

Mery Natali Silva Abreu<sup>I,II</sup>

Arminda Lucia Siqueira<sup>III</sup>

Waleska Teixeira Caiiffa<sup>I,II</sup>

# Ordinal logistic regression in epidemiological studies

---

## ABSTRACT

Ordinal logistic regression models have been developed for analysis of epidemiological studies. However, the adequacy of such models for adjustment has so far received little attention. In this article, we reviewed the most important ordinal regression models and common approaches used to verify goodness-of-fit, using R or Stata programs. We performed formal and graphical analyses to compare ordinal models using data sets on health conditions from the National Health and Nutrition Examination Survey (NHANES II).

**DESCRIPTORS:** Statistics as Topic. Logistic Models. Regression Analysis. Epidemiologic Methods.

---

## INTRODUCTION

Ordinal logistic regression models have been applied over the last few years for analyzing data, the response or outcome of which is presented in ordered categories. Ordered information in score-form has been increasingly used in epidemiological studies, such as quality of life in interval scales, health condition indicators and even for indicating the seriousness of illnesses.<sup>1</sup> Depending on the study's purpose, these models also allow the odds ratio (OR) statistic or the probability of the occurrence of an event to be calculated.<sup>1</sup>

There are several ordinal models, such as the proportional odds, partial proportional odds, continuation-ratio and stereotype logistic models. Despite this diversity and the great variety of studies<sup>1-4,7-15</sup> on the subject their use in the public health area is still rare.<sup>1</sup> This may be attributed not only to their complexity, but specially to the difficulty encountered when it comes to validating their assumptions.<sup>11</sup> Another factor that may be related to the limited use of these models is the reduced number of modeling options offered in the commercial statistical packages used in the public health area, examples being SPSS and Minitab. Even if other, more complex packages are used, such as SAS and Stata, it is frequently difficult to select the appropriate commands and to interpret the results.<sup>3</sup> Added to these difficulties is the considerable cost of the majority of the commercial statistical packages available, because access to their licenses is highly restricted.

One statistical package that has become more and more popular is the free software R,<sup>6,a</sup> which is distributed under a general public license. The package contains a variety of statistical techniques, including several ordinal logistic regression models, which allow them to be tested and adjustments to be compared.

<sup>I</sup> Programa de Pós-graduação em Saúde Pública. Faculdade de Medicina (FM). Universidade Federal de Minas Gerais (UFMG). Belo Horizonte, MG, Brasil

<sup>II</sup> Grupo de Pesquisa em Epidemiologia e Observatório de Saúde Urbana. FM-UFMG. Belo Horizonte, MG, Brasil

<sup>III</sup> Departamento de Estatística. Instituto de Ciências Exatas. UFMG. Belo Horizonte, MG, Brasil

### Correspondence:

Mery Natali Silva Abreu  
Av. Alfredo Balena, 190, 6º andar  
Sala 625 Santa Efigênia  
31130-100 Belo Horizonte, MG, Brasil  
E-mail: merynatali@yahoo.com.br

Received: 5/10/2007  
Revised: 5/23/2008  
Approved: 6/11/2008

---

<sup>a</sup> The R Project for Statistical Computing [internet]. Viena: Viena University of Economics and Business Administration, [s.d.] [cited 2008 Sep 7]. Available from: <http://www.r-project.org/about.htm>

The aim of this article was to analyze the adjustment and adaptation of the main ordinal regression models and show the commands used in the R software. In addition to this assessment/analysis the partial proportional odds model was adjusted, using Stata software, because it has not yet been included in the R software.

To provide examples for the methods analyzed, some of the data from the well-known Second National Health and Nutrition Examination Survey (NHANES II) will be used. This survey is available on the internet<sup>a</sup> and it is widely used as examples in statistical and epidemiological studies. The survey includes demographic, anthropometric, nutritional history, health and hematology information. Information in the database about children, extracted from 10,337 interviews with people between the ages of 20 and 74, was excluded.

## UNIVARIATE ANALYSIS

As in any analytical procedure that uses regression models, multiple analysis using ordinal models must always be preceded by comparing each covariable with the event of interest being investigated. By means of this analysis, which is known as univariate, it is possible to select the factors that will be introduced into the regression model.

The chi-squared test for trend is one of those that is suitable for selecting principal effects, since it considers the ordinal nature of the response variable. Normally, a conservative level of significance is used (generally between 10% and 25%) for entering the covariables in the model.<sup>10</sup>

Furthermore, OR can be estimated considering one response variable category as a reference point and comparing it with the others or grouping the larger categories and comparing them with the smaller categories.

## ORDINAL REGRESSION MODELS

After univariate analysis, the final multiple regression model should be constructed to control possible confusion factors. As the event of interest is ordinal, an ordinal logistic regression model must be used.

Let  $Y$  be the response variable with  $k$  categories codified as  $1, 2, \dots, k$ .  $\mathbf{X} = (x_1, x_2, \dots, x_p)$  and the vector of the explanatory variables or covariables. The  $k$  categories of  $Y$  that are conditional on  $\mathbf{X}$  the values occur with probabilities  $p_1, p_2, \dots, p_k$ , in other words  $p_j = P(Y = j)$ , for  $j = 1, 2, \dots, k$ . The term  $\alpha$  refers to the intercept of the model and  $\beta$  corresponds to the effects of the covariables on the response variable. Table 1 gives the forms

of the main models, an indication of their use and their commands in the R and Stata packages.

### Proportional odds model (POM)

In the MOP ( $k - 1$ ) cut-off points of the categories are considered, with the  $j$ th ( $j = 1, \dots, k - 1$ ) cut-off point being based on a comparison of the accumulated probabilities, as shown in Table 1.

The term  $\alpha_j$  varies for each of the  $k$  categories and each  $\beta$  does not depend on the  $j$  index, implying that the relation between  $\mathbf{X}$  and  $Y$  is independent of the category.

Therefore, the model has a proportional odds assumption around the ( $k - 1$ ) cut-off points, also called the parallel regression assumption, which is assumed for each covariable included in the model. This assumption must be tested for each covariable separately and in the final model, using for example the score test.<sup>10</sup>

This model is suitable for analyzing ordinal variables arising from a continuous variable, which in turn has been grouped.

### Partial proportional odds model (PPOM)

As the proportional odds assumption is difficult to achieve in practice, the PPOM may be used as an alternative.<sup>13</sup> This model allows some covariables with the proportional odds assumption to be modeled, but for those variables in which this assumption is not satisfied it is increased by a coefficient ( $\gamma$ ), which is the effect associated with each  $j$ th cumulative logit, adjusted by the other covariables.<sup>10</sup> The general form of the model is the same as the previous one, but now the coefficients are associated with each category of the response variable.

It is normally expected that there will be a type of linear trend between each OR of the specific cut-off points and the response variable.<sup>1</sup> If there is then a set of restrictions ( $\gamma_{kj}$ ) may be included in the model to clarify this linearity (Table 1). When these restrictions are included this model is called the restricted partial proportional odds model.

The  $\tau_j$  parameters are fixed scale parameters which take the form of restrictions allocated to the parameters. In this case for a given covariable  $X_m$ ,  $\alpha_m$  does not depend on the cut-off points, but is multiplied by  $\tau_j$  for each  $j$ th logit.<sup>11</sup>

### Continuous ratio model (CRM)

This model allows for a comparison to be made between the probability of a response equal to the category with

<sup>a</sup> National Center for Health Statistics. Publications and Information Products: NHANES II public-use and data files. [cited 2008 Nov 11] Available from: [http://www.cdc.gov/nchs/products/elec\\_prods/subject/nhanesii.htm](http://www.cdc.gov/nchs/products/elec_prods/subject/nhanesii.htm)

**Table 1.** Information about ordinal logistic regression models, indication of use, and commands when applying the R and Stata, using NHANES II as an example.

Model	Functional form of the model	Use indication	R commands*	Stata commands
POM	$\lambda_j(\underline{x}) = \ln \left\{ \frac{\Pr(Y = 1   \underline{x}) + \dots + \Pr(Y = j   \underline{x})}{\Pr(Y = j + 1   \underline{x}) + \dots + \Pr(Y = k   \underline{x})} \right\} = \ln \left\{ \frac{\sum_{j=1}^j \Pr(Y = j   \underline{x})}{\sum_{j=j+1}^k \Pr(Y = j   \underline{x})} \right\}$ $\lambda_j(\underline{x}) = \alpha_j + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p), \quad j = 1, \dots, k-1$	Variable response, originally continuous and then grouped and proportional chances assumption is valid	<pre>&gt; library (Design) &gt; mcp=lm (health~age + diabetes + color_ skin,nhanes,x,TRUE,y=TRUE) &gt; mcp &gt; par(mfrow=c(1,5)) &gt; residuals (mcp, type = 'score.binary', pl=TRUE) &gt; residuals (mcp, type= 'partial', pl=TRUE)</pre>	<pre>.ologit Health age diabetes color_ skin</pre>
PPOM-NR	$\lambda_j(\underline{x}) = \ln \left\{ \frac{\Pr(Y = 1   \underline{x}) + \dots + \Pr(Y = j   \underline{x})}{\Pr(Y = j + 1   \underline{x}) + \dots + \Pr(Y = k   \underline{x})} \right\} = \ln \left\{ \frac{\sum_{j=1}^j \Pr(Y = j   \underline{x})}{\sum_{j=j+1}^k \Pr(Y = j   \underline{x})} \right\}$ $\lambda_j(\underline{x}) = \alpha_j + [(\beta_1 + \gamma_1)x_1 + \dots + (\beta_q + \gamma_q)x_q + (\beta_{q+1}x_{q+1}) + \dots + (\beta_p x_p)], j = 1, \dots, k-1$	When proportional chances assumption is not valid	-	<pre>.gologit2 Health age color_skin diabetes Autofit lforce</pre>
PPOM-R	$\lambda_j(\underline{x}) = \ln \left\{ \frac{\Pr(Y = 1   \underline{x}) + \dots + \Pr(Y = j   \underline{x})}{\Pr(Y = j + 1   \underline{x}) + \dots + \Pr(Y = k   \underline{x})} \right\} = \ln \left\{ \frac{\sum_{j=1}^j \Pr(Y = j   \underline{x})}{\sum_{j=j+1}^k \Pr(Y = j   \underline{x})} \right\}$ $\lambda_j(\underline{x}) = \alpha_j + \tau_j [(\beta_1 + \gamma_1)x_1 + \dots + (\beta_q + \gamma_q)x_q] + (\beta_{q+1}x_{q+1}) + \dots + (\beta_p x_p), j = 1, \dots, k-1$	Proportional chances assumption is not valid and there is a linear relation between OR of a covariable and the response variable	-	-
CRM	$\lambda_j(\underline{x}) = \ln \left\{ \frac{\Pr(Y = j   \underline{x})}{\Pr(Y = j + 1   \underline{x}) + \dots + \Pr(Y = k   \underline{x})} \right\} = \ln \left\{ \frac{\Pr(Y = j   \underline{x})}{\sum_{j=j+1}^k \Pr(Y = j   \underline{x})} \right\}$ $\lambda_j(\underline{x}) = \alpha_j + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{pp} x_p), \quad j = 1, \dots, k$	There is an intrinsic interest in a specific category of the response variable	<pre>&gt; cr=cr.setup (nhanes\$health) Health.cr=cr\$cohort Cohort=cr\$cohort Age.cr=age cr\$subs Diabetes.cr=diabetes cr\$subs Color.cr=color_skin cr\$subs Crm=lm(health.cr+diabetes.cr+color.cr+cohort) &gt; crm</pre>	<pre>.oratio Health age diabetes color_skin</pre>
SE	$\lambda_j(\underline{x}) = \ln \left\{ \frac{\Pr(Y = j   \underline{x})}{\Pr(Y = 0   \underline{x})} \right\}$ $\lambda_j(\underline{x}) = \alpha_j + \omega_j (\beta_1 x_1 + \dots + \beta_p x_p), \quad j = 1, \dots, k$	Discrete ordinal response variable that does not come from any grouped continuous variable	<pre>&gt; install.packages ("VGAM",response="http/www.stat.auckland. Ac.nz/~yee") &gt; library(VGAM) Ms=mvglm(health~age+diabetes+color_skin, Nhanes,multinomial) &gt; summary(ms)</pre>	<pre>.logit Health age diabetes color_skin</pre>

POM= Proportional odds model; PPOM-NR = Non-restricted proportional odds model; PPOM-R= Restricted proportional odds model; CRM=Continuous ratio model; SE = Stereotype model  
 \*NHANES study – variable response: health; covariables: age, diabetes and color\_skin  
 NHANES = Second National Health and Nutrition Examination Survey

a certain score, let us say  $y_p, Y = j$ , with the probability of a greater response,  $Y > y_p$ , as indicated in Table 1.

This model has different intercepts and coefficients for each comparison and can be adjusted for  $k$  binary logistic regression models.<sup>11</sup> It is more suitable when there is an intrinsic interest in a specific category of the response variable.<sup>1</sup>

### Stereotype model (SM)

The SM can be considered an extension of the multinomial regression model.<sup>10</sup> It compares each category of the response variable with a reference category, normally the first or last category. But, due to the ordinal nature of the data a linear structure is imposed on  $\beta_{jl}$  ( $j=1, \dots, k$  e  $l=1, \dots, p$ ), in other words, weights ( $\omega_j$ ) are attributed to the coefficients.<sup>11</sup>

The weights ( $\omega_j$ ) of the model are directly related to the effect of the covariables. Because of this the *OR* will tend to grow, since the weights are normally constructed in an ordered manner

$$(0 = \omega_1 \leq \omega_2 \leq \dots \omega_k)$$

This model should be used when the response variable is an ordinal variable with discrete categories.

In all the ordinal models mentioned the significance of the coefficients should be tested using the Wald test.<sup>10</sup> In the exercise presented it was calculated using approximation by the normal standardized distribution.

### CHECKING THE QUALITY OF THE ADJUSTMENT OF THE ORDINAL MODELS

As in any type of regression analysis, it is important to assess the quality of the adjustment of the ordinal logistic regression models, because failure to adjust may, for example, lead to a bias in the estimation of the effects. Assessment of the adjustment may detect: important covariables; interactions that were omitted; cases in which the linking function (logit) was not appropriated; cases in which the functional form of the modeling of the covariables is not correct; and finally, cases in which there has been a violation of the proportional odds assumption.<sup>4</sup>

Although many methods have been developed for evaluating the adjustment of binary logistic regression models, few of these methods have been extended to ordinal response data.<sup>10</sup> Normally, the quality of the adjustment of ordinal models is checked using the Pearson or deviance tests. These tests involve the constitution of a contingency table in which the lines comprise all the possible configurations of the covariables of the model and the columns are the categories of the ordinal response.<sup>14</sup>

The expected counts of this table are expressed as  $E_{lj} = \sum_{i=1}^{N_l} \hat{P}_{lj}$ , where  $N_l$  is the total number of individuals classified in line  $l$  and  $\hat{P}_{lj}$  represents the probability of an individual in line  $l$  having the response  $j$  calculated from the model adopted.<sup>14</sup> The Pearson test for evaluating the suitability of the adjustment compares these expected counts with those actually observed, using the formula:

$$\chi^2 = \sum_{l=1}^L \sum_{j=1}^k \frac{(O_{lj} - E_{lj})^2}{E_{lj}} \quad (1)$$

The deviance statistic also compares observed and expected counts, but using the formula:

$$D^2 = 2 \sum_{l=1}^L \sum_{j=1}^k O_{lj} \log \frac{O_{lj}}{E_{lj}} \quad (2)$$

Tests used for evaluating the quality of the adjustment of the model are based on an approximation of the statistics (1) and (2) for the chi-squared distribution with  $(L-1)(k-1)p$  degrees of freedom ( $L$  is the number of lines,  $k$  is the number of columns in the contingency table and  $p$  is the number of covariables in the model). A significant p-value leads to the conclusion of a lack of adjustment of the model to the data being studied.<sup>14</sup>

Pulkstenis & Robinson<sup>14</sup> (2004) report that statistics (1) and (2) do not provide a good approximation of the chi-squared distribution when continuous covariables are adjusted. They suggest small modifications in this case.

In this study, in all the models considered, Pearson or deviance adjustment quality tests were used, since they are found in the usual statistical packages.

As the literature on adjustment diagnosis or evaluation tools for ordinal models is relatively scarce, Hosmer & Lemeshow<sup>10</sup> (2000) suggest the use of binary regressions, separated for each cut-off point, thus creating diagnosis statistics for the ordinal models. Residual graphs are normally constructed for proportional odds models using the adjustment of these models to predict a series of binary events  $Y > j, j=1, 2, \dots, k$ . Therefore, for the indicator variable  $[Y > j]$ , the residual score for case  $i$  and covariable  $p$  is given by:<sup>9</sup>

$$U_{ip} = X_{ip} ([Y_i \geq j] - \hat{P}_{ij}),$$

$$\hat{P}_{ij} = \frac{1}{1 + \exp[-(\hat{\alpha}_j + \underline{x} \hat{\beta})]} \quad (3)$$

In residual score graphs, the mean  $\bar{U}_{.p}$  and the respective reliability intervals are placed along the vertical axis, with the response variable categories along the horizontal axis. If the proportional odds assumption is valid for each covariable, the reliability intervals for

each category of the response variable should have a similar appearance.<sup>9</sup>

Partial residuals are also widely used for checking if all the covariables of the model have linear behavior. In the context of ordinal regression, it is necessary to calculate binary logistic regression models for all the cut-off points of the response variable *Y*, with the partial residual for each case *i* and the covariable *p* being defined in the following way:<sup>9</sup>

$$r_{ip} = \hat{\beta}_p X_{ip} + \frac{[Y_i \geq j] - \hat{P}_{ij}}{\hat{P}_{ij}(1 - \hat{P}_{ij})}$$

$$\hat{P}_{ij} = \frac{1}{1 + \exp[-(\hat{\alpha}_j + \underline{x}\hat{\beta})]} \quad (4)$$

The partial residual graphs provide estimates of how each covariable (*x*) relates to each category of response variable (*Y*).<sup>9</sup>

So partial residuals are used to check the need for changes in the covariables (linearity) or even the vali-

dity of the proportional odds assumption (parallelism of the curves).<sup>9</sup>

### COMMANDS USED IN THE MODELS

In this section, the steps for adjusting the models in Section 2 in the R or Stata software that was summarized in Table 1 will be shown.

The commands were illustrated with data taken from NHANES II and the variables were called:

Response variable: *health*

Covariables: *age, diabetes, skin color*

### Adjustment of the Models in R software

#### A) Proportional odds model

In the R software, the POM can be adjusted using the command *lrm*, developed by Harrell and forming part of the Design package (Table 1). This command adjusts

**Table 2.** Association between diabetes and health status, according to OR estimates and 95% confidence intervals.

Variable	Health status				
	Excellent	Good	Average	Reasonable	Poor
Diabetes					
No	2383 (24%)	2546 (26%)	2805 (29%)	1508 (15%)	594 (6%)
Yes	24 (5%)	45 (9%)	133 (27%)	162 (32%)	135 (27%)
OR (95% CI)*	1.0	1.8 [1.1;2.9]	4.7 [3.0;7.3]	10.7 [6.9;16.5]	22.6 [14.5;35.2]
Diabetes					
No	2383	2546	2805	1508	594
Yes	24	45	133	162	135
OR (95% CI)**	1.0		6.3 [4.2;9.6]		
Diabetes					
No	2383	2546	2805	1508	594
Yes	24	45	133	162	135
OR (95% CI)**		1.0		6.3 [4.8;8.1]	
Diabetes					
No	2383	2546	2805	1508	594
Yes	24	45	133	162	135
OR (95% CI)**		1.0		5.4 [4.7;7.2]	
Diabetes					
No	2383	2546	2805	1508	594
Yes	24	45	133	162	135
OR (95% CI)**			1.0		5.8 [4.7;7.2]

\* OR considering one category as a reference point (where *Y* = state of health; *YD* = excellent category, *x(a)* = no diabetes.

$$OR_j = \frac{P(Y = Y_0 | x^{(A)})}{P(Y > Y_0 | x^{(A)})} \bigg/ \frac{P(Y = Y_0 | x^{(B)})}{P(Y > Y_0 | x^{(B)})}$$

*X*(*b*) = suffering from diabetes

\*\* OR for ordinal data – (where *Y* = state of health, *j*= each category of *Y*, *x(a)* = no diabetes, *x(b)* = suffering from diabetes)

$$OR_j = \frac{P(Y \leq Y_j | x^{(A)})}{P(Y > Y_j | x^{(A)})} \bigg/ \frac{P(Y \leq Y_j | x^{(B)})}{P(Y > Y_j | x^{(B)})}$$

**Table 3.** Results\* of the final proportional odds model\*\*, according to the health status response (NHANES II)

Covariable	$\beta$	EP( $\beta$ )	OR	P-value score test***
Diabetes				
No			1.00	0.60
Yes	1.26	0.09	3.35	
Skin color				
Non-white			1.00	0.01
White	-0.86	0.14	0.42	
Other	-0.64	0.06	0.53	
Age (in years)	0.04	0.01	1.04	<0.01

\* In all models all the variables were significant at the 1% level

\*\* Final p-value score test (value-p<0.001)

\*\*\* Deviance test (value-p<0.001)

binary and ordinal proportional odds models using the maximum verisimilitude method or alternatively the penalized maximum verisimilitude method.<sup>9</sup>

The arguments used are: formula, in other words the terms to be included in the model (variable response and covariables) and file name, for the data to be used, etc. The outcomes shown after using the commands are: expression used, the frequency table for the response, vector with some important statistics, estimates of the coefficients, vector of the first estimates derived from the verisimilitude function log and the deviance of the model.

### B) Continuous ratio model

The CRM may be implemented in R software by restructuring the data, which is done using the *cr.setup* command, which forms part of the Design package. This command makes it possible to create new variables from the response variable y, which will be used to adjust the continuous ratio model.<sup>1</sup>

Four new variables are added with this command:

y – new binary variable that will be used as a response in adjusting the binary logistic regression model;

cohort – a vector indicating which cut-off point (two comparisons of the CRM) was applied;

subs – a vector used to replicate the other variables (explanatory) in the same way that y was replicated;

reps – a variable that specifies how many times each original observation was replicated.

The model is obtained by adjusting a binary logistic regression in the restructured data with a new dichotomous response (y) as a dependent variable, including the created covariable (cohort), which indicates the level of the cut-off point, and restructuring the covariables by the vector (subs), as shown in Table 1.<sup>1</sup>

The assumption of the heterogeneity of the cut-off points can be tested by including an interaction term in the model between the interest declaration and the indicator variable of the cut-off point (cohort). This is called the saturated model. The log value of the verisimilitude function of the models can be compared both with and without the interaction term.

### C) Stereotype model

The SM can be adjusted using generalized linear models that have estimated restriction matrixes. The weights (restrictions) are estimated as additional parameters of the model, using the multinomial family for the adjustment.

In the R software, the SM can be adjusted using the command *rrvglm* that forms part of the VGAM package.<sup>17</sup>

### Residual analysis

The Design package function *residuals.lrm* is used to construct residual graphs after adjusting the POM in the R software, as shown in Table 1.

In the score residual graph (*score.binary*), if the proportional odds assumption is valid, it is expected that for each covariable, the trend of the response variable categories will be constant on the horizontal line. In the partial residual graph (*partial*), on the other hand, in a well-adjusted model it is expected that the curves will be both linear and parallel.<sup>9</sup>

### Stata software

#### Adjustment of the partial proportional odds model

So far the PPOM is not available in R software, but it can be adjusted in Stata 9.0 using the command *gologit2* developed by Williams<sup>16</sup> (2006). This command made

it possible to test the proportional odds assumption, using the *autofit* option and adjusting coefficients for the various variable categories in which this assumption is violated.

## APPLICATION EXAMPLE - NHANES II

As a dependent variable, state of health classified into five categories was considered: (1=poor, 2=reasonable, 3=average, 4=good and 5=excellent). Didactically, a model was constructed using three explanatory variables: a quantitative variable – age (in years), a categorical binary variable – diabetes (no; yes) and a variable with more than two categories – skin color (white; non-white; other). For the skin color variable indicator variables were created, considering non-white as a point of reference.

Table 2 was constructed as a didactic way of showing how OR can be obtained, considering a category as the reference or grouping the categories. In the first calculation “excellent” state of health was considered as the reference point and each of the subsequent categories was compared with it separately, as was done in the SM. The value of OR is seen to increase as the state of health deteriorates.

In the second calculation, OR was calculated in accordance with the POM equation, in which smaller or equal values are compared with a given category with larger values (Table 1). When compared with the “excellent” state of health, the OR of the states “good” and “poor” (OR=6.3) was equal to when we compared “excellent” and “good” states of health with “average”, “reasonable” and “poor” states. This case is an example of the proportional odds model, in other words, with an OR similar for all categories compared, which is the main assumption of the POM.

The POM results are shown in Table 3. The test score suggests there is violation of the proportional odds assumption for the variables “skin color” and “age”, in isolation, and also for the multiple model. Furthermore, the deviance test indicated that the model lacked adjustment.

The residual graphs (score and partial) for evaluating the suitability of the POM are shown in Figures 1 and 2, respectively. In Figure 1 (residual score), the results reinforce the conclusion of the score test, because the curves for the “diabetes” variable showed a horizontal format close to zero. However, the “age” variable behavior oscillated for the “good” and “average” states of health categories and is well below the line of the zero residual. The same happened with the covariable “skin color”. But in this case, the biggest oscillation was seen in the “reasonable” and “poor” states of health categories for the classification “others” and in the “average” category for the color “white”.

In the partial residual graphs (Figure 2), the assumption of parallel regression seemed very acceptable due to its linear aspect and because of the approximately parallel straight lines for the variable “diabetes”. In the “others” category graph of the variable “skin color”, on the other hand, despite its linear behavior, the curves crossed, therefore violating the assumption of parallelism. There was no linear behavior with the covariable “age”, which might contribute to the lack of adjustment of the model. Even when higher degree terms for “age” were included, the deviance test continued to indicate poor adjustment.

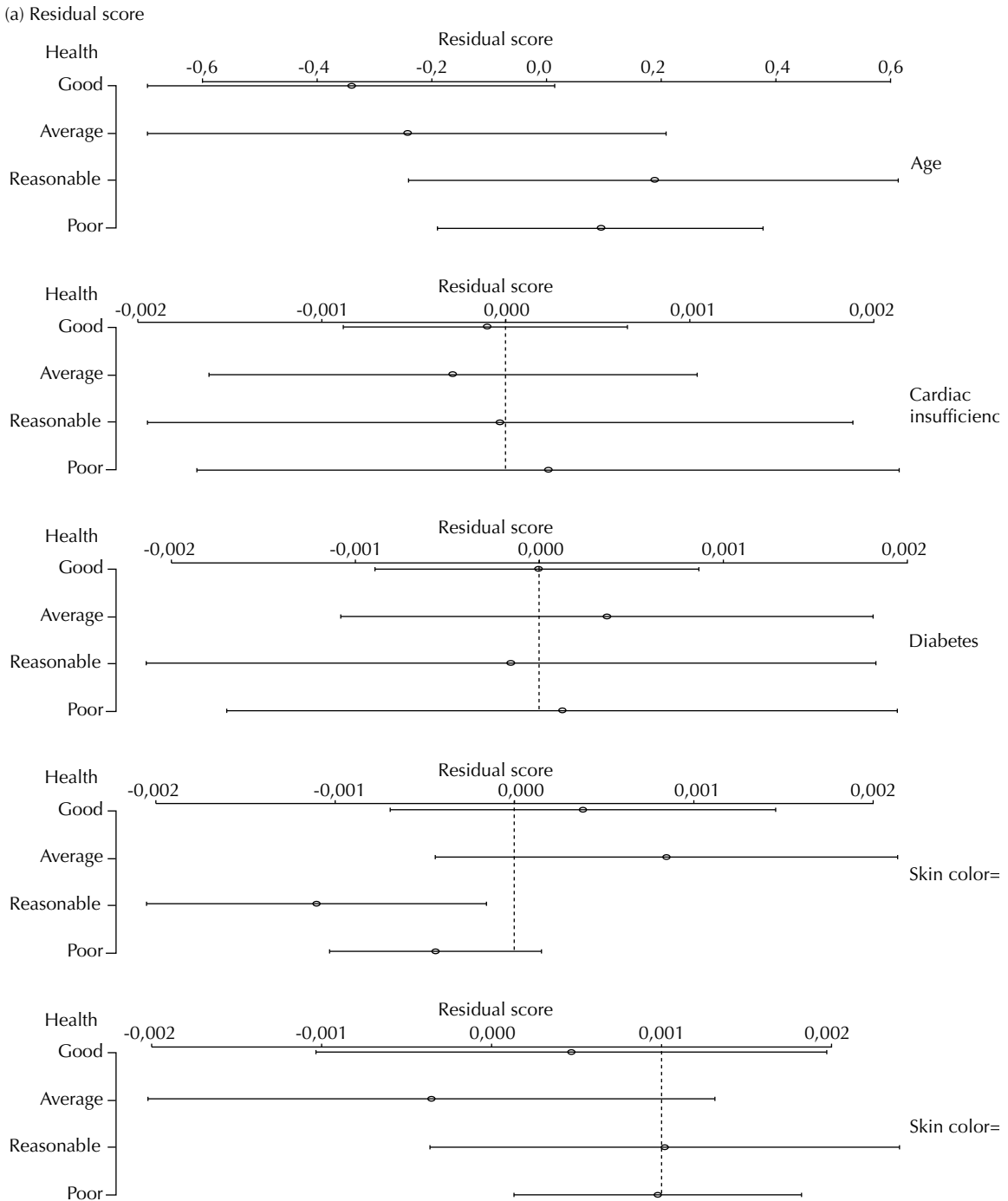
Adjustment of the PPOM is shown in Table 4. In this model, the effects were significant for the four comparisons and the coefficients did not vary for the variable “diabetes”, indicating that an individual with a worse state of health has 3.39 times more chance of being diabetic, when compared with an individual whose state of health is better.

Compared to the “non-white” skin color, there was no variation in the coefficients in the various comparisons for skin color “white”, while for “other” skin color there was a variation. As worse states of health were evaluated, there was an increase in the protection effect, i.e., there was a reduction in the absolute value of OR. For the covariable “age”, a change in the size of OR was also seen in various categories when comparing states of health. For every additional year in age, the chance of the state of health moving from good to poor was 1.03 times higher than that with a person in an excellent state of health. This chance can reach 1.05 times when someone with a poor state of health is compared with someone with excellent health moving to having reasonable health.

As for SM, in all comparisons, the effect of the covariables was significant (value- $p < 0,01$ ), and the deviance test indicated there was good adjustment of the model (Table 5). In other words, people with poor health have ten times more chance of being diabetic than those who are in excellent health. The magnitude of this association reduces as the state of health gets close to excellent, reaching 1.43 in the comparison of good health with excellent health.

## COMMENTS

In the example presented in this paper, POM did not provide a good adjustment and the residual graphs showed non-parallel straight lines for some covariables, indicating violation of the main premise. Considering that any inferences based on this model may not be correct the PPOM was alternatively presented with an estimate of OR for each of the comparisons. However, the SM was the one that best adjusted to the data analyzed, according to the results of the deviance test.

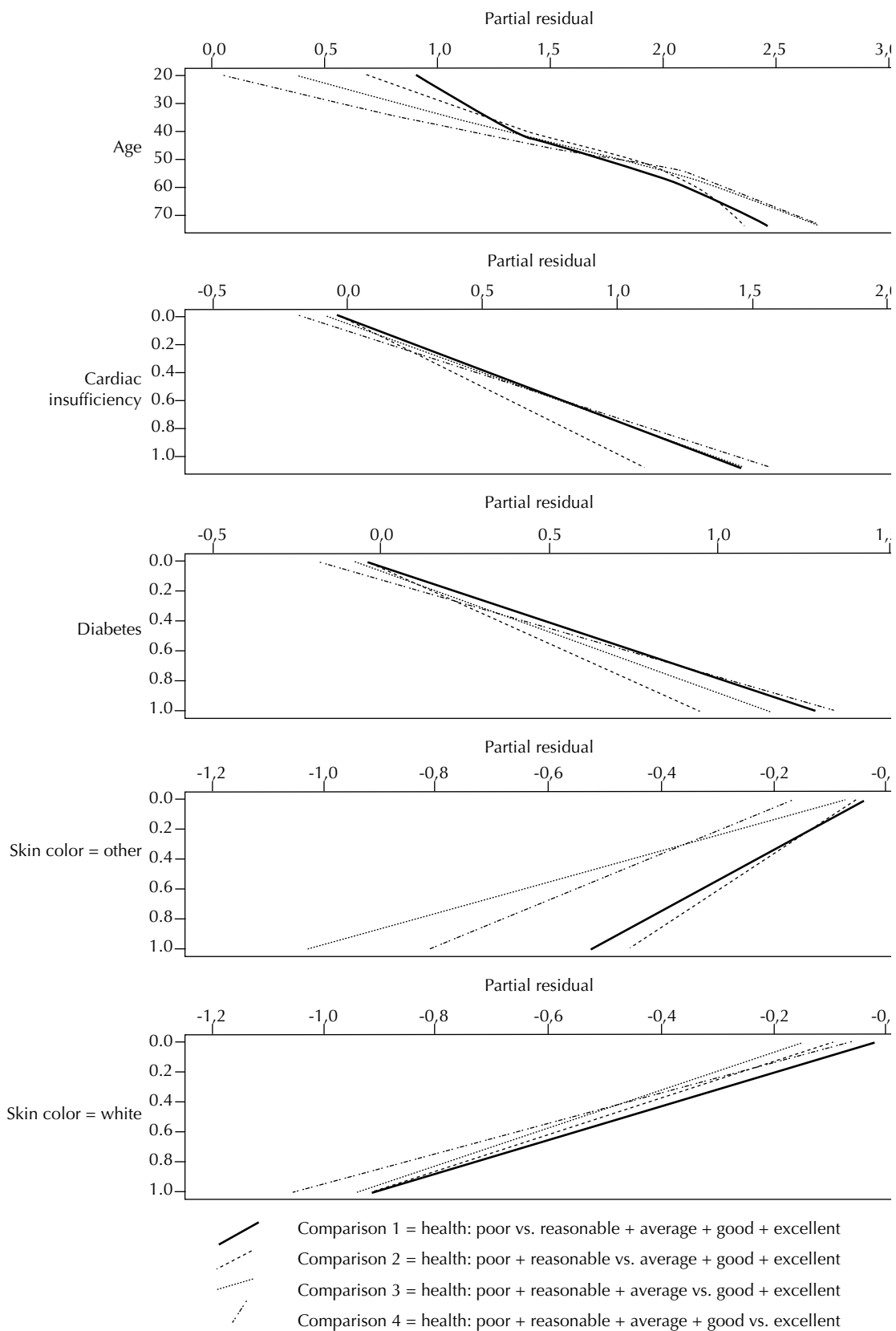


**Figure 1.** Residual graphs (score and partial) for the covariables included in the model, according to the health status response.

Generally speaking, ordinal logistic regression models are recommended for analyzing ordinal data.<sup>1,3,8,10,11</sup> Ananth & Kleinbaum<sup>1</sup> (1997) report that the POM and CRM are the most widely used in epidemiological and biomedical applications in relation to PPOM and SM. But these models lead to strong assumptions that, if they are not valid, may lead to incorrect interpretations, as occurred in the example used.<sup>1</sup>

Other authors<sup>11</sup> state that the type of model used depends on the character of the ordinal response variable, i.e., whether this variable was ordered starting with a regrouped continuous variable or if it originated from a discrete variable. In the first case, the POM is the most indicated when the premise of parallel straight lines is not violated. In the second case, the SM is the most indicated, as in this study, in which the response





**Figure 2.** Residual graphs (score and partial) for the covariables included in the model, according to the health status response.

**Table 4.** Results\* of the final partial proportional odds model, according to the health status response (NHANES II).

Covariable	Comparisons							
	Excellent vs. (good+average+reasonable+poor)		Excellent+good vs. (average+reasonable+poor)		Excellent+good+average vs. (reasonable+poor)		Excellent+good+average+reasonable vs. (poor)	
	$\hat{\beta}_1$	$\hat{OR}_1$	$\hat{\beta}_2$	$\hat{OR}_2$	$\hat{\beta}_3$	$\hat{OR}_3$	$\hat{\beta}_4$	$\hat{OR}_4$
Diabetes								
No		1.00		1.00		1.00		1.00
Yes	1.22	3.39	1.22	3.39	1.22	3.39	1.22	3.39
Skin color								
Non-white		1.00		1.00		1.00		
White	-0.88	0.42	-0.88	0.42	-0.88	0.42	-0.88	0.42
Other	-0.54	0.58	-0.44	0.64	-1.04	0.35	-1.14	0.32
Age (in years)	0.03	1.03	0.04	1.04	0.05	1.05	0.05	1.05

\* In all models all the variables were significant at the 1% level

variable, defined as state of health in the NHANES II study was treated with discrete categories.

Even in the presence of an ordinal response, other multivariate analysis options should be considered. One way is to use the decision-tree method,<sup>5</sup> which is a more descriptive method that considers a greater number of variables in the final model. Other linking functions of the model may also be used, like probit analyses and complementary log-log. However, ordinal regression is a parametric technique that, by imposing a rigid, more conservative and economical structure on the model, allows the reliability intervals for the parameters to be quantified; this makes it easier to interpret the OR. While alternative approaches are worth analyzing and discussing, they will not be discussed in this article.

In constructing ordinal models, Hosmer & Lemeshow<sup>10</sup> (2000) propose strategies like those used in the example

in this article. These authors initially recommend carrying out a univariate analysis for selecting the principal effects and including in the model just significant variables that have a prefixed level of significance. Then, the model should be adjusted, its suitability checked using appropriate tests and residual graphs and finally the model should be interpreted by estimating the OR.

However, there are few methods for checking the adjustment of ordinal models. In the literature available we have so far found no technique for checking the adjustments of the SM. The existing diagnosis statistics proposed by Harrel<sup>9</sup> (2002) and applicable to the POM, are merely graphs taken from binary regressions, separated for the ordinal variable cut-off points. Although they are incomplete, these techniques are extremely important when it comes to getting an indication of the quality of the adjustment of ordinal models. Analysis of partial residuals, even when graphic, is considered

**Table 5.** Results\* of the final stereotype model\*\*, according to the health status response (NHANES II)

Covariable	Comparisons of state of health							
	Poor vs. excellent		Reasonable vs. excellent		Average vs. excellent		Good vs. excellent	
	$\hat{\beta}_1$	$\hat{OR}_1$	$\hat{\beta}_2$	$\hat{OR}_2$	$\hat{\beta}_3$	$\hat{OR}_3$	$\hat{\beta}_4$	$\hat{OR}_4$
Diabetes								
No		1.00		1.00		1.00		1.00
Yes	2.25	9.45	1.69	5.43	0.97	2.64	0.36	1.43
Skin color								
Non-white		1.00		1.00		1.00		1.00
White	-1.64	0.19	-1.23	0.29	-0.71	0.49	-0.26	0.77
Other	-1.30	0.27	-0.98	0.38	-0.56	0.57	-0.21	0.81
Age (years)	0.08	1.08	0.06	1.06	0.03	1.03	0.01	1.01

\* In all models all the variables were significant at 1% level.

\*\* Deviance test (value-p<0.001)

very useful for ordinal models, because they check both linearity, indicating possible changes that should be used, and the proportional odds assumption.

On the other hand, care must be taken when interpreting residuals, principally considering once again the lack of information of how to do it. Sometimes, the graphs may suggest confusing information and make it difficult to take decisions as to whether the proportional odds assumption has been violated or not. An alternative is to use residual analysis along with the score test, because when there is any doubt as to the format of the graph this test may contribute to a final conclusion being reached.

Ordinal logistic regression models have shown to be suitable for analyzing data with ordinal response. The choice of the best model depends on the character of the ordinal variable, adaptation of the model to the assumptions, the quality of the adjustment and the capacity it has for coming up with a good explanation with a reduced number of parameters to be estimated.

Finally, when it comes to using ordinal models, good computer skills and mastery of the commands are essential, not only for choosing the most suitable model but also for making comparisons between models. For this reason, the R program, which includes various models and some diagnosis graphs, is an important tool.<sup>9</sup>

## REFERENCES

- Ananth CV, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. *Int J Epidemiol.* 1997;26(6):1323-33. DOI: 10.1093/ije/26.6.1323
- Anderson JA. Regression and ordered categorical variables. *J R Statist Soc B.* 1984;46(1):1-30.
- Bender R, Benner A. Calculating ordinal regression models in SAS and S-Plus. *Biometrical J.* 2000;42(6):677-99. DOI: 10.1002/1521-4036(200010)42:6<677::AID-BIMJ677>3.0.CO;2-O
- Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics.* 1990;46(4):1171-8. DOI: 10.2307/2532457
- Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. New York: Chapman & Hall; 1984.
- Colin RB. Bioestatística usando R: apostila para biólogos. Bragança; 2004.
- Fienberg SE. Fixed Margins and Logit Models. In: Fienberg SE. The analysis of cross-classified categorical data. Cambridge, MA: MIT Press; 1980. p.110-6.
- Greenland S. Alternative models for ordinal logistic regression. *Stat Med.* 1994;13(16):1665-77. DOI: 10.1002/sim.4780131607
- Harrell Jr FE. Regression modelling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2002.
- Hosmer DW, Lemeshow S. Applied logistic regression. 2. ed. New York: John Wiley & Sons; 2000.
- Lall R, Campbell MJ, Walters SJ, Morgan K. A review of ordinal regression models applied on health-related quality of life assessments. *Stat Methods Med Res.* 2002;11(1):49-67. DOI: 10.1191/0962280202sm271ra
- McCullagh P. Regression models for ordinal data. *J R Statist Soc B.* 1980;42(2):109-42.
- Peterson BL, Hanrel FE. Partial proportional odds models for ordinal response variables. *Appl Statistic.* 1990;39(2):205-17. DOI: 10.2307/2347760
- Pulkstenis E, Robinson TJ. Goodness-of-fit tests for ordinal response regression models. *Stat Med.* 2004;23(6):999-1014. DOI: 10.1002/sim.1659
- Walker SH, Ducan DB. Estimation of the probability of an event as a function of several independent variables. *Biometrika.* 1967;54(1):167-79.
- Williams R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata J.* 2006;6(1):58-82.
- Yee TW, Hastie TJ. Reduced-rank vector generalized linear models. *Statist Model.* 2003;3(1):15-41. DOI: 10.1191/1471082X03st045oa

---

Caiiffa WT is supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; research productivity scholarship); Abreu MNS was supported by pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes; master's scholarship).