ratio

Results for the conditional model are presented in Table 4. Comparing the first and the second models in this table, the LRT was statistically significant (p-value< 0.0001). This fact is an indication that the variance of the random effect is greater than 0. Using AIC, the second model is chosen among the three adjusted models. Thus, the final model in this case would be the conditional model with a random intercept.

Table 4. Parameter estimates for the conditional proportional odds model in the study of analgesia in childbirth.

	GLM			Conditional random intercept			Conditional random intercept and slope		
Variable	$\widehat{\beta}$	SE	p-value	$\widehat{\beta}$	SE	p-value	$\widehat{\beta}$	SE	p-value
γ_1	0.171	0.401	-	0.432	0.745	-	0.428	0.732	-
γ_2	1.785	0.412	-	2.352	0.757	-	2.345	0.744	-
Group (Remi)	-1.078	0.205	<0.001	-1.322	0.375	<0.001	-1.318	0.371	<0.001
Age	-0.046	0.018	0.010	-0.057	0.032	0.076	-0.056	0.031	0.072
Time				-0.005	0.004	0.234	-0.005	0.004	0.235
$g_{11} = Var(b_{1i})$				0.961			0.928		
$g_{22} = Var(b_{2i})$							0.516		
LRT*						<0.001			0.99
AIC		849.79			820.53			823.34	

* on H$_0$ $\sigma^2_{b_{li}} = 0, \quad l = 1, 2$

For the chosen conditional model, it was observed that the effect of group was the only one being statistically significant. For this variable we have that, given two parturients

with the same random effect, the odds to feel pain is lower ($e^{-1.322} = 0.27$) among those who used reminfetanil when compared to those who opted for epidural analgesia. And just like in the marginal model, the negative effect on age variable may be indicative of spurious results, since women who received epidural anesthesia are slightly older than women who received reminfetanil. Moreover, conditioned on the random effects, the odds of experiencing pain decreases ($e^{-0.057} = 0.94$) for each year that adds the age of the parturient.

5. Discussion

Ordinal responses modeling has been the subject of increasing interest in recent years, and many methods have been proposed in the literature. The proportional odds model McCullagh (1980) is the most known model. It assumes that the odds in all categories of the ordinal response are the same. Unlike the model for nominal polytomous responses, the interpretations of the parameters are not obtained directly, considering one of the response categories as a reference. In the proportional odds model, the reference category changes according to the logit that is being analyzed. This happens because the calculated probabilities are accumulated. In this model, it is important to test the assumption of proportionality of odds, so that it can be assured that the estimates are consistent. In situations where this assumption is violated, it is recommended to search for another suitable methodology for ordinal responses. Existing models include the partial proportional odds model (Peterson and Harrell, 1990), which is a generalization of the model proposed by McCullagh (1980). This model allows a subset of predictor variables to violate the proportionality assumption, with their effects varying according to the category of the ordinal response. The model also includes a subset of predictor variables that do not violate the proportionality assumption.

Studies that discuss the modeling of ordinal responses mostly consider just the theoretical aspects of the proportional odds models. The emphasis on this methodology might be due to its availability in most statistical software. Little attention, however, is given to the interpretation of the fitted model, which could be helpful to improve understanding of these approaches. In practice, the ordinal response is often recategorized as a binary variable and used in models in such a way that have direct interpretation.

Currently few studies addressing the diagnostic techniques for models for ordinal responses are found in the literature. O'Connell and Liu (2011) provided a review of diagnostic methods that may be useful in investigating model assumptions and in identifying unusual cases for proportional and partial proportional odds models. A common method used to test model fit for categorical responses is to compare observed frequencies and estimated expected frequencies under the assumed model via chi-squared-type goodness-of-fit statistics. Lipsitz et al. (1996) generalized the popular Hosmer-Lemeshow statistic for the binary logistic regression model to the situation with ordinal responses. Fagerland and Hosmer (2013) examined goodness-of-fit tests for the proportional odds model by deriving test statistics based on the Hosmer-Lemeshow test for binary logistic regression. Graphical and numerical methods for verifying the adequacy of the proportional odds model have been proposed, focusing on the assessment of bad functional specification to the effect of specific covariates using observed cumulative residuals (Arbogast and Lin, 2005; Liu et al., 2009). Nevertheless, all aforementioned methodologies were discussed for analysis of cross-sectional ordinal data.

In longitudinal studies, the usual models (marginal models and conditional) can be extended to accommodate the ordinal nature of the response. It is important to emphasize that the choice of the most appropriate methodology depends on the type of question that the researcher wants to answer. Thus, such models are not comparable because the focus

of the analysis differs for different models, and different models are interpreted in different ways. The marginal model, for example, has a focus on inferences of population mean, and it is one of the most popular methods in longitudinal modeling because it is based exclusively on assumptions about the mean response. In contrast, the conditional method focuses on the individual and on the introduction of a random effect at the individual level, which has important implications for estimation and interpretation of the regression coefficients. However, the choice of the best model depends on the purpose of the study.

An additional important issue for analysis of longitudinal data is the potential of incomplete information over time. In longitudinal analysis missing data is very common and in some applications inevitable. Thus, the methodologies for analysis of longitudinal data deal with missing data and still provide consistent estimates depending upon the missing data mechanism. In case of GEE, estimates are consistent under missing completely at random (MCAR) assumption. Usually MAR is more genuine assumption about the missing mechanism, but it cannot be checked from the data. For fitting GEE to our data we used geepack package in R, which assumes MCAR. Multilevel models have the advantage of being applicable under weaker assumption regarding the missing mechanism than GEE. In general, potential bias is avoided with maximum likelihood inference under other mechanism than MNAR (missing not at random) (Agresti, 2010). Particularly regarding our data, there are a few children with missing data regarding the outcome in some time points in the first application, but the missing pattern seems to be at random. For the second application, the women only has information up to delivery. Thus, if a woman delivered in 30 minutes, she will not have measurement of pain intensity in the later time points. In our data 3 women delivered up to 30 minutes, with the fastest having a baby in 20 minutes. On the other hand, there are 31 (63.3%) women with complete data, i.e., information up to 90 minutes. In both cases, however, we have strong evidence to believe that the missing data patterns are not MNAR.

Of the two applications discussed in the paper, the assumption of proportionality of odds was violated only in the first study, and, therefore, the partial proportional odds model was used in that case. The two applications differ in sample size, in the balance of data, in the number of repeated measurements per individual, and also in the equidistance of the time measurement. Considering the results from our applications, we observed that estimates of both models tend to be similar in the second application, which has a larger number of measurements per subject. Further research and simulation studies are still required to evaluate if the number of repeated longitudinal measures may lead to closer conditional and marginal estimates. Such a definitive conclusion can not be drawn based on only two examples. It is important to mention that there are no large computational cost in the fit of both models. The longer time (no more than 5 minutes) is spent in fitting the conditional model depending on the number of points used in the adaptive Gaussian quadrature.

In our applications both data got some limitations. One of the limitations is the potential of several confounding factors that were not taken into account into these analyses. For the first application about the effect of rainwater harvesting systems on children's infections, factors such access to other water facilities, children's immune factors, hygiene practices (such as the use of boiled water), sanitation conditions, among others, may affect our ability to conclude about the intervention effect. Moreover, some children have moved from an area with cistern to other without cistern during the study, and vice-versa. For the second application about the effect of analgesia in labor pain, factors such as number of previous births (primiparous vs multiparous mothers), educational level, pre-natal care, among others, may also be important for this evaluation. Another limitation is the sample size, notably regarding the second data. Nevertheless, since the major goal of the paper is to systematize and summarize statistical methodologies for modeling ordinal polytomous

responses in longitudinal analysis, these applications are mainly included for illustration of the methodologies, particularly regarding interpretation of parameter estimates and model assumptions.

Discussion about sample size and/or statistical power is available in the literature for ordinal logistic models for independent observations (Abreu et al., 2008; Capuano et al., 2007; Fagerland and Hosmer, 2013; Liu and Agresti, 2005; Rabbee et al, 2003; Walters, 2004). Particularly regarding the tests for evaluation of proportional odds assumption, these tests have been criticized for having a tendency to reject the null hypothesis in cases with minimum deviation from proportionality for analysis of independent observations (Agresti, 2010). In such cases, some authors recommend plotting the log odds generated by each cut-point (Capuano et al., 2007); others recommend to use p-value < 0.10 to indicate that the models need further scrutiny. There is no doubt that fitting ordered logistic regression using maximum likelihood estimates requires sufficient sample size. However, there is still no consensus about how big the sample sizes should be for analysis of longitudinal ordinal data. In our second application, even though we have only 49 subjects, there are 408 observations being analyzed due to its longitudinal feature, with mean number of observations per subject being 8.3 ($min = 4, max = 9$). Research is still required to extend known results from settings with independent observations to longitudinal designs.

In this paper, we aimed helping researchers to improve their understanding on available methodology for analysis of longitudinal ordinal responses, emphasizing interpretation through discussion of real data applications.

Further research and discussion is still needed for the modeling of ordinal responses in longitudinal studies. One idea for future work would be to explore diagnostic techniques for ordinal responses, a topic not yet explored in studies with this type of data. Furthermore, other models could be used for analysis of ordinal responses in longitudinal studies, such as the class of models known as transition models (Azzalini, 1994; Lee and Daniels, 2007).

Acknowledgements

This work was partially supported by CNPq, CAPES and FAPEMIG grants. Analgesia data set was kindly supplied by E. S. C. Soares.

References

Abreu, M.N.S., Siqueira, A.L., Cardoso, C.S. and Caiaffa, W.T. 2008. Ordinal logistic regression models: application in quality of life studies. Cadernos de Saúde Pública, 24, 581-591.

Agresti, A. 2010. Analysis of ordinal categorical data. Wiley, New York.

Ananth, C.V. and Kleinbaum, D.G. 1997. Regression models for ordinal responses: a review of methods and applications. International Journal of Epidemiology, 26, 1323-1333.

Arbogast, P.G. and Lin, D.Y. 2005. Model-checking techniques for stratified case-control studies. Statistics in Medicine, 24, 229-247.

Azzalini, A. 1994. Logistic regression for autocorrelated data with application to repeated measures. Biometrika, 81, 767-775.

Capuano, A.W., Dawson, J.D. and Gray, G.C. 2007. Maximizing power in seroepidemiological studies through the use of the proportional odds model. Influenza and Other Respiratory Viruses, 1, 87-93.

Carey, V., Zeger, S.L. and Diggle, P. 1993. Modelling multivariate binary data with alternating logistic regressions. Biometrika, 80, 517-526.

Clayton, D. 1992. Repeated ordinal measurements: a generalised estimating equation approach. Medical Research Council Biostatistics Unit Technical Report.

Fagerland, M.W. and Hosmer, D.W. 2013. A goodness-of-fit test for the proportional odds regression model. Statistics in Medicine, 32, 2235-2249.

Fitzmaurice, G.M., Laird, N.M. and Ware, J.H. 2011. Applied Longitudinal Analysis. Wiley, New York.

Fonseca, J.E. 2012. Implantação de cisternas para armazenamento de água de chuva e seus impactos na saúde infantil: Um estudo de coorte em Berilo e Chapada do Norte, MG. Master Thesis, Department of Sanitary Engineering, Federal University of Minas Gerais, Brazil.

Heagerty, P.J. and Zeger, S.L. 1996. Marginal regression models for clustered ordinal measurements. Journal of the American Statistical Association, 91, 1024-1036.

Hedeker, D. and Gibbons, R.D. 1994. A random-effects ordinal regression model for multilevel analysis. Biometrics, 50, 933-944.

Hedeker, D. and Gibbons, R.D. 2006. Longitudinal data analysis. Wiley, New York.

Hedeker, D. and Mermelstein, R.J. 1998. A multilevel thresholds of change model for analysis of stages of change data. Multivariate Behavioral Research, 33, 427-455.

Hedeker, D. and Mermelstein, R.J. 2000. Analysis of longitudinal substance use outcomes using ordinal random-effects regression models. Addiction, 95, 381-394.

Lee, K. and Daniels, M.J. 2007. A class of Markov models for longitudinal ordinal data. Biometrics, 63, 1060-1067.

Liang, K.-Y. and Zeger, S.L. 1986a. Longitudinal data analysis using generalized linear models. Biometrika, 73, 13-22.

Lipsitz, S.R., Laird, N.M. and Harrington, D.P. 1991. Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association. Biometrika, 78, 153-160.

Lipsitz, S.R., Fitzmaurice, G.M. and Molenberghs, G. 1996. Goodness-of-fit tests for ordinal response regression models. Applied Statistics, 45, 175-190.

Liu, I., Mukherjee, B., Suesse, T., Sparrow, D., Park, S.K. and Harrington, D.P. 2009. Graphical diagnostics to check model misspecification for the proportional odds regression model. Statistics in Medicine, 28, 412-429.

Liu, I. and Agresti, A. 2005. The analysis of ordered categorical data: An overview and a survey of recent developments. Test, 14, 1-73.

McCullagh, P. 1980. Regression models for ordinal data. Journal of the Royal Statistical Society, Series B, 42, 109-142.

O'Connell, A.A. and Liu, X. 2011. Model Diagnostics for proportional and partial proportional odds models. Journal of Modern Applied Statistical Methods, 10, 1-15.

Parsons, N.R., Costa, M.L., Achten, J. and Stallard, N. 2009. Repeated measures proportional odds logistic regression analysis of ordinal score data in the statistical software package R. Computational Statistics & Data Analysis, 53, 632-641.

Peterson, B. and Harrell Jr, F.E. 1990. Partial proportional odds models for ordinal response variables. Journal of Applied Statistics, 39, 205-217.

Prentice, R.L. 1988. Correlated binary regression with covariates specific to each binary observation. Biometrics, 44, 1033-1048.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. URL: http://www.R-project.org.

Rabbee, N., Coull, B.A., Mehta, C., Patel, N. and Senchaudhuri, P. 2003. Power and sample size for ordered categorical data. Statistical Methods in Medical Research, 12, 73-84.

SAS Institute Inc. 2002. SAS/STAT Software. URL: http://www.sas.com.

Soares, E.S.C. 2013. Estudo comparativo entre analgesia venosa com reminfetanil em infusão contínua e analgesia peridural intermitente com cateter para alívio da dor durante o trabalho de parto. Master Thesis, Medicine Faculty, Federal University of Minas Gerais, Brazil.

Stiger, T.R., Barnhart, H.X. and Williamson, J.M. 1999. Testing proportionality in the proportional odds model fitted with GEE. Statistics in Medicine, 18, 1419-1433.

Walker, S.H. and Duncan, D.B. 1967. Estimation of the probability of an event as a function of several independent variables. Biometrika, 54, 167-179.

Walters, S.J. 2004. Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the SF-36. Health and Quality of Life Outcomes, 2, 1-26.

Zeger, S.L. and Liang, K.-Y. 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics, 42, 121-130.