AN APPLICATION OF MIXED EFFECTS MIXED DISTRIBUTION (MEMD) MODEL FOR LONGITUDINAL PATIENT-REPORTED OUTCOMES DATA WITH MANY ZEROS

Helen Q. Huang, M.S.[1], Mark F. Brady, Ph.D.[2], Randoph Carter, Ph. D.[3]

1. Biostatistician, Gynecologic Oncology Group Statistical and Data Center, Roswell Park Cancer Institute, Buffalo, NY 14263

2. Director of Statistics, Gynecologic Oncology Group Statistical and Data Center, Roswell Park Cancer Institute, Buffalo, NY 14263

3. Professor, Department of Biostatistics, State University of New York at Buffalo, Buffalo, NY .

Corresponding Author: Helen Q. Huang
Gynecologic Oncology Group Statistical and Data Center
Roswell Park Cancer Institute
Carlton & Elm Streets
Buffalo, NY 14263
Phone: (716) 845-2318; Fax: (716) 845-8393
Email: hhuang@gogstats.org

Reprint Information: Ms. Denise Mackey
GOG Administrative Office
Four Penn Center
1600 John F. Kennedy Boulevard, Suite 1020
Philadelphia, PA 19103
Phone: 215-854-0770
Email: dmackey@gog.org

ABSTRACT

The Mixed Effects Mixed Distribution (MEMD) model has been proposed by Tooze and Olsen to deal with longitudinal or repeated measures with clumping at zero. This model contains two components with one modeling the probability of a nonzero value and another modeling the mean of nonzero values. The repeated measures are handled with random effects and correlation between the random effects is allowed to connecting the two components. This paper presents an application of the MEMD model on the patient-reported peripheral neuropathy in a cancer clinical trial which are characterized by repeated measures with many zeroes.

1. INTRODUCTION

Measurement of patient-reported outcomes (PRO) has grown to be a focus in recent years in evaluating treatment benefits in randomized clinical trials and other clinical studies. Most instruments that are used to assess PRO are in a questionnaire format that outlines several questions for patients to answer. Each question is followed by a few choices with a pre-defined scale, e.g. 0 represents no symptoms and a higher score represents a worse scenario. Even after aggregating items (e.g. by summation of item rating), many zeros often remain. The distribution of this type of data is characterized by a spike or discrete probability mass at zero, followed by a heavily skewed bump or ramp describing the positive values. This type of data occurs in many applications in the medical sciences in biometrics, such as health policy research, epidemiology, substance abuse studies, and quality of data research and clinical trials.

Analyzing data with clumping at zeros and non-zero values extremely skewed pose an important challenge for biostatisticians. If the data are treated as if they come from a normal distribution, the clumping at zero is ignored and so is the skewness of the non-zero data. Clearly a normal distribution is an incorrect specification in such cases. If a nonparametric approach is employed by using the distribution of the ranks, a large number of ties will exist due to the zero values and the distribution will not be symmetric. In addition, it is not possible to obtain estimation and prediction with a nonparametric approach. One way to analyze this type of data is to treat the zeros and nonzeros separately as in the "two-part model" used for cross-sectional data in econometrics [1, 2]. This method, however, does not account for the relationship that might exist between the probability of a nonzero value and the magnitude of the nonzero value. In addition,

many studies involving PRO assessment include repeated assessments over time. The correlation among repeated measurements on the same person must be accounted for.

The most challenging aspect of analyzing repeated measures with clumping at zero is how to deal with the zeros. One way as suggested by Berk [3, 4] is to regard some zeros as the result of left censoring. This approach assumes an association between zero value and low values for the nonzero data. Another way is to consider zeros as "true" zeros and handle the "true" zeros with logistic model as proposed by Tooze [5, 6] and Olsen [6]. The latter one is similar to the 'two-part model' by combining models for the probability of occurance of a nonzero value and for the probability a)distribution of the nonzero values. However, they allow for a connection between the two parts by way of the correlation between the random effects in each model. This paper represents an application of the mixed-effects mixed-distribution (MEMD) model proposed by Tooze to a patient-reported neurotoxicity data.

2. MIXED EFFECT MIXED DISTRIBUTION (MEMD) MODEL

Let y_{ij} be an observation from the jth measurement on the ith subject, where the y_{ij} are all nonnegative. Let p_{ij} be the probability for $y_{ij} > 0$ and $f\left(y_{ij} \mid y_{ij} > 0\right)$ be the p.d.f of y_{ij} given $y_{ij} > 0$. Then the p.d.f. of y_{ij} [5] is

$$f\left(y_{ij}\right) = \begin{cases} \Pr\left(y_{ij} = 0\right) = 1 - p_{ij} & \text{if } y_{ij} = 0 \\ \Pr\left(y_{ij} > 0\right)f\left(y_{ij} \mid y_{ij} > 0\right) = p_{ij}f\left(y_{ij} \mid y_{ij} > 0\right) & \text{if } y_{ij} > 0 \end{cases}$$

And the unconditional expectation and variance of y_{ij} [7] can be calculated by:

$$E\left(y_{ij}\right) = p_{ij}\mu_{y_{ij} \mid y_{ij} > 0} \tag{1}$$

$$Var\left(y_{ij}\right) = p_{ij}\sigma^2_{y_{ij} \mid y_{ij} > 0} + p_{ij}\left(1 - p_{ij}\right)\mu^2_{y_{ij} \mid y_{ij} > 0}$$

where $\mu_{y_{ij}|y_{ij}>0}$ and $\sigma^2_{y_{ij}|y_{ij}>0}$ are the mean and variance of y_{ij} given $y_{ij}>0$ respectively.

Let r_i be the random effect that accounts for the correlations due to the repeated measures on the same subject. Given the random effect r_{1i} for the ith subject, we assume a logistic model for the probability of being a nonzero value ($y_{ij}>0$) such that

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = X'_{1ij}\beta_1 + r_{1i} \qquad (2)$$

where β_1 is a vector of fixed effect parameters and X_{1ij} is a vector of covariates for the probability of a nonzero value.

Under the condition of being a nonzero value ($y_{ij}>0$) and random effect r_{2i} for the ith subject we assume a lognormal model for a nonzero value so that

$$\log(y_{ij} \mid y_{ij}>0, r_{1i}) \sim N(X'_{2ij}\beta_2 + r_{2i}, \sigma^2_e) \qquad (3)$$

where β_2 is a vector of fixed effects and X_{2ij} is a vector of covariates for a nonzero value.

The random effects r_{1i} and r_{2i} in the logistic model and lognormal model are allowed to be correlated by assuming that

$$\begin{bmatrix} r_{1i} \\ r_{2i} \end{bmatrix} \sim BVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_1 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma^2_2 \end{bmatrix}\right)$$

Under above assumptions, the subject-specific expectation of a nonzero y_{ij} given $y_{ij}>0$ and random effect r_{2i} is

$$E(y_{ij} \mid y_{ij} > 0, r_{2i}) = \mu_{y_{ij} \mid y_{ij} > 0, r_{2i}} = \exp\left(X'_{2ij}\beta_2 + r_{2i} + \frac{\sigma_e^2}{2}\right) \qquad (4)$$

and the variance[7] is

$$Var(y_{ij} \mid y_{ij} > 0, r_{2i}) = \sigma^2_{y_{ij} \mid y_{ij} > 0, r_{2i}} = \mu^2_{y_{ij} \mid y_{ij} > 0, r_{2i}}\left(\exp(\sigma_e^2) - 1\right)$$

The marginal expectation of y_{ij} given y_{ij} >0 is

$$E(y_{ij} \mid y_{ij} > 0) = \mu_{y_{ij} \mid y_{ij} > 0} = \exp\left(X'_{2ij}\beta_2 + \frac{\sigma_2^2}{2} + \frac{\sigma_e^2}{2}\right) \qquad (5)$$

and the variance is

$$Var(y_{ij} \mid y_{ij} > 0) = \sigma^2_{y_{ij} \mid y_{ij} > 0} = \mu^2_{y_{ij} \mid y_{ij} > 0}\left(\exp(\sigma_2^2 + \sigma_e^2) - 1\right)$$

Let z_{ij} be an indicator variable such that

$$z_{ij} = \begin{cases} 1 & \text{if } y_{ij} > 0 \\ 0 & \text{if } y_{ij} = 0 \end{cases}$$

Given the random effects, the p.d.f. for y_{ij} is

$$f(y_{ij} \mid r_{1i}, r_{2i}) = \left[(1 - p_{ij})\right]^{(1 - z_{ij})} \times \left[p_{ij} \times \phi\left(\frac{\ln(y_{ij}) - \mu_{y_{ij} \mid y_{ij} > 0, r_{2i}}}{\sigma_e}\right)\right]^{z_{ij}}$$

Here ϕ is the standard normal density. The density function of y_{ij} is the product of two terms, where the first corresponds to the zero values and the second to the nonzero values.

The conditional likelihood for a single subject is obtained as a product

$$L_i(y_i \mid r_{1i}, r_{2i}) = \prod_j f(y_{ij} \mid r_{1i}, r_{2i}).$$

Letting $g(r_{1i}, r_{2i})$ be the bivariate normal density of the random effects, the marginal likelihood for a single subject is

$$L(y_i) = \int_{r_{1i}} \int_{r_{2i}} \prod_j f\left(y_{ij} \mid r_{1i}, r_{2i}\right) g\left(r_{1i}, r_{2i}\right) dr_{1i} dr_{2i}$$

Finally, the likelihood to be maximized is

$$L = \prod_i \int_{r_{1i}} \int_{r_{2i}} \prod_j f\left(y_{ij} \mid r_{1i}, r_{2i}\right) g\left(r_{1i}, r_{2i}\right) dr_{1i} dr_{2i}$$

3. **APPLICATION**

A randomized clinical trial was conducted by Gynecological Oncology Group (GOG) to determine whether the addition of paclitaxel to standard doxorubicin/cisplatin chemotherapy (TAP) produces improvement in the survival of patients with advanced endometrial carcinoma compared to standard treatment (AP) alone [8]. A potential adverse effect of this trial was chemotherapy-induced peripheral neuropathy. To measure the patient-reported neurotoxicity symptoms and concerns, the neurotoxicity subscale (NTX subscale) of FACT/GOG-NTX (Functional Assessment of Cancer Therapy/Gynecologic Oncology Group-Neurotoxicity) was administered to all qualified patients at baseline and prior to each subsequent course of chemotherapy. The maximum of numbers of courses was seven. The NTX subscale had 11 items when it was initially applied to this study. Each item uses 0-4 rating scale, with 0 = no symptoms at all, 1 = a little bit, 2 = somewhat, 3 = quite a bit, and 4 = very much. A scale score was obtained by summating the rating over all items to produce a total score for analysis. It has been validated and suggested that the 11 item subscale could be reduced to 4 items without compromising its performance of psychometric properties [9]. Thus, the total score of the reduced NTX sucscale (NTX4) ranges from 0 to 16 for the reduced subscale. A higher score indicates worse toxicity and the nonzero (>0) score

indicates the onset of neurotoxicity. The baseline neurotoxicity assessment is completed at the time of patient entry. Six follow-up assessments are conducted prior to the 2nd to 7th course of chemotherapy. A total of 263 patients participated in the study. Of them, 238 (91%) patients completed the assessment prior to cycle 1 (considered as baseline). The completion rates for follow-up assessments were 83%, 77%, 70%,65%, 59% respectively at the 2nd, 3rd, 4th, 5th, 6th assessment and 52% at the last assessment. The distributions of the NTX4 reported at follow-up assessments are displayed in Figure 1 by assessment time points. Among these assessments, about 34% to 62% of patients scored zero on the reduced neurotoxicity subscale. Furthermore, nonzero NTX4 scores were extremely skewed to the right.

For this NTX4 data, we applied the mixed-effect mixed-distribution (MEMD) model defined above to assess the effects of TAP on both the development of neurotoxicity and the severity of the toxicity. Specifically, the logistic model was used to model the probability of reporting a nonzero NTX4 score and a lognormal model is applied in modeling the magnitude of a nonzero NTX4 score when reported.

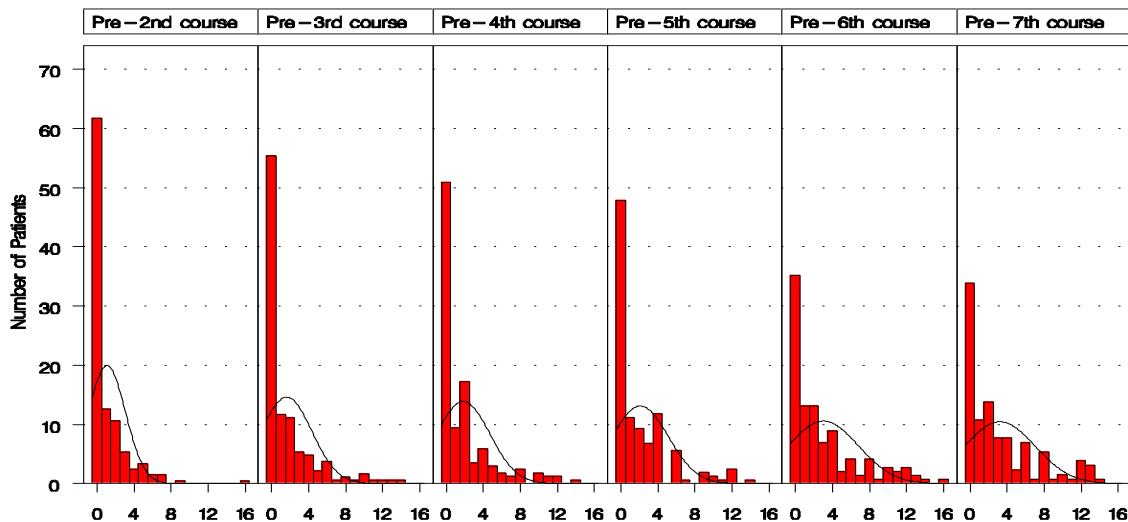## Distribution of Neurotoxicity Scores (NTX4)

Figure 1.

In the analysis, the effect of treatment, assessment time, and the interaction between treatment and assessment time were examined for both the logistic and lognormal models. The baseline NTX4 score was included as a covariate in both models. The treatment effect is coded as 0 for the control group (AP) and 1 for the experiment group (TAP). Assessment time points are coded as 1, 2, … ,6 for assessments prior to the 2nd , 3rd, …. , 7th cycle of chemotherapy. To account for the repeated measures on the same subject, a random effect was incorporated to both logistic and lognormal models and the two random effects were allowed to be correlated.

This MEMD model with correlated random effects was fitted using SAS PROC NLMIXED. The PROC NLMIXED in SAS software enables you to specify a conditional distribution of the data given the random effects. It is also flexible in specifying a standard distribution such as normal, binomial, and Poisson or a general distribution that you code using SAS programming statements. Parameter estimates for logistic part and lognormal part of the MEMD model are given in Table 1.

Table 1.   **Parameter estimates for MEMD model fitting NTX4 data**

Parameters	Logistic model		Lognormal model	
	Estimate (S.E)	P > | t |	Estimate (S.E)	P > | t |
Intercept	-2.839(0.445)	<0.001	0.078 (0.136)	0.565
Treatment	0.871(0.569)	0.127	0.131 (0.153)	0.394
Assessment time	0.275(0.087)	0.002	0.051 (0.026)	0.051
Treatment*time	0.531(0.134)	0.001	0.134 (0.032)	<0.001

Baseline NTX4 scores	1.267(0.208)	<0.001	0.174 (0.038)	<0.001
Var. of Random effect	5.912(1.290)	<0.001	0.335 (0.059)	<0.001
Var. of Random error	-	-	0.269	<0.001
Correlation between random effects		0.742 (0.08)		p<0.001

The significant random effects in the logistic model (variance=5.912, p<0.001) and the lognormal model (variance=0.335, p<0.001) constitute clear evidence of substantial heterogeneity among patients in perceiving neurotoxicity symptoms. A positive correlation (r=0.74, p<0.001)) between the random effects indicated that, after adjusting for the baseline scores, the patients with a higher tendency to report neurotoxicity symptoms also tend to report a greater neurotoxicity score.

After adjusting for the baseline scores, the significant interactions between treatment and time in both components of the model suggest that the addition of paclitaxel to the standard chemotherapy (TAP) in treating advanced endometrial cancer would not only increased the odds of developing neurotoxicity symptoms but also increased, among those who do develop neurotoxicity, the reported severity of toxicity during the treatment, when compared with AP alone.

4. MODEL CHECKING

Our specification of the MEMD model assumes that the random effects in the logistic and lognormal models are bivariate normal and independent of errors in the lognormal model, which is assumed normally distributed. The residuals for the lognormal model were calculated as $\ln(y_{ij}) - (X'_{2ij}\hat{\beta}_2 + \hat{r}_{2i})$. Quantile-quantile plots of these random

effects (Figure 2,3) and residuals (Figure 4) didn't reveal severe violation of the assumptions.
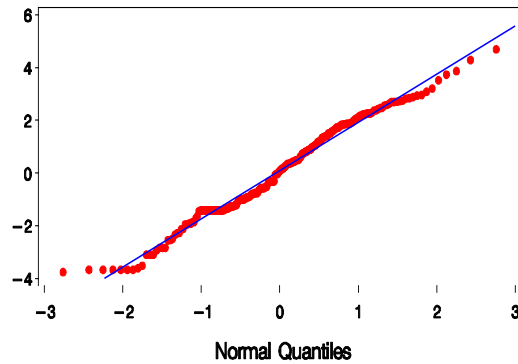
The Random Effect in the Logistic Model

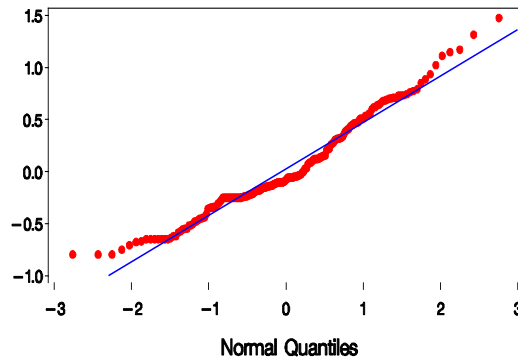Figure 2.

The Random Effect in the Lognormal Model

Figure 3.
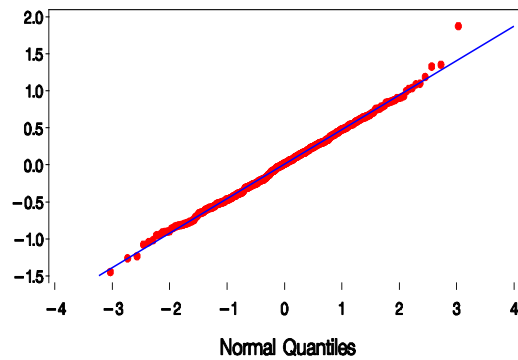
The Residuals in the Lognormal Model

Figure 4.

5. PREDICTIONS FOR THE PROBABILITY OF A NONZERO NTX4 SCORE AND MEAN NTX4 SCORES

Under the MEMD model defined by (1) – (4), the predicted probability of a nonzero NTX4 score is

$$\hat{p}_{ij} = \log it^{-1}(X'_{1ij}\hat{\beta}_1 + \hat{r}_{1i})$$

The predicted NTX4 scores given the NTX4>0 is calculated as

$$\hat{y}_{ij} = \exp(X'_{2ij}\hat{\beta}_2 + \hat{r}_{2i} + \frac{\hat{\sigma}_e^2}{2})$$

and the predicted unconditional NTX4 score is calculated as

$$\tilde{y}_{ij} = \hat{p}_{ij} \times \hat{y}_{ij}$$

The observed and predicted probabilities ($\hat{p}_{ij}$) of reporting a nonzero NTX4 score and

the observed and predicted nonzero NTX4 scores ($\hat{y}_{ij}$) given a nonzero score for two

treatment groups are displayed in Table 2 and Table 3 respectively. The fitted MEMD

model estimates showed a good prediction for the probability of a nonzero NTX4 score

and the mean of nonzero NTX4 scores. The mean of predicted NTX4 score ($\tilde{y}_{ij}$) were

plotted together with the mean of observed NTX scores that include zeros in the Figure 5

and presents a good prediction for the self-assessed neurotoxicity score.

Table 2. **Observed and predicted probability of a nonzero NTX4 score**

Assessment Time	AP		TAP	
	Observed Prob.	Predicted Prob.	Observed Prob.	Predicted Prob.
Pre-2nd course	0.30 (30/101)	0.27	0.47 (49/105)	0.46
Pre-3rd course	0.35 (33/93)	0.31	0.54 (52/96)	0.55
Pre-4th course	0.29 (24/82)	0.34	0.68 (60/88)	0.65
Pre-5th course	0.30 (23/76)	0.35	0.72 (61/85)	0.74
Pre-6th course	0.47 (32/68)	0.42	0.81 (62/77)	0.81
Pre-7th course	0.46 (27/59)	0.45	0.83 (59/71)	0.86

Table 3. **Observed and predicted means of nonzero NTX4 score**

Assessment Time	AP		TAP	
	Observed Means	Predicted Means	Observed Means	Predicted Means
Pre-2nd course	2.70	2.45	2.87	3.09
Pre-3rd course	2.30	2.62	4.47	3.63

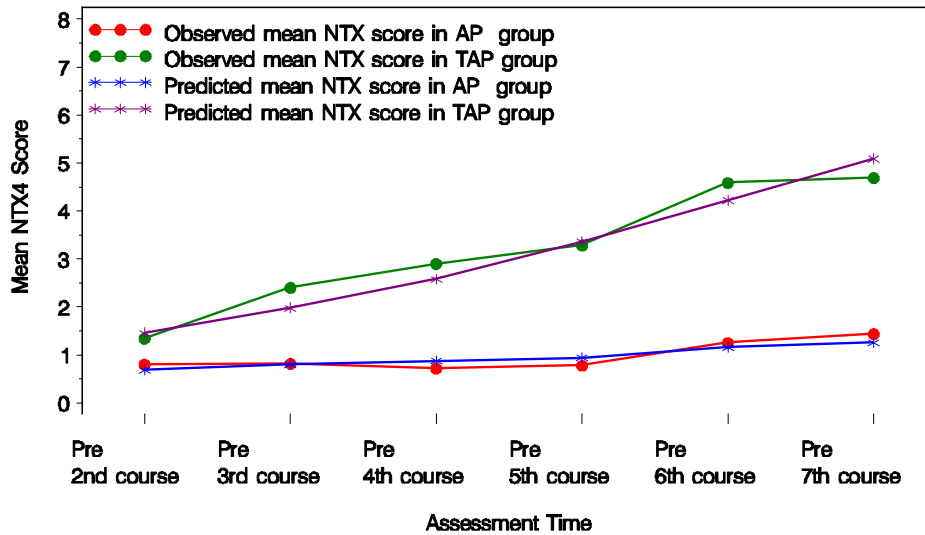|				
Pre-4th course	2.46	2.75	4.27	3.97
Pre-5th course	2.57	2.60	4.57	4.63
Pre-6th course	2.66	2.70	5.69	5.26
Pre-7th course	3.14	2.84	5.64	5.95

Figure 5.

6. DISCUSSION

The Mixed-Effect Mixed Distribution (MEMD) model with correlated random effects demonstrated to be a reasonable model for longitudinal patient-reported outcomes data with clumping at zero. This MEMD model handles the excess zeros by modeling the probability of a nonzero rating and modeling the magnitude of a nonzero rating separately and using random effects to account for the correlations among the repeated measures. The key feature of this method lies in that the two separate models are linked by allowing the random subject effects to be correlated with each other such that the parameter estimates are computed by means of models whose random components are intercorrelated [10].

A zero score in this study is treated as a real zero, implying no neurotoxicity perceived by the patient. This is considered reasonable for patient-reported neurotoxicity. Although, sometimes a zero might not be a true zero but left censored values that would not be zero if using a more sensitive scale, or a mixture of true zeros and left censored values, the MEMD model is still applicable. In these cases, the method can be modified by adding a left-censored lognormal to the probability of a zero value [3, 4].

This model can also be applied to repeated measures data with many 'ceiling' values [11]. For, example, suppose the PRO data are collected from extremely sick patients. Many might report highest scores for the worst symptoms or side effects. In this case, the 'ceiling' value is considered right-censored and a right-censored lognormal can be added to the density of a nonzero score.

The correlations among the longitudinal PRO data were dealt with by assuming random intercepts in both components of the MEMD model. It is possible that there is substantial heterogeneity in rate of change in PRO score. Inclusion of random coefficients to the random effect could be a topic of further attention for longitudinal data that include many zero values.

REFERENCE

1.	Lachenbruch, P.A., *Analysis of data with clumping at zero.* Biometrische Zeitschrift, 1976. **18**(5): p. 351-356.

2.	Lachenbruch, P.A., *Power and sample size requirements for two-part models.* Stat Med, 2001. **20**(8): p. 1235-8.

3.   Berk, K.N. and P.A. Lachenbruch, *Repeated measures with zeros.* Stat Methods Med Res, 2002. **11**(4): p. 303-16.

4.   Moulton, L.H. and N.A. Halsey, *A mixture model with detection limits for regression analyses of antibody response to vaccine.* Biometrics, 1995. **51**(4): p. 1570-8.

5.   Tooze, J.A., G.K. Grunwald, and R.H. Jones, *Analysis of repeated measures data with clumping at zero.* Stat Methods Med Res, 2002. **11**(4): p. 341-55.

6.   Olsen, M.K. and J.L. Schafer, *A two-part random-effects model for semicontinuous longitudinal data.* Journal of the American Statistical Association, 2001. **96**(454): p. 730-745.

7.   Aitchison, J., *On the distribution of a positive random varible having a discrete probability mass at the origin.* Journal of the American Statistical Association, 1955. **50**: p. 901-908.

8.   Fleming, G.F., et al., *Phase III trial of doxorubicin plus cisplatin with or without paclitaxel plus filgrastim in advanced endometrial carcinoma: a Gynecologic Oncology Group Study.* J Clin Oncol, 2004. **22**(11): p. 2159-66.

9.   Huang, H.Q., et al., *Validation and reduction of FACT/GOG-Ntx subscale for platinum/paclitaxel-induced neurologic symptoms: a gynecologic oncology group study.* Int J Gynecol Cancer, 2007. **17**(2): p. 387-93.

10.   Delucchi, K.L. and A. Bostrom, *Methods for analysis of skewed data distributions in psychiatric clinical studies: working with many zero values.* Am J Psychiatry, 2004. **161**(7): p. 1159-68.

11.   Austin, P.C. and M.D. Escobar, *The use of finite mixture models to estimate the distribution of the health utilities index in teh presence of a ceiling effect.* Journal of Applied Statistics, 2003. **30**(8): p. 909-923.