Describing the severity of mottling in a community: a different approach

William Butler*, Vincent Segreto** and Edwin Collins**

*Department of Biostatistics, University of Michigan, Ann Arbor, Michigan and **Dental School, University of Texas Health Science Center, San Antonio, Texas, USA

Butler W, Segreto V, Collins E: Describing the severity of mottling in a community: a different approach. *Community Dent Oral Epidemiol* 1985; 13: 277-80.

Abstract – Dean's community mottling index, though a useful summary measure, suffers from the use of arbitrary scores and the inability to describe the frequency in each of the levels of mottling. An alternative statistical approach, based on the proportional odds model, is used here to describe the severity of mottling in a community. The use of this model is demonstrated on prevalence data collected on 16 Texas communities.

Key words: dental enamel; mottled enamel; public health dentistry

W. Butler, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, USA

Accepted for publication 1 July 1984

A standard classification scheme for ordering the various degrees of severity of mottled enamel was developed by DEAN (1-3). This consists of a six-point scale ranging from "normal" to "severe".

DEAN and others proposed that, in addition to indicating an individual's severity of mottling, this classification scheme be used to compute an index to summarize the extent of mottling in a community. This community mottling index is calculated by assigning scores to each of the six categories of mottling and computing a weighted average. The community mottling index would then be used "for epidemiologic purposes and subsequent correlation with chemical and other studies" (4). This index is attractive since it has been found to be a useful summary of the severity of mottling in a community and exhibits a simple linear relationship with the natural logarithm of fluoride in the drinking water (5).

Major drawbacks of the community mottling index include its dependence on an imposed set of scores and its inability to describe the distribution of the population across the six categories of mottling. This latter point is considered crucial since it is the proportion of the population which falls in the different mottling categories, particularly moderate and severe, which is considered relevant in balancing the health costs and benefits of fluoride in drinking water. For example, DEAN describes a mottling classification scheme for communities in terms of the proportion of the population in the various mottling categories, not on the community mottling index (4). Similarly, in

estimating "optimal" levels of fluoride in drinking water, RICHARDS *et al.* (6) concentrated on the proportion of the population in the moderate and severe categories, not the community mottling index.

In this report, a proportional odds model (7) is used to describe the severity of mottling in a community. This statistical model provides a description of the distribution of the population across mottling categories based on DEAN's scheme. This approach does not require the assignment of scores to mottling categories, utilizes the qualitative ordering in severity of the mottling categories and provides an intuitively appealing description of the prevalence of mottling in the population.

The fit of the model to community mottling data collected on 16 Texas communities in 1980 (8) is examined. Based on previous work (9), consideration of predictor variables will be restricted to level of fluoride in the drinking water, mean maximum temperature, total dissolved solids, zinc, race and presence of air conditioning.

Material and methods

In this report, DEAN's six mottling categories are reduced to three levels: (1) normal/questionable; (2) very mild/mild; (3) moderate/severe. The collapsing of "moderate" and "severe" to a single category is reasonable since these are the two levels considered to be cosmetically objectionable. Also, in the data set to be analyzed, there is only a single individual with severe mottling, a frequency too low for proper statistical analysis. The grouping of the normals with the questionables

and the very milds with the milds was done for purposes of simplicity and is based on the similarity in the statistical characteristics of the frequencies of these categories. Thus, a trichotomous, qualitatively ordered response variable will be examined. The statistical methods to be discussed, however, are general enough to analyze a six-level response variable if desired.

The proportional odds model was proposed by MCCULLAGH (7) for the analysis of a qualitatively ordered response variable and categorical or continuous predictor variables. If $P_j(X)$ represents the probability of the response variable being less than or equal to j for a set of predictor variables X , then the proportional odds model is written

$$P_j(X) = \frac{1}{1 + K_j \exp(\beta'X)} = \frac{1}{1 + \exp(\alpha_j + \beta'X)}$$

Note that if there are J levels of the response variable, j ranges only from 1 to $J-1$ because of the constraint $P_J(X) = 1$. The constant terms will follow the ordering $\alpha_1 \leq \alpha_{j+1}$ due to the inherent ordering of the levels of the response variables.

The odds of an observation being less than j is written

$$\gamma_j(X) = P_j(X)/(1 - P_j(X)) = \exp(-(\alpha_j + \beta'X)),$$

and the ratio of corresponding odds is written

$$OR_j(X_1, X_2) = \gamma_j(X_1)/\gamma_j(X_2) = \exp(\beta'(X_2 - X_1)).$$

Positive values for the coefficients β_j thus indicate that higher values of X_j are as-

Table 1. Texas communities to be used for examining fit of proportional odds model, along with their prevalence of mottling and community-level predictor variables

Community	Sample size	Normal/questionable	Mottling		Fluoride (ppm)	*Factor-of-optimal	Predictors		Zinc (ppm)
			Very mild/mild	Moderate/severe			Temp (°F)	TDS (ppm)	
New Braunfels	102	93	9	0	0.2	0.3	78.6	415	18.4
San Marcos	197	180	17	0	0.2	0.3	78.4	491	284.0
San Antonio	114	113	1	0	0.3	0.4	78.8	398	2.1
Kingsville	314	181	132	1	0.7	1.0	83.0	1260	3.6
Alvin	193	139	52	2	1.0	1.3	76.6	785	22.3
Angleton	179	119	58	2	1.0	1.3	76.6	751	14.1
Kerrville	118	98	20	0	1.1	1.4	77.1	590	0.8
Alpine	22	6	13	3	1.8	2.3	76.1	245	192.8
Littlefield	101	24	61	16	1.9	2.3	72.8	550	68.1
Fort Stockton	277	96	171	10	1.9	2.5	79.0	1870	50.2
Hillsboro	192	16	169	7	2.1	2.7	76.8	1340	58.2
Monahan	163	39	102	22	2.1	2.7	77.6	550	11.0
Perryton	83	16	62	5	2.3	2.7	69.7	550	16.0
Gatesville	111	26	80	5	2.3	3.1	78.0	1340	3.5
Abernathy	64	4	38	22	2.4	2.9	73.2	550	373.0
Taylor	174	8	108	58	3.3	4.3	78.4	960	3.0
Total	2404	1158	1093	153					

*Calculated according to GALAGAN & VERMILLION (11).

Table 2. Comparison of observed and fitted percentages for mottling categories for 16 Texas communities. Fitted values obtained from model which uses a predictor's race, presence of air conditioning and indicator variables to indicate town

		Normal/questionable	Very mild/mild	Moderate/severe
New Braunfels	Obs	91.2	8.8	0.0
(102)	Fit	91.2	8.6	0.2
San Marcos	Obs	91.4	8.6	0.0
(197)	Fit	91.4	8.4	0.2
San Antonio	Obs	99.1	0.9	0.0
(114)	Fit	99.1	0.9	0.0
Kingsville	Obs	57.6	42.0	0.3
(314)	Fit	58.1	40.5	1.4
Alvin	Obs	72.0	26.9	1.0
(193)	Fit	71.9	27.3	0.8
Angleton	Obs	66.5	32.4	1.1
(179)	Fit	66.3	32.6	1.0
Kerrville	Obs	83.1	16.9	0.0
(118)	Fit	83.1	16.5	0.4
Alpine	Obs	27.3	59.1	13.6
(22)	Fit	21.5	71.7	6.8
Littlefield	Obs	23.8	60.4	15.8
(101)	Fit	17.4	73.9	8.7
Fort Stockton	Obs	34.7	61.7	3.6
(277)	Fit	34.7	61.6	3.7
Hillsboro	Obs	8.3	88.0	3.6
(192)	Fit	15.0	74.4	10.6
Monahan	Obs	23.9	62.6	13.5
(163)	Fit	19.1	72.9	8.0
Perryton	Obs	19.3	74.7	6.0
(83)	Fit	20.3	72.4	7.3
Gatesville	Obs	23.4	72.1	4.5
(111)	Fit	24.5	69.6	5.9
Abernathy	Obs	6.3	59.4	34.4
(64)	Fit	4.0	63.2	32.8
Taylor	Obs	4.6	62.1	33.3
(174)	Fit	4.1	63.1	32.9

Town (sample size). Obs = observed; Fit = fitted.

sociated with a shift in response toward the more severe mottling categories; that is, predictors with positive coefficients indicate risk factors for mottling.

For the case of a dichotomous response variable, the proportional odds and the logistic regression (10) models are identical. As with the logistic regression model, the existence of an underlying continuous distribution of tolerances can be used to motivate the proportional odds model (7).

The fit of this model to the prevalence of mottling in 16 Texas communities is demonstrated. A complete description of the study design and subject selection is given elsewhere (8, 9). Briefly, in these 16 communities, the school-age, life-time residents who agreed to participate were given dental exams which included an evaluation of each tooth for enamel mottling and complete visual tactile caries assessment of each tooth surface. Demographic, dental health practice and socioeconomic information on each child was obtained by questionnaires administered to a parent of the child. Fluoride in drinking water and temperatures are known to affect the prevalence of mottling (11). Previous work with this data set has indicated that race and presence of air conditioning are individual-level predictors associated with mottling (9). In addition, total dissolved solids and zinc are community-level predictors found to be associated with mottling (9). Only these six

variables will be considered in this report. Only the 2404 children from the 16 communities for whom race and air conditioning status are known will be considered here.

Estimates for the parameters of the proportional odds model are obtained using standard maximum likelihood methods and asymptotic test statistics are constructed (7). Adequacy of the fit of the proportional odds model is assessed by visual inspection of the observed and fitted values.

Results

The prevalence of mottling and values for community level predictor variables for the 16 Texas communities are presented in Table 1. The expected shift of individuals to the higher mottling categories with increasing fluoride levels in the drinking water is seen. It should be noted that with increasing fluoride levels from 0.2 to 3.3 there is first an increase in the frequency in the very mild/mild category with no cases of moderate/severe mottling; subsequently, at higher fluoride levels, a greater number of moderate/severe cases appear. The increase in the frequency of these two categories is accompanied by a decrease in the frequency in the normal/questionable category. This is the relationship to be described by the proportional odds model.

There are substantial differences in prevalence even among communities with the same fluoride level or the same factor-of-optimal of fluoride. For example, the communities of Hillsboro and Monahan have the same fluoride level and essentially the same mean maximum temperatures but substantially different prevalences of the three levels of mottling ($\chi^2 = 31.781$, 2 DF, P -values < 0.001). Also, Littlefield has a fluoride level lower than that for Hillsboro, Monahan and others but a substantially greater number in the very mild/mild and moderate/severe categories. It appears that predictors other than fluoride and mean maximum temperature need to be included.

The first model fit includes the variables race and presence of air conditioning, the two individual-level predictors considered here. Indicator variables were also included to indicate the 16 communities (10). This model will allow an assessment of the goodness of fit of the proportional hazards model to the data.

Table 3. Parameter estimates for proportional odds model used in obtaining fitted values presented in Table 4

Variable	Coefficient (SE)	P-value (one-sided)
Intercept (α_1)	12.673(2.311)	< 0.0001
Intercept (α_2)	16.500(2.330)	< 0.0001
Fluoride	3.150(0.312)	< 0.0001
Fluoride ²	-0.358(0.079)	< 0.0001
Temperature	0.107(0.029)	0.0001
Total dissolved solids	-0.00047(0.00016)	0.002
Zinc	0.00214(0.00062)	0.0003
Air (1 = present, 2 = absent)	0.335(0.096)	0.0002
Race (1 = white, 2 = other)	0.520(0.105)	< 0.0001

Table 4. Comparison of observed and fitted percentages for mottling categories for 16 Texas communities. Fitted values were obtained from model described in Table 3

		Normal/ questionable	Very mild/ mild	Moderate/ severe
New Braunfels	Obs	91.2	8.8	0.0
(102)	Fit	92.8	7.1	0.2
San Marcos	Obs	91.4	8.6	0.0
(197)	Fit	86.3	13.4	0.3
San Antonio	Obs	99.1	0.9	0.0
(114)	Fit	89.8	9.9	0.2
Kingsville	Obs	57.6	42.0	0.3
(314)	Fit	70.2	28.9	0.9
Alvin	Obs	72.0	26.9	1.0
(193)	Fit	72.6	26.6	0.8
Angleton	Obs	66.5	32.4	1.1
(179)	Fit	72.1	27.1	0.9
Kerrville	Obs	83.1	16.9	0.0
(118)	Fit	62.9	35.8	1.3
Alpine	Obs	27.3	59.1	13.6
(22)	Fit	14.1	73.9	12.0
Littlefield	Obs	23.8	60.4	15.8
(101)	Fit	25.1	68.6	6.3
Fort Stockton	Obs	34.7	61.7	3.6
(277)	Fit	25.2	68.5	6.3
Hillsboro	Obs	8.3	88.0	3.6
(192)	Fit	21.7	70.5	7.8
Monahan	Obs	23.9	62.6	13.5
(163)	Fit	15.6	73.3	11.1
Perryton	Obs	19.3	74.7	6.0
(83)	Fit	28.1	66.5	5.4
Gatesville	Obs	23.4	72.1	4.5
(111)	Fit	20.2	71.6	8.2
Abernathy	Obs	6.3	59.4	34.4
(64)	Fit	7.4	70.4	22.2
Taylor	Obs	4.6	62.1	33.3
(174)	Fit	4.5	62.3	33.3

The fitted proportions are given in Table 2.

Overall, the fit is remarkably close. Only three towns are seen to deviate from a good fit: Littlefield, Hillsboro and Monahan. This lack of fit is probably due to either not including important individual-level predictors other than race and presence of air conditioning or to an inadequacy of the proportional odds model. Because of the excellent fit for the other 13 communities, the inclination is to act as if the model is adequate and to

expect to be able eventually to identify additional predictor variables. McCULLAGH has shown how the proportional odds model can be adjusted to improve the fit by including scale parameters (7). This adjustment is not considered here.

The community level predictors were next included in the model, replacing the dummy variables used to indicate community. The community level predictors considered here were selected on the basis of previous work (9). The estimated coefficients and their standard errors are

presented in Table 3 and the fitted frequencies are presented in Table 4.

All the coefficients are statistically significant and associations are in the expected direction (Table 3). For example, higher values for fluoride, mean maximum temperature, zinc, air conditioning and race are associated with higher levels of mottling. Higher values of total dissolved solids are associated with lower levels of mottling. This is consistent with previous analysis (9).

The fit of this model is substantially poorer than that presented in Table 2. San Antonio, Kingsville, and Kerrville are each poorly fit by this model in addition to Littlefield, Hillsboro and Monahan.

The deterioration in fit between models in Tables 3 and 4 indicates that additional community level predictors need to be identified. Successful investigations which find these predictors could result in predicted fits as close as those presented in Table 2.

Discussion

The proportional hazards model was used to describe the distribution of school-age, lifetime residents of 16 Texas communities across mottling categories defined in terms of DEAN's scheme. This model potentially provides for an intuitively appealing description of the prevalence of mottling and avoids the need to impose scores on each of the categories.

The fit of the model to the 16 com-

munities, after controlling for race and presence of air conditioning, was quite good, fitting observed frequencies almost exactly in 13 of the towns. This indicates that the basic form of the model can potentially do quite well in describing the prevalence of mottling in communities. Use of the water qualities measures, in addition to race and air conditioning, resulted in a substantially poorer fit; thus, additional predictors still need to be identified.

In addition to level of fluoride in drinking water and mean maximum temperature, total dissolved solids, zinc, race and presence of air conditioning

were found to be related to the prevalence of mottling. The association with presence of air conditioning could be explained by its effect on amount of water consumed or could be due to socioeconomic status. This latter explanation might also hold for the observed association with race. Educational levels of father and mother were not associated with the prevalence of mottling. The inverse association with total dissolved solids could be due to the fluoride binding with other materials or to the water being less palatable. The positive association with zinc is less clear but is consistent with laboratory findings (12).

References

1. DEAN HT, DIXON R, COHEN C. Mottled enamel in Texas. *Public Health Rep* 1935; 50: 424-42.
2. DEAN HT, ELVOVE E. Mottled enamel in South Dakota. *Public Health Rep* 1939; 54: 212-28.
3. DEAN HT. The investigation of physiological effects by the epidemiologic method. In: MOULTON FR, ed. *Fluorine and health*, Washington DC: American Association for the Advancement of Science, 1942.
4. DEAN HT. Chronic endemic dental fluorosis (mottled enamel). In: GORDON M, ed. *Dental science and dental art*. Philadelphia, PA: Lee and Febiger, 1938.
5. HODGE HC, SMITH FA. *Fluoride chemistry*, Vol. IV. New York: Academic Press, 1965.
6. RICHARDS L, WESTMORELAND W, TASHIRO M, MCKAY C, MORRISON J. Determining optimum fluoride levels for community water supplies in relation to temperature. *JADA* 1967; 74: 389-97.
7. McCULLAGH P. Regression models for ordinal data (with discussion). *J R Stat Soc B* 1980; 42: 109-42.
8. SEGRETO V, COLLINS E, CAMANN D, KIMBALL, K. *Water fluoride levels in the critical zone: dental effects*. Report to the Health Effects Laboratory USEPA, Cincinnati, Ohio, 1983.
9. BUTLER W, SEGRETO V, COLLINS E. Prevalence of moderate and severe mottling in school-aged lifetime residents of sixteen Texas Communities. *J Am Public Health Assoc* 1985: in press.
10. COX DR. *The analysis of binary data*. London: Chapman-Hall, 1970.
11. GALAGAN D, VERMILLION J. Determining optimum fluoride concentration. *Public Health Rep* 1957; 72: 491-3.
12. EISENMANN D, YAEGER J. Alterations in the formation of rat dentin and enamel induced by various ions. *Arch Oral Biol* 1969; 14: 1045-6.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.