

# DIAGNOSTICS FOR STRATIFIED CLINICAL TRIALS IN PROPORTIONAL ODDS MODELS

Ivy Liu and Dong Q. Wang

School of Mathematics, Statistics and Computer Science

Victoria University of Wellington

New Zealand

Corresponding author: Ivy Liu

iliu@mcs.vuw.ac.nz

Key Words: clinical trials; cumulative odds ratio; diagnostics; influence measure; Mantel-Haenszel estimator; ordinal response; proportional odds model; random effects; and strata deletion.

## ABSTRACT

For stratified clinical trials with an ordinal response, this article considers two diagnostics strategies for evaluating the heterogeneity of the ordinal odds ratios across multi-centers in proportional odds models. The first strategy assumes multi-centers as a sample from a population and then uses a proportional odds model allowing random effects to model the heterogeneity among the ordinal log odds ratios. It gives an overall view of the heterogeneity across multi-centers. The second strategy applies the influence measure to the Mantel-Haenszel type estimators of the ordinal odds ratios. It shows the detail of heterogeneity in each of the multi-centers. At the end, this article uses a multi-center clinical trials example to illustrate both strategies.

## 1. INTRODUCTION

This article considers diagnostics methods for stratified multi-center clinical trials when the response variable has a natural ordering (e.g., better, unchanged, and worse). For a clinical trial study, the odds ratio is commonly used to describe the relationship between treatments and response. When the response is ordinal, the ordinal odds ratios that sum-

marize the relationship across different centers are considered based on a proportional odds model.

Diagnostic methods play an important role in categorical data analysis. For instance, Grizzle and Williams (1972) explored a general diagnostic approach to the analysis of categorical data for fitting a loglinear model. Pregibon (1981) developed diagnostic measures for a maximum likelihood fit of a logistic regression model. Andersen (1992) discussed diagnostics for categorical data in the Goodman association model. Wang, Critchley and Liu (2004) considered the diagnostic method for a contingency table assuming a clustered sampling model. A well known diagnostic measure, Cook distance, is widely used for linear regression models (Cook and Weisberg, 1982) and multivariate analysis (Wang and Critchley, 2000 and 2003).

This article uses two diagnostics strategies for evaluating the heterogeneity of the ordinal odds ratios across multi-centers in proportional odds models. The strategies are the following: (1) Using a proportional odds model (McCullagh, 1980) allowing random effects that assumes the multi-centers as a sample from a population (Hartzel, Liu and Agresti, 2001). The model permits heterogeneity in the conditional associations between treatments and response. The estimate of the mean for the association effects describes the average relationship between treatments and response across centers. The variance estimate for the association effects describes the heterogeneity of the relationship across centers. (2) Using Mantel-Haenszel (Mantel and Haenszel, 1959) type estimators to evaluate the ordinal odds ratios in a proportional odds model (Liu and Agresti, 1999; Liu, 2003). This technique summarizes the association effects across the centers using ordinal odds ratios. Then, we propose an influence measure to indicate the detail of the heterogeneity for the ordinal odds ratio in each of the centers. When each center has very few observations, the second strategy is better than the first one, because the model fitting algorithm for the proportional odds model with random effects often has computational problems for sparse data. However, the first strategy gives an overall view of the heterogeneity while the second gives the detail of the heterogeneity in each of the centers.

Section 2 introduces the proportional odds model used for fitting stratified clinical trials data with an ordinal response. It also shows the Mantel-Haenszel type estimators of the ordinal odds ratios that describe the conditional associations between treatments and response. Section 3 discusses the first strategy that determines the level of heterogeneity using a proportional odds model allowing random effects. Section 4 proposes the second strategy where an influence measure is used to explore the detail of heterogeneity for the ordinal log odds ratios across centers. Section 5 gives an example using these strategies.

## 2. ORDINAL ODDS RATIOS IN A PROPORTIONAL ODDS MODEL

When the response is ordinal (e.g., better, unchanged, and worse), the most popular model uses logits of cumulative probabilities. The model is often called the *proportional odds model* (McCullagh, 1980). For instance, in a clinical trial study, one can compare  $r$  treatments on a  $c$ -level ordinal response for data from  $K$  clinical centers using the model. Let  $\pi_{ijk}$  denote the probability that the response is at level  $j$  when a patient at center  $k$  received treatment  $i$ , where  $\pi_{i1k} + \dots + \pi_{ick} = 1$  for all  $i = 1, \dots, r$  and  $k = 1, \dots, K$ . The cumulative probabilities are denoted by  $\pi_{ijk}^* = \pi_{i1k} + \dots + \pi_{ijk}$  for all  $i = 1, \dots, r$ ,  $j = 1, \dots, c - 1$ , and  $k = 1, \dots, K$ . The proportional odds model considered has the form

$$\log\left(\frac{\pi_{ijk}^*}{1 - \pi_{ijk}^*}\right) = \alpha_j + \gamma_k + \beta_i \quad i = 1, \dots, r, \quad j = 1, \dots, c - 1, \quad k = 1, \dots, K, \quad (1)$$

with a constraint  $\gamma_K = \beta_r = 0$ . The model is simply a logit model when the response is binary.

This article uses an example provided by Merck Research Laboratories for a double-blind, parallel-group preliminary clinical study conducted at 21 centers, where patients suffering from asthma were randomly assigned to 3 different treatments (2mg active drug, 10mg active drug, and placebo). At the end of the study, the doctors described the patients' change in condition using an ordinal scale from better to worse (1 to 4). Table 1 shows the results of the doctors' evaluations associated with the treatments and is presented as 21 separate  $3 \times 4$  tables. Such a study might use many clinics because of the time it takes each clinic center to recruit many patients. Therefore, the three-way table might then have many strata but

few observations per stratum. It leads to a sparse data set.

In the proportional odds model (1), parameters  $\{\beta_i\}$  describe the treatment effects given clinics. For each clinic center, the odds that the evaluation for treatment  $i$  falls below any fixed level are  $\exp(\beta_i)$  times the odds for treatment  $r$ . We refer to the ordinal odds ratio  $\exp(\beta_i)$  as the *cumulative* odds ratio, since the model is based on the cumulative probabilities. When fitting the model, the maximum likelihood (ML) estimators of  $\{\beta_i\}$  perform badly for sparse data. Liu and Agresti (1996) and Liu (2003) have shown that the ML method tends to overestimate the effects. This is not surprising, because it also happens for a binary response case (Andersen 1980, p. 244 and Ghosh, 1995). The standard asymptotic properties of ML estimators fail when the number of nuisance parameters (such as  $\{\gamma_k\}$ ) grows at the same rate as the sample size (Neyman and Scott, 1948).

Alternatively, the Mantel-Haenszel (MH) type method provides better estimators for  $\{\beta_i\}$  (Liu and Agresti, 1996; Liu, 2003). Liu (2003) extended the ordinary MH estimator (Mantel and Haenszel, 1959) for a common odds ratio for several  $2 \times 2$  tables to the case of several  $r \times c$  tables. Like the ordinary MH estimator the extended MH-type estimator is consistent under both the *ordinary asymptotics* in which the number of strata is fixed and also the *sparse asymptotics* in which the number of strata grows with the sample size. The estimator is referred to a *dually* consistent estimator.

We assume that each  $r \times c$  table is formed by  $r$  independent multinomial samples  $(X_{1jk}, X_{2jk}, \dots, X_{rjk}, j = 1, \dots, c)$  with sample sizes denoted by  $(n_{1k}, n_{2k}, \dots, n_{rk})$ . The total sample size in each stratum is denoted by  $N_k = n_{1k} + \dots + n_{rk}$ , for  $k = 1, \dots, K$ . Let cumulative counts be  $X_{ijk}^* = X_{i1k} + \dots + X_{ijk}$  and let  $R_{jk}^{ih} = X_{ijk}^*(n_{hk} - X_{hjk}^*)/N_k$  and  $S_{jk}^{ih} = X_{hjk}^*(n_{ik} - X_{ijk}^*)/N_k$ . Liu (2003) proposed the MH type estimator of  $\beta_i$  as

$$\bar{L}_i = \frac{1}{r} \left( \sum_{h=1, h \neq i}^r L_{ih} - \sum_{h=1}^{r-1} L_{rh} \right), \quad i = 1, \dots, r-1, \quad (2)$$

where

$$L_{ih} = \log \left( \sum_{k=1}^K \sum_{j=1}^{c-1} R_{jk}^{ih} \right) - \log \left( \sum_{k=1}^K \sum_{j=1}^{c-1} S_{jk}^{ih} \right), \quad i = 1, \dots, r, \quad h = 1, \dots, r, \quad \text{and } i \neq h.$$

Furthermore, the dually consistent variance and covariance estimators for  $\{\bar{L}_i\}$  follow immediately from Greenland (1989) and Liu (2003) as

$$\begin{aligned}\hat{\text{Cov}}(\bar{L}_i, \bar{L}_h) &= \frac{1}{r^2}(U_{ih}^+ - U_{ir}^+ - U_{rh}^+ + U_{rr}^+) \\ \hat{\text{Var}}(\bar{L}_i) &= \frac{1}{r^2}(U_{i++} - 2U_{ir}^+ + U_{r++}),\end{aligned}\tag{3}$$

where  $U_{ihg} = 0$  if  $i = h$  or  $g$ ,  $U_{ih}^+ = U_{i++}$  if  $i = h$ , and  $U_{ih}^+ = U_{hi}^+ = U_{+ih} - U_{ih+} - U_{hi+} + U_{ihh}$  if  $i \neq h$ . The form of  $U_{ihg}$  is given in Appendix A. Liu (2003) gave the proof of dual consistency for the variance and covariance estimators. For sparse data such as that in Table 1, the MH-type estimators (2) of the cumulative log odds ratios  $\{\beta_i\}$  perform better than the ML estimators.

### 3. PROPORTIONAL ODDS MODEL WITH RANDOM EFFECTS

When comparing treatments on an ordinal response with stratified data, it is common to use Model (1) that assumes a lack of interaction between the treatment effects and strata. That is, the cumulative odds ratios used to describe the conditional association between treatment and response are the same for each clinic center. To check the homogeneity of cumulative odds ratios, we might use ordinary Pearson or likelihood-ratio goodness-of-fit tests for the corresponding model. Such tests are appropriate only under the ordinary asymptotics, which is not the case for Table 1.

In a clinical trial conducted to compare treatments among several clinical centers, the true treatment effects might vary due to some other factors, such as age among subjects at different centers. Therefore, it is plausible to estimate the degree of the heterogeneity for the cumulative odds ratios across the centers. Hartzel, Liu and Agresti (2001) used random effects terms to describe the variability for the cumulative log odds ratios for two treatments. When the strata are a sample, such as a sample of clinical centers, it is natural to treat the strata effects or the cumulative log odds ratios as random effects across strata. Consider a proportional odds model allowing random effects terms as

$$\log\left(\frac{\pi_{ijk}^*}{1 - \pi_{ijk}^*}\right) = \alpha_j + c_k + b_{ik} \quad i = 1, \dots, r, \quad j = 1, \dots, c - 1, \quad k = 1, \dots, K,\tag{4}$$

where  $b_{rk} = 0$  for all  $k$  and  $\{b_{1k}, \dots, b_{r-1,k}\}$  are a vector of correlated random effects. In the model,  $\{c_1, \dots, c_K\}$  are independent observations from a  $N(\gamma, \sigma_c)$  and  $\{b_{i1}, \dots, b_{iK}\}$  are independent observations from a  $N(\beta_i, \sigma_{\beta_i})$  for all  $i = 1, \dots, r-1$ . The model is a special case of a multivariate generalized linear mixed model for ordinal responses (Tutz and Hennevogl, 1996).

To obtain the likelihood function we construct the usual product of multinomials and then integrate out the random effects with respect to the random effects distribution. Let  $\mathbf{u}_k' = (c_k, b_{1k}, \dots, b_{r-1,k})$  be the random effects having a multivariate normal distribution with mean  $\boldsymbol{\beta} = (\gamma, \beta_1, \dots, \beta_{r-1})$  and the covariance matrix  $\boldsymbol{\Sigma}$ . That is,  $\mathbf{u}_k \sim N[\boldsymbol{\beta}', \boldsymbol{\Sigma}]$ . Let  $g(\mathbf{u}; \boldsymbol{\beta}, \boldsymbol{\Sigma})$  denote the multivariate normal density function with mean  $\boldsymbol{\beta}$  and covariance matrix  $\boldsymbol{\Sigma}$ . The likelihood function has the form

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = \prod_{k=1}^K \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[ \prod_{i=1}^r \prod_{j=1}^c (\pi_{ijk})^{X_{ijk}} \right] g(\mathbf{u}_k; \boldsymbol{\beta}, \boldsymbol{\Sigma}) d\mathbf{u}_k,$$

where

$$\begin{aligned} \pi_{i1k} &= \frac{1}{1 + \exp[-(\alpha_1 + c_k + b_{ik})]} \\ \pi_{ijk} &= \frac{1}{1 + \exp[-(\alpha_j + c_k + b_{ik})]} - \frac{1}{1 + \exp[-(\alpha_{j-1} + c_k + b_{ik})]}, \quad j = 2, \dots, c-1 \\ \pi_{ick} &= 1 - \frac{1}{1 + \exp[-(\alpha_{c-1} + c_k + b_{ik})]} \end{aligned}$$

To fit Model (4), one can consider Hedeker and Gibbons (1994) procedures. They maximized the likelihood after approximating the integrals by Gauss-Hermite quadrature. Alternatively, the procedure PROC NLMIXED in SAS for fitting generalized linear mixed models using adaptive Gauss-Hermite quadrature is also available to fit the proportional odds models with random effects. See Hartzel, Liu, and Agresti (2001) for the details of SAS codes. However, fitting the proportional odds model with random effects may not be suitable for highly sparse data. The fitting algorithm often fails to converge when the data are sparse and when there are many random effects. In our experience, the fitting algorithm for SAS PROC NLMIXED is often successful when the number of random effects is less than 3.

For Model (4), we describe the conditional association between treatments and response using the estimate of means  $\{\beta_1, \beta_2, \dots, \beta_{r-1}\}$ . The degree of the heterogeneity for the cumulative log odds ratios across the centers is given by the estimate of standard deviations  $\{\sigma_{\beta_1}, \sigma_{\beta_2}, \dots, \sigma_{\beta_{r-1}}\}$ .

#### 4. A DELETION INFLUENCE MEASURE

There are several ways to check for the homogeneity of cumulative odds ratios or to describe the heterogeneity among them. For instance, we might use an overall goodness-of-fit test to check the homogeneity. In the proportional odds model with random effects (4), we can use the variance estimators to indicate the level of heterogeneity among the cumulative log odds ratios across centers. When the homogeneity doesn't hold, the next question that arises is to identify which pieces of data contribute to the heterogeneity most. Sometimes, even though the heterogeneity is minor in magnitude and perhaps not even significant in the sample according to a statistical test, it is always interesting to get the detail of the heterogeneity for the cumulative log odds ratios among multi-centers.

Suppose we replace  $\{\beta_i\}$  by  $\{\beta_{ik}\}$  in the proportional odds model (1). This model allows the cumulative odds ratio to vary among centers. The variability in  $\{\beta_{i1}, \beta_{i2}, \dots, \beta_{iK}\}$  indicates the heterogeneity of the conditional associations between treatments and response across centers for all  $i = 1, \dots, r - 1$ . However, when there are few observations in a particular center (say center  $k$ ), the ML estimate of  $\{\beta_{1k}, \beta_{2k}, \dots, \beta_{r-1,k}\}$  is not reliable. The estimate can approach to either negative infinity or positive infinity because of the many zero counts within the center. A similar problem occurs if we use the Mantel-Haenszel type estimator (2) to estimate the cumulative odds ratios for each center. For the example in Table 1, each center has very few patients. It is not appropriate to estimate the center-specific cumulative odds ratios in the sense that the treatment effects obtained separately from each center provide little information. Consequently, the detail of heterogeneity is not available based on the estimates of the center-specific cumulative odds ratios.

Instead, this article proposes an influence measure that determines the strength of the heterogeneity of the cumulative log odds ratios across different centers based on the MH type

estimators (2). The proposed influence measure has a similar form as the Cook's distance (Cook and Weisberg, 1982) for the ordinary linear regression models. For the linear regression models, the Cook's distance is a common diagnostics method used to detect a goodness-of-fit of the model and also to identify possibly aberrant data points in multivariate analysis. In general, the Cook's distance  $D_k$  has the form

$$D_k = (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{(k)})^T \text{Cov}(\hat{\boldsymbol{\theta}})^{-1} (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{(k)})$$

where  $\hat{\boldsymbol{\theta}}$  is the vector of estimates based on all data points and  $\hat{\boldsymbol{\theta}}_{(k)}$  is the vector of estimates when the  $k$ th data point is deleted.

For the stratified clinical trials, we are interested which clinical center has a strong influence to the treatment effects. A clinical center is said to be influential if its deletion has a relatively large influence on the inference measure. For the multi-center clinical trials, if one center (e.g.,  $k$ th) is deleted, the effect of deleting the  $k$ th center on the estimates of  $\{\bar{L}_i\}$  is to yield the estimates of  $\{\bar{L}_{i(k)}\}$ . The proposed influence measure is defined as

$$C_k = (\bar{\mathbf{L}} - \bar{\mathbf{L}}_{(k)})^T \text{Cov}(\hat{\bar{\mathbf{L}}})^{-1} (\bar{\mathbf{L}} - \bar{\mathbf{L}}_{(k)}), \quad k = 1, \dots, K, \quad (5)$$

where the vector  $\bar{\mathbf{L}} = (\bar{L}_1, \bar{L}_2, \dots, \bar{L}_{r-1})$  is the MH estimates of the cumulative log odds ratios and  $\text{Cov}(\hat{\bar{\mathbf{L}}})$  is the estimate of the covariance matrix of  $(\bar{L}_1, \bar{L}_2, \dots, \bar{L}_{r-1})$ . The form of  $\{\bar{L}_i\}$  is given on (2) and their variance and covariance estimates are given on (3). The effect  $C_k$  of deletion is a quadratic form measuring the overall difference between the vector  $\bar{\mathbf{L}}$  and the  $\bar{\mathbf{L}}_{(k)}$ . A relatively large value of  $C_k$  implies that the  $k$ th clinical center has a high influence on the cumulative log odds ratios.

## 5. EXAMPLE

For the example in Table 1, we use an MH type method to estimate the cumulative odds ratios. We also describe the heterogeneity for the cumulative odds ratios among centers based on the two strategies discussed in Sections 3 and 4.

### 5.1. Cumulative Odds Ratios Estimates in a Proportional Odds Model



Using the MH type estimates (2) to describe the conditional relationship between the treatments and the patients' change in condition given each center, we have  $\bar{L}_{13} = 0.640$  and  $\bar{L}_{23} = 1.063$  with standard error estimates of 0.333 and 0.357 respectively. That is, the odds of improving using the active drug with dose 2mg (or 10mg) are estimated to be  $\exp(0.640) = 1.90$  (or  $\exp(1.063) = 2.90$ ) times the odds for the placebo, given each center. In comparison, the ordinary ML estimates of  $\hat{\beta}_1$  and  $\hat{\beta}_2$  for model (1) equal 0.797 with an estimated standard error of 0.343, and 1.099 with an estimated standard error of 0.355, respectively. We tend to obtain a larger parameter estimate from the ML fitting when the number of observations in each center is small.

## 5.2. First Diagnostics Strategy Using a Random Effects Model

Using the proportional odds model with random effects (4), the fitting algorithm fails to converge because of the sparseness of the data and more than 2 random effects in the model. To illustrate the strategy discussed in Section 3 in this example, we combine the response scales 3 and 4 to make the data less sparse. Since  $r = 3$  in this example, there are 3 random effects in model (4), including  $(c_k, b_{1k}, b_{2k})$ . To reduce the number of random effects, we let  $\{b_{2k}\}$  degenerate to its mean  $\beta_2$ . We treat  $\{c_1, c_2, \dots, c_K\}$  as independent observations from a  $N(\gamma, \sigma_c)$  and  $\{b_{11}, b_{12}, \dots, b_{1K}\}$  as independent observations from a  $N(\beta_1, \sigma_{\beta_1})$ . The random effects  $(c_k, b_{1k})$  are correlated with covariance  $\sigma_{c\beta}$ . Using SAS PROC NLMIXED, the estimate of  $\beta_1$  is 0.517 with a standard error of 0.286 and the estimate for the variability of  $\{b_{11}, b_{12}, \dots, b_{1K}\}$  is  $\hat{\sigma}_{\beta_1} = 0.690$  with a standard error of 0.595. Using the proportional odds model with random effects, the odds of improving using the active drug with dose 2mg are estimated to be  $\exp(0.517) = 1.68$  times the odds for the placebo across centers. Although the standard deviation  $\sigma_{\beta_1}$  is not significantly different from zero, the strength of heterogeneity is large compared to the level of  $\hat{\beta}_1$ . The estimate of  $\beta_2$  is 1.185, with a standard error of 0.325, that is, the odds of improving using the active drug with dose 10mg are estimated to be  $\exp(1.185) = 3.27$  times the odds for the placebo. Comparing to the MH type estimates of the cumulative odds ratio, the model (4) produces similar estimates of the conditional associations between treatments and response. Furthermore, the random effects

model gives the overall estimate for the level of heterogeneity in the treatment effects, that is  $\hat{\sigma}_{\beta_1}$ .

### 5.3. Second Diagnostics Strategy Using the Influence Measure

On the other hand, we can consider the influence measure (5) for the cumulative log odds ratios in each center. Table 2 lists the influence measure  $C_k$  and the estimates of the cumulative log odds ratios when each of the centers is deleted. The notations  $\{\bar{L}_{i(k)}\}$  denote the cumulative log odds ratios when the  $k$ th center is deleted for  $i = 1, 2$  and  $k = 1, \dots, 21$ . It is clear from Figure 1 that clinical centers 13, 15, 3, and 17 have a larger influence measure compared to the others. This might suggest that the cumulative log odds ratios for those centers are different from the others. It could be due to error, or some other factors (e.g., age), which is worth investigation.

### ACKNOWLEDGMENT

We would like to thank Megan Clark for helpful comments.

## BIBLIOGRAPHY

- Andersen, E. B. (1980). *Discrete statistical methods with social science applications*. New York: North-Holland.
- Andersen, E. B. (1992). Diagnostics in categorical data analysis. *J. R. Statist. Soc. B* 54(3):781-791.
- Atkinson, A. C. (1985). *Plot, transformations and regression*. Oxford: Clarendon.
- Cook, R. D. and Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.
- Ghosh, M. (1995). Inconsistent MLE for the Rasch model. *Statist. Probab. Lett.* 23:165-170.
- Greenland, S. (1989). Generalized Mantel-Haenszel estimators for  $K \times J$  tables. *Biometrics* 45:183-191.
- Hartzel, J., Liu, I-M. and Agresti, A. (2001). Describing heterogeneous effects in stratified ordinal contingency tables, with application to multi-center clinical trials. *Comput. Statist. Data Anal.* 35:429-449.
- Hedeker, D. and Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics* 50:933-944.
- Liu, I-M. and Agresti, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal responses. *Biometrics* 52:1223-1234.
- Liu, I. (2003). Describing ordinal odds ratios for stratified  $r \times c$  tables. *Biometrical J.* 45:730-750.
- Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl. Cancer Inst.* 22:719-748.
- McCullagh, P. (1980). Regression models for ordinal data. *J. R. Statist. Soc. B* 42:109-142.

- Neyman, J. and Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrika* 16:1-22.
- Pregibon, D. (1981). Logistic regression diagnostics. *Ann. Statist.* 9:705-724.
- Tutz, G. and Hennevogl. W. (1996). Random effects in ordinal l regression models. *Comput. Statist. Data Anal.* 22:537-557.
- Wang, D. Q. and Critchley, F. (2000). Multiple deletion measures and conditional influence in regression models. *Commun. Statist. – Theory Meth.* 29:2391-2404.
- Wang, D. Q., Critchley, F. and Liu, I. (2004). Diagnostics analysis and perturbations in a clustered sampling model. *Commun. Statist. – Theory Meth.* 33:2709-2722.
- Wang, D. Q., Critchley, F. and Smith, P. J. (2003). The multiple sets of deletion measurer and masking in regression. *Commun. Statist. – Theory Meth.* 32:407-413.

## APPENDIX A

The form of  $U_{ihg}$  is as follows:

$$U_{ihg} = \begin{cases} \frac{1}{\hat{\theta}_{ih}^2} \frac{\sum_{k=1}^K \sum_{j=1}^{c-1} \hat{\phi}_{jk}^{ih}(\hat{\theta}_{ih}) + 2 \sum_k \sum_{j < s}^{c-1} \hat{\phi}_{j sk}^{ih}(\hat{\theta}_{ih})}{\left[ \sum_{k=1}^K \sum_{j=1}^{c-1} S_{jk}^{ih} \right]^2}, & \text{if } i \neq h = g \\ \frac{1}{\hat{\theta}_{ih} \hat{\theta}_{ig}} \frac{\sum_{k=1}^K \sum_{j=1}^{c-1} \hat{\phi}_{jk}^{ihg}(\hat{\theta}_{ih}) + \sum_{k=1}^K \sum_{j \neq s}^{c-1} \hat{\phi}_{j sk}^{ihg}(\hat{\theta}_{ih}, \hat{\theta}_{ig})}{\left( \sum_{k=1}^K \sum_{j=1}^{c-1} S_{jk}^{ih} \right) \left( \sum_{k=1}^K \sum_{j=1}^{c-1} S_{jk}^{ig} \right)}, & \text{if } i \neq h \neq g, \end{cases}$$

where  $\hat{\theta}_{ia} = (\sum_k \sum_j R_{jk}^{ia}) / (\sum_k \sum_j S_{jk}^{ia})$  for all  $a \neq i$  and

$$\hat{\phi}_{jk}^{ih}(\theta) = \frac{1}{N_k^2} [(n_{ik} - X_{ijk}^*) X_{hjk}^{*2} \theta^2 + (n_{ik} - X_{ijk}^*) (n_{hk} - X_{hjk}^*) \\ (X_{ijk}^* + X_{hjk}^*) \theta + (n_{hk} - X_{hjk}^*) X_{ijk}^*]$$

$$\hat{\phi}_{j sk}^{ih}(\theta) = \frac{1}{N_k^2} [(n_{ik} - X_{isk}^*) X_{hjk}^{*2} \theta^2 + (n_{ik} - X_{isk}^*) (n_{hk} - X_{hsk}^*) \\ (X_{ijk}^* + X_{hjk}^*) \theta + (n_{hk} - X_{hsk}^*) X_{isk}^*]$$

$$\hat{\phi}_{jk}^{ihg}(\theta) = \frac{1}{N_k^2} [(n_{gk} X_{ijk}^* X_{hjk}^* - n_{ik} X_{hjk}^* X_{gjk}^*) \theta \\ + (n_{hk} n_{gk} X_{ijk}^* - n_{gk} X_{ijk}^* X_{hjk}^*)]$$

$$\hat{\phi}_{j sk}^{ihg}(\theta_{ih}, \theta_{ig}) = \begin{cases} \frac{n_{ik} \theta_{ih}}{N_k^2} [X_{hjk}^* (n_{gk} - X_{gsk}^*)] & \text{if } j < s \\ \frac{n_{ik} \theta_{ig}}{N_k^2} [X_{gsk}^* (n_{hk} - X_{hjk}^*)] & \text{if } j > s. \end{cases}$$

Figure 1: Influence Measure for Table 1

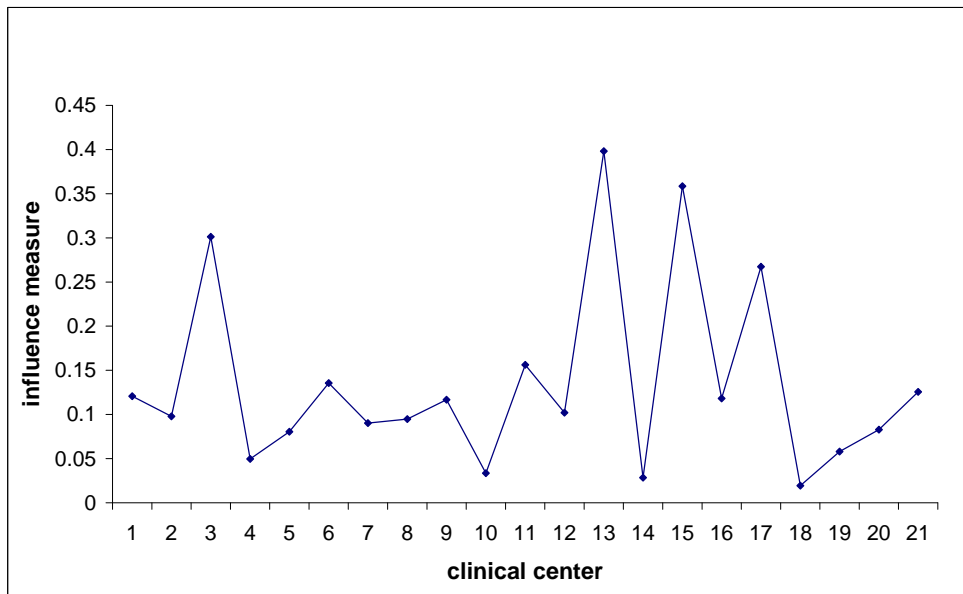


Table 1: Doctors' Evaluations of patients suffering from asthma

Center	Drug	Response				Center	Drug	Response			
		1	2	3	4			1	2	3	4
1	2mg	0	1	2	1	2	2mg	0	0	1	1
	10mg	0	2	0	2		10mg	0	0	1	2
	Placebo	0	0	0	4		Placebo	0	1	1	1
3	2mg	0	0	2	2	4	2mg	0	2	3	1
	10mg	0	0	1	0		10mg	0	2	2	2
	Placebo	0	1	4	1		Placebo	0	0	1	1
5	2mg	0	1	1	0	6	2mg	0	2	0	0
	10mg	1	0	0	2		10mg	0	1	0	0
	Placebo	1	0	0	2		Placebo	0	0	0	2
7	2mg	0	0	2	2	8	2mg	1	0	0	1
	10mg	0	0	2	1		10mg	0	3	0	0
	Placebo	0	0	2	1		Placebo	0	0	1	1
9	2mg	0	0	2	1	10	2mg	0	2	1	1
	10mg	1	0	2	0		10mg	0	1	0	0
	Placebo	0	0	1	0		Placebo	0	1	0	1
11	2mg	0	1	1	0	12	2mg	1	0	0	0
	10mg	0	1	1	1		10mg	1	1	0	0
	Placebo	0	0	0	3		Placebo	0	2	0	0
13	2mg	0	1	1	2	14	2mg	1	1	3	0
	10mg	0	1	0	1		10mg	1	0	1	0
	Placebo	0	0	0	5		Placebo	0	0	1	0
15	2mg	0	0	3	2	16	2mg	0	2	2	1
	10mg	0	2	3	0		10mg	2	1	2	0
	Placebo	0	1	2	1		Placebo	1	1	1	1
17	2mg	0	1	0	0	18	2mg	1	1	1	1
	10mg	1	1	1	2		10mg	0	1	0	0
	Placebo	0	1	1	3		Placebo	1	0	1	3
19	2mg	0	3	0	0	20	2mg	0	0	2	1
	10mg	0	1	3	0		10mg	0	1	0	3
	Placebo	0	1	1	1		Placebo	0	1	1	0
21	2mg	0	1	1	2						
	10mg	1	1	0	1						
	Placebo	0	1	0	0						

Note: Response is scaled from better (1) to worse (4).

Table 2: Influence Measure

center $k$	$C_k$	$\bar{L}_{1(k)}$	$\bar{L}_{2(k)}$
1	0.12077054	0.5282153	0.9743305
2	0.09793579	0.7408020	1.1446891
3	0.30112743	0.7972340	1.2365672
4	0.04966114	0.5705335	1.0466812
5	0.08045233	0.6477367	1.1530978
6	0.13563627	0.5512686	1.0898087
7	0.09029371	0.6998710	1.0233371
8	0.09484786	0.5829268	1.1082610
9	0.11666712	0.6872194	0.9950723
10	0.03345109	0.6582769	1.0201875
11	0.15629901	0.5518062	0.9246463
12	0.10196999	0.5997148	1.1298268
13	0.39812785	0.5163253	0.8389362
14	0.02842851	0.6215989	1.1008199
15	0.35854960	0.7749191	1.0047930
16	0.11812639	0.7500024	1.1544630
17	0.26737486	0.5764530	0.8815465
18	0.01933118	0.6169018	1.0863282
19	0.05794393	0.5611022	1.0292188
20	0.08288072	0.7351112	1.1035303
21	0.12551658	0.7508712	1.0878349