Community Dent Oral Epidemiol 1998; 26: 1–6
Printed in Denmark . All rights reserved

COMMUNITY DENTISTRY AND ORAL EPIDEMIOLOGY
ISSN 0301-5661

Increased efficiency of analyses: cumulative logistic regression *vs* ordinary logistic regression

George W. Taylor and Mark P. Becker
University of Michigan School of Dentistry and University of Michigan School of Public Health, Ann Arbor, USA

Taylor GW, Becker MP: Increased efficiency of analyses: cumulative logistic regression *vs* ordinary logistic regression. Community Dent Oral Epidemiol 1998; 26: 1–6. © Munksgaard, 1998

Abstract – The common practice of collapsing inherently continuous or ordinal variables into two categories causes information loss that may potentially weaken power to detect effects of explanatory variables and result in Type II errors in statistical inference. The purpose of this investigation was to illustrate, using a substantive example, the potential increase in power gained from an ordinal instead of a dichotomous specification for an inherently continuous response. Ordinary (OLR) and cumulative logistic regression (CLR) modeling were used to test the hypothesis that the risk of alveolar bone loss over 2 years is greater for subjects with poorer control of non-insulin-dependent diabetes mellitus (NIDDM) than for those who do not have diabetes or have better controlled NIDDM. There were 359 subjects; 21 of whom had NIDDM. Analysis of main effects using OLR for the dichotomous outcome (no change in radiographic bone loss *vs* any change) produced parameter estimates for better control and poorer control that were not statistically significant. CLR analysis of main effects using a 4-category ordinal specification for radiographic bone loss also produced a parameter estimate for better control that was not statistically significant, but which estimated poorer control to have a significant effect. The fit of this CLR model was significantly better at $P<0.05$ than that for the OLR. While an OLR model testing the interaction between age and control status did not converge after 100 iterations, the CLR interaction model converged without difficulty and estimated a significant effect for interaction between age and poorer control. Results from the CLR analysis, in contrast to the OLR model, would lead one to conclude that the risk for more severe bone loss progression after 2 years is greater in subjects with poorer controlled NIDDM and that subjects with better controlled NIDDM may not have greater risk of bone loss progression than those without diabetes. The use of an ordinal instead of a dichotomous specification for an inherently continuous response provided increased power, more precise parameter estimates, and a significantly better fitting model. In estimating parameter estimates for odds ratios or risks, it is important to consider using ordinal logistic regression where the response is inherently continuous or ordinal.

Key words: adults; diabetes mellitus; epidemiology; non-insulin-dependent diabetes mellitus; periodontal diseases; statistical methods

George W. Taylor, Department of Cariology, Restorative Sciences, and Endodontics, University of Michigan School of Dentistry, 1011 N. University, Ann Arbor, MI 48109-1078, USA
Tel: +1-313-764-1737
Fax: +1-313-936-1597
e-mail: gwt@umich.edu

Accepted for publication 3 September 1997

Many types of oral health status data are continuous or ordinal but, when evaluated as outcomes, often dichotomized for analysis using standard techniques for binary data, e.g. contingency table analysis or logistic regression. For example, periodontal disease status may be measured using probing depth or clinical attachment level. Though both are continuous variables they are often summarized as dichotomous responses, classifying periodontitis as present or absent, evaluated at, for example, a cut-point for clinical loss of attachment of 4 mm or more at any tooth or probing depth of 5 mm or more in 2 or more teeth. Similarly, other oral health status measures are ordinal; pain might be measured on a scale using ''absent'', ''mild'', ''moderate'', or '' severe'' as response categories.

These ordinal data are also often dichotomized for analysis, as in classifying pain as absent or present. Such an approach is valid and useful when the relationship of interest is defined exclusively in terms of a dichotomy, such as the presence or absence of a condition or response above or below a threshold, but in general alternative approaches that take account of the ordinal nature of the response should be considered. Dichotomizing inherently continuous or ordinal variables can result in information loss, and subsequently in a loss of statistical power for detecting relationships between the response and explanatory variables. That is, the major limitation arising from the loss of information is the potential weakening of power to detect effects of explanatory variables on responses and committing Type II errors in statistical inference (i.e., failing to reject the null hypothesis when it is false). An ordinal specification of an ordinal or continuous response uses more of the available information derived from measuring the response. Use of regression models for ordinal responses can result in a meaningful gain in statistical power that is reflected in more efficient estimates of coefficients of explanatory variables than the estimates derived from models limited to analysis of dichotomous responses. Armstrong & Sloan describe, in detail, the relationships among loss of information, collapsing ordinal or continuous responses into two categories, and the relative efficiency of ordinary logistic regression and cumulative odds models (1).

Several types of regression models have been developed to analyze ordinal responses with more than two categories (2, 3), thus providing ways to eliminate this information loss in truly ordinal responses and greatly reduce it in continuous variables by allowing for collapsing into more ordinal categories than the often used dichotomy. The most frequently encountered extensions of ordinary logistic regression for the analysis of ordinal responses are adjacent-categories logit models, continuation-ratio logit models, and cumulative odds models (2). McCullagh (3) has been influential in establishing cumulative odds models with a proportional odds assumption as appropriate and useful for the analysis of ordinal responses, particularly when there is an (assumed, but in some cases not explicit) underlying continuous variable for which the observed manifestation is the ordinal response variable. Such models provide a means to estimate parameters useful in calculating odds ratios that quantify the effects of multiple explanatory variables on an ordinal response. The remainder of this discussion focuses on proportional cumulative odds, and their extension when the proportionality assumption is violated. Model fitting, parameter interpretation, and testing for the proportionality assumption are addressed.

The purpose of this paper is to illustrate, using a substantive example, the potential increase in power that can be attained by using an ordinal instead of a dichotomous specification for a response that is inherently continuous. This increase in power is achieved by using cumulative logistic regression (to estimate the proportional odds model) in a case where ordinary logistic regression did not estimate a statistically significant effect (at $P<0.05$) for the exposure of interest.

Methods

The data for this analysis came from 359 subjects who were part of a longitudinal study of non-insulin-dependent diabetes mellitus (NIDDM) and periodontal disease (as measured by radiographic bone loss over two years) in the Gila River Indian community; 21 subjects had NIDDM.

Full details describing the subjects and variables used in this analysis have been described elsewhere (4). Briefly, data were derived from 359 subjects, aged 15–57, of whom 338 were free of diabetes at the beginning and did not develop NIDDM during a 2-year follow-up period. The other 21 subjects had NIDDM at baseline. The subjects selected were all those who had 20 or more teeth, lost no teeth during the study, and had less than 25% radiographic bone loss at baseline. The hypothesis tested in this analysis is: the risk of alveolar bone loss over 2 years is greater for subjects with poorer control of NIDDM, as measured by glycosylated hemoglobin values ≥9% (HbA1 ≥9%), than for those without diabetes (NoNIDDM) or with better controlled NIDDM (HbA1 <9%). The response, change in radiographic bone score, was determined from panoramic radiographs, using measures of interproximal bone loss throughout the dentition, with a modified Schei technique (5). The percentage of bone loss for each tooth was recorded on a 0–4 ordinal scale for that tooth, the score corresponding to percentage of radiographic bone loss. Bone scores (scale 0–4) from the panoramic radiographs corresponded to bone loss of 0%, 1–24%, 25–49%, 50–74%, or 75% and greater. Change in bone score category was computed as the change in worst bone score (WBS) reading after

Table 1. Dichotomous response distribution for change in worst bone score at follow-up (percentage of subjects in each row in parentheses)

	Degree of change in WBS at follow-up		
Diabetes status	None	Some	Total
NoNIDDM	201 (59.5)	137 (40.5)	338
NIDDM: better control	6 (42.9)	8 (57.1)	14
NIDDM: poorer control	1 (14.3)	6 (85.7)	7

Table 2. Ordinal response categories distribution for change in worst bone score at follow-up (percentages of subjects in rows in parentheses)

Diabetes status	Degree of change in WBS at follow-up – number of categories changed				
	None	1	2	3–4	Total
NoNIDDM	201 (59.5)	126 (37.3)	7 (2.1)	4 (1.2)	338
NIDDM: better control	6 (42.9)	6 (42.9)	2 (14.3)	0	14
NIDDM: poorer control	1 (14.3)	4 (57.1)	1 (14.3)	1 (14.3)	7

2 years. Change in WBS category, the outcome of interest in this analysis, was specified in two ways: (a) a dichotomy for no change *vs* any change and (b) a 4-category ordinal scale representing no change, a 1-category increase, a 2-category increase, or a 3- or 4-category increase over baseline WBS.

The covariates used were glycemic control status, time to follow-up examination, baseline age, baseline worst bone score, and calculus index. Age, calculus index, and time to follow-up examination were used in the models as continuous covariates. Baseline worst bone score was specified as a dichotomous variable (0% or 1–24%) and control status was specified using two dummy variables, better control and poorer control, with NoNIDDM as the reference group. In the specification for control status, better control was HbA1 $<9\%$ with NoNIDDM and poorer control as the referent categories, and poorer control was HbA1 $\geq 9\%$ with NoNIDDM and better control as the referent categories. Logistic regression models, using both dichotomous and ordinal responses for WBS, were developed to test the effects of other covariates on the WBS-glycemic control association. We used SAS Proc Logistic (6) to perform the regression analyses, testing both main effects and interaction models. Further details of these modeling procedures have been described previously (4). The discussion here is limited to the final models selected at the completion of the detailed modelling procedures.

Table 3. Summary table of change in worst bone score for ordinary and cumulative logistic regression models (OLR and CLR, respectively). Number of subjects, $n=359$

	OLR: dichotomous outcome		CLR: main effects model		CLR: interaction model	
		P-value		*P*-value		*P*-value
Model evaluation statistics						
L.R. CHI-SQ. STAT.[a]	72.3 (6 df)	0.0001	79.4 (6 df)	0.0001	89.34 (7 df)	0.0001
SCORE TEST P.O.A.[b]	n.a.		14.27 (12 df)	0.2841	27.15 (14 df)	0.0184
Maximum likelihood estimates	Coefficient estimate (s.e.[d])	*P*-value[c]	Coefficient estimate (s.e.[d])	*P*-value[c]	Coefficient estimate (s.e.[d])	*P*-value[c]
Main effects terms						
Intercept 1	−3.460 (0.562)	0.0001	−3.385 (0.513)	0.0001	−3.612 (0.526)	0.0001
Intercept 2	n.a.		−6.699 (0.647)	0.0001	−7.039 (0.673)	0.0001
Intercept 3	n.a.		−7.886 (0.753)	0.0001	−8.260 (0.780)	0.0001
Better control	0.543 (0.632)	0.3904	0.769 (0.567)	0.1750	0.742 (0.569)	0.1919
Poorer control	2.158 (1.133)	0.0568	2.438 (0.785)	0.0019	9.131 (3.133)	0.0036
Age	0.100 (0.020)	0.0001	0.104 (0.018)	0.0001	0.118 (0.019)	0.0001
Follow-up time	0.358 (0.112)	0.0014	0.306 (0.106)	0.0038	0.293 (0.107)	0.0061
Baseline WBS	−1.746 (0.305)	0.0001	−1.749 (0.292)	0.0001	−1.799 (0.295)	0.0001
Calculus index	1.193 (0.301)	0.0001	1.087 (0.281)	0.0001	1.048 (0.282)	0.0002
Interaction term						
Poorer contrl*age	n.a.		n.a.		−0.253 (0.126)	0.0447

[a] Likelihood ratio chi-square statistic.
[b] Score test for the proportional odds assumption.
[c] **P*-value of the Wald chi-square statistic with respect to a chi-square distribution with one degree of freedom.

Table 4. Change in maximum worst bone score, by diabetes status and baseline age, for subjects included in the cumulative logistic regression analysis for the period, baseline to 2-year follow-up examination. Percentage of subjects in rows in parentheses

Diabetes status	Baseline age (years)	Degree of change in WBS at follow-up: number of categories changed				Number of subjects
		None	1	2	3–4	*n*
NoNIDDM	15–19	82 (71.9)	30 (26.3)	2 (1.75)	0	114
	20–34	110 (56.7)	78 (40.2)	4 (2.1)	2 (1.03)	194
	35–57	9 (30.0)	18 (60.0)	1 (3.3)	2 (6.7)	30
	Total	201 (59.5)	126 (37.3)	7 (2.1)	4 (1.2)	338
NIDDM: better control	15–19	1 (100.0)	0	0	0	1
	20–34	4 (40.0)	6 (60.0)	0	0	10
	35–57	1 (33.3)	0	2 (66.7)	0	3
	Total	6 (42.9)	6 (42.9)	2 (14.3)	0	14
NIDDM: poorer control	15–19	0	1 (100.0)	0	0	1
	20–34	0	3 (60.0)	1 (20.0)	1 (20.0)	5
	35–57	1 (100.0)	0	0	0	1
	Total	1 (14.3)	4 (57.1)	1 (14.3)	1 (14.3)	7

Table 5. Constituent tables of the proportional odds model

Diabetes status	Degree of change in WBS at follow-up								
	Cut-point 1			Cut-point 2			Cut-point 3		
	0	*vs*	1, 2, 3–4	0, 1	*vs*	2, 3–4	0, 1, 2	*vs*	3–4
NoNIDDM	201		137	327		11	334		4
NIDDM: better control	6		8	12		2	14		0
NIDDM: poorer control	1		6	5		2	6		1

Results

The data in Tables 1 and 2 reflect the dichotomous and ordinal specifications, respectively, of change in worst bone score category, crudely classified by glycemic control status. Table 1 shows an increasingly higher proportion of subjects with any increase in WBS at follow-up as glycemic control worsens. Table 2, showing the ordinal change in worst bone score, provides additional information about severity of progression of radiographic bone loss. The data in Table 2 suggest that poorer glycemic control is associated with more severe progression of radiographic bone loss, as reflected in the tendency for higher proportions of subjects to have greater changes in WBS as glycemic control worsens. These two-way contingency tables provide useful information about the bivariate relationships between glycemic control status and change in radiographic bone score; however, they are limited in not allowing for simultaneous evaluation of other variables that might confound or modify the effects seen.

The multivariate analyses, using logistic regression models for both dichotomous and ordinal responses, provided the final models shown in Table 3. The analysis of main effects using ordinary logistic regression for the dichotomous outcome (no change in WBS *vs* any change), as shown in Table 3, produced parameter estimates for better control (HbA1 $<9\%$) and poorer control (HbA1 $\geq 9\%$) that were not statistically significant at the $P<0.05$ level ($\beta=0.543$ and 2.16; s.e.[d]$=0.633$ and 1.13; P-values$=$ 0.3904 and 0.057 respectively). Further, the 95% confidence intervals for the estimated odds ratios included the value 1 (0.5–5.9 and 0.9–79.7, respectively) thus suggesting no effect. We were not able to test for significant interaction because an interaction model testing for the interaction between age and control status did not converge after 100 iterations; hence no parameters were estimated. Results from this analysis using ordinary logistic regres-

sion might lead to the conclusion that there is no difference in the effect of poorer control of NIDDM on any bone loss progression after 2 years.

The analysis of main effects using cumulative logistic regression (Table 3) produced a parameter estimate for better control that was not statistically significant ($\beta=0.769$; s.e.[d]$=0.567$; P-value$=0.17$), but estimated poorer control to have a significant effect ($\beta=2.438$; s.e.[d]$=0.785$; P-value$=0.0001$). The odds ratio for poorer control was 11.45 and the 95% confidence interval (2.46–53.35) did not include the null value. The interaction model converged without difficulty and estimated a significant effect for interaction between age and poorer control. The score test for the proportional odds assumption, ($P=0.0184$), suggests that separate parameters for each cut-point may be required in the interaction model (6), though this violation of the proportional odds assumption may result from sparse data in the oldest ages, particularly among those with poor control, as shown in Table 4. Results from this analysis using cumulative logistic regression for the ordinal response specification would lead one to conclude that the risk of more severe bone loss progression after 2 years is greater in subjects with more poorly controlled NIDDM and that subjects with better controlled NIDDM may not have a greater risk of bone loss progression than those without diabetes. An approach to fitting non-proportional odds models using the SAS system is described in the *Discussion* section.

Discussion

This analysis attained increased analytic power by using cumulative logistic regression (with a 4-category ordinal response) in a case where ordinary logistic regression (with a dichotomous response) did not estimate a statistically significant effect (at $P<0.05$) for the exposure of interest, poor glycemic control. This increase in power will not be attained in all instances where there is a statistically insignificant effect of an explanatory variable on a continuous or ordinal response that has been collapsed into a dichotomous variable. It should also be noted that inferences based on cumulative logistic regression, like other regression methodologies, will be influenced by outlying observations. One or more outlying observations may cause Type I errors (rejecting a true null hypothesis). In those cases where there are outliers on the ordinal response scale it is prudent to perform separate analyses with and without the outliers present. Differences in substantive findings between the analyses require a qualified reporting of results. It may also be desirable in such cases to explore recodings of the outlying observations. Care must be taken to establish the substantive justification and consequences of recoding responses. Nevertheless, it is worthwhile to consider using cumulative logistic regression where ordinal specifications can be used for dichotomous responses derived from continuous or ordinal variables.

The proportional odds models model for cumulative probabilities, i.e. the probability of all response categories up to and including the response category of interest *vs* the probability of all response categories beyond the response category of interest (hence the term cumulative logistic regression). This model allows the calculation of the odds ratio of cumulative probabilities and is interpreted as the odds of being in response category greater than or equal to j compared to all categories less than j or, because it is palindromically invariant, the odds of being in response category less than or equal to j compared to all categories greater than j. The proportional odds model can be written:

$$L_j(\mathbf{x}) = \log\left[\frac{(\Pr\ Y\geq j|\mathbf{x})}{(\Pr\ Y<j|\mathbf{x})}\right] = A_j+B_1x_1+B_2x_2+\ldots+B_nx_n$$

where $\Pr(Y\geq j|\mathbf{x})$ is the cumulative probability for response category j, when the explanatory variables take a particular set of values, x; A_j are the cut-point parameters (equivalent to separate intercepts for each cut-point); and $B_{i(i=1,\ldots,n)}$ are the coefficients estimated for each explanatory variable, x_n. The score test for the proportional odds assumption compares this model (i.e., as a null model) to one (the alternative model) where a separate set of regression coefficients is estimated for each cut-point j. A significant test result implies that the above model should be rejected in favor of one where the values of at least some of the coefficients vary across two or more cut-points.

Table 5 presents the constituent tables of the cumulative odds model used for the comparisons made in this analysis. Note that while there are four response categories, three cut-points are created. Application of this model to analyze the effect of glycemic control on radiographic bone loss progression involved comparing the odds of having a 1-, 2,- or 3–4-category change in WBS *vs* no change; a 2- or 3–4-category change in WBS *vs* no

change or a 1-category change; and a 3–4-category change *vs* none, 1-, or 2-category change. If we consider NoNIDDM as the baseline category against which glycemic control status is being compared, B_1 is the coefficient for better control and B_2 the coefficient for poorer control, then, for subjects with NIDDM, e^{B1} would be an estimate of effect of better glycemic control, expressed as an odds ratio, and e^{B2} would be an estimate of the effect of poorer glycemic control, also interpreted as an odds ratio. A key assumption made with this model is the proportional odds assumption, i.e., the effect of each explanatory variable is the same for all the cut-points forming the cumulative logits (3). The appropriateness of the proportional odds assumption in modeling change in WBS as an ordinal outcome in this analysis is supported by the score test for the proportional odds assumption (14.27, 12 df, $P=$ 0.2841) in the main effects model; note, the score test for the proportional odds assumption is routinely included in the Proc Logistic output. The results from this test support using the same coefficient for covariates at each cut-point of the ordinal response (change in WBS in this case) to estimate the cumulative odds of being in the response category of interest compared to being in all categories greater than (or less than) the response category of interest. The proportional odds assumption is a simplifying one in that it allows comparison of the effect of different values of a covariate on the ordinal outcome without having to estimate a separate parameter for each value of the response (cut-point).

Models that do not make the proportional odds assumption generally do not fit automatically within the SAS system, though the clever programmer can make use of the ''BY'' command to do so in certain limited circumstances (which are beyond the scope of this paper). One approach to fitting non-proportional odds models using SAS, or any other software package that can be used to fit OLR models, is to fit a separate OLR for each of the cut-points. In the case of our example this means fitting three OLR models, one each for the three tables reported in Table 5. The net result is that there is a separate set of regression coefficients for each of the cut-points, rather than the single set that applies across all cut-points in a proportional odds model.

The use of an ordinal instead of a dichotomous specification for an inherently continuous response provided increased power, more precise parameter estimates, and a significantly better fitting model. By using cumulative logistic regression, we were able to detect an effect for the exposure of interest that would have been missed (or, at best minimized) if we had used only the ordinary logistic regression approach. In estimating parameter estimates for odds ratios or risks, it is important to consider using ordinal logistic regression when the response is inherently continuous.

Acknowledgments

This study was supported by NIDR grant numbers DE 06514 and DE 07157 and the Michigan Diabetes Research and Training Center.

References

1. Armstrong BG, Sloan M. Ordinal regression models for epidemiologic data. Am J Epidemiol 1989;129:191–204.
2. McCullagh P. Regression models for ordinal data. J R Stat Soc 1980;42:109–42.
3. Agresti A. Categorical data analysis. New York: J Wiley; 1990.
4. Taylor GW, Burt BA, Becker MP, Genco RJ, Shlossman M. Glycemic control and alveolar bone loss progression in Type II diabetes. Ann Periodontol. In press 1997.
5. Shlossman M, Knowler WC, Pettitt DJ, Genco RJ. Type 2 diabetes mellitus and periodontal disease. J Am Dent Assoc 1990;121:532–6.
6. SAS Institute Inc. SAS/STAT® software: changes and enhancements through release 6.11. Cary (NC): SAS Institute; 1996.