International Journal of Epidemiology
Vol. 26, No. 6

Printed in Great Britain

Regression Models for Ordinal Responses: A Review of Methods and Applications

CANDE V ANANTH* AND DAVID G KLEINBAUM†

Ananth C V (The Center for Perinatal Health Initiatives, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, 125 Paterson St, New Brunswick, NJ 08901–1977, USA) and Kleinbaum D G. Regression models for ordinal responses: A review of methods and applications. *International Journal of Epidemiology* 1997; **26:** 1323–1333.

Background. Epidemiologists are often interested in estimating the risk of several related diseases as well as adverse outcomes, which have a natural ordering of severity or certainty. While most investigators choose to model several dichotomous outcomes (such as very low birthweight versus normal and moderately low birthweight versus normal), this approach does not fully utilize the available information. Several statistical models for ordinal responses have been proposed, but have been underutilized. In this paper, we describe statistical methods for modelling ordinal response data, and illustrate the fit of these models to a large database from a perinatal health programme.

Methods. Models considered here include (1) the cumulative logit model, (2) continuation-ratio model, (3) constrained and unconstrained partial proportional odds models, (4) adjacent-category logit model, (5) polytomous logistic model, and (6) stereotype logistic model. We illustrate and compare the fit of these models on a perinatal database, to study the impact of midline episiotomy procedure on perineal lacerations during labour and delivery. Finally, we provide a discussion on graphical methods for the assessment of model assumptions and model constraints, and conclude with a discussion on the choice of an ordinal model. The primary focus in this paper is the formulation of ordinal models, interpretation of model parameters, and their implications for epidemiological research.

Conclusions. This paper presents a synthesized review of generalized linear regression models for analysing ordered responses. We recommend that the analyst performs (i) goodness-of-fit tests and an analysis of residuals, (ii) sensitivity analysis by fitting and comparing different models, and (iii) by graphically examining the model assumptions.

Keywords: Multinomial probabilities, ordinal models, cumulative logit model, continuation-ratio model, partial-proportional odds model, polytomous logistic model, stereotype logistic model

In the study of the dependence of a response variable on a set of independent variables, the choice of a model is largely determined by the scale of measurement of the response.[1] Epidemiologists are often interested in estimating the risk of adverse events originally measured on an interval scale (such as birthweight), but they often choose to divide the outcome into two or more categories in order to compute an estimate of effect (risk or odds ratio). Similarly, response variables originally measured on an ordinal scale (e.g. severity of pre-eclampsia: none, mild, severe) are often categorized into several binary variables during statistical analysis.

Consider, as a motivating example, the data set described in Table 1. This data is derived from a clinical trial of a single-dose, post-operative analgesic clinical trial.[2] A series of four drugs, denoted by C15, C60, Z100, and EC4 were randomized to patients. The patient responses to the drug were recorded on a five-level ordinal scale (poor, fair, good, very good and excellent). Counts for the most favourable responses, 'very good' and 'excellent', were amalgamated into one category ('very good') due to sparse cell counts. The two drugs Z100 and EC4 were found to be quite similar and together rated better than the pair C15 and C60.[2] The drugs C15 and C60 are the same drug, but vary in their potency. Usually, such data are analysed by creating dichotomies among the levels of the response variable. Possible dichotomies include comparing 'very good' to 'poor', 'good' to 'poor', and so on. Standard regression procedures, such as the logistic regression, can then be utilized for data analyses.

Although such approaches are not incorrect, they often result in a loss of information due to collapsing (or ignoring) some categories of the response (unless

* The Center for Perinatal Health Initiatives, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, 125 Paterson St, New Brunswick, NJ 08901–1977, USA.
† Department of Epidemiology, The Rollins School of Public Health, Emory University, Atlanta, GA, USA.

TABLE 1 *Ratings of drugs on a single-dose, post-operative analgesic clinical trial*[a]

Drug	Rating of the drugs				
	Poor	Fair	Good	Very good[b]	Total
C15 & C60	17	18	20	5	61
Z100 & EC4	10	4	13	34	60
Log-odds ratio[c], $\hat{\beta}$		0.7013	1.5476	2.6384	
Standard error $\widehat{se}$ $(\hat{\beta})$		0.4491	0.4016	0.5535	
Odds ratios, $e^{\hat{\beta}}$	1.0	2.0	4.7	13.9	

[a] Source: ref.[2]
[b] Favourable responses very good and excellent are amalgamated.
[c] Log-odds ratios involve comparisons between ('poor' versus ⩾ 'fair'), (⩽ 'fair' versus ⩾ 'good'), and (⩽ 'good' versus 'very good').

there is perfect homogeneity within the categories being collapsed), typically resulting in a considerable loss of statistical power. Although several statistical models for ordinal responses have been proposed, their utilization in the epidemiological and biomedical literature has been minimal.

The purpose of this paper is twofold: first, to provide a synthesized review of models for analysing data with ordinal responses, and second, to evaluate their usefulness in epidemiological research, with particular emphasis on model formulation, interpretation of model coefficients, and their implications. Ordinal models that are considered include (1) cumulative logit or the 'grouped continuous' model,[3,4] (2) continuation-model,[5] (3) constrained and unconstrained partial proportional odds models,[6] (4) polytomous logistic model,[7,8] (5) adjacent-category logistic model,[9] and (6) stereotype logistic model.[10] The development of each model is described in detail, with analysis examples using a data set from a perinatal health programme. We then describe briefly the statistical software that were used to fit these models. Finally, the paper concludes with a discussion on the choice of ordinal model.

REGRESSION MODELS FOR ORDINAL RESPONSES

1. Cumulative Logit Model

Attempts to extend the logistic regression model for binary responses to allow for ordinal responses have often involved modelling cumulative logits. Consider a multinominal response variable Y with categorical outcomes, denoted by $1,2,\ldots,k$, and let x_i denote a p-dimensional vector of covariates. When no confusion arises, the subscript 'i' will be dropped. The cumulative logit model was originally proposed by Walker and Duncan[3] and later called the proportional odds model by McCullagh.[4] The dependence of Y on x for the proportional odds model has the following representation:

$$Pr(Y \leqslant y_j \mid \mathrm{x}) = \frac{\exp(\alpha_j - \mathrm{x}'\beta)}{1 + \exp(\alpha_j - \mathrm{x}'\beta)}, \; j = 1,2,\ldots,k \qquad (1)$$

or equivalently can be re-expressed in logit form as

$$\begin{aligned} \operatorname{logit}(\Pi_j) &= \log\left[\frac{\Pi_j}{1-\Pi_j}\right] \\ \log\left[\frac{Pr(Y \leqslant y_j \mid \mathrm{x})}{Pr(Y > y_j \mid \mathrm{x})}\right] &= \alpha_j - \mathrm{x}'\beta, \; j = 1,2,\ldots,k \end{aligned} \qquad (2)$$

where $\prod_j = Pr(Y \leqslant y_j)$ is the cumulative probability of the event $(Y \leqslant y_j)$. α_j are the unknown intercept parameters, satisfying the condition $\alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_k$, and $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ is a vector of unknown regression coefficients corresponding to x.

The regression coefficient, β_l, for a binary explanatory variable x_l is the log-odds ratio for the Y by x_l association, controlling for other covariates in model (2). Notice that the regression coefficient vector, β, does not depend on j, implying that model (2) assumes that the relationship between x_l and Y is independent of j. McCullagh[4] calls this assumption of identical log-odds ratios across the k-cut points, the proportional odds assumption, and hence the name 'proportional odds' model. The validity of this assumption can be checked based on a χ^2 Score test.[11] A model that relaxes the proportional odds assumption can be represented as $\operatorname{logit}(\prod_j) = \alpha_j - \mathrm{x}'\beta_j$, where the regression parameter vector β is allowed to vary with j. The usefulness of this latter model is to test the assumption of proportionality

in the log-odds ratio (β), and may formally be stated as a test of the null hypothesis $H_0 : \beta_1 = \beta_2 = ,\ldots, = \beta_k$.

The proportional odds model is invariant when the codes for the response Y are reversed[4,12] (i.e. y_1 recoded as y_k, y_2 recoded as y_{k-1}, and so on), resulting only in a reversal of the sign of the regression parameters. Secondly, the proportional odds model is invariant under collapsability of the categories of the ordinal response.[11] This property implies that when the categories of Y are deleted or collapsed, the coefficients β will remain unchanged, although the intercept parameters α will be affected. The collapsibility property of the proportional odds model enables one to model an ordinal outcome Y which may be continuous. Greenland[12] provides a more detailed review of these properties.

Based on the fit of model (2), the cumulative odds ratio, Ψ_l, for the *lth* binary covariate, x_l, can be obtained by the following relationship:

$$\Psi_{\mathrm{P}} = \frac{Pr(Y \leqslant y_j \,|\, x_l^{(1)})}{Pr(Y \leqslant y_j \,|\, x_l^{(0)})} = \exp\left\{-\beta_l(x_l^{(1)} - x_l^{(0)})\right\} \tag{3}$$

2. Continuation-Ratio Model

Feinberg[5] proposed an alternative method (to the proportional odds model) for the analysis of categorical data with ordered responses. When the cumulative probabilities, $\prod_j = Pr(Y \leqslant y_j)$, of being in one of the first j categories in the cumulative logit model (model 2) is replaced by the probability of being in category j [i.e. $\theta_j = Pr(Y = y_j)$] conditional on being in categories greater than j [i.e. $(1 - \prod_j)$], this results in the continuation-ratio model. Define $\delta_j = \theta_j/(1 - \prod_j)$. The continuation-ratio model can then be formulated as:

$$\begin{aligned} \mathrm{logit}(\delta_j) &= \log\left[\frac{\delta_j}{1-\delta_j}\right] \\ \log\left[\frac{Pr(Y = y_j \,|\, \mathrm{x})}{Pr(Y > y_j \,|\, \mathrm{x})}\right] &= \alpha_j - \mathrm{x}'\beta, \quad j = 1,2,\ldots,k \end{aligned} \tag{4}$$

and could essentially be viewed as the ratio of the two conditional probabilities, $Pr(Y = y_j \,|\, \mathrm{x})$ and $Pr(Y > y_j \,|\, \mathrm{x})$. This model of conditional odds has been referred to as the 'continuation-ratio' model.[5] When the 'logit' link is replaced by the 'complimentary log–log' link function in model (4), the resulting model is

$$\log\,[-\log\,(\delta_j)] = \alpha_j - \mathrm{x}'\beta \tag{5}$$

which is the Cox proportional-hazards model[13] for survival data in discrete time.[13–15] Läära and Mathews[16] explicitly prove that when the complimentary log-log link is used, the proportional odds and the continuation-ratio models are identical. A more detailed discussion can be found in McCullagh[4] and McCullagh and Nelder.[14]

The odds ratio's, Ψ_{C}, based on continuation-ratios for the *lth* covariate x_l can be obtained directly from model (4) as follows:

$$\Psi_{\mathrm{C}} = \frac{Pr(Y = y_j \,|\, x_l^{(1)})/Pr(Y > y_j \,|\, x_l^{(1)})}{Pr(Y = y_j \,|\, x_l^{(0)})/Pr(Y > y_j \,|\, x_l^{(0)})} = \exp\left\{-\beta_l(x_l^{(1)} - x_l^{(0)})\right\}$$

The continuation-ratio model is best suited to circumstances where the individual categories of the response variable are of intrinsic interest, and are not merely an arbitrary grouping of an underlying continuous variable.[14] Unlike the proportional odds model (model 2), the continuation-ratio model (model 4) is neither preserved by a reversal of the codes for the ordinal response nor under collapsibility of the categories of Y.[12]

3. Partial-Proportional Odds Model

The primary motivation for the development of the partial-proportional odds model[6] was to relax the strong assumption of identical log-odds ratio for the Y by x_1 association, in the proportional odds model. Violation of the assumption of identical log-odds could lead to the formulation of an incorrect or misspecified model. A situation under which this assumption does not hold is illustrated below.

Analgesic trial data. For purposes of illustration, consider the analgesic trial data[2] described in Table 1. The estimated log-odds ratios [$\hat{\beta}$], and their estimated standard errors [$\widehat{\mathrm{se}}(\hat{\beta})$], for the logits are presented in Table 1, for comparisons between the drugs Z100 and EC4 versus C15 and C60. The results indicate that the log-odds ratio is largest ($\hat{\beta} = 2.6384$) when the rating of the drug is dichotomized at $Y = 4$ ' less than very good' ($\leqslant 3$) versus 'very good' (4); the dichtomization for the next largest ($\hat{\beta} = 1.5476$) being at $Y = 3$, 'poor or fair' versus 'good or very good' ($Y \geqslant 3$), and the log-odds ratio is smallest ($\hat{\beta} = 0.7013$), when the dichtomization is made at $Y = 2$, 'poor' versus 'above fair' ($Y \geqslant 2$). These data are suggestive of a trend in the log-odds ratios. Fitting a proportional odds model (model 2) to the above data resulted in a log-odds ratio (se) of

TABLE 2 *Results of fitting partial proportional odds models: Analgesic trial data*[a]

Variable	$\hat{\beta} \pm \hat{se}\ (\hat{\beta})$	χ^2	*P*-value
Constrained model			
Drug	0.6899 ± 0.4495	2.3	0.1248
Constraint	0.9216 ± 0.2661	12.0	0.0005
Drug (2 d.f.)		31.1	0.0001
Likelihood ratio test (2 d.f.)		37.0	0.0001
Goodness-of-fit : linear constraint–drug		19.1	0.0001
Test for proportional odds		0.2	0.7114
Unconstrained model			
Drug	0.7013 ± 0.4491	2.4	0.1184
Response ⩾ good	0.8463 ± 0.3552	6.4	0.0116
Response ⩾ v. good	0.9272 ± 0.5891	10.7	0.0011
Likelihood ratio test (3 d.f.)		37.2	0.0001
Test for nonproportional odds (2 d.f.)		11.8	0.0028
Drug (3 d.f.)		30.3	0.0001

[a] Source: ref.[2] Data illustrated in Table 1.

1.7710 (0.3625), with the assumption of the proportionality of odds being violated ($P < 0.001$). Hence, fit of a proportional odds model to the data is inappropriate, and may result in incorrect or misleading inferences.

3a. Unconstrained partial-proportional odds model. The partial-proportional odds model model[6] permits non-proportional odds for a subset q of the p-predictors ($q < p$). In addition, the assumption of proportional odds can be tested for the sub-set q. With Y being an ordinal variable with k categories, and x being a p-dimensional vector of covariates, the model suggested for the cumulative probabilities is

$$Pr(Y \leqslant y_j \mid \mathrm{x}) = \frac{\exp(-\alpha_j - \mathrm{x}'\beta - \mathrm{t}'\gamma_j)}{1 + \exp(-\alpha_j - \mathrm{x}'\beta - \mathrm{t}'\gamma_j)},\ j = 1,2,\ldots,k \quad (6)$$

where t is a $(q \times 1)$ vector, $q < p$, of a subset q-covariates for which the proportional odds assumption either is not assumed *a priori* or is to be tested; γ_j is a $(q \times 1)$ vector of regression coefficients associated with the q-covariates in t, so that $\mathrm{t}'\gamma_j$ is the increment associated only with the *jth* cumulative logit ($1 \leqslant j \leqslant k$), and $\gamma_1 = 0$. We will henceforth refer to this model as the 'unconstrained model'. When $\gamma_j = 0$ for all j, model (6) reduces to the proportional odds model (model 1). A test of the proportional odds assumption for the q-covariates in t is based on the null hypothesis $H_0 : \gamma_j = 0$, for all j ($2 \leqslant j \leqslant k$). Notice that since $\gamma_1 = 0$, the model uses only $(\alpha + \mathrm{x}'\beta)$ to estimate the odds ratio associated with the dichotomization of Y into $y_j = 1$ versus $y_j > 1$. However, estimation of odds ratios associated with the remaining cumulative probabilities involve incrementing $(\alpha + \mathrm{x}'\beta)$ by $\mathrm{t}'\gamma_j$.

3b. Constrained partial-proportional odds model. Peterson and Harrell,[6] in addition to the partial-proportional odds model, propose another model called the 'constrained partial-proportional odds model'. In the analgesic trial example (Table 1), we noted the existence of a linear relationship in the log-odds ratios between the drugs and the response. Although fitting model (6) to the data in Table 1 will require two γ_{jl} parameters, a model constraining the γ_{jl} to account for the linearity (in log-odds ratios) in j would require an additional parameter in the model. Such a model has the form described below.

$$Pr(Y \leqslant y_j \mid \mathrm{x}) = \frac{\exp(-\alpha_j - \mathrm{x}'\beta - \mathrm{t}'\gamma\Gamma_j)}{1 + \exp(-\alpha_j - \mathrm{x}'\beta - \mathrm{t}'\gamma\Gamma_j)},\ j = 1,2,\ldots,k \quad (7)$$

where Γ_j's are pre-specified, fixed scalars, and $\Gamma_1 = 0$. The new parameter, γ, is a vector of length q, and is not indexed by j. Although γ is not dependent on j, it is multiplied by the fixed scalar constant Γ_j in the computation of the *jth* cumulative logit.

Results of fitting the unconstrained (model 6) and constrained partial-proportional odds (model 7) models to the analgesic trial data (Table 1) are contrasted in Table 2. It was demonstrated earlier that the data did not satisfy the proportional odds assumption, and that a monotonically increasing trend in the log-odds ratios was observed in the different logits. Hence, a constrained (partial-proportional

odds) model was fit, with the specification of the following constraints: $\Gamma_1 = 0$, $\Gamma_2 = 1$, and $\Gamma_3 = 2$.

The log-odds ratio when the response is dichotomized at ($y_j = 1$) is $\hat{\beta}$ (0.6899), while the log-odds ratios associated with the second and third cumulative logits are $\hat{\beta} + \hat{\Gamma}_2\hat{\gamma}$ (0.6899 + 0.9216) and $\hat{\beta} + \hat{\Gamma}_3\hat{\gamma}$ (0.6899 + 2 * 0.9216), respectively. A simultaneous two degrees of freedom test of $H_0 : \beta = 0,\ \gamma = 0$, was rejected ($\chi^2 = 31.1$, 2 d.f.). However, since the goodness-of-fit of the linearity constraint[6] was not satisfied ($\chi^2 = 19.9$, 1 d.f., $P < 0.0001$), the model was rejected in favour of the unconstrained model.

The fit of the unconstrained model implies that no constraints are placed in the estimation of the log-odds ratios. Hence, instead of using one constrained parameter in the model (as in model 6), $2 - \gamma_j$ parameters associated with the response are used for the second and third cumulative logits. The estimated log-odds ratios are 0.7013, 0.7013 + 0.8463 and 0.7013 + 0.9272 for the three cumulative logits, respectively.

4. Polytomous Logistic Model

The polytomous logistic model[7,8] is a straight forward extension of the logistic model for binary responses, to accommodate multinomial responses. Unlike the models discussed above, the polytomous logistic model does not impose any restrictions on the ordinality of the response. The model has the following representation:

$$Pr(Y = y_j \mid \mathrm{x}) = \frac{\exp(\alpha_j + \mathrm{x}'\beta_j)}{\sum_{l=1}^{k} \exp(\alpha_l + \mathrm{x}'\beta_l)}, \quad j = 1,2,\ldots,k \qquad (8)$$

where $\alpha_k = 0$ and $\beta_k = 0$. The parameter vector $\beta = (\beta_1,\beta_2,\ldots,\beta_k)'$ corresponds to the regression coefficients for the log-odds of ($Y = y_j$), relative to the referent category ($Y = y_k$), and there are ($k - 1$) intercept parameters α_j. Notice that unlike the models described above, the regression coefficient β_j, in the polytomous model depends on j. Exponentiating the regression coefficient β_l, for the *lth* covariate x_l will result in the odds ratio comparing ($Y = y_j$) versus ($Y = y_k$) for a unit increase in x_l.

5. Adjacent-Category Logistic Model

The adjacent-category logistic model[9] involves modelling the ratio of the two probabilities, $Pr(Y = y_j)$ and $Pr(Y = y_{j+1})$, ($j = 1,2,\ldots,k$). The model has the following representation:

$$\log\left[\frac{Pr(Y = y_j \mid \mathrm{x})}{Pr(Y = y_{j+1} \mid \mathrm{x})}\right] = \alpha_j - \mathrm{x}'\beta_j, j = 1,2,\ldots,k \qquad (9)$$

where $\alpha_k = 0$ and $\beta_k = 0$. The parameter β_1 corresponds to the regression coefficient for the log-odds of ($Y = y_1$) relative to ($Y = y_2$); β_2 corresponds to the log-odds of ($Y = y_2$) relative to ($Y = y_3$), and so on, and there are ($k - 1$) intercept parameters α_j. Exponentiating the regression coefficient β_l, for the *lth* covariate x_l will result in the odds ratio comparing ($Y = y_j$) versus ($Y = y_{j+1}$), for a unit increase in x_l.

6. Stereotype Logistic Model

The most flexible model for analysing an ordinal response is the polytomous logistic model, where β represents the log odds ratio for ($Y = y_j$) versus ($Y = y_0$) per unit change in x_l. The polytomous logistic model, however, fails to utilize the ordering of the response categories of *Y*. Anderson[10] proposed modelling the regression coefficients, β_j, by imposing the (linear) relationship

$$\beta_j = -\phi_j \beta \quad j = 1,2,\ldots,k \qquad (10)$$

where ϕ_j may be thought of as 'scores' assigned to the response y_j.[12] Note that since $\beta_k = 0$, we have $\phi_k = 0$, and a further constraint, $\phi_1 = 1$ (in order to uniquely identify the parameters when using estimated scores[12]). Substituting equation (10) in the polytomous logistic model (8) yields the stereotype model:

$$Pr(Y = y_j \mid \mathrm{x}) = \frac{\exp(\alpha_j - \mathrm{x}'\phi_j\beta)}{\sum_{l=1}^{k} \exp(\alpha_l - \mathrm{x}'\phi_l\beta)}, \quad j = 1,2,\ldots,k \qquad (11)$$

Anderson[10] further imposed an additional order constraint on the ϕ's with $1 = \phi_1 > \phi_2 > \ldots > \phi_k = 0$. Under this model, the odds ratio relating $Y = y_j$ versus $Y = y_k$ for the *lth* covariate x_l is given by

$$\Psi_S = \frac{Pr(Y = y_j \mid x_l^{(1)})/Pr(Y = y_k \mid x_l^{(1)})}{Pr(Y = y_j \mid x_l^{(0)})/Pr(Y = y_k \mid x_l^{(0)})}$$

$$= \exp\left\{-\phi_j\beta(x_l^{(1)} - x_l^{(0)})\right\}$$

The stereotype models described thus far relate to situations in which the response *Y* is considered one-dimensional. Consider a situation where one is interested in modelling a response *Y* that is constructed by merging multiple factors, or is a multidimensional construct such as the Apgar score. The Apgar score (ranging from 0 to 10) is constructed by amalgamating the scores of five individual measurements (Heart rate, Respiratory effect, Colour, Muscle tone, and Reflex response) on a newborn infant, with each recorded on

TABLE 3 *Distribution of perineal lacerations in relation to episiotomy: Nova Scotia, Canada, 1992–93*

Episiotomy	Degree of laceration					
	None	1°	2°	3°	4°	Total
None	9238	140	71	131	37	9617
Midline	1204	8	8	89	38	1347

a three-point scale (0 = no response, 1 = intermediate response, and 2 = full response to the function). Anderson[10] extended the stereotype model to capture the multidimensional structure (such as Apgar scores) of the underlying response *Y*. He proposed a two-dimensional extension to equation (10) as

$$\beta_j = \varnothing_j \beta - \varphi_j \gamma \quad j = 1,2,\ldots,k \tag{12}$$

with $\varnothing_k \equiv \varphi_k \equiv 0$. Equation (12) can be extended further to allow for higher dimensions.

Model Fitting and Statistical Software

With the exception of the Stereotype logistic model, all models described above were fit to the data of Table 3 using the SAS system (SAS Institute, Cary, NC). The proportional odds and continuation ratio models were fit using the LOGISTIC procedure with the *logit* (logistic) and the *cloglog* (complimentary log-log) link function specifications, respectively. The polytomous and adjacent category logistic models were fit using the CATMOD procedure with the *logit* and the *alogit* specifications for link functions, respectively. The LOGIST procedure of SAS's supplemental library (version 5.18) was utilized to fit both the constrained and the unconstrained partial proportional-odds models. Peterson and Harrell[11] describe the SAS code required to fit the partial proportional odds models. All models were fit through the procedure of maximum likelihood estimation, while the adjacent category model was fit using the weighted least squares procedure. Peterson and Harrell[6,12] however, warn against the use of the Score test for assessing the proportional odds and parallel slopes assumptions due to its extreme anti-conservatism. Hence, we used graphical methods to assess the validity of these assumptions. Although we have illustrated the general form of Anderson's Stereotype logistic model, we did not fit this model to our data due to the lack of availability of statistical software.

APPLICATION

The data described in Table 3 were derived from the Nova Scotia Atlee perinatal database of the Reproductive Care Program of Nova Scotia, Canada. The ordinal response variable 'degree of laceration' refers to the wounds of the perineum as a consequence of performing an episiotomy. Episiotomy is an obstetric surgical procedure for the enlargement of the vaginal opening just prior to delivery.[17] There are two commonly accepted techniques for episiotomy: midline and mediolateral procedures. The midline procedure is performed by making a midline perineal incision and directed toward the rectum. The mediolateral procedure is performed by making a perineal incision at the midline directed obliquely away from the rectum.[17] The response variable, laceration, is coded on a five-point ordinal scale, classified as '1°' (least severe) to '4°' (most severe), the classification based on the amount of tissue damage involvement, and a fifth group consisting of women free of any laceration.[17,18] For the purpose of this paper, we will restrict our analysis to midline episiotomy, coded as '0' indicating the absence of midline episiotomy and '1' if the procedure was performed.

The proportional odds (PO) and continuation-ratio (CR) models (models 2 and 4, respectively) were fit to the data described in Table 3, and their results summarized in Table 4. Midline episiotomy carried with it a relative risk of 2.1 (95% CI : 1.8–2.5) compared to no episiotomy for 'any laceration' by the PO model, and a relative risk of 1.4 (95% CI : 1.3–1.5) by the CR model. It is important to note that the assumptions of the underlying models differ: the PO model assumes that the relative risk associated with 'any laceration' is equivalent when comparing 4° versus none to 1° – 3° (combined), 3° – 4° (combined) versus none plus 1° – 2° (combined), and so on. In constrast, the CR model assumes that the relative risk associated with 'any laceration' is equivalent to 4° versus 3°, 3° – 4° (combined) versus 2°, and so on. The likelihood ratio test of H_0: $\beta = 0$ is rejected both for the PO and the CR models, implying that midline episiotomy is a strong predictor of lacerations during pregnancy, although both models violated the proportional odds and parallel slopes assumptions (discussed later).

The results of fitting a partial proportional odds model to the laceration data are summarized in Table 5a. Based on the fit of an unconstrained model (model 6), the estimated log odds ratio comparing women with 'any laceration' to no laceration in relation to midline episiotomy is β (0.7125), whereas the log odds comparing 2° – 4° to none plus 1°, 3° – 4° to none plus 1° – 2°, and 4° to none plus 1° – 3° are $\beta + \gamma_2$, $\beta + \gamma_3$, and $\beta + \gamma_4$, respectively. The corresponding log-odds ratios with the standard errors are presented in Table 5. A

TABLE 4 *Maximum likelihood estimates: Results of fit of proportional odds (PO model) and continuation-ratio (CR model) models*[a]

Variable	PO model $\hat{\beta} \pm \hat{se}(\hat{\beta})$	*P*-value	CR model $\hat{\beta} \pm \hat{se}(\hat{\beta})$	*P*-value
Intercept$_1$ (α_1)	–2.8452 ± 0.0336		–3.9387 ± 0.0484	
Intercept$_2$ (α_2)	–3.0363 ± 0.0363		–0.8260 ± 0.0213	
Intercept$_3$ (α_3)	–3.1680 ± 0.0383		–0.8260 ± 0.0213	
Intercept$_4$ (α_4)	–5.0027 ± 0.0883		–0.8260 ± 0.0213	
Midline episiotomy	0.7423 ± 0.0946	0.0001	0.2968 ± 0.0345	0.0001
Model fit				
Likelihood Ratio test, χ^2_1	52.9	0.0001	74.7	0.0001

[a] Response variable is degree of laceration: no laceration, 1°, 2°, 3°, and 4°.
Model: logit[$Pr(Y \leq y_j)$] = $\alpha_j - \beta$ (Episiotomy).

TABLE 5 *Maximum likelihood estimates: Results of fit of partial-proportional odds models*[a]

Variable	Parameter	Estimate ± SE	χ^2	*P*-value
a. Unconstrained model				
Intercept$_1$	α_1	–2.8431 ± 0.0336		
Intercept$_2$	α_2	–3.0543 ± 0.0370		
Intercept$_3$	α_3	–3.1972 ± 0.0395		
Intercept$_4$	α_4	–5.1914 ± 0.1034		
Midline episiotomy	β	0.7125 ± 0.0946	56.7	0.0001
Laceration ⩾ 2°	γ_2	0.1471 ± 0.0278	28.0	0.0001
Laceration ⩾ 3°	γ_3	0.2225 ± 0.0393	32.1	0.0001
Laceration ⩾ 4°	γ_4	0.9394 ± 0.1738	29.1	0.0001
Likelihood ratio test	χ^2_4		89.8	0.0001
Midline episiotomy	χ^2_4		116.0	0.0001
b. Constrained model[b]				
Intercept$_1$	α_1	–2.8432 ± 0.0336		
Intercept$_2$	α_2	–3.0455 ± 0.0363		
Intercept$_3$	α_3	–3.2026 ± 0.0393		
Intercept$_4$	α_4	–5.1510 ± 0.0939		
Midline episiotomy	β	0.7160 ± 0.0946	57.3	0.0001
Constraint parameter	γ	0.1126 ± 0.0165	46.6	0.0001
Likelihood ratio test	χ^2_2		85.6	0.0001
Midline episiotomy	χ^2_2		108.1	0.0001

[a] Response variable is degree of laceration: no laceration, 1°, 2°, 3°, and 4°.

Model (a): $\log\left[\frac{Pr(Y \leq y_j)}{Pr(Y > y_j)}\right] = \alpha_j - \beta$(Episiotomy) – γ_2(Episiotomy:Lacr ⩾2°) – γ_3(Episiotomy:Lacr ⩾3°) – γ_4(Episiotomy:Lacr ⩾4°)

Model (b): $\log\left[\frac{Pr(Y \leq y_j)}{Pr(Y > y_j)}\right] = \alpha_j - \beta$(Episiotomy) – $\gamma\Gamma_j$(Episiotomy) .

Constraints: $\Gamma_1 = 0$, $\Gamma_2 = 1$, $\Gamma_3 = 2$, and $\Gamma_4 = 7$.

TABLE 6 *Results of fit of polytomous logistic (PL model) and adjacent category (AC model) logistic models*[a]

Variable	PL model $\hat{\beta} \pm \hat{se}(\hat{\beta})$	*P*-value	AC model $\hat{\beta} \pm \hat{se}(\hat{\beta})$	*P*-value
Intercept$_1$ (α_1)	–4.5479 ± 0.0771		–5.0987 ± 0.1013	
Intercept$_2$ (α_2)	–5.0987 ± 0.1013		–0.5508 ± 0.1268	
Intercept$_3$ (α_3)	–3.3328 ± 0.0425		1.7659 ± 0.1093	
Intercept$_4$ (α_4)	–5.1404 ± 0.1013		–1.8076 ± 0.1113	
Midline episiotomy				
1°	–0.4661 ± 0.3628	0.1988	–0.4661 ± 0.3630	0.1991
2°	0.0847 ± 0.3687	0.8182	0.5508 ± 0.5158	0.2856
3°	0.7280 ± 0.1178	0.0001	0.6433 ± 0.3849	0.0947
4°	1.6845 ± 0.1945	0.0001	0.9565 ± 0.2235	0.0001

[a] Response variable is degree of laceration: none, 1°, 2°, 3°, and 4°.

likelihood ratio test of H_0: $\beta = \gamma_j = 0$ resulted in a $\chi^2 = 89.8$ (4 d.f.), a significant improvement over the proportional odds model.

We also fit a constrained partial proportional odds model (Table 5b) to the data, with the following constraints specified *a priori*: $\Gamma_1 = 0$, $\Gamma_2 = 1$, $\Gamma_3 = 2$, and $\Gamma_4 = 7$. These resulted in the following log-ratios: β, $\beta + \gamma$, $\beta + 2\gamma$, and $\beta + 7\gamma$, for the four logits, as described earlier. Our choice of constraints were based on examining the log odds ratios from the observed data, which were derived by constructing four 2×2 tables, with episiotomy (yes/no) as the two rows, and lacerations 'any' versus 'none' as the columns for the first table; 2° – 4° versus none plus 1° for the second table, and so on. A simultaneous test of H_0: $\beta = 0$, $\gamma = 0$ (based on 2 d.f.) resulted in a $\chi^2 = 85.6$, implying good fit. Notice that the choice of different constraints will produce different parameter estimates and standard errors.

The results of fitting the polytomous logistic and the adjacent-category logistic models are summarized in Table 6. The formulation of these models is more flexible when compared to the proportional odds and continuation-ratio models, in that the regression coefficients corresponding to a covariate (such as midline episiotomy) is allowed to vary by every level of the ordinal response. The regression coefficient corresponding to the polytomous model for the first logit comparison (1° versus no laceration) is negative (–0.4661), resulting in a relative risk estimate of 0.6 (95% CI : 0.3–1.3). However, the coefficients for other logit comparisons are all positive, implying that midline episiotomy increases the likelihood of lacerations of ⩾2°. Women with midline episiotomy are 5.4 (95% CI : 3.7–7.9) times at greater risk of a 4° laceration relative to women with no episiotomy. Notice the monotonic increase in the regression coefficients for midline episiotomy. The results of fitting an adjacent-category logit model are somewhat similar to that of the polytomous model, although the underlying model assumption differ; the latter model involves comparing women with 1° to no laceration (first logit), 2° to 1° (second logit), and so on, while the logit comparisons for the polytomous model correspond to each category of laceration versus a baseline category (i.e. women with no laceration).

DISCUSSION

Choice of an Ordinal Model

The choice between the cumulative logit and continuation-ratio models merit further discussion. Armstrong and Sloan[15] argue that when the cumulative logit model is valid implying that the cumulative log-odds ratios are a constant, say β^*, then the continuation-ratio model will begin at $\beta_j = \beta^*$, but will approach 0 as j increases. This argument has led to the cumulative logit model being proposed as an adjunct to the Cox's proportional hazards model[13] i.e. continuation-ratio model for survival data when hazard rates of groups are thought likely to converge with time (see McCullagh[4] and McCullagh and Nelder[14] for a more thorough discussion).

The choice of a simple logistic model to an ordered response has its own disadvantages. Based on extensive simulations, Armstrong and Sloan[15] conclude that the logistic model attains only between 50 and 75% efficiency, relative to the cumulative logit model for a five-level ordered response. Moreover, efficiency is based on the assumption that the dichotomization for the logistic model is made close to the optimal point. In reality, dichotomization can be somewhat arbitrary

which may violate the underlying assumptions of the model. The arbitrariness is worsened in situations when a logistic model is fit to a response that has many ordered categories.

A more intuitive choice between the proportional odds and continuation-ratio models can be based on the goals of the statistical analysis. Assuming that both models are valid, if an *a priori* interest is to estimate the risk of 4° laceration relative to other groups (none and 1° – 3° combined), then the cumulative logit model is the obvious model choice. On the contrary, if the analyst is interested in estimating the risk by comparing 4° laceration to 3°, then the continuation-ratio model is the preferred model. In general, the choice of a model depends on how the logits are formulated, *a priori*. The analyst should, however, be wary of departures from the underlying model assumptions (proportional odds and parallel slopes assumptions), as was the case in our data. Tests for model assumptions could also be viewed as goodness-of-fit tests of the link functions; Holtbrugge and Schumacher[19] provide a detailed review of such tests.

Since the formulation of the logit functions in the proportional odds and the partial-proportional odds model are identical (i.e. 4° versus 1° – 3° plus no laceration, 3° – 4° versus 1° – 2° plus no laceration, etc), the overall fit of these models are comparable. The proportional odds model can be viewed as a model 'nested' within the unconstrained partial-proportional odds model. The deviance[14] (defined as the difference in the likelihood ratios between two nested models) is $\chi^2 = 36.9$ (89.8 – 52.9) with 2 d.f. (4 – 2), favouring the unconstrained partial-proportional odds model as a better fit to the data. Applying the same argument, the deviance comparing the likelihood ratios between the unconstrained and the constrained partial proportional odds models is $\chi^2 = 4.2$ (89.8 – 85.6) with 2 d.f. (4 – 2), $0.1 \leqslant P \leqslant 0.25$. This suggests that the unconstrained partial proportional odds model, once again, fits the data better than the constrained partial proportional odds model.

Graphical Methods of Assessing Model Fit and Model Constraints

The assumptions of proportional odds and parallel slopes in the proportional odds and continuation ratio models, respectively, were examined by graphical methods. First, consider the fit of the proportional odds, unconstrained, and constrained partial proportional odds models to the laceration data. Relative risks estimated from each of these three models were contrasted to those based on the observed data (Figure 1). Clearly, the fit of a proportional odds model, constraining the relative risk to be 2.1 performs the least satisfactorily, while the fit of a constrained and unconstrained partial proportional odds models are almost identical, but an improvement over the proportional odds model. Note, however, that based on comparing the likelihood ratios between these two models the unconstrained partial proportional odds model performs better than the constrained model.

Similarly, the fit of a continuation-ratio model was compared to the observed data (Figure 2). Once again, the continuation ratio model does not adequately fit the data for the first continuation-ratio (relative risk = 2.7), whereas the model fits well for the other three continuation-ratios. On the other hand, when the ordering of the response variable is subjective, then a polytomous logistic model is more preferable than fitting $(k - 1)$ simple logistic models. If the ordering is valid, then an appropriate ordinal model must be chosen from the class of ordinal models, the choice made based on the goals of statistical analyses.

Finally, Koch *et al.*[20] developed a two-stage procedure called as Functional Assymptotic Regression Methodology (FARM) for fitting a partial proportional odds model based on the weighted least squares estimation procedure. Although we did not consider fitting this model, the interested reader is referred to the papers by Koch *et al.*,[20] and Peterson and Harrell[6] for thorough review and discussion.

CONCLUSIONS

This paper presents a synthesized review of generalized linear regression models for analysing ordered responses. The cumulative logit and the continuation-ratio models for ordinal responses have been the primary focus in epidemiological and biomedical applications,[1–10,15,19–23] while other models for the analysis of ordinal outcomes have received less attention. However, since all these models are developed under rather strong assumptions (such as the proportional-odds assumption), departures from these assumptions may well result in the incorrect model formulation. It is imperative that the analyst performs (i) goodness-of-fit tests and an analysis of residuals, and (ii) sensitivity analysis by fitting and comparing different models.[1]

Notice that when the number of response levels is two, all models discussed above reduce to the simple logistic model for binary responses:

$$Pr(Y = 1 \mid \mathrm{x}) = \frac{\exp(\alpha + \mathrm{x}'\beta)}{1 + \exp(\alpha + \mathrm{x}'\beta)} \qquad (13)$$

Our illustration of the models for ordered responses was based on a single covariate. Our intent here was to

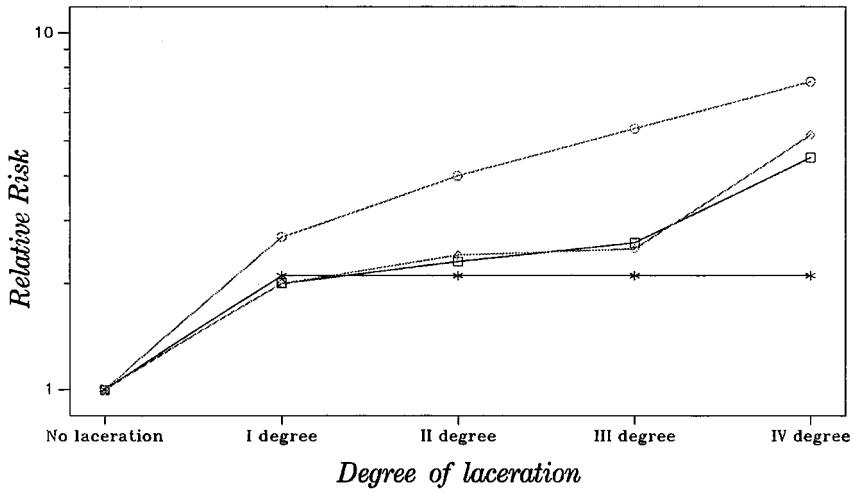

FIGURE 1 *Relative risks for lacerations by midline episiotomy based on observed data (o), proportional odds (⋆), constrained (□) and unconstrained (◇) partial proportional odds models*

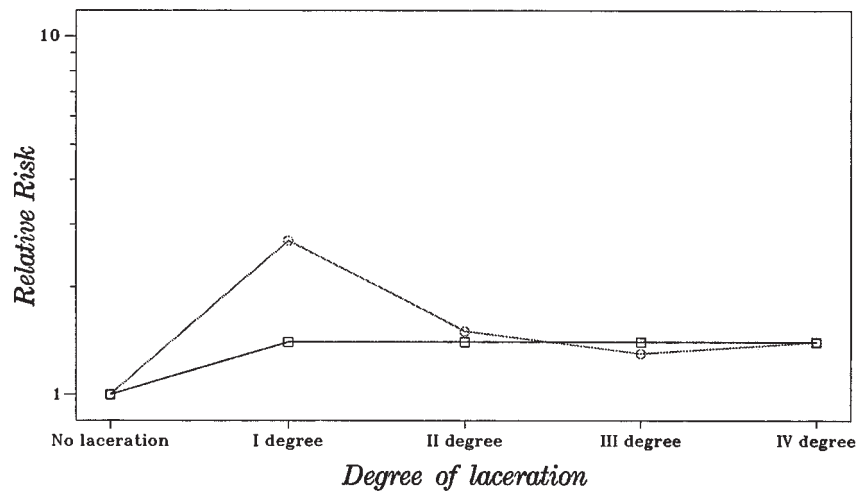

FIGURE 2 *Relative risks for lacerations by midline episiotomy based on observed data (o), and the continuation-ratio model (□)*

clearly illustrate the formulation and interpretation of every model using a practical, but simple data set. All the models described in this paper have much wider applications to situations involving several covariates under a variety of sampling situations (see for example, the work by Greenland[12] and Anderson and Philips[23]).

ACKNOWLEDGEMENTS

The Center for Perinatal Health Initiatives is supported, in part, by a grant (#029553) from the Robert Wood Johnson Foundation, NJ.

We thank Professors David A Savitz, John M Thorp, Jr (University of North Carolina at Chapel Hill) and John C Smulian (University of Medicine and Dentistry of New Jersey) for reviewing the manuscript, and Edwin R Luther (Dalhousie University, Nova Scotia, Canada) for graciously allowing us to utilize the Nova Scotia Atlee perinatal database cited in the 'Applications' section. This paper formed part of the masters thesis of Cande Ananth in the Department of Biostatistics at the University of Carolina at Chapel Hill.

REFERENCES

[1] Greenland S. An application of logistic models to the analysis of ordinal responses. *Biometrical J* 1985; **27:** 189–97.

[2] Cox C, Chuang C. A comparison of Chi-square partitioning and two logit analyses of ordinal pain data from a pharmaceutical study. *Stat Med* 1984; **3:** 273–85.

[3] Walker S H, Duncan D B. Estimation of the probability of an event as a function of several independent variables. *Biometrika* 1967; **54:** 167–79.

[4] McCullagh P. Regression models for ordinal data (with discussion). *J R Statis Soc Series B* 1980; **42:** 109–42.

[5] Feinberg B. *Analysis of Cross-Classified Data. Second Edition*. Cambridge: Massachusetts Institute of Technology Press, 1980.

[6] Peterson B L, Harrell F E. Partial proportional odds models for ordinal response variables. *Appl Stat* 1990; **39:** 205–17.

[7] Engel J. Polytomous logistic regression. *Stat Neerlandica* 1988; **42:** 233–52.

[8] Hosmer D W, Lemeshow D. *Applied Logistic Regression*. New York: John Wiley, 1989.

[9] Agresti A. *Analysis of Ordinal Categorical Data*. New York: John Wiley and Sons, 1984.

[10] Anderson J A. Regression and ordered categorical variables (with discussion). *J R Statis Soc Series B* 1984; **46:** 1–30.

[11] Peterson B, Harrell F E. 'Partial proportional odds models and the LOGIST procedure'. Paper presented at SAS Users Group International, thirteenth annual conference, Orlando FL, 1988.

[12] Greenland S. Alternative models for ordinal logistic regression. *Stat Med* 1994; **13:** 1665–77.

[13] Cox D R. Regression models and life tables (with discussion). *J R Statis Soc Series B* 1972; **34:** 187–220.

[14] McCullagh P, Nelder J H. *Generalized Linear Models. Second Edition*. London, UK: Chapman and Hall, 1989.

[15] Armstrong B G, Sloan M. Ordinal regression models for epidemiologic data. *Am J Epidemiol* 1989; **129:** 191–204.

[16] Läära E, Mathews J N S. The equivalence of two models for ordinal data. *Biometrika* 1985; **72:** 206–07.

[17] Thorp J M Jr. Episiotomy. In: Repke J T (ed.). *Intrapartum Obstetrics*. New York: Churchill Livingstone, 1996.

[18] Green J R, Soohoo S L. Factors associated with rectal injury in spontaneous deliveries. *Obstet Gynecol* 1989; **73:** 732–38.

[19] Holtbrugge W, Schumacher M. A comparison of regression models for the analysis of ordered categorical data. *Appl Stat* 1991; **40:** 249–59.

[20] Koch G G, Amara I A, Singer J M. A two-stage procedure for the analysis of ordinal categorical data. In: Sen P K (ed.). *Statistics in Biomedical, Public Health, and Environmental Science*. North Holland: Elsevier Science Publishers, 1985, pp. 357–87.

[21] Ashby D, Pocock S J, Shaper A G. Ordered polytomous regression: An example relating serum biochemistry and haematology to alcohol consumption. *Appl Stat* 1986; **35:** 289–301.

[22] Greenwood C, Farewell V. A comparison of regression models for ordinal data in an analysis of transplanted-kidney function. *Can J Stat* 1988; **16:** 325–35.

[23] Anderson J A, Philips P R. Regression, discrimination and measurement models for ordered categorical variables. *Appl Stat* 1981; **30:** 22–31.

(Revised version received February 1997)