5

Statistical Models for Dental Caries Data

David Todem
Division of Biostatistics, Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI, USA

1. Introduction

Tooth decay is ubiquitous among humans and is one of the most prevalent oral diseases. Although this condition is largely preventable, more than half of all adults over the age of eighteen present early signs of the disease, and at some point in life about three out of four adults will develop the disease. Tooth decay is also common among children as young as five and remains the most common chronic disease of children aged five to seventeen years. It is estimated that tooth decay is four times more prevalent than asthma in childhood (Todem, 2008). Tooth decay and its correlates such as poor oral health place an enormous burden on the society. Poor oral health and a propensity to dental caries have been related to decreased school performance, poor social relationships and less success later in life. It is estimated that about 51 million school hours per year are lost in the U.S. alone because of dental-related illness. In older adults, tooth decay is one of the leading causes of tooth loss which has a dramatic impact on chewing ability leading to detrimental changes in food selection. This, in turn, may increase the risk of systemic diseases such as cardiovascular diseases and cancer.

The etiology of dental caries is well established. It is a localized, progressive demineralization of the hard tissues of the crown and root surfaces of teeth. The demineralization is caused by acids produced by bacteria, particularly mutans Streptococci and possibly Lactobacilli, that ferment dietary carbohydrates. This occurs within a bacteria-laden gelatinous material called dental plaque that adheres to tooth surfaces and becomes colonized by bacteria. Thus, dental caries results from the interplay of three main factors over time: dietary carbohydrates, cariogenic bacteria within dental plaque, and susceptible hard tooth surfaces. Dental caries is also a dynamic process since periods of demineralization alternate with periods of remineralization through the action of fluoride, calcium and phosphorous contained in oral fluids.

The evaluation of the severity of tooth decay is often performed at the tooth surface level. According to the World Health Organization, both the shape and the depth of a carious lesion at the tooth surface level can be scored on a four-point scale, D_1 to D_4. Level D_1 refers to clinically detectable enamel lesions with non-cavitated surfaces; D_2 for clinically detectable cavities limited to the enamel; D_3 for clinically detectable lesions in dentin; and finally D_4 for lesions into the pulp. Despite these detailed tooth-level data, most epidemiological studies often rely on the decayed, missing and filled (DMF) index,

developed in the 1930s by Klein *et al.* (see for example Klein and Palmer, 1938). This index is applied to all the teeth (DMFT) or to all surfaces (DMFS), and represents the cumulative severity of dental caries experience for each individual. These scores have well documented shortcomings regarding their ability to describe the intra-oral distribution of dental caries (Lewsey and Thomson, 2004). But they continue to be instrumental in evaluating and comparing the risks of dental caries across population groups. Most importantly, they remain popular in dental caries research for their ability to conduct historical comparisons in population-based studies.

Statistical analysis of dental caries data relies heavily on the research question under study. These questions can be classified into two groups. The first group represents questions that can be answered using mouth-level outcomes generated using aggregated scores such as the DMF index. The second group refers to questions that necessitate the use of tooth or tooth-surface level outcomes. A very important issue to address for the data analyst is the modeling strategy to adopt for the response variable under investigation. Broadly, two fairly different views are advocated. The first view, supported by large-sample properties, states that normal theory should be applied as much as possible, even to non-normal data such as counts (Verbeke and Molenberghs, 2000). This view is strengthened by the notion that, normal models, despite being a member of the generalized linear models (GLIM), are much further developed than any other GLIM (e.g. model checks and diagnostic tools), and that they enjoy unique properties (e.g., the existence of closed form solutions, exact distributions for test statistics, unbiased estimators, etc...). Although this is correct in principle, it fails to acknowledge that normal models may not be adequate for some types of data. As an example, the abundance of zeros in DMF scores rules out any attempt to use normal models, such as linear models, even after a suitable transformation. While a transformation may normalize the distribution of nonzero response values, no transformation could spread the zeros (Hall, 2000). A different modeling view is that each type of outcome should be analyzed using tools that exploit the nature of the data. For dental data, features to be accommodated include the discrete nature of the data (count responses for mouth-level data and binary response for intra-oral data), the abundance of zeros for example in the DMF/S scoring, and the clustering in intra-oral responses. The clustering of participants as a result of the study design is another important feature.

This chapter reviews common statistical parametric models to answer questions that arise in dental caries research, with an eye to discerning their relative strengths and limitations. Missing data problems arising in caries dental reasrch will also be discussed but touched on briefly.

2. Statistical models for mouth-level caries data

Mouth-level data, resulting from the DMF index, are typically analyzed as unbounded or bounded counts. For unbounded counts, a Poisson regression model or its extension the negative binomial regression model that accounts for overdispersion in the data, are often used. A binomial regression model for bounded counts is often advocated.

For unbounded counts, these models assume that the basic underlying distribution for the data is either a Poisson or a negative binomial distribution. The Poisson model is the simplest distribution for nonnegative discrete data, and is entirely specified by a positive

parameter the mean. This mean is often related to potential explanatory variables using a log link function. Specifically, let Y define the outcome variable and X the set of explanatory variables. A Poisson regression model for the mean is defined as $E(Y|X) = e^{\alpha+X\beta}$, where α and β are the intercept and the regression parameter vector associated with X. The probability mass function of Y is given by: $P(Y = y|X) = \frac{e^{-\mu_x}\mu_x{}^y}{\Gamma(y+1)}$, y=0,1,…, where $\mu_x = E(Y \mid X)$ is the conditional mean which depends on covariates.

One major restriction of the Poisson regression model is that its mean is equal to its variance. For dental caries data, however, it is not uncommon for the variance to be much greater than the mean. For such data, a negative binomial regression model has been advocated as an alternative to Poisson regression models. It is typically used when the variability in the data cannot be properly captured by Poisson regression models. The negative binomial model is a conjugate mixture distribution for count data (Agresti, 2002). It is entirely specified by two parameters, its mean and the overdispersion parameter. Similarly to the Poisson regression model, the mean is related to potential explanatory variables using a log link function. However, the probability mass function of Y is given by:

$$P(Y = y|X) = \frac{\Gamma(y + \kappa^{-1})}{\Gamma(\kappa^{-1})\Gamma(y + 1)} \left(\frac{\kappa^{-1}}{\mu_x + \kappa^{-1}}\right)^{\kappa^{-1}} \left(1 - \frac{\kappa^{-1}}{\mu_x + \kappa^{-1}}\right)^{y}, \qquad y = 0,1, \ldots$$

where $\mu_x = E(Y|X) = e^{\alpha+X\beta}$ is the conditional mean which depends on covariates, and κ is the overdispersion parameter. This distribution has variance $\mu_x + \kappa\mu_x{}^2$. Parameter κ is typically unknown and estimated from data to evaluate the extent of overdispersion in the data. When κ tends to zero, the negative binomial model converges to a Poisson process (Agresti, 2002).

The presence of an upper bound for possible values taken by DMF scores suggests a model based on the binomial rather than the Poisson distribution (Hall, 2000). Data are then viewed as being generated from a binomial process with m trials and success probability π_X. Here m represents the maximum number of teeth or tooth surfaces in the mouth susceptible to decay, and π_X the probability for a tooth or tooth surface to present a sign of decay. The binomial model is given by:

$$P(Y = y|X) = \frac{\Gamma(m + 1)}{\Gamma(m - y + 1)\Gamma(y + 1)} \ (\pi_X)^y \ (1 - \pi_X)^{m-y}, y = 0,1, \ldots m,$$

where the success probability is related to covariates as $\pi_X = \frac{e^{\alpha+X\beta}}{1+e^{\alpha+X\beta}}$, with α and β being the intercept and the regression parameter vector associated with X. One should note however that Poisson and negative binomial distributions provide a reasonable approximation to the binomial distribution in dental caries research.

Dental caries data with excess zeros are common in statistical practice. For example, in young children, DMF scores generally generate an excessive number of zeros in that many children do not experience dental caries. This is typically due to a short exposure time to caries development. The limitations of Poisson and negative binomial regression models to analyze such data are well established (see, for example, Lambert, 1992; and Hall, 2000). One approach to analyze count data with many zeros is to use zero-inflated models. This class of

models views the data as being generated from $P(Y = y|X)$ a mixture of a zero point mass and a non-degenerate homogenous discrete distribution $P_1(Y = y|X)$ as follows:

$$P(Y = y|X) = \begin{cases} \omega + (1-\omega)P_1(Y = y|X), & y = 0 \\ (1-\omega)P_1(Y = y|X), & y > 0, \end{cases}$$

where $0 \leq \omega \leq 1$ represents the mixing probability that captures the heterogeneity of zeros in the population. The choice of the homogenous distribution $P_1(Y = y|X)$ for the most part depends on the nature of counts under consideration. For bounded counts, a binomial distribution is typically used (Hall, 2000). Poisson and negative binomial distributions are the standard for unbounded counts (Bohning et al., 1999). Ridout, Demetrio and Hinde (1998) provide an extensive review of this literature. In real applications of these models in dental caries research, the mixing probability is often related to covariates using for example a logistic model.

We illustrate below how some of these simple models can be applied to dental caries scores data generated from a survey designed to collect oral health information on low-income African American children (0-5 years), living in the city of Detroit (see Tellez et al., 2006). This study aimed at promoting oral health and reducing its disparities within this community through the understanding of determinants of dental caries. Dental caries were measured using DMF scores which represent the cumulative severity of the disease for each surveyed participants. Possible covariates include the study participant's age (AGE) and his/her sugar intake (SI). In Table 1, we present the fitted regression models applied to children's data. For these data, the mean structure of the homogeneous model is specified as $E(Y|X) = e^{\alpha+X\beta}$, where $X = (AGE, SI, AGE * SI)$, with $AGE * SI$ being a multiplicative interaction, and $\beta = (\beta_1, \beta_2, \beta_3)'$. Parameter κ of the Negative Binomial model captures overdispersion in data.

Parameter	Homogeneous Poisson	Homogeneous Negative Binomial	Zero-inflated Negative Binomial (mixing weight depends on covariates)
α	1.3994(0.0209)*	1.3484(0.0725)*	2.0158(0.0676)*
β_1	0.6981(0.0193)*	0.9188(0.0861)*	0.2350(0.0679)*
β_2	0.2696(0.0203)*	0.2378(0.0853)*	0.0573(0.0695)
β_3	-0.2790(0.0219)*	-0.3314(0.0877)*	-0.0728(0.0739)
γ_0	-	-	-0.6131(0.1595)*
γ_1	-	-	-1.7191(0.2276)*
γ_2	-	-	-0.2226(0.1509)
γ_3	-	-	0.3163(0.2022)
κ	-	2.6178(0.1753)*	0.9295(0.1058)*
-2logLik	8455.7	4059.1	3815.4
AIC	8463.7	4069.1	3833.4

*: p-value<0.05

Table 1. Parameter estimates and (Standard errors) from a homogeneous Poisson model, a homogeneous Negative Binomial model, and a zero-inflated Negative Binomial model with covariate dependent mixing weights applied to DMF scores

As a basic starting model, a homogeneous Poisson regression model is fit and compared to a homogeneous Negative Binomial model. In view of the AIC, the homogeneous Negative Binomial model provides a reasonably good fit compared to the Poisson model. This result is consistent with overdispersion parameter κ in the homogeneous Negative Binomial model being statistically significant at 5%, suggesting that overdispersion cannot be ignored in these data. As a result, the standard errors of parameter estimates in the mean model under the homogeneous Negative Binomial model are larger compared to those of the homogeneous Poisson model. The homogeneous Negative Binomial model is further compared to a zero-inflated Negative Binomial model which potentially accommodates extra zeros in the data. In the latter model, the mixing weight ω is related to covariates as, $\omega = \{1 + e^{-\gamma Z}\}^{-1}$, where $Z = (\text{AGE}, SI, AGE * SI)$ and $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)'$. In view of the AIC, this model provides a better representation of the data compared to the homogeneous Negative Binomial model. This is consistent with findings from the literature dental caries in young children typically exhibit overdispersion in addition to zero-inflation (Bohning et al, 1999).

The zero-inflated regression models provide an interesting parametric framework to accommodate heterogeneity in a population. A prevailing concern, however, is that these models only accommodate an inflation of zeros in the population. Inflation and deflation at zero often arise in various practical applications. Homogeneous models (Poisson and negative binomial regression models) when applied to data from the Detroit study typically reveal an inflation of zeros (few children with no dental caries predicted than observed) for younger children and deflation of zeros (more children with no dental caries predicted than observed) for older children. For such data, a model that captures only inflation of zeros may fail to properly represent heterogeneity in the population. This then necessitates the use of models that can accommodate both inflation and deflation in the population. A good example of such models is the two-stage model also known as the Hurdle model (Mullahy, 1986). An alternative approach is to use the marginal distribution derived from the mixture distribution:

$$\mathrm{P}(\mathrm{Y}=\mathrm{y}|\mathrm{X}) = \begin{cases} \omega + (1-\omega)\mathrm{P}_1(\mathrm{Y}=\mathrm{y}|\mathrm{X}), & \mathrm{y}=0 \\ (1-\omega)\mathrm{P}_1(\mathrm{Y}=\mathrm{y}|\mathrm{X}), & y>0, \end{cases}$$

where $\frac{-\mathrm{P}_1(\mathrm{Y}=0|\mathrm{X})}{1-\mathrm{P}_1(\mathrm{Y}=0|\mathrm{X})} \leq \omega \leq 1$. Note here that the constraints on the mixing weights are obtained only by imposing that, $0 \leq \mathrm{P}(\mathrm{Y}=\mathrm{y}|\mathrm{X}) \leq 1$ for all y. The mixing weight is potential negative to accommodate deflation in the data. For this class of models, the marginal mixture model maintains his hierarchical representation only if the mixing weight are bounded between 0 and 1. When the mixing weight is negative, the marginal mixture model then loses its hierarchical representation.

Finally, the models described above are basic starting models and should be extended to accommodate unique features of the data under consideration. For example, it is often the case that the sampling design used to recruit study participants leads to clustered data. In survey research, sampled subjects living in the same neighborhood are more likely to share common, typically unmeasured, predispositions or characteristics that lead to dependent data. This therefore necessitates the use of models for clustered or correlated data. An example of such models is described by Todem et al. (2010) for the analysis of dental caries for low-income African American children under the age of six living in the city of Detroit. These authors extended the family of Poisson and negative binomial models to derive the

joint distribution of clustered counted outcomes with extra zeros. Two random effects models were formulated. The first model assumed a shared random effects term between the logistic model of the conditional probability of perfect zeros and the conditional mean of the imperfect state. The second formulation relaxed the shared random effects assumption by relating the conditional probability of perfect zeros and the conditional mean of the imperfect state to two correlated random effects variables. Under the conditional independence assumption and the missing data at random assumption, a direct optimization of the marginal likelihood and an EM algorithm were proposed to fit the proposed models.

3. Statistical models for intra-oral caries data

Although many dental studies provide detailed tooth-level data on caries activity, most analyses still rely on aggregated scores such as the DMF index. These scores summarize at mouth level caries information for each individual typically recorded at the tooth level or tooth-surface level. They have therefore been instrumental in evaluating and comparing the risks for dental caries among population groups. Despite these advances in the etiology of dental caries, there are still some fundamental questions regarding the spatial distribution of dental caries in the mouth that remain unanswered. The intra-oral spatial distribution of dental caries can help answer questions on whether the disease develops symmetrically in the mouth, and whether different types of teeth (Incisors, Canines and Molars) and tooth surfaces (Facial, Lingual, occlusal, Mesial, Distal, and incisal surfaces) are equally susceptible to dental caries. It is well recognized that the different morphology of the pit-and-fissure surfaces of teeth makes them more susceptible to decay than the smooth surfaces. Thus, it is no surprise that the posterior molar and premolar teeth that have pit-and-fissure surfaces are more susceptible to decay than the anterior teeth.

The analysis of intra-oral data poses a number of difficulties due to inherent spatial association of teeth and tooth-surfaces in the mouth. It is well known, for example, that the multiplicity of outcomes recorded on the same unit necessitates the use of methods for correlated data. This section reviews some of the commonly used statistical techniques to analyze such data. A focus will be on parametric models, namely the class of generalized linear mixed effects models and the class of generalized estimating equation models. These regression models take into account the unique spatial structure of teeth and tooth-surfaces in the mouth.

i. Generalized linear mixed effects models

Generalized linear mixed-effects models constitute the broader class of mixed-effects models for correlated continuous, binary, multinomial and count data (Breslow and Clayton, 1993). They are likelihood-based and often are formulated as hierarchical models. At the first stage, a conditional distribution of the response given random effects is specified, usually assumed to be a member of the exponential family. At the second stage, a prior distribution (typically normal) is imposed on the random effects. The conditional expectations (given random effects) are made of two components, a fixed effects term and a random effects term. The fixed effects term represents covariate effects that do not change with the subject. Random effects represent subject-specific coefficients viewed as deviations from the fixed effects (average) coefficients. Most importantly, they account for the within-mouth correlation

under the conditional independence assumption. In dental caries research, data collected at the tooth level or tooth-surface level are typically binary outcomes representing the presence or absence of decay. For such data, a logistic regression model with random effects is typically used. In this class of models, fixed-effects regression parameters have a subject-specific interpretation, conditional on random effects (Verbeke and Molenberghs, 2000). That is, they have a direct and meaningful interpretation only for covariates that change within the cluster level (subject's mouth) such as the location of a tooth or a tooth-surface in the mouth. The probabilities of tooth and tooth-surface decay are conditional given random effects and can be used to capture changes occurring within a particular subject's mouth. To assess changes across all subjects' mouths, the modeler is then required to integrate out the random effects from the quantities of interest. Generalized linear mixed effects models are likelihood-based and therefore can be highly sensitive to any distribution misspecification. But they are known to be robust against less restrictive missing data mechanisms (Little and Rubin, 1987).

ii. Generalized estimating equations models

Although there are a variety of standard likelihood-based models available to analyze data when the outcome is approximately normal, models for discrete outcomes (such as binary outcomes) generally require a different methodology. Kung-Yee Liang and Scott Zeger (1986) have proposed the so-called Generalized Estimating Equations-GEE model, which is an extension of generalized linear models to correlated data. The basic idea of this family of models is to specify a function that links the linear predictor to the mean response, and use a set of estimating functions with any working correlation model for parameter estimation. A sandwich estimator that corrects for any misspecification of the working correlation model is then used to compute the parameters' standard errors. GEE-based models are very popular as an all-round technique to analyze correlated data when the exact likelihood is difficult to specify. One of the strong points of this methodology is that the full joint distribution of the data does not need to be fully specified to guarantee asymptotically consistent and normal parameter estimates. Instead, a working correlation model between the clustered observations is required for estimation. GEE regression parameter estimates have a population-averaged interpretation, analogous to those obtained from a cross-sectional data analysis. This property makes GEE-based models desirable in population-based studies, where the focus is on average affects accounting for the within-subject association viewed as a nuisance term.

The GEE approach has several advantages over a likelihood-based model. It is computationally tractable in applications where the parametric approaches are computationally very demanding, if not impossible. It is also less sensitive to distribution misspecification as compared to full likelihood-based models. A major limitation of GEE-based models at least in their 1986 original formulation is that they require a more stringent missing data mechanism to produce valid inferences.

4. Conclusion

As the search for effective measures for the prevention and treatment of dental caries continues, it is essential that we have effective, robust and rigorous statistical methods to help our understanding of the condition. This chapter has reviewed common statistical

models to answer questions involving intra-oral and mouth-level outcomes, with an eye to discerning their relative strengths and limitations. Models for mouth-level data such as the DMF scores are basically count regression techniques. These models are often extended to two-component distributions when there are excess zeros. This class of models views the data as being generated from a mixture of a zero point mass and a non degenerate discrete distribution. Models for intra-oral outcomes are primarily correlated models for binary data, such as generalized linear mixed effects models and generalized estimating equations models. These models can account for the multilevel data structure (e.g., teeth within a quadrant and quadrants within the mouth) which generate a very complex and unique correlation structure (Zhang et al., 2010). Despite the relative merits of these models to account for the correlation structure, they need to be adapted to accommodate other unique features of intra-oral caries data. Intra-oral data present a unique set of challenges to statistical analysis which includes, but are not limited to, large cluster sizes and informative cluster sizes (Leroux *et al.*, 2006). More generally, models for intra-oral and mouth-level outcomes need to be adapted to the study design. For example, when the study design involves a longitudinal component, the model needs to be adapted accordingly.

Another important issue that needs to be accounted for is that of missing data. This problem is commonly encountered throughout statistical work and is almost ever present in the analysis of dental caries data. Incomplete data can have a dramatic impact on inferences if they are not properly investigated. Using terminology from Little and Rubin (1987), missing data mechanisms are classified as missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR), if missingness is allowed to depend (1) none of the outcomes, (2) the observed outcomes only, or (3) unobserved outcomes as well, respectively. GEE-based models at least in their 1986 original formulation require the more stringent MCAR mechanism to produce valid inferences. Weighted GEE-based models have been proposed to accommodate a less stringent missing data mechanism, the missing data at random process (Robins et al., 1995). Likelihood-based models such as generalized linear mixed effects models are known to be robust against the less restrictive MAR mechanism. When the missingness mechanism depends on the unobserved outcomes, these two classes of regression models are likely to produce biased inferences. For example, missing dental caries data generated from missing teeth are likely to be informative in that a missing tooth may be an indication of the severity of the decay for that particular tooth prior to the loss. For such data, ignoring missing data may lead to biased inferences. When a MNAR mechanism is suspected, a model that incorporates both the information from the outcome process and the missing data process into a unified estimating function was advocated (Diggle and Kenward, 1994 and Molenberghs et al. 1997). Such an approach has provoked a large debate about the role for such models in understanding the true data generating mechanism. The original enthusiasm was followed by skepticism about the strong and untestable assumptions on which this type of models rests (Verbeke et al., 2001). Specifically, joint models for the outcomes and missing data are typically not identifiable from observed data at hand. One then has to impose quantitative restrictions to recover identifiability. Conventional restrictions result from considering a minimal set of parameters, called sensitivity parameters, conditional upon which the remaining parameters are assumed identifiable. This method therefore produces a range of models which forms the basis of sensitivity analysis (Vach And Blettner, 1995).

Well established parametric models for dental caries data can be fit with most common statistical software including but not limited to SAS, Splus, R and SPSS. Options are however limited for newly developed models that have emerged in the literature. For recent statistical models in dental caries research to be accepted and used widely, there should be reliable and user-friendly software, readily available to perform regression analysis routinely. The software should be time-efficient, well-documented and most importantly should have a friendly interface, features that are of course closely related to the requirement of being user-friendly. Once these regression models are implemented, this will help answer both mouth-level and questions in population-based research.

Keywords: Generalized estimating equation models, Generalized linear mixed effects models, Negative Binomial models, Poisson models, Zero-inflated models for count data

5. References

Agresti, A. (2002), *Categorical Data Analysis*, Second Edition, New York: John Wiley & Sons.

Breslow, N. E. and Clayton, D. G. Approximate inference in generalized linear mixed models. *J. Amer. Statist. Assoc.*, 88:9–25, 1993.

Bohning, D., Dietz, E., Schlattmann, P., Mendonca, L. and Kirchner, U. (1999). The zero-inflated poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162, 195–209.

Diggle, P. & Kenward, M. G. (1994). Informative dropout in longitudinal data analysis (with discussion). *Applied Statistics* 43, 49–93.

Hall, D. B. (2000). Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics,* 56, 1030–1039.

Klein H, Palmer C (1938). Studies on dental caries vs. familial resemblance in the caries experience of siblings. *Pub Hlth Rep,* 53:1353-1364.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34, 1–14.

Leroux, B. (2006). Analysis of correlated dental data: challenges and recent developments. *Statistical Methods for Oral Health Research, JSM 2006.*

Liang, K. Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* 73, 13–22.

Little, R. and Rubin, D. (1987). *Statistical analysis with missing data*. New York: John Wiley and Sons.

Lewsey, J. D. and Thomson, W. M. (2004), The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status, Community Dentistry and Oral Epidemiology, 32:183—189

Molenberghs, G., Kenward, M. G. & Lesaffre, E. (1997). The analysis of longitudinal ordinal data with non-random dropout. *Biometrika* 84, 33–44.

Mullahy, J. 1986. Specification and testing of some modified count data models. Journal of Econometrics 3: 341–365.

Ridout, M., Hinde, J. and Demetrio, C. G. B. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics* 57, 219–223.

Robins, J., Rotnitzky, A. and Zhao, L.P. (1995). Analysis of semiparametric regression models for repeated outcomes under the presence of missing data. *Journal of the American Statistical Association*, 90,106

Tellez, M., Sohn, W., Burt, B.A., & Ismail A. I. (2006). Assessment of the relationship between neighborhood characteristics and dental caries severity among low-income African-Americans: a multilevel approach., *J Public Health Dent*.66:30-6.

Todem, D. (2008). Oral Health. In Sarah Boslaugh (ed.) Encyclopedia of Epidemiology, Sage Publications (vol 2, pp: 762-764), Los Angeles

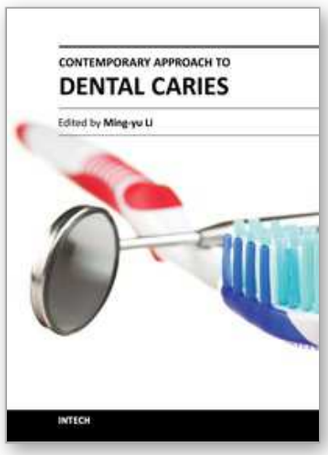
Todem, D, Zhang, Y., Ismail, A., and Sohn, W. (2010). Random effects regression models for count data with excess zeros in caries research, Journal of Applied Statistics, 37(10): 1661 - 1679

Vach, W. & Blettner, M. (1995). Logistic regression with incompletely observed categorical covariates – Investigating the sensitivity against violation of the missing at random assumption. *Statistics in Medicine* 14, 1315–1329.

Verbeke, C. and Molenberghs, G. (2000) *Linear mixed models for longitudinal data.* New York: Springer-Verlag.

Verbeke, G., Molenberghs, G., Thijs, H., Lesaffre, E. & Kenward, M. (2001). Sensitivity analysis for nonrandom dropout: A local influence approach. *Biometrics* 57, 7–14.

Zhang, Y., D. Todem, K. Kim and E. Lesaffre (2011). Bayesian Latent Variable Models for Spatially Correlated Tooth-level Binary Data in Caries Research, Statistical Modelling, 11(1):25-47



Contemporary Approach to Dental Caries
Edited by Dr. Ming-Yu Li

ISBN 978-953-51-0305-9
Hard cover, 488 pages
Publisher InTech
Published online 14, March, 2012
Published in print edition March, 2012

With an update of the recent progress in etiology, pathogenesis, diagnosis, and treatment of caries, it may be said that the final defeat of dental caries is becoming possible soon. Based on the research in this area in recent decades, "Contemporary Approach to Dental Caries" contained the caries in general, the diagnosis of caries, caries control and prevention, the medical treatment of caries, dental caries in children and others such as secondary caries. This book provides the reader with a guide of progress on the study of dental caries. The book will appeal to dental students, educators, hygienists, therapists and dentists who wish to update their knowledge. It will make you feel reading is profitable and useful for your practice.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen