Truncated Estimate in
Log-Binomial Model: Algorithm and Simulation

[1]Luyan Dai, [2,3]Yang Li and [4]Ye Shen
[1]Department of Biometrics and Data Management,
Boehringer Ingelheim Pharmaceuticals, Ridgefield, CT, USA
[2]Data Mining Center, Renmin University of China, Beijing, P.R. China
[3]Division of Biostatistics, Yale School of Public Health, New Haven, CT, USA
[4]Department of Epidemiology and Biostatistics,
College of Public Health, University of Georgia, Athens, GA, USA

Abstract: Problem statement: Relative risk has concrete meanings of comparing two groups and measuring the association between exposures and outcomes in medical and public health studies. Log-binomial model, using a log link function on binary outcomes, is straightforward to estimate risk ratios, whereas generates boundary problems. When the estimates are located near the boundary of constrained parameter space, common approaches or procedures using software such as R or SAS fail to converge. **Approach:** In this study we proposed a truncated algorithm to estimate relative risk using the log-binomial model. We used simulation studies on both single and multiple covariates models to investigate its performance and compare with other similar methods. **Results:** Our algorithm was shown to outperform other methods regarding precision, especially in high dimensional predictor space. **Conclusion:** The truncated IWLS method solves the slow convergence problem and provides valid estimates when previously proposed methods fail.

Key words: Generalized linear model, iterative weighted least square, maximum likelihood estimator, odds ratio, relative risk, truncation

INTRODUCTION

For datasets with binary responses, there has been tremendous research study done by statisticians on prediction and inference. So far, the logistic model has been among the most popular models and widely used in the fields of medical and public health studies due to its pleasant characteristics linking to estimation of the odds ratio (McCullagh and Nelder, 1999). Nevertheless, odds ratio may occasionally not be scientifically appropriate to measure the association between exposures and outcomes. Relative risk, which is a ratio of the probability of the event occurring in the exposed group versus a non-exposed group, has more concrete meanings in the sense of comparing two groups (Holford, 2002; Robbins *et al*., 2002; Spiegelman and Hertzmark, 2005). If we further explore the literatures, we found out that many studies used odds ratios obtained from logistic models to approximate the relative risks. Neglect of the fact that such approximation merely retains its validity under the rare disease assumption will produce largely biased estimates away from the null hypothesis and consequently lead to false positive conclusions in many scenarios (Zhang and Yu, 1998; McNutt *et al*., 2003; Chu and Cole, 2010). As an alternative, log-binomial models may be more favorable with direct connection of coefficient estimates to the risk ratio. However, the algorithm of log-binomial models sometimes fails to converge and produces an invalid Maximum Likelihood Estimate (MLE) (Baumgarten *et al*., 1989; Petersen and Deddens, 2010), attributable to the constrained space of the linear predictors using log link. The problem becomes severe when multiple covariates are considered (Lumley *et al*., 2006; Williamson, 2011). Many researchers have been working on the log-binomial model and offered several virtuous estimation methods. Lumley *et al*. (2006) had a comprehensive discussion on the issues of log-binomial model.

Among all, the methods in favor include: Maximum Likelihood Estimation; Nonlinear Least Squares; Scaling by the Average Prevalence; Duplication of Cases. Whereas some of these methods don't fit in the regression scenario. For instance, the

Corresponding Author: Ye Shen, Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, GA, USA

Scaling by the average prevalence method requires that the model has the sole binary predictor, which renders the method limited in general. In contrast, MLE based algorithm is prevalent among the regression studies owing to the solid characteristics of inference. Major MLE based algorithms include: Searching the boundary; COPY method; Truncating fitted values (We also name it as "Truncated IWLS Method (T-IWLS)"); (Wacholder, 1986; Deddens and Petersen, 2003) The Searching the boundary algorithm may get to be intractable due to the difficulty in finding the boundaries in high dimensional parameter space. In this article, we focus on the discussions on COPY and Truncated IWLS methods. A modified algorithm of Truncated IWLS method is proposed.

MATERIALS AND METHODS

COPY method: Deddens and Petersen (2003) proposed a COPY method which could possibly obtain estimates close to the maximum likelihood estimator. The idea of COPY is to take C copies of the data and 1 copy with response Y set to 1-Y and fit the relative risk model to these modified data. They did not achieve clarity in several issues: (a) for the example they illustrated, they did not provide detailed proof showing the mechanism of why and how the method works; (b) further simulation results indicated that COPY method may not necessarily study well when C is very large and how to adjust for large C remains unsolved; (c) the performance of COPY method in models with multiple covariates is not yet promising according to our simulation results.

SAS solutions: Here we use SAS 9.2 PROC GENMOD to explore the illustrative example of Deddens and Petersen (2003). We fit the data using 3 models: (a) log link with binomial distribution; (b) log link with Poisson distribution; (c) log link with no specific distribution.

From the table we observed that estimates from the log-binomial model failed. SAS 9.2 tried to fix the boundary problem by unspecification of the distribution. The estimates of-2.1538 and 0.2351 are not far from the exact MLE, although still outperformed by COPY Method. Moreover, estimates from the model with unspecified distribution lack interpretability in practical use.

Truncated IWLS method: The log-binomial model belongs to the class of Generalized Linear Models (GLM). It is known that the Iterative Weighted Least Square method (IWLS) (Wolke and Schwetlick, 1988) is frequently used to estimate GLM parameters. We

start with illustrating the algorithm of IWLS, particularly for log-binomial model and then introduce the proposed the method in subsequent context.

IWLS method: Assume Y is Bernoulli distributed, we can write the p.d.f. as:

$$f(y) = p^y(1-p)^{1-y} = \exp\{y\log\frac{p}{1-p} + \log(1-p)\}$$

where, p=E(Y). Set the linear predictor as:

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_{q-1} X_{q-1} = X\beta$$

Log-link or $\eta = \log(p)$ or $p = e^\eta$ is adopted. Based on GLM theory on exponential distribution family, we can write:

$$\theta = \log\frac{p}{1-p}, b(\theta) = \log(1+e^\theta), v_i = b''(\theta) = p_i(1-p_i)$$

Clearly, $\partial\eta_i\,|\partial p_i = 1|p_i$. Let W=diag(w_i) and Z=($z_1,\ldots,z_n$)', where:

$$w_i = v_i^{-1}\left(\frac{\partial\eta_i}{\partial p_i}\right)^{-2} = \frac{p_i}{1-p_i} \text{ and } Z_i = \eta_i + \frac{\partial\eta_i}{\partial p_i}(y_i - p_i)$$

We can get the MLE of β's using the following equation interatively until converge Eq. 1:

$$X^T W_{\beta^{old}} X\beta^{new} = X^T W_{\beta^{old}} Z_{\beta^{old}} \text{ or } \beta^{new} \text{ or}$$
$$= (X^T W_{\beta^{old}} X)^{-1} X^T W_{\beta^{old}} Z_{\beta^{old}} \tag{1}$$

However, the convergence problem occurs occasionally in estimation and leads to un-estimable model parameters using the above IWLS method. In SAS, we will receive warnings indicating the algorithm fails to converge.

Truncated IWLS method: We lands on the method of truncating fitted values to get MLEs for log-binomial models to solve the boundary problem. The idea is very straightforward. $p = E(Y) = e^{X\beta}$, which may exceed the parameter space [0, 1] in the iterative steps when values hitting the boundary, i.e., X β > 0. To solve the problem, we propose a constraint on Xβ in each iterative step. As Lumley *et al.* (2006) pointed out: An approximate MLE for the log-binomial model is to simply truncate the range of ρ (Wacholder, 1986). Usually we pick a threshold near 1, such as 0.999 and ρ

is set to min (p, 0.999) after each iteration for the purpose of computing working residuals and working weights for the next iteration. In the i-th iterative step, we update β_{new} using β_{dd} when $X\beta_{old}<0$, otherwise, $\beta_{dd}=T$ is used in (1). It is equivalent to construct a constrain:

$$P_i = \min (T, \exp(X\beta))$$

Here, $0<T<1$. It can be conjectured that T are preferred to be some large value close to 1 when the MLE is on the boundary of parameter space. The estimator improves the robustness although scarifying some degree of un-biasness in these extreme scenarios.

At the same time, we notice that iterations may be stuck within parameter areas which are impossible for convergence. To improve the efficiency of the algorithm, we add another step to assist iterative estimates in jumping out of such area. Meanwhile, this "jumping" step will help to solve the problem of instability and diverged estimates due to starting values. The idea is intuitive: if the steps fail to converge after a significant number of iterations, we replace ρ_i by independent sample from the interior of the parameter space. For instance, we can sample $\rho_i \sim$ Uniform (0.1, 0. 9) .There is no definitive choice of the range for uniform distribution. The purpose is to start with new values when iterations rush to the wrong direction against convergence. This step makes this method favorable and greatly improves the truncated IWLS algorithm.

We can oversee its benefits when generalizing the application to multiple covariate cases. Consider the following simple example with 2 quantitative covariates.

For the log-binomial model, the parameter space is:

$$\{(\beta_0,\beta_1,\beta_2): 0 \leq \exp(\beta_0 + \beta_1 X_1 + \beta_1 X_2)$$
$$\leq 1 \quad \text{for} \quad 1 \leq X_1 \leq 5, \quad 2 \leq X_2 \leq 10\}$$

Figure 1 illustrates the surface of the boundary which has consists of 4 segments. Using Truncated IWLS algorithm without the jumping step, we experienced cases when iterations kept bouncing back and forth in one particular corner. Adding this step avoids those redundant iterations.

Some issues still remain open to discussion: (1) the choice of T and its sensitivity; (2) the performance in multiple covariates model; (3) comparison with COPY and other methods. These issues will be addressed in the remaining context of this study. We begin with the illustrative example by Deddens and Petersen (2003) and explore the properties of the method using simulated datasets in single and multiple covariate models with dichotomous outcomes. The conclusion will be drawn based upon the comparison in the end.
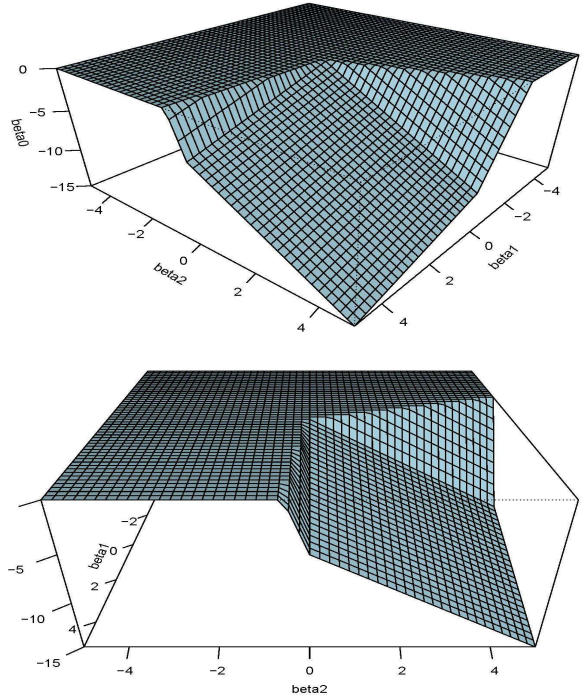

Fig. 1: Parameter space for Illustrative example in Table 2

Table 1: SAS results from proc GENMOD on Illustrative example by Deddens and Petersen (2003)

	β_0	SE (β_0)	β_1	SE (β_1)
Binomial	-0.8273	0.5892	0.0827	0.1272
Poisson	-2.8839	1.5956	0.3251	0.1964
Distribution unspecified	-2.1538	0.6655	0.2351	0.0765
Exact MLE	-2.0936	1.0208	0.2094	0.1021
COPY method	-2.0913	1.0197	0.2091	0.1020

Table 2: Illustrative example of 2-quantitative covariates

X_1	1	1.5	2	2.5	3	3.5	4	4.5	5
X_2	3	10.0	5	4.0	6	7.0	8	2.0	9
Y	0	00.0	0	0.0	1	0.0	0	1.0	1

Table 3: Illustrative example of 1-quantitative covariate by Deddens and Petersen (2003)

X	1	2	3	4	5	6	7	8	9	10
Y	0	0	0	0	1	0	1	1	1	1

Illustrative example by Deddens and Petersen (2003) revisited: In this example, 10 pairs of observations are generated (Table 3). Y is the response variable with possible values 0 and 1. Variable X is the predictor in the range of (1, 10).

We use R to implement the algorithm of the proposed method. We compare it with Exact MLE, COPY method and log-binomial and log-poisson models from SAS PROC GENMOD (Table 1).

Table 4: Results of Illustrative Example by Deddens and Petersen (2003)

	β_0	SE (β_0)	β_1	SE (β_1)
Truncated IWLS (T = 0.2)	-4.5309	2.6207	0.7864	0.4160
Truncated IWLS (T = 0.4)	-3.4217	1.5731	0.4639	0.2207
Truncated IWLS (T = 0.6)	-2.7877	1.0670	0.3454	0.1345
Truncated IWLS (T = 0.8)	-2.4243	0.8971	0.2744	0.1073
Truncated IWLS (T = 0.9)	-2.2447	0.7864	0.2490	0.0875
Truncated IWLS (T = 0.999)	-2.0946	0.7917	0.2097	0.0793
Truncated IWLS (T = 0.9999)	-2.0935	0.7928	0.2094	0.0792

Table 5: True values of model parameter

Prevalence at	Zero slope		Medium slope		High slope	
X = 5	β_0	β_1	β_0	β_1	β_0	β_1
0.1	-2.3026	0.00	-3.3026	0.20	-4.3026	0.40
0.3	-1.2040	0.00	-1.7040	0.10	-2.2040	0.20
0.5	-0.6931	0.00	-0.9431	0.05	-1.1931	0.10
0.7	-0.3567	0.00	-0.5067	0.03	-0.6567	0.06
0.9	-0.1054	0.00	-0.1554	0.01	-0.2054	0.02

Table 6: Average slope and average estimated SE (β_1)

Prevalence at		Zero slope		Medium slope		High slope	
X = 5	Method	β_1	s.e	β_1	s.e.	β_1	s.e.
0.1	T-IWLS (T = 0.999)	-0.0072	0.1152	0.206	0.103	0.417	0.0920
	T-IWLS (T = 0.9999)	-0.0071	0.1163	0.211	0.103	0.418	0.0910
	GEN+COPY1000	-0.0020	0.1050	0.218	0.111	0.422	0.1050
0.3	T-IWLS (T = 0.999)	0.0020	0.0550	0.100	0.051	0.205	0.0440
	T-IWLS (T = 0.9999)	0.0010	0.0550	0.100	0.051	0.201	0.0440
	GEN+COPY1000	-0.0020	0.0500	0.103	0.051	0.203	0.0480
0.5	T-IWLS (T = 0.999)	0.0010	0.0350	0.050	0.034	0.101	0.0290
	T-IWLS (T = 0.9999)	0.0000	0.0350	0.050	0.033	0.100	0.0290
	GEN+COPY1000	-0.0000	0.0320	0.050	0.032	0.102	0.0310
0.7	T-IWLS (T = 0.999)	0.0010	0.0230	0.031	0.021	0.059	0.0175
	T-IWLS (T = 0.9999)	-0.0010	0.0230	0.030	0.021	0.060	0.0174
	GEN+COPY1000	-0.0010	0.0210	0.029	0.039	0.059	0.0390
0.9	T-IWLS (T = 0.999)	0.0000	0.0110	0.010	0.010	0.019	0.0080
	T-IWLS (T = 0.9999)	0.0000	0.0110	0.010	0.010	0.019	0.0080
	GEN+COPY1000	0.0010	0.0110	0.010	0.010	0.018	0.0090

We select the candidates of T in {0.2, 0.4, 0.6, 0.8, 0.9, 0.999, 0.9999}. The estimates of β_0 and β_1 are summarized in Table 4.

The results confirm our conjecture that larger T could enhance the estimation. With large enough T, e.g., 0.999, 0.9999, the truncated method provides reasonable estimates compared with Exact MLE. The difference between the choice of T = 0.999 and 0.9999 is almost neglectable. It also outperforms the COPY method regarding precision. We tried different sets of the starting values; all of them converge to the estimates listed in Table 4. However, we cannot neglect the tendency of underestimating standard errors using truncated IWLS. Scale factors in GLM adjust the effect of over-dispersion. Originally, the scale factor for binomial model is 1. However, the distribution achieved by truncated algorithm is not exactly binomial since we add a constraint on the boundary. Thus, some correction factor is required. Observed from later simulation, the magnitude of the impact by truncation can partially measured by the prevalence at median values of covaraites. When the prevalence is low, the underestimation of the standard errors is trivial and almost ignorable. We will leave the discussion of the problem open to further investigation.

RESULTS

Single covariate model: We used same simulation scheme as Deddens and Petersen (2003) to compare with the proposed method. Data were generated from the log-binomial model. Covariate X ranging from 0-10 is generated uniformly. Given the prevalence at X = 5 varying among {0.1, 0.3, 0.5, 0.7, 0.9), three values are selected for β_1 namely 0, medium and large. True values of β_1 and corresponding β_0 are listed in Table 5. Note that β_0 = log (prevalence) $-5 \times \beta_1$. Each simulation involves 1,000 replications (same X's, different Y's).

Table 6 displays the average slope and average estimated standard errors of slope from the combined method of PROC GENMOD and COPY method (details in Deddens and Petersen, 2003) and the proposed method using T = 0.999 and 0.9999.

Comparing to true population values (Table 5), the truncated IWLS method estimates are all as close as the combined method and indeed verify that truncated method is exact or at least asymptotic MLE. Using different values of T yield very similar precision and MSE (Table 7), indicating the truncation IWLS method estimates is very stable regarding the choice of T. This characteristic can make the method more attractive and realistic in application since T = 0.9999 is sufficient in dealing with various scenarios.

Multiple covariate model: Deddens and Petersen (2003) claimed that the simulations with 2 independent variables are consistent with the results for one dependent variable using COPY method.

However, COPY method also may fail to converge when the dimension of covariates goes higher. The parsimonious strategy is to make more copies, which heavily increase the burden of the computation. Our method is quietly efficient without much adjustment. One remaining concern is the choice of truncation threshold T. The following simulation is mainly designed to address the issue.

Two sets of simulations are conducted. The first one is based on 2-covaraite model with covariates X_1 ranging from 0-10 and X_2 among [0, 5]. Given the prevalence at $(X_1, X_2) = (5, 2.5)$, we select relatively bigger and smaller values of β_1 respectively. The specified values are listed in Table 8. We then compare the truncated IWLS with T = 0.999 and T = 0.9999 regarding the precision and MSE. Each simulation contains 1,000 replications (same covariates and different Y). Table 8 summarizes the population values of $(\beta_0, \beta_1, \beta_2)$ in these 2-covariate models.

The second set is based on 3-covariate model with covariates X_1 among [0, 10], X_2 among [0, 5] and X_3 from 0-2. We cover 3 combinations (e.g., large β_1, medium β_2 and small β_3) and replicate 1,000 samples to compare proposed method using T=0.999 and 0.9999 respectively. The population values of $(\beta_0, \beta_1, \beta_2, \beta_3)$ are listed in Table 9.

The results of simulations from 2-covaraite models are summarized in Table 10. The different truncation values yield very close results. Both obtain the estimates of parameters close to population values (Table 8) up to 3 decimal points. The MSE of estimates from the two truncation values are very small.

Table 7: MSE of estimated β_1 using truncated IWLS

Prevalence at X=5	Method	Zero	Medium	High
0.1	T-IWLS (T = 0.999)	0.0146	0.0127	0.0103
	T-IWLS (T = 0.9999)	0.0220	0.0128	0.0104
0.3	T-IWLS (T = 0.999)	0.0032	0.0026	0.0020
	T-IWLS (T = 0.9999)	0.0031	0.0027	0.0021
0.5	T-IWLS (T = 0.999)	0.0013	0.0444	0.0009
	T-IWLS (T = 0.9999)	0.0013	0.0011	0.0009
0.7	T-IWLS (T = 0.999)	0.0006	0.0006	0.0003
	T-IWLS (T = 0.9999)	0.0006	0.0006	0.0003
0.9	T-IWLS (T = 0.999)	0.0001	0.0001	0.0001
	T-IWLS (T = 0.9999)	0.0001	0.0001	0.0000

Table 8: True values of 2-covariate model parameters

Prevalence at $(X_1, X_2) = (5, 2.5)$	0.3	0.5	0.7
$(\beta_0, \beta_1, \beta_2)$	(-1.829, 0.1, 0. 05)	(-1.318, 0.1, 0. 05)	(-0.982, 0.1, 0. 05)
$(\beta_0, \beta_1, \beta_2)$	(-1.704, 0.05, 0. 1)	(-1.193, 0.05, 0. 1)	(-0.857, 0.05, 0. 1)

Table 9: True Values of 3-covariate model parameters

Prevalence at $(X_1, X_2, X_3)=(5,2.5,1)$	0.3	0.5	0.7
$(\beta_0, \beta_1, \beta_2, \beta_3)$	(-1.804, 0.10, 0.02, 0.05)	(-1.293, 0.10, 0.2, 0.05)	(-0.957, 0.10, 0.20, 0.05)
$(\beta_0, \beta_1, \beta_2, \beta_3)$	(-1.849, 0.10, 0.05, 0.02)	(-1.338, 0.10, 0.05, 0.02)	(-1.00, 0.10, 0.05, 0.02)
$(\beta_0, \beta_1, \beta_2, \beta_3)$	(-1.529, 0.02, 0.05, 0.10)	(-1.018, 0.02, 0.05, 0.10)	(-0.682, 0.02, 0.05, 0.10)

Table 10: Average and MSE of Estimated (β_1, β_2) in 2-Covarate Models (Table 8)

Prevalence at $(X_1, X_2) = (5, 2.5)$	Method	$(\beta_1\beta_2) = (0.1, 0. 05)$		$(\beta_1,\beta_2) = (0,05,0,1)$	
		$\beta1$	$\beta2$	$\beta1$	$\beta2$
		$[MSE (\beta_1)]$	$[MSE (\beta_2)]$	$[MSE (\beta_1)]$	$[MSE (\beta_2)]$
0.3	T-IWLS (T=0.999)	0.104 [0.007]	0.043 [0.023]	0.058 [0.005]	0.109 [0.024]
	T-IWLS (T=0.999)	0.105	0.047	0.053	0.109
		[0.006]	[0.025]	[0.005]	[0.023]
0.5	T-IWLS (T=0.999)	0.1	0.05	0.052	0.103
		[0.002]	[0.005]	[0.003]	[0.006]
	T-IWLS (T=0.999)	0.1	0.05	0.052	0.102
		[0.002]	[0.004]	[0.003]	[0.005]
0.7	T-IWLS (T=0.999)	0.101	0.049	0.049	0.101
		[0.001]	[0.002]	[0.001]	[0.002]
	T-IWLS (T=0.999)	0.1	0.05	0.05	0.101
		[0.001]	[0.002]	[0.001]	[0.002]

Although the results are not listed in this article, results of simulations from 3-covariate models are consistent with those in 2-covaraites models. In fact, we do increase the number of covariates up to 8 in exploratory studies. The results support our conclusion that selection of T is trivial in truncated IWLS method once T is large enough. The suggested value in application is 0.9999.

DISCUSSION

In this study we discussed the importance of applying log-binomial models instead of logistic models to epidemiological studies when the rare disease assumption is invalid. In all, the truncated IWLS method study reasonably in both single and multiple covariates models. Our algorithm solves the slow convergence problem and provides valid estimates when previously proposed methods fail. Simulation results also show that the algorithm is not sensitive when the threshold in the truncation is selected close enough toward 1. *Ad hoc* methods such as multiple endpoints investigation could be used to obtain a working threshold, but more objective methods are in demand. Another potential research topic is to study the convergence performance in the truncated IWLS algorithm so that we can obtain further information on the convergence rate. Simulation studies to compare the convergence rate between the truncated IWLS algorithm and the algorithm of its corresponding untruncated model is also suggested.

CONCLUSION

The boundary problems in log-binomial models were solved by a newly developed truncated IWLS method. The proposed method outperformed the existing COPY method when multiple covariates co-exist in a log-binomial model and is therefore of practical value. Such models and algorithms can be widely used in high-prevalence disease modeling, such as diabetes and cardiovascular diseases.

REFERENCES

Baumgarten, M., P. Seliske and M.S. Goldberg, 1989. Warning regarding the use of GLIM macros for the estimation of risk ratios. Am. J. Epidemiol., 130: 1065-1065.

Chu, H. and S.R. Cole, 2010. Estimation of risk ratios in cohort studies with common outcomes: A Bayesian approach. Epidemiology, 6: 855-862. DOI: 10.1097/EDE.0b013e3181f2012b

Deddens, J.A. and M.R. Petersen, 2003. Estimation of prevalence ratios when PROC GENMOD does not converge. Proceedings of the 28th Annual SAS Users Group International Conference, (SASUGIC' 03), Cary NC, SAS Institute Inc.

Holford, T.R., 2002. Multivariate Methods in Epidemiology. 1st Edn., Oxford University Press, Oxford, ISBN-10: 0195124405, pp: 408.

Lumley, T., R. Kronmal and S. Ma, 2006. Relative risk regression in medical research: Models, contrasts, estimators and algorithms. University of Washington.

McCullagh, P. and J.A. Nelder, 1999. Generalized Linear Models. 2nd Edn., Chapman and Hall/CRC, Boca Raton, ISBN-10: 0412317605, pp: 511.

McNutt, L.A., C. Wu, X. Xue and J.P. Hafner, 2003. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am. J. Epidemiol., 10: 940-943. DOI: 10.1093/aje/kwg074

Petersen, M.R. and J.A. Deddens, 2010. Maximum likelihood estimation of the log-binomial model. Commun. Stat. Theory Methods, 5: 874-883. DOI: 10.1080/03610920902807879

Robbins, A.S., S.Y. Chao and V.P. Fonseca, 2002 What's the relative risk? A method to directly estimate risk ratios in cohort studies of common outcomes. Ann. Epidemiol., 12: 452-454. PMID: 12377421

Spiegelman, D. and E. Hertzmark, 2005. Easy SAS calculations for risk or prevalence ratios and differences. Am. J. Epidemiol., 162: 199-200. DOI: 10.1093/aje/kwi188

Wacholder, S., 1986. Binomial regression in GLIM: Estimating risk ratios and risk differences. Am. J. Epidemiol., 123: 174-84.

Williamson, T.S., 2011. Log-binomial models: Maximum Likelihood and failed convergence. Ph.D. Thesis, University of Calgary.

Wolke, R. and H. Schwetlick, 1988. Iteratively Reweighted least squares: Algorithms, convergence analysis and numerical comparisons. SIAM J. Sci. Stat. Comput., 9: 907-921. DOI: 10.1137/0909062

Zhang, J. and K.F. Yu, 1998. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. J. Am. Med. Assoc., 280: 1690-1691.