HHS Public Access

Author manuscript

J Comput Graph Stat. Author manuscript; available in PMC 2018 February 06.

Published in final edited form as:

J Comput Graph Stat. 2017 ; 26(3): 734–737. doi:10.1080/10618600.2017.1321552.

One-Step Generalized Estimating Equations with Large Cluster Sizes

Stuart Lipsitz,
Brigham & Women's Hospital, Boston, MA

Garrett Fitzmaurice,
Harvard Medical School, Boston, MA

Debajyoti Sinha,
Florida State University, Tallahassee, FL

Nathanael Hevelone,
Brigham & Women's Hospital, Boston, MA

Jim Hu, and
Cornell Medical College, New York, NY

Louis L. Nguyen
Brigham & Women's Hospital, Boston, MA

Abstract

Medical studies increasingly involve a large sample of independent clusters, where the cluster sizes are also large. Our motivating example from the 2010 Nationwide Inpatient Sample (NIS) has 8,001,068 patients and 1049 clusters, with average cluster size of 7627. Consistent parameter estimates can be obtained naively assuming independence, which are inefficient when the intra-cluster correlation (ICC) is high. Efficient generalized estimating equations (GEE) incorporate the ICC and sum all pairs of observations within a cluster when estimating the ICC. For the 2010 NIS, there are 92.6 billion pairs of observations, making summation of pairs computationally prohibitive. We propose a one-step GEE estimator that 1) matches the asymptotic efficiency of the fully-iterated GEE; 2) uses a simpler formula to estimate the ICC that avoids summing over all pairs; and 3) completely avoids matrix multiplications and inversions. These three features make the proposed estimator much less computationally intensive, especially with large cluster sizes. A unique contribution of this paper is that it expresses the GEE estimating equations incorporating the ICC as a simple sum of vectors and scalars.

Keywords

clustered data; efficient estimation; exchangeable correlation; fully-iterated; intra-cluster correlation

1 Introduction

Healthcare studies increasingly involve a large sample of independent clusters, where the cluster sizes are also large. The clusters are often households or patients from the same

hospital. When estimating the regression parameters of a generalized linear model for for clustered data with large cluster sizes, for reasons of computational feasibility, the most popular approach is to naively assume the observations within a cluster are independent to obtain consistent estimates (Liang and Zeger, 1986); a consistent estimate of the covariance matrix of these regression parameter estimates can be obtained using a 'sandwich estimator'. These estimates, obtained under naive independence, can be inefficient when the intracluster correlation (ICC) is relatively large; number of the loss of efficiency could be important for regression models with many covariates or interaction terms.

Gains in efficiency can be achieved via the use of generalized estimating equations (GEE) (Liang and Zeger, 1986), incorporating the ICC under an exchangeable correlation structure. However, for large clusters sizes, GEE typically sums over all pairs of outcomes within a cluster to estimate the ICC. Further, it requires inversion of matrices of the same dimensions as the cluster size, which is computationally demanding for large cluster sizes. In this paper, we propose an efficient and computationally feasible estimator of the regression parameters for GEE with an exchangeable correlation when there are a large number of clusters and large cluster sizes. This is achieved by constructing a one-step GEE estimator that 1) matches the asymptotic efficiency of the fully-iterated GEE while reducing the computational burden; 2) uses a simpler formula to estimate the ICC that avoids summing over all pairs; and 3) completely avoids matrix multiplications and inversions (both of which are computationally intensive). The key contribution is that we express the GEE as a simple sum of vectors and scalars. A SAS macro that implements the proposed one-step GEE estimator can be obtained from the authors. Our motivating example, from the U.S. 2010 Nationwide Inpatient Sample (NIS), encompasses over 8 million acute hospital stays from 1049 hospitals, with average cluster size of 7627 patients.

2 Generalized estimating equations

Suppose there are $i = 1, \ldots, N$ clusters and $j = 1, \ldots, n_i$ subjects within cluster i, with outcome variable Y_{ij} and a vector $\mathbf{x}_{ij} = (1, x_{ij1}, \ldots, x_{ijK})'$ of K covariates (including a constant for the intercept). Let $\mu_{ij} = E(Y_{ij}|\mathbf{x}_{ij}, \boldsymbol{\beta})$ denote the expectation of the outcome Y_{ij} given the covariates and regression coefficients $\boldsymbol{\beta}$ Here,

$$E(Y_{ij}|\mathbf{x}_{ij},\boldsymbol{\beta})=\mu_{ij}=\mu_{ij}(\boldsymbol{\beta})=g(\mathbf{x}_{ij}^{'}\boldsymbol{\beta}), \quad (2.1)$$

where $g(\cdot)$ is a known link function. The variance of Y_{ij} has general form

$$v_{ij}=Var(Y_{ij}|\mathbf{x}_{ij},\boldsymbol{\beta})=\phi v(\mu_{ij}), \quad (2.2)$$

where $v(\mu_{ij})$ can be any function of μ_i that is always positive and ϕ is a scale parameter.

We let $\mathbf{Y}_i = [Y_{i1}, \ldots, Y_{in_i}]'$ be an $n_i \times 1$ vector containing the outcomes for the n_i subjects in cluster i; $X_i = [\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}]'$ represents the $n_i \times K$ covariate matrix and $\boldsymbol{\mu}_i = [\mu_{i1}, \ldots, \mu_{in_i}]'$ is

an $n_i \times 1$ mean vector. We assume the correlation between any two observations in the same cluster is exchangeable, i.e., $\rho = Corr(Y_{ij}, Y_{ik}/X_i)$; ρ is often referred to as the ICC. The exchangeable correlation matrix of $\mathbf{Y}_i$ is

$$R_i = Corr(\mathbf{Y}_i) = \rho I_i + (1-\rho)\mathbf{J}_i\mathbf{J}_i'$$

where I_i is an $n_i \times n_i$ identity matrix and $\mathbf{J}_i$ is an $n_i \times 1$ vector of 1's. In this case, the covariance matrix of $\mathbf{Y}_i$ is $V_i = A_i^{1/2} R_i A_i^{1/2}$, where A_i is a diagonal matrix, with diagonal elements $v(\mu_{ij})\phi$, with $v(\mu_{ij})$ specified entirely by the marginal distributions, i.e., by $\boldsymbol{\beta}$.

To estimate $\boldsymbol{\beta}$, consider GEE (Liang and Zeger, 1986) of the form

$$\mathbf{u}(\hat{\boldsymbol{\beta}}) = \sum_{n=1}^{N}\sum_{j=1}^{n_i} \hat{D}_i' \hat{V}_i^{-1}[\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})] = \mathbf{0}, \tag{2.3}$$

where $D_i = \frac{d[\boldsymbol{\mu}_i(\boldsymbol{\beta})]'}{d\boldsymbol{\beta}}$, and V_i (described above) is a function of ρ which must be estimated; however, the scale parameter ϕ can be ignored when solving for $\hat{\boldsymbol{\beta}}$.

We first simplify (2.3) by simplifying V_i^{-1}. Note that $V_i^{-1} = A_i^{-1/2} R_i^{-1} A_i^{-1/2}$. To avoid matrix inversions in GEE, Qu et al. (2000) proposed using

$$R_i^{-1} = \frac{1}{(1-\rho)} I_i - \frac{\rho}{(1-\rho)[(1-\rho)+n_i\rho]} \mathbf{J}_i\mathbf{J}_i',$$

and thus

$$V_i^{-1} = \frac{1}{(1-\rho)} A_i^{-1} - \frac{\rho}{(1-\rho)[(1-\rho)+n_i\rho]} (A_i^{-1/2}\mathbf{J}_i)(A_i^{-1/2}\mathbf{J}_i)'.$$

However, in addition to avoiding inversion of large matrices, we show that by multiplying out the terms analytically, multiplication of large matrices can be completely by-passed using this expression for V_i^{-1}. Thus, as shown below, a unique contribution of this paper is that it shows how (2.3) can be expressed as a simple sum of vectors and scalars. That is, the estimating function in (2.3) becomes

$$\mathbf{u}(\boldsymbol{\beta}) = \frac{1}{(1-\rho)} \sum_{i=1}^{N} D_i' A_i^{-1}[\mathbf{Y}_i - \boldsymbol{\mu}_i] - \sum_{i=1}^{N} \frac{\rho}{(1-\rho)[(1-\rho)+n_i\rho]} D_i' (A_i^{-1/2}\mathbf{J}_i)(A_i^{-1/2}\mathbf{J}_i)' [\mathbf{Y}_i - \boldsymbol{\mu}_i].$$

Further, without loss of generality since we are setting this equal to 0 and solving for $\hat{\boldsymbol{\beta}}$, we can multiply the estimating function by $(1-\rho)$ to obtain

$$\mathbf{u}(\boldsymbol{\beta})=\sum_{i=1}^{N} D_i' A_i^{-1}[\boldsymbol{Y}_i-\boldsymbol{\mu}_i]-\sum_{i=1}^{N}\frac{\rho}{[(1-\rho)+n_i\rho]} D_i'(A_i^{-1/2}\mathbf{J}_i)(A_i^{-1/2}\mathbf{J}_i)'[\boldsymbol{Y}_i-\boldsymbol{\mu}_i]. \quad (2.4)$$

The second sum in the estimating function can be simplified further by noting that

$$(\mathbf{A}_i^{-1/2}\mathbf{J}_i)'[\boldsymbol{Y}_i-\boldsymbol{\mu}_i]=\sum_{j=1}^{n_i}(Y_{ij}-\mu_{ij})/\sqrt{v_{ij}} \quad \text{and} \quad D_i'(A_i^{-1/2}\mathbf{J}_i)=\sum_{j=1}^{n_i}\mathbf{d}_{ij}/\sqrt{v_{ij}}$$

where $\mathbf{d}_{ij}=\frac{d[\mu_{ij}(\boldsymbol{\beta})]}{d\boldsymbol{\beta}}$ is a given column of D_i. Then, (2.4) becomes

$$\mathbf{u}(\boldsymbol{\beta})=\sum_{i=1}^{N}\sum_{j=1}^{n_i}\mathbf{d}_{ij}(Y_{ij}-\mu_{ij})/v_{ij}-\sum_{i=1}^{N}\frac{\rho}{[(1-\rho)+n_i\rho]}\left[\sum_{j=1}^{n_i}\mathbf{d}_{ij}/\sqrt{v_{ij}}\right]\sum_{j=1}^{n_i}(Y_{ij}-\mu_{ij})/\sqrt{v_{ij}}. \quad (2.5)$$

Thus (2.3) has now been expressed as a simple sum of vectors and scalars. The first sum,

$$\mathbf{u}_I(\boldsymbol{\beta})=\sum_{i=1}^{N}\sum_{i=1}^{n_i}\mathbf{d}_{ij}(Y_{ij}-\mu_{ij})/v_{ij} \quad (2.6)$$

is the GEE under naive independence, and yields consistent, but possibly inefficient estimators of $\boldsymbol{\beta}$. Thus, the second sum in (2.5) is where efficiency is gained when incorporating the ICC. However, ρ must be estimated and plugged into (2.5) to realize potential efficiency gains. Any consistent estimator of ρ will yield the same asymptotic efficiency of the resulting estimator of $\boldsymbol{\beta}$. In the following section, we consider a very simple estimator that is computationally feasible with large clusters.

Using Taylor series expansions similar to Liang and Zeger (1986) and Prentice (1988), assuming that the regression for $\boldsymbol{\mu}_{ij}$ is correctly specified, the solution $\hat{\boldsymbol{\beta}}$ for $\mathbf{u}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ is consistent for $\boldsymbol{\beta}$; in addition, $N^{1/2}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ has an asymptotic distribution which is multivariate normal with mean vector $\mathbf{0}$. The asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}$ can be consistently estimated by the "sandwich estimator"

$$\left[\sum_{i=1}^{N}\hat{W}_i\right]^{-1}\left[\sum_{i=1}^{N}\mathbf{u}_i(\hat{\beta})\mathbf{u}_i(\hat{\beta})'\right]\left[\sum_{i=1}^{N}\hat{W}_i\right]^{-1}, \quad (2.7)$$

where $\mathbf{u}_i(\boldsymbol{\beta})=\sum_{j=1}^{n_i}\mathbf{d}_{ij}(Y_{ij}-\mu_{ij})/v_{ij}-\frac{\rho}{[(1-\rho)+n_i\rho]}\left[\sum_{j=1}^{n_i}\mathbf{d}_{ij}/\sqrt{v_{ij}}\right]\sum_{j=1}^{n_i}(Y_{ij}-\mu_{ij})/\sqrt{v_{ij}}$ is the sum of the score vectors from the subjects in cluster i and

$$W_i=W_i(\boldsymbol{\beta},\rho)=-E\left[\frac{d[\mathbf{u}_i(\boldsymbol{\beta})]'}{d\boldsymbol{\beta}}\right]$$
$$=\sum_{j=1}^{n_i}\mathbf{d}_{ij}\mathbf{d}_{ij}'/v_{ij}-\frac{\rho}{[(1-\rho)+n_i\rho]}\left[\sum_{j=1}^{n_i}\mathbf{d}_{ij}/\sqrt{v_{ij}}\right]\left[\sum_{j=1}^{n_i}\mathbf{d}_{ij}/\sqrt{v_{ij}}\right]'. \quad (2.8)$$

Also, W_i and $\mathbf{u}_i(\boldsymbol{\beta})$ are evaluated at $\hat{\boldsymbol{\beta}}$ and a consistent estimate of ρ.

3 Estimating the ICC

Denoting the true residual for the jth subject from the ith cluster by $e_{ij}=(Y_{ij}-\mu_{ij})/\sqrt{v_{ij}}$, then by definition, for $j\neq j'$, $E(e_{ij}e_{ij'})=\rho$. This suggests that a consistent method of moments estimator of ρ (Liang and Zeger, 1986) can be obtained via

$$\hat{\rho}=\left[\sum_{i=1}^{N}n_i(n_i-1)/2\right]^{-1}\sum_{i=1}^{N}\sum_{j<j'}\hat{e}_{ij}\hat{e}_{ij'} \quad (3.9)$$

where $\hat{e}_{ij}=(Y_{ij}-\hat{\mu}_{ij})/\sqrt{\hat{v}_{ij}}$. The estimator given by (3.9) requires the sum of $\sum_{i=1}^{N}n_i(n_i-1)/2$ pairs. However, the square of a summation can be written as (Parzen, 1960),

$$\left(\sum_{j=1}^{n_i}\hat{e}_{ij}\right)^2=\sum_{j=1}^{n_i}\hat{e}_{ij}^2+2\sum_{j<j'}\hat{e}_{ij}\hat{e}_{ij'}.$$

Thus,

$$2\sum_{j<j'}\hat{e}_{ij}\hat{e}_{ij'}=\left(\sum_{j=1}^{n_i}\hat{e}_{ij}\right)^2-\sum_{j=1}^{n_i}\hat{e}_{ij}^2$$

which requires the sum of $2n_i$ terms instead of $n_i(n_i-1)/2$ terms. Thus, we suggest using the following formulation to obtain an identical estimate to that given in (3.9),

$$\hat{\rho}=\left[\sum_{i=1}^{N} n_i(n_i-1)/2\right]^{-1}\sum_{i=1}^{N}\left[\left(\sum_{j=1}^{n_i}\hat{e}_{ij}\right)^2-\sum_{j=1}^{n_i}\hat{e}_{ij}^2\right]; \quad (3.10)$$

this expression has $2\sum_{i=1}^{N} n_i$ (2 times the total sample size) terms instead of $\sum_{i=1}^{N} n_i(n_i-1)/2$ terms. For the NIS data, this translates into approximately 16 million terms using our proposed approach instead of 92.6 billion terms using all pairs.

4 One-step estimator

Typically, one iterates between solving $\mathbf{u}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ in (2.5) for $\hat{\boldsymbol{\beta}}$ (given the current estimate of ρ) and estimating ρ with (3.10) (given the current estimate of $\boldsymbol{\beta}$) until convergence. If the estimate of $\boldsymbol{\beta}$ under naive independence, say $\hat{\boldsymbol{\beta}}_I$ is initially used to estimate ρ, and the resulting estimate of ρ is plugged back into (2.5) and $\mathbf{u}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ is solved for $\hat{\boldsymbol{\beta}}$, this yields an asymptotically equivalent estimator as the fully-iterated GEE estimator. The proof is similar to Lehmann (1983) for creating a one-step asymptotically efficient estimator from a consistent estimator (in this case $\hat{\boldsymbol{\beta}}_I$). A one-step GEE estimator is much less computationally intensive than a fully-iterated GEE for large cluster sizes.

In particular, a one-step estimator of $\boldsymbol{\beta}$ is formed by using one iteration of a Fisher scoring algorithm for obtaining a solution to $\mathbf{u}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$, with $\hat{\boldsymbol{\beta}}_I$ as the starting value, e.g.,

$$\hat{\boldsymbol{\beta}}_{1GEE}=\hat{\boldsymbol{\beta}}_I+[W(\hat{\boldsymbol{\beta}}_I,\hat{\rho})]^{-1}\mathbf{u}(\hat{\boldsymbol{\beta}}_I), \quad (4.11)$$

where $\hat{\boldsymbol{\beta}}_{1GEE}$ is the one-step GEE estimator and $\hat{\rho}$ is calculated by using $\hat{\boldsymbol{\beta}}_I$ to estimate $\hat{\mu}_{ij}$ and v_{ij} in $\hat{e}_{ij}$ in (3.10), and $W(\boldsymbol{\beta},\rho)=\sum_{i=1}^{N} W_i(\boldsymbol{\beta},\rho)$. Since $\mathbf{u}_I(\hat{\boldsymbol{\beta}}_I) = 0$, then $\mathbf{u}(\hat{\boldsymbol{\beta}}_I)$ in (4.11) reduces to

$$\begin{aligned}\mathbf{u}(\hat{\boldsymbol{\beta}}_I)&=\mathbf{u}_I(\hat{\boldsymbol{\beta}}_I)-\sum_{i=1}^{N}\frac{\hat{\rho}}{[(1-\hat{\rho})+n_i\hat{\rho}]}\left[\sum_{j=1}^{n_i}\hat{\mathbf{d}}_{ij}/\sqrt{\hat{v}_{ij}}\right]\sum_{j=1}^{n_i}(Y_{ij}-\hat{\mu}_{ij})/\sqrt{\hat{v}_{ij}}\\&=-\sum_{i=1}^{N}\frac{\hat{\rho}}{[(1-\hat{\rho})+n_i\hat{\rho}]}\left[\sum_{j=1}^{n_i}\hat{\mathbf{d}}_{ij}/\sqrt{\hat{v}_{ij}}\right]\sum_{j=1}^{n_i}(Y_{ij}-\hat{\mu}_{ij})/\sqrt{\hat{v}_{ij}},\end{aligned}$$

so that (4.11) simplifies to

$$\hat{\boldsymbol{\beta}}_{1GEE}=\hat{\boldsymbol{\beta}}_I-[W(\hat{\boldsymbol{\beta}}_I,\hat{\rho})]^{-1}\sum_{i=1}^{N}\frac{\hat{\rho}}{[(1-\hat{\rho})+n_i\hat{\rho}]}\left[\sum_{j=1}^{n_i}\hat{\mathbf{d}}_{ij}/\sqrt{\hat{v}_{ij}}\right]\sum_{j=1}^{n_i}(Y_{ij}-\hat{\mu}_{ij})/\sqrt{\hat{v}_{ij}}. \quad (4.12)$$

The asymptotic covariance of $\hat{\boldsymbol{\beta}}_{1GEE}$ is consistently estimated by (2.7) evaluated at $\hat{\boldsymbol{\beta}}_{1GEE}$:

5 Application to 2010 Nationwide Inpatient Sample

The binary outcome of interest is a patient complication within the first 48 hours post surgery (1 if complication, 0 if none). We fit a logistic regression model, where the main covariate was U.S. payer type, with 5 categories: Medicare, Medicaid, private insurance, self-pay, and uninsured/other. The other covariates were: race (1 if white, 0 otherwise), age (in years), advanced cancer stage (1 if yes, 0 if no), AIDS (1 if yes, 0 if no), renal failure (1 if yes, 0 if no) and number of other comorbidities.

Table 1 gives the GEE estimates of $\boldsymbol{\beta}$ under naive independence, and exchangeable (ICC) correlation (the proposed one-step and fully iterated). The estimated standard errors are from the sandwich variance estimator. An estimate of the asymptotic relative efficiency (ARE) can be obtained by comparing the variance estimates of $\hat{\boldsymbol{\beta}}$. From Table 1, we see that the estimated standard errors of $\hat{\boldsymbol{\beta}}$ under exchangeable correlation are much smaller than those under independence. Even with a relatively small estimated ICC, $\hat{\rho} = 0.0078$, the gains in efficiency appear appreciable. For example, for the payer effects, the estimated AREs range from 50% to 60%. The proposed one-step and fully iterated GEE give very similar results.

Next we compare estimators in terms of computation times (real not CPU). Standard logistic regression maximum likelihood estimation under naive independence without a robust variance in SAS PROC GENMOD (SAS Institute, 2015) is the fastest (1.1 minutes); logistic regression under naive independence but with a sandwich variance estimate (PROC GENMOD) takes 2.5 minutes. Thus, calculation of the sandwich variance estimate for standard logistic regression takes an additional 1.4 minutes. Our SAS macro for the proposed one-step approach takes 2.1 minutes, as opposed to the fully-iterated estimation (in PROC GENMOD) with exchangeable correlation, which takes 8.1 hours. We note that use of only a single iteration of PROC GENMOD with exchangeable correlation should be comparable to our one-step approach because PROC GENMOD uses the logistic regression estimates under naive independence as starting values. Thus, direct comparison of our one-step approach (2.1 minutes) to one-step of PROC GENMOD (1.6 hours) emphasizes the potential advantage of the proposed method. We attribute the advantage in computation time to the following: 1) we have expressed the GEE in a form that does not require inversion or multiplication of any matrices; 2) we use a simple formula to estimate the ICC that does not require summing over all pairs of outcomes within a cluster; and 3) we only use a one-step GEE estimator instead of fully-iterating. A unique contribution of this paper is that it expresses the GEE estimating equations with an exchangeable correlation given by (2.3) as a simple sum of vectors and scalars.

We note that we fit the models on a 64-bit PC workstation with an Intel Xeon CPU E5-2630 0 @ 2.30GHz processor with 16.0 GB of RAM and an SSD hard drive; this PC workstation is faster than a typical desktop PC. Finally, although the results presented here were obtained using SAS, we also attempted to obtain the fully-iterated GEE estimates using both gee in R (Carey, 2002) and the xtgee command in Stata (StataCorp, 2015); both program were unable to produce estimates of the model parameters due to insufficient memory.

Finally, we note that clustering can also arise in longitudinal studies with repeated measures on the same subjects. For such studies an exchangeable correlation structure is often not appropriate; for example, Toeplitz, autoregressive, *m*-dependent, or even unstructured correlation patterns may be preferred. However, typical cluster sizes for most longitudinal studies tend to be relatively small, say less than 15–20, so that the computationally feasible methods similar to those discussed here are not required.

Acknowledgments

We are grateful for the support provided by grant CA60679 from the U.S. National Institutes of Health.

References

Carey, V. gee: Generalized Estimation Equation Solver. R Package Version 4.13-10; Ported from S-PLUS to R by Thomas Lumley (versions 3.13 and 4.4) and Brian Ripley (version 4.13). 2002.

Lehmann, E. Theory of Point Estimation. John Wiley & Sons; 1983.

Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika. 1986; 73:13–22.

Parzen, E. Modern probability theory and its applications. John Wiley & Sons; 1960.

Prentice RL. Correlated binary regression with covariates specific to each binary observation. Biometrics. 1988; 44:1033–1048. [PubMed: 3233244]

Qu A, Lindsay B, Li B. Improving generalised estimating equations using quadratic inference functions. Biometrika. 2000; 87:823–836.

SAS Institute. SAS/STAT Software, Version 9.4. Cary, NC: 2015.

StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP; 2015.

Table 1

Comparison of logistic regression parameter estimates for the post-operative surgical complications data

Effect	Approach	Estimate	*SE*	*Z*-statistic	P-value
Intercept	IND-Robust	−2.8382	0.0426	−66.56	<.0001
	Proposed 1-step	−2.5203	0.0290	−86.94	<.0001
	Fully Iterated GEE	−2.4005	0.0356	−67.35	<.0001
Medicare	IND-Robust	−0.0513	0.0236	−2.18	0.0296
	Proposed 1-step	−0.0493	0.0183	−2.69	0.0071
	Fully Iterated GEE	−0.0428	0.0175	−2.45	0.0143
Medicaid	IND-Robust	−0.0307	0.0254	−1.21	0.2257
	Proposed 1-step	−0.0271	0.0193	−1.41	0.1593
	Fully Iterated GEE	−0.0274	0.0185	−1.48	0.1384
Private	IND-Robust	−0.0936	0.0217	−4.32	<.0001
	Proposed 1-step	−0.0741	0.0158	−4.68	<.0001
	Fully Iterated GEE	−0.0736	0.0152	−4.84	<.0001
Self-pay	IND-Robust	0.1848	0.0277	6.67	<.0001
	Proposed 1-step	0.1571	0.0219	7.17	<.0001
	Fully Iterated GEE	0.1496	0.0211	7.08	<.0001
White	IND-Robust	−0.0106	0.0136	−0.77	0.4389
	Proposed 1-step	−0.0233	0.0074	−3.16	0.0016
	Fully Iterated GEE	−0.0244	0.0072	−3.39	0.0007
# Comorbidities	IND-Robust	0.2897	0.0030	96.80	<.0001
	Proposed 1-step	0.2800	0.0027	105.04	<.0001
	Fully Iterated GEE	0.2767	0.0027	102.64	<.0001
Cancer	IND-Robust	0.5291	0.0093	56.98	<.0001
	Proposed 1-step	0.4972	0.0098	50.63	<.0001
	Fully Iterated GEE	0.4935	0.0098	50.42	<.0001
AIDS	IND-Robust	0.1696	0.0396	4.28	<.0001
	Proposed 1-step	0.1790	0.0307	5.83	<.0001
	Fully Iterated GEE	0.1687	0.0298	5.67	<.0001
Renal Fail	IND-Robust	0.8258	0.0097	85.20	<.0001

Effect	Approach	Estimate	*SE*	Z-statistic	P-value
	Proposed 1-step	0.8333	0.0100	83.49	<.0001
	Fully Iterated GEE	0.8383	0.0103	81.03	<.0001
Age	IND-Robust	0.0248	0.0006	40.19	<.0001
	Proposed 1-step	0.0239	0.0004	63.39	<.0001
	Fully Iterated GEE	0.0232	0.0003	66.38	<.0001