



Published in final edited form as:

Stat Biopharm Res. 2009 May 1; 1(2): 201–212. doi:10.1198/sbr.2009.0026.

Analysis of Binary Adherence Data in the Setting of Polypharmacy: A Comparison of Different Approaches

Denise A. Esserman, PhD¹, Charity G. Moore, PhD, MSPH², and Mary T. Roth, PharmD, MHS, FCCP³

¹Department of Medicine, Division of General Medicine and Epidemiology, University of North Carolina at Chapel Hill

²Department of Medicine, Division of General Internal Medicine, University of Pittsburgh

³School of Pharmacy, Division of Pharmaceutical Outcomes and Policy, University of North Carolina at Chapel Hill

Abstract

Older community dwelling adults often take multiple medications for numerous chronic diseases. Non-adherence to these medications can have a large public health impact. Therefore, the measurement and modeling of medication adherence in the setting of polypharmacy is an important area of research. We apply a variety of different modeling techniques (standard linear regression; weighted linear regression; adjusted linear regression; naïve logistic regression; beta-binomial (BB) regression; generalized estimating equations (GEE)) to binary medication adherence data from a study in a North Carolina based population of older adults, where each medication an individual was taking was classified as adherent or non-adherent. In addition, through simulation we compare these different methods based on Type I error rates, bias, power, empirical 95% coverage, and goodness of fit. We find that estimation and inference using GEE is robust to a wide variety of scenarios and we recommend using this in the setting of polypharmacy when adherence is dichotomously measured for multiple medications per person.

Keywords

Beta-binomial; Correlated binomial data; Generalized estimating equations; Polypharmacy

1. INTRODUCTION

Medication regimen adherence has been defined as the “extent to which patients take medications as prescribed by their health care providers” (Osterberg and Blaschke 2005). Adherence to a *single* prescribed medication is often measured as the percentage of the medication taken by the patient over some period of time (Osterberg and Blaschke 2005). However, adults aged 50 and over often take *multiple* medications for numerous chronic diseases and the likelihood that they will be prescribed *multiple* medications significantly increases with age (Murray et al. 2004; Vik, Maxwell and Hogan 2004). Measuring,

Corresponding Author: Denise Esserman, PhD 5034B Old Clinic Building Department of Medicine Division of General Medicine and Epidemiology University of North Carolina School of Medicine Chapel Hill, NC 27599 Phone: 919-843-2887 Fax: 919-843-4031 esserman@med.unc.edu Charity G. Moore, PhD, MSPH Center for Research on Health Care Data Center 200 Meyran Ave., Suite 300 University of Pittsburgh Pittsburgh, PA 15213 Phone: 412-246-6961 moorecg@upmc.edu Mary Roth, PharmD, MHS, FCCP Pharmaceutical Outcomes and Policy Kerr Hall, School of Pharmacy University of North Carolina at Chapel Hill Chapel Hill, NC 27599 Phone: 919-843-8083 mroth@unc.edu.

modeling and determining the factors which predict adherence in this older adult population is an extremely important area of research since proportions of hospitalizations as high as 11% have been attributed to medication non-adherence (Vik et al. 2004).

Numerous methods have been used to measure adherence in the older community-dwelling population, including but not limited to biological assays, pill counts, electronic monitoring, pharmacy records, prescription claims, third-party assessment, and self-report; however, to date, there has been no “gold” standard (Vik et al. 2004). For a discussion of the different methods used to measure adherences along with the pros and cons of these methods see Vik et al. (2004). A recent community-based study in North Carolina assessing the quality of medication use among older adults used a clinical pharmacist to evaluate one aspect of quality, adherence (i.e. Adherent, non-Adherent), for each medication a person was currently taking (NIH Grant 5K23AG024229). In meeting with each older individual, the pharmacist had the individual explain how he or she used each medication and asked the individual a series of questions to determine whether the individual was likely adhering to the medication as prescribed. Taking all the information gathered at the interview, the pharmacist arrived at an assessment of adherence (i.e. Adherent or non-Adherent) for each medication the individual was taking. This resulted in *multiple* dichotomous responses for each older adult with the total number of responses (medications) varying across individuals. Just as with this method, oftentimes the estimates of adherence are reported as binary variables (Vik et al. 2004).

Previous researchers (Lee, Grace and Taylor 2006) working in the setting of polypharmacy have summed these binary variables and defined adherence as the proportion of adherent medications out of the total medications prescribed per person, and analyzed the data using linear regression models to assess person-level characteristics (Weisberg 1985). However, due to the varying number of medications taken per individual, and the inability of linear models to guarantee a predicted value between zero and one (Weisberg 1985; Fleiss, Levin and Paik 2003), we hypothesize that the data may more appropriately be analyzed using logistic regression (Hosmer and Lemeshow 1989) with an extension to models which adjust for clustering within individuals, such as generalized estimating equations (GEE) (Liang and Zeger 1986) and beta-binomial (BB) regression (Williams 1975; Prentice 1986). These models account for the original binary nature of the outcome at the medication level (Adherent vs. Non Adherent) and the intra-individual correlation (adherence statuses of medications taken by the same individual are related). Ignoring the within-person correlation among adherence responses could lead to invalid inferences about the rates of adherence (Diggle, Heagerty, Liang and Zeger 2002; Hu, Goldberg, Hedeker, Flay and Pentz 1998; Fitzmaurice, Laird and Rotnitsky 1993; Fleiss et al. 2003).

In Section 2 we discuss six methods (naïve logistic regression; GEE; BB regression; standard, unadjusted, and weighted least squares linear regression) for analyzing binary adherence data in the setting of polypharmacy, focusing on methods that are readily available to implement in standard statistical software (contact the corresponding author for code pertaining to the analysis presented in this article). In Section 3, we compare the six methods using a sub-sample of data collected on 200 community-dwelling older adults, by investigating how well the methods performed when comparing rates of adherence between African American (AA) older adults and their white peers. We present simulation results in Section 4, where we studied the influence of sample size, intra-individual correlation among multiple medications, and the average number of total medications per person on Type I error rates, the bias of the parameter estimates, empirical 95% coverage, and power for the six different methods. Since we were concerned about the appropriateness of linear regression in this setting, we conducted further simulations to examine some aspects of the goodness of fit of this model. We provide a discussion in Section 5 and make

recommendations about the most appropriate method for analyzing this type of data in Section 6.

2. METHODS FOR ANALYZING ADHERENCE FOR MULTIPLE MEDICATIONS

Let Y_{ij} , $i=1, \dots, N$ and $j=1, \dots, n_i$, be binary adherence responses (i.e. Adherent ($Y_{ij}=1$); non-Adherent ($Y_{ij}=0$)), where n_i is the number of current medications for the i^{th} individual and N is the total number of individuals in the cohort. For all analyses proposed, we assume that the covariates of interest (\mathbf{X}) are observed at the individual level (demographics, health status, disease, etc), not at the medication level (type of medication, medication indication, medication dose, etc.). For this study, each medication contributes equally towards the adherence rate estimation. For example, a person's adherence to daily aspirin for MI prophylaxis would count the same in the analysis as adherence to insulin for diabetes. This assumption may not be valid in studies where adherence to certain drugs is more important (over the counter vs. not, by disease, etc.). In these scenarios, more "weight" could be placed on certain medications when calculating individual adherence rates. Exploring methodologies in these scenarios is beyond the scope of this article.

Standard naïve logistic regression (Hosmer and Lemeshow 1989) can be used to analyze the data if we assume an individual's adherence statuses for multiple medications, Y_{ij} 's, are independent, identically distributed Bernoulli random variables with $P[Y_{ij}=1|\mathbf{X}_i] = p_{ij}$, the rate of adherence for the i^{th} individual, and that being adherent to one medication does not affect the adherence status of another medication. If we assume that the probability of being adherent to a medication is the same for each medication an individual is currently taking ($P[Y_{ij}=1|\mathbf{X}_i] = p_i$), this probability can be parameterized by the logistic response function

$$p_i = \frac{\exp(\mathbf{X}_i' \boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i' \boldsymbol{\beta})}, \quad (1)$$

where $\boldsymbol{\beta}$ is a vector containing the regression coefficients for the covariates of interests, \mathbf{X}_i . Also, under these assumptions the total number of medications in which the person is

adherent, $S_i = \sum_{j=1}^{n_i} Y_{ij}$, follows a binomial distribution with mean $n_i p_i$ and variance $n_i p_i (1 - p_i)$. Thus, when covariates are considered at the individual level as opposed to the medication level, this is equivalent to modeling the proportion adherent for a given individual (S_i/n_i : Number of adherent medications/total number of medications for the i^{th} individual) via a logistic regression model. However, when adherence statuses for multiple medications are taken from the same individual, it may be naïve to assume these repeated observations are independent. Although the estimates of the regression coefficients and rates of adherence will be unbiased if the dependence (intra-individual correlation) is ignored, the variability of the coefficients will be underestimated if the correlation is positive (as was observed in the example data set presented in Section 3) leading to improper inferences about the importance of factors (i.e. demographics) used in the model to predict adherence (Diggle et al. 2002; Hu et al. 1998; Fitzmaurice et al. 1993; Fleiss et al. 2003; Stokes, Davis and Koch 2000).

We consider two extensions of the standard logistic regression, which can model the correlation that exists among the repeated measures within individuals: beta binomial (BB) regression and generalized estimating equations (GEE). The BB regression model extends

the naive logistic approach by modeling the correlation through an additional parameter which accounts for positive correlation among the multiple medications for a given individual. This model can only accommodate correlation structures for which all responses for a given individual are assumed to be equally correlated (i.e. the exchangeable structure; Neuhaus 1992). With the BB model, rates of adherence across individuals (p_i), are assumed to be randomly distributed from a beta distribution, $B(a, b)$, with $a = p_i/\lambda$ and $b = (1-p_i)/\lambda$, where p_i is defined as in equation (1) and λ is a dispersion parameter such that the within-person correlation coefficient, ρ , is defined by $\rho = \lambda/(1+\lambda)$ ($\rho \geq 0$). Thus, S_i , the total number of adherent medications for individual i , follows a BB distribution (Skellam 1948; Johnson, Kemp and Kotz 2005) with mean $n_i p_i$ and variance

$$n_i p_i (1 - p_i) \left\{ 1 + (n_i - 1) \frac{\lambda}{1 + \lambda} \right\}.$$

The mean is identical to that of the standard naïve logistic regression, but the variance has a multiplier, $1 + (n_i - 1) (\lambda / (1 + \lambda))$, which models the overdispersion due to positive intra-individual correlation (Johnson et al. 2005). While it is possible to observe negative correlations in this setting, based on our example data set, we do not anticipate negative correlations, and thus do not consider it further. However, the correlated-binomial model proposed by Kupper and Haseman (1978) is capable of handling negative correlations.

While BB regression is a fully parameterized method of accounting for the intra-individual correlation, the semi-parametric GEE method incorporates the dependence by robustly estimating the variance. The “working” or approximate covariance matrix for $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{i n_i})$ is given by

$$\mathbf{V}_i = \mathbf{A}_i^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_i^{1/2},$$

where \mathbf{A}_i is a diagonal matrix of the marginal variance functions $\text{var}[Y_{ij}]$ (equal to $p_i(1-p_i)$ in the binomial case where p_i is defined as in equation (1)), and $\mathbf{R}(\boldsymbol{\alpha})$ is the “working” correlation matrix (given by the investigator). Adherence statuses for the same individual are assumed to be correlated while observations for different individuals are assumed to be independent. The parameter $\boldsymbol{\alpha}$ is a vector of the correlation parameters for a given specification of the correlation matrix. A number of structures can be specified for the correlation matrix, including but not limited to independent, exchangeable, unstructured and autoregressive. If $\mathbf{R}(\boldsymbol{\alpha})$ is correctly specified and represents the true correlation matrix for the \mathbf{Y}'_i , then \mathbf{V}_i will be the true covariance of \mathbf{Y}_i instead of just the “working” or approximate covariance matrix (Liang and Zeger 1986). However, it is important to note that the variance estimates and the regression parameters, $\widehat{\boldsymbol{\beta}}$, will be consistent even if the “working” correlation matrix, and thus the “working” covariance matrix, is misspecified, as long as the model for the mean $\mu_i(\boldsymbol{\beta}) = E[Y_{ij}]$ is correctly specified (Liang and Zeger 1986; Fitzmaurice et al. 1993; Hu et al. 1998). This is made possible through the use of a sandwich estimator of the covariance; the empirical estimates of the covariance (obtained from making use of the observed correlations) are “sandwiched” between the model-based covariance estimates (obtained assuming working correlation assumptions are correct) to give a robust estimate of the covariance (Liang and Zeger 1986; Kleinbaum and Klein 2002; Dunlop 1994; Fitzmaurice et al. 1993).

While GEE can be used for a wide variety of outcome distributions, for this paper, we are solely concerned with binary responses, since adherence status is often measured as a binary

variable (or reduced to one). We use the logit link, as with logistic regression, where the $p_i = \mu_i(\beta) = E[Y_{ij}]$ is defined as in Equation (1). For this type of data, where the covariates of interest are at the individual level and not the medication level, the exchangeable matrix is most appropriate and will be the “working” correlation matrix for all data analyzed using a GEE model below. This implies that observations for a given individual share the same correlation and this correlation is common for all individuals in the analysis. Thus, α consists of only one parameter, the intra-individual correlation, ρ . An autoregressive correlation matrix was not appropriate in this setting since there was no time structure or distance between measurements. An unstructured correlation matrix was not considered because the type and the number of medications vary across individuals.

Each of the three models discussed above uses a logistic link function to model the probability of adherence, which treats the outcome variable as categorical, and thus the coefficients in the models can be used to estimate odds ratios (OR). However, if we were to define adherence for an individual as a continuous random variable (number of medications in which the person is adherent divided by the total number of medications in which the person is taking, $p_i^* = S_i/n_i$), as has been done previously (Lee et al. 2006), we could use standard linear regression to model adherence as a function of person level characteristics (i.e. demographics, co-morbidities, etc.), $p_i^* = \mathbf{X}_i' \boldsymbol{\beta}^*$. We could then compare subpopulations using the estimated difference in adherence rates (DAR). For example, holding all other variables constant, we could calculate the difference in adherence for two groups as

$\text{DAR} = \beta_1^* = \bar{p}_{x=1} - \bar{p}_{x=0}$. The assumptions of this model are that the p_i^* are independently and identically distributed normal random variables with mean $\mathbf{X}'\boldsymbol{\beta}^*$ and constant variance σ^2 . In using this model, the user falsely assumes the variability will be constant, since values of p_i^* closer to zero or one would naturally be less variable.

In the standard linear model (referred to as linear), the number of medications an individual is currently taking is not taken into account. Thus, an individual adherent to two out of three medications would contribute the same amount of information to the model estimation as an individual adherent to eight out of twelve medications. Therefore, in addition to the standard linear model we considered two variations: an adjusted linear model, in which we control for the total number of medications a person is taking (referred to as adjusted); and a weighted linear model, where the weight is the number of medications taken per individual (referred to as weighted). These latter two models are only appropriate if the number of medications differs across individuals and are considered here because intuitively, these are possible models investigators may use in order to control for the number of medications an individual is taking. Lee *et al.* (2006) adjusted for the number of medications when using linear models to model the proportion of adherent medications.

In this article, we do not consider Poisson and negative binomial regression (Cameron and Trivedi 1998), as these are not valid models for this type of data in which there are a finite number of successes (adherent medications) bounded by a finite number of trials (total number of medications per person). Using these models violates the assumption of the number of trials being “essentially” infinite and the number of successes being allowed to be indefinitely large. In addition, as was noted by Kupper and Haseman (1978), the Poisson model, and thus by extension the negative binomial model, does not account for the number of medications, nor is the assumption of the probability of adherence following a gamma distribution theoretically justified.

Of note, the models discussed in this paper are population-averaged approaches in which the focus is on making inferences about group differences. Random intercept models (Singer and Willet 2003) constitute a subject-specific approach in which inferences about individual differences are of primary interest. Discussion of these models is beyond the scope of this

article. For further comparisons of population-average versus subject-specific approaches see Hu et al. (1998).

3. EXAMPLE

For our example, we used data from an ongoing study assessing the quality of medication use among community-dwelling older adults (NIH Grant 5K23AG024229). Participants met the following inclusion criteria: (a) age ≥ 60 years, (b) residing independently in the community setting; and (c) taking ≥ 3 regularly scheduled medications. Patients were excluded if they had cognitive impairment (made ≥ 3 errors on a cognitive screening instrument). Baseline information was obtained on 200 older adults (100 White; 100 AA) during home interviews by a trained clinical pharmacist. Information was collected on demographics, medication history and current use (prescription; over the counter; dietary; alternative or complementary medicine), drug therapy concerns, functional status, health literacy, and quality medication use.

During the baseline home interview, the clinical pharmacist conducted a comprehensive medication review, which included the assessment of adherence status for *each* medication an individual was taking, including prescription, over-the-counter, and herbal therapies. The pharmacist classified the individual as “Adherent” or “non-Adherent” for each particular medication using the information provided by the patient along with her clinical judgment. For example, if an individual was taking an opioid medication used twice daily as needed for pain, and the individual had not experienced any pain over the past week and therefore had not required the use of the medication, then the patient was considered “adherent” with this medication. Following the interview with the patient, the pharmacist did have access to the patient’s medical record and used this information as well in determining medication adherence. At this time, this method of measuring adherence has not been validated; however, work is currently being done to assess the validity and reproducibility of this measure. Although this method of measuring adherence has not been validated, these methods could be used with other measures of adherence in which the outcome is binary at the medication level (i.e., Adherent if $> 80\%$ pills taken).

The study is ongoing with follow-up data collections planned for 6 and 12 months. We only used the baseline information for our methods demonstration. For all participants ($N=200$), the average (standard deviation (SD)) number of current medications per person (n_i) was 10.68 (4.61) and ranged from 3 to 27; the mean age was 77 years (range 60-96); 77% were female; and the average (SD) proportion adherent was 84.0% (21.2%). Since the results of this study have not yet been published, we focused on the difference between white and AA community-dwelling older adults and took a random sample of 100 individuals, 50 white and 50 AA, for demonstration purposes. The average number (SD) of current medications was 11.28 (4.51) and 9.88 (4.43) and the average (SD) proportion adherent was 84.9% (19.9%) and 80.0% (24.8%), for whites and AA’s, respectively, in the sub-sample.

The linear, adjusted, and weighted models estimated that older white adults have approximately a 5-7% higher rate of adherence than older AA adults as indicated by the DAR ranging from 0.05 ($0.05*100\% = 5\%$ difference) to 0.07 in the three models. However, the confidence interval (CI) contains zero (for all three models) and thus, we are unable to conclude that the rate is higher in the white group compared to the AA group (see Table 1). Using the results of the naïve logistic regression, we would conclude older white adults have a greater odds of being adherent compared to older AA adults (Table 1; OR=1.57, 95% CI 1.14-2.16). According to the results of the BB regression and the GEE analysis with exchangeable correlation structure, we conclude there is no significant difference in adherence between older white and older AA adults. Note, the parameter estimates are

similar for the three models but the standard error (SE) estimate for the logistic model is 40-50% smaller than the SE from the GEE (obtained by using the robust sandwich covariance estimator) and BB regressions.

4. SIMULATIONS

4.1 Type I error, Bias, Empirical 95% Coverage and Power

In order to compare the models discussed in Section 2, we conducted a simulation study modeled after the example dataset, in which we examined the effects of cohort size ($N=100$; $N=200$), intra-individual correlation (ρ) in adherence among medications taken by the same individual, and the total number of medications per person (n_i). Data were generated under two different methods: the BB distribution (fully parametric model) and the shared response model (Lunn and Davies 1998; Pang and Kuk 2005). We expected the BB model to perform well when data were generated from a BB distribution, but we also wanted to explore how robust this model was to more general correlated binomial data. Intra-individual correlations were varied from 0 to 0.5 for both a fixed number ($n=10$) and varying number ($n_i \sim$ negative binomial with $\mu=10$ and $\theta=10$, for $\sigma^2 = \mu + \mu^2/\theta$) of total medications per individual. When the randomly generated n_i resulted in a zero value (no medications for a given individual), the 0 was replaced with a regenerated value of n_i ; this was an extremely rare event, occurring approximately 0.1% of the time. A dummy group covariate ($X=0$ or $X=1$) was generated with each level containing an equal number of individuals (50 per group for $N=100$; 100 per group for $N=200$). Ten-thousand datasets were generated for analysis.

The evaluation of the simulations was based on Type I error rates, bias, power and empirical 95% coverage. The parameter of interest for the linear models is the DAR: $\beta_j^* = p_i^*|X_i=1 - p_i^*|X_i=0$; the parameter of interest for the naïve logistic, GEE, and BB models is the natural log of the odds ratio ((lnOR); β_j). For the BB and GEE models we were also interested in the bias in estimating the intra-individual correlation. The adherence probability was modeled using the logistic response function defined by Equation (1) with the probability of being adherent in the reference group set at $p=0.731$ ($\beta_0=1$). To explore Type I error rates, we set the difference between groups to be zero ($\beta^*_j=\beta_j=0$; OR=1). It is important to note that for all models, we are interested in being able to unbiasedly estimate the difference between the two groups with a Type I error rate close to the expected Type I error rate (nominal 0.05 level). The robust covariance estimator with an exchangeable correlation structure was used to estimate all GEE standard errors.

Figures 1(a) and 1(b) present Type I error rates (the nominal error rate is $\alpha=0.05$) when the number of total medications is the same across individuals ($n_i=n=10$) for data generated under the BB distribution and the shared response model. Adjusted and weighted linear regression models were not applied to this scenario, since the number of medications is constant across individuals. Our results show that the linear, BB, and GEE models have Type I error rates close to the nominal value of 0.05 regardless of the amount of intra-individual correlation. In contrast, the logistic model has highly inflated Type I error rates; even for the smallest amount of intra-individual correlation ($\rho=0.05$), the Type I error rate is already two times larger than expected. This rate only increases as the intra-individual correlation increases. Regardless of N and the data generation method, the pattern of the Type I error rate is approximately the same for all four models presented (linear; BB; GEE; logistic). The biases of the parameter estimates for these models are negligible ranging from -0.6% to 0.6% (data not shown) regardless of the method of data generation with the range decreasing as the cohort size increases.

The results of the simulations with varied number of medications across individuals (Figures 2(a) and 2(b)) showed the same patterns as those when the number of medications was the

same for each person ($n=10$). The biases of the parameter estimates for the six models presented (linear; adjusted; weighted; logistics; GEE; BB) are negligible ranging from -0.6% to 0.8% (data not shown) regardless of the method of data generation with the range decreasing as the cohort size increases. However, the weighted model has a slightly inflated Type I error rate compared to the linear and adjusted models, as well as the BB and GEE models, and on average, its 95% confidence interval does not cover 0.05 (data not shown). The generation method and the cohort size do not seem to impact the Type I error rates.

As can be seen in Figure 3(a) for fixed n , for data generated under a BB distribution, the correlation is estimated well by both the GEE and BB models regardless of cohort size. However, when the data are generated under the shared response model (Figure 3(b)), the BB model tends to underestimate the intra-individual correlation with the bias increasing as the true correlation increases. The same pattern can be seen when the total number of medications varied across individuals (Figures 3(c) and 3(d)).

To explore power, data were generated under the methods described above but with β_1 values ranging between 0 and 1.5. This range in β_1 values sets the probability of adherence in the nonreference group to range from 73.1% to 92.4% (OR to range from 1 to 4.48; β_1^* to range from 0 to 0.193). As demonstrated in Figure 4, for data generated under a shared response model with random n_i , the power curves are similar for all models. The logistic model is not presented here since the power will be inflated due to the inflated Type I error rates we observed (Figures 1 and 2). The BB model tends to have slightly larger power than the linear, adjusted, weighted, and GEE models, and this difference increases as the intra-individual correlation increases. For all models, as the cohort size (N) increases, the power increases; and as the intra-individual correlation increases, the power decreases, reflecting the loss of information for a fixed sample size when observations within individuals are highly correlated. Similar results were observed for data generated under a BB distribution with random n_i and both the shared response model and BB distribution for fixed n (results not shown).

The empirical 95% coverage is presented in Table 2 for the same scenarios presented for the power study (β_1 values ranging between 0 and 1.5; data generated under shared response model with random n_i ; $\rho=0.1, 0.3, 0.5$; $N=100, 200$). The linear, adjusted, GEE, and BB models all have coverage close to the nominal 0.95 value for varying values of β_1 and ρ . The empirical coverage for the weighted linear model tends to be slightly lower than the nominal 0.95 value, while the coverage for the naïve logistic model is grossly lower than the nominal 0.95 value and is strongly influenced by the strength of the correlation.

4.2 Goodness of Fit of Linear Model

In addition to exploring the bias and accuracy of the linear model, we also explored goodness of fit using two criteria: (1) the average proportion of times the linear model predicts values of the probability of being adherent outside of the range of a legitimate probability (0-1); and (2) how often the GEE predicted values closer to the “true” probability of being adherent compared to the linear model. (Note: We chose to focus on the GEE here since it has the same expected mean as the naïve logistic and BB models, but performed as well or better than these two models during the simulations presented in Section 4.1.)

Data were generated under the shared response model with fixed sample size ($n=10$), fixed cohort size ($N=200$), varied intra-individual correlation ($\rho = 0.1, 0.3, 0.5$), and p_i ranging from 0.01 to 0.99. Ten-thousand data sets were generated. Over these data sets, the average probability of being adherent was 0.73 with a range from 0.01 to 0.99; on average, the 10th percentile of individual adherence was 0.28, the 25th percentile was 0.59, the median was

0.84, the third quartile was 0.94, and the 90th percentile was 0.97. For all three intra-individual correlations, approximately 15% of the values predicted by the linear model were outside of the 0-1 range. On average 85%, 79% and 75% of the predicted adherence probabilities were closer (as measured by absolute distance) to the “true” adherence probabilities for the GEE compared to the linear model for $\rho = 0.1, 0.3$ and 0.5 , respectively.

5. DISCUSSION

The results from the example dataset demonstrated that the naïve logistic regression would have led us to declare higher adherence among whites compared to AA older adults while results from the other models would not have led to this conclusion. The results of the simulations demonstrate that even when the smallest amount of correlation is present among adherence statuses of multiple medications taken by the same individual, the naïve logistic model has an inflated Type I error rate. This rate of inflation increased with increased intra-individual correlation. Although the estimates for the rates of adherence using this model will be unbiased, the standard errors are severely underestimated, leading to this inflated Type I error rate and incorrect inference (Diggle et al. 2002; Hu et al. 1998; Fitzmaurice et al. 1993; Stokes et al. 2002). Thus, in any dataset in which the assumption of independent and identically distributed binary responses could be violated (i.e. repeated measures from the same individual), naïve logistic regression should not be the method of choice for the analysis.

All of the linear models appeared to perform fairly well, although the weighted linear model had a slightly inflated Type I error rate (on average, the 95% confidence interval did not cover 0.05) and a slightly lower 95% empirical coverage; the adjusted and standard linear models had Type I error rates close to the nominal value of 0.05 across increasing intra-individual correlation. The estimates of the DAR had negligible to no bias. In addition, the power curves of these models were similar to that of the GEE and BB models and the 95% empirical coverage was close to the nominal value of 0.95. Due to these results, researchers may be tempted to use these models to analyze binary adherence (or any data presented as a percentage, i.e. percentage of medication taken), especially since they are easy to implement and interpret; however, these models, in theory, are not appropriate for this type of data. First, the outcome, adherence rate, is a probability with a restricted range of zero to one. When using linear models, there is no restriction placed on the probability of adherence such that predicted probabilities and their corresponding confidence intervals will fall into this range (Weisberg 1985; Fleiss et al. 2003). As was shown in our simulations, approximately 15% of the predicted values fell outside of the zero to one range. In using linear regression to analyze proportion data for four example datasets, Zhao, Chen and Schaffner (2001) found that between 21% and 32% of the values were predicted outside of the zero to one range. They also observed that model predictions are especially poor when the observed values are close to zero and one. Second, the model assumes that the rates of adherence across individuals are normally distributed with constant variance. The distribution of the probability of adherence will be close to the normal distribution if the probability lies between 0.1 and 0.9 (Fleiss et al. 2003); however, in Vik et al.’s (2004) review, they reported estimates of individual medication adherence ranging between 43.7% and 100%. In our example, we saw individual adherence rates ranging between 0 and 100%. Thus, there is no guarantee that adherence probabilities will remain between the “normal” range. In addition, the variance of the probability of adherence ($p_i(1-p_i)$) is dependent on the covariates (\mathbf{X}) since the p_i is dependent on \mathbf{X}_i (see Equation (1)), and thus the assumption of constant variance is violated (Zhao et al. 2001). Third, more than 75% of the time, compared to the GEE, the linear model predicted probabilities farther (in absolute distance) from the “true” value.

With the standard linear model, we assume that each individual makes the same contribution to the model regardless of the number of medications being taken (i.e., a person who is adherent to four out of five medications, $p_i^*=0.80$, is equivalent to a person who is adherent to 12 out of 15, $p_i^*=0.80$, of their medications.); with adjusted linear regression we assume that the number of medications is linearly related to the proportion of adherence; and with weighted linear regression we assume that each individual should be given a weight that is proportional to the number of medications the individual is taking. The most appropriate of these assumptions seems to be the weighted linear model; however, it was the model that performed the worst of the three in terms of Type I error rate and empirical coverage. In addition, none of these models allows estimation of the intra-individual correlation which could be informative to investigators designing interventions to increase rates of adherence for general medication use, nor do they allow medication-specific covariates to be factored into the model which could be important if the type of medication impacts adherence. Since the probability of adherence is calculated by summing across all of the medications of an individual, if an individual is more likely to be adherent to insulin for diabetes than daily aspirin for MI prophylaxis, these medication-level covariates could not be accounted for in a linear model nor could variations in correlation between the medications if we had a more structured dataset (i.e. medication level covariates).

As expected, the BB model performed extremely well when the data were generated under a BB distribution, but did not perform as well when the data were generated under a shared response setting. Although the Type I error rates and empirical coverage probabilities were close to the nominal 0.05 and 0.95 values, respectively, under both circumstances, the intra-individual correlation was underestimated resulting in slightly higher power than the three linear models and the GEE model. The BB can estimate the intra-individual correlation, but is important to note that the BB model is limited to positive correlations and correlation structures which assume responses within an individual share the same correlation (i.e., exchangeable; Neuhaus 1992). Therefore, we would expect that the BB model would not perform as well if the true correlation structure differed from an exchangeable matrix. And just as with the linear model, BB regression is unable to account for medication-level covariates (Neuhaus 1992). In comparison, the GEE is capable of handling medication-level covariates, a wide variety of correlation structures with both positive and negative correlations, and performed extremely well under all circumstances of data generation with negligible bias in the estimates of the regression parameters and the intra-individual correlation. The Type I error rates and empirical coverage probabilities were close to the nominal values of 0.05 and 0.95, respectively, as well. The variance estimate of the GEE is considered a “robust” estimator of the variance because the estimates of the regression parameters and their variances are consistent even if the “working” correlation matrix is misspecified as long as the model for the mean is correctly specified (Zeger and Liang 1989; Dunlop 1994; Stokes et al. 2000). However, a caveat of the GEE is that it does not perform as well when the number of individuals (the number of medications per individual is not important here, only the cohort size) is less than 50 (Mancl and DeRouen 2001), especially if the intra-individual correlations are high (Stokes et al. 2000). We do not expect small sample sizes in the community level setting similar to our example but investigators working with smaller sample sizes should be aware of the small-sample properties of GEE.

6. CONCLUSIONS

Measuring medication adherence in the setting of polypharmacy is a complex issue and one that we anticipate will become more prevalent in research regarding the quality of medication use. We recommend using the GEE approach for analyzing adherence data measured dichotomously in the setting of polypharmacy. The GEE is more robust and can accommodate a wider variety of correlation structures than the BB model for situations

where the dataset is more structured with respect to specific medications, as well as being able to handle negative correlations. In addition, GEE can incorporate medication-level covariates when researchers are interested in adherence differences across types of medications (Prentice 1988), or when the intra-individual probability of adherence ($p_{ij} \neq p_i$, for at least one $j, j=1, \dots, n_i$) is not constant across all medications, a property both the BB and linear models (as used in these circumstances, where the probability is a function of the total number of medications) lack. If in fact the $p_{ij} \neq p_i$ for at least one j , then the unexplained error would most likely increase in both the linear and BB models due the introduction of another level of variability (between the p_{ij} 's) and the inability of both models to account for the variability because of their use of summary measures in their modeling. The GEE can also be extended to situations in which adherence is measured for each medication with a nominal or ordinal scale having more than two levels, or even as a continuous measure (i.e., using pill counts to assess the percentage of each medication taken). While the linear regression model in which the probability of adherence is treated continuously performed well in terms of bias, Type I error, and power, we warn investigators about this model, especially since as we have shown through simulations and Zhao et al. (2001) have shown in four example datasets, there is no guarantee that the model will only predict valid probabilities, nor is it possible to guarantee that adherence rates will remain within the “normal” range and not violate the linear model assumptions.

Acknowledgments

The authors would like to thank James T. Peterson, Russell F. Thurnow, and John W. Guzevich for providing sample code in R for the beta-binomial regression. This work was supported in part by grants from the National Institute of Health (K12 RR023248 (PI: Orringer); K30 RR022267 (PI: Ransohoff); K23 AG024229 (PI: Roth); UL1RR025747 (PI: Pisano)) and an American College of Clinical (ACCP) Frontiers Research and Career Development Award (PI: Roth). The authors also thank Dr. Larry Kupper and Dr. Morris Weinberger for reading this article and providing helpful insights and suggestions.

REFERENCES

- Cameron, AC.; Trivedi, RK. Regression analysis of count data. Cambridge University Press; Cambridge: 1998.
- Davis, ME.; Davis, CS.; Koch, GG. Categorical Data Analysis Using the SAS®System. Second Edition. SAS Institute Inc; Cary, NC: 2002.
- Diggle, PJ.; Heagerty, P.; Liang, KY.; Zeger, SL. Analysis of longitudinal data. 2nd ed. Oxford University Press; Oxford: 2002.
- Dunlop DD. Regression for longitudinal data: a bridge from least squares regression. *The American Statistician* 1994;48:299–303.
- Fitzmaurice GM, Laird NM, Rotnitzky AG. Regression models for discrete longitudinal responses. *Statistical Science* 1993;8:284–309.
- Fleiss, JL.; Levin, B.; Paik, MC. Statistical Methods for Rates and Proportions. 3rd ed. Wiley; Hoboken, NJ: 2003.
- Hosmer, DW., Jr; Lemeshow, S. Applied logistic regression. Chapman and Hall; New York: 1989.
- Hu FB, Goldberg J, Hedeker D, Flay BR, Pentz MA. Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American Journal of Epidemiology* 1998;147:694–703. [PubMed: 9554609]
- Johnson, NL.; Kemp, AW.; Kotz, S. Univariate discrete distributions. 3rd ed. Wiley; Hoboken, NJ: 2005.
- Kleinbaum, DG.; Klein, M. Logistic Regression [electronic resource]: A Self-Learning Text. 2nd ed. Springer-Verlag New York, Inc; New York, NY: 2002.
- Kupper LL, Haseman JK. The use of a correlated binomial model for the analysis of certain toxicological experiments. *Biometrics* 1978;34:69–76. [PubMed: 564719]

- Lee JK, Grace KA, Taylor AJ. Effect of pharmacy care program on medication adherence and persistence, blood pressure, and low-density lipoprotein cholesterol: A randomized controlled trial. *The Journal of the American Medical Association* 2006;296:2563–2571.
- Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13–22.
- Lunn AD, Davies SJ. A note on generating correlated binary variables. *Biometrika* 1998;85:487–90.
- Mancl LA, DeRouen TA. A covariance estimator for GEE with improved small-sample properties. *Biometrics* 2001;57:126–134. [PubMed: 11252587]
- Murray MD, Morrow DG, Weiner M, Clark DO, Tu W, Deer MM, Brater DC, Weinberger M. A conceptual framework to study medication adherence in older adults. *The American Journal of Geriatric Pharmacotherapy* 2004;2:36–43. [PubMed: 1555477]
- Neuhauser JM. Statistical methods for longitudinal and clustered designs with binary responses. *Statistical Methods in Medical Research* 1992;1:249–273. [PubMed: 1341660]
- Osterberg L, Blaschke T. Adherence to medication. *New England Journal of Medicine* 2005;353:487–497. [PubMed: 16079372]
- Pang Z, Kuk AYC. A shared response model for clustered binary data in developmental toxicity studies. *Biometrics* 2005;61:1076–84. [PubMed: 16401281]
- Prentice RL. Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *The Journal of the American Statistical Association* 1986;81:321–327.
- Prentice RL. Correlated binary regression with covariates specific to each binary observation. *Biometrics* 1988;44:1033–48. [PubMed: 3233244]
- Singer, JD.; Willet, JB. *Applied longitudinal data analysis*. Oxford University Press; Oxford: 2003.
- Skellam JG. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. *Journal of the Royal Statistical Society, Series B* 1948;10:257–61.
- Vik SA, Maxwell CJ, Hogan DB. Measurement, correlates, and health outcomes of medication adherence among seniors. *The Annals of Pharmacotherapy* 2004;38:303–312. [PubMed: 14742770]
- Weisberg, S. *Applied linear regression*. 2nd ed. Wiley; New York: 1985.
- Williams DA. The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics* 1975;31:949–52. [PubMed: 1203435]
- Zhao L, Chen Y, Schaffner DW. Comparison of logistic regression and linear regression in modeling percentage data. *Applied and Environmental Microbiology* 2001;67:2129–2135. [PubMed: 11319091]

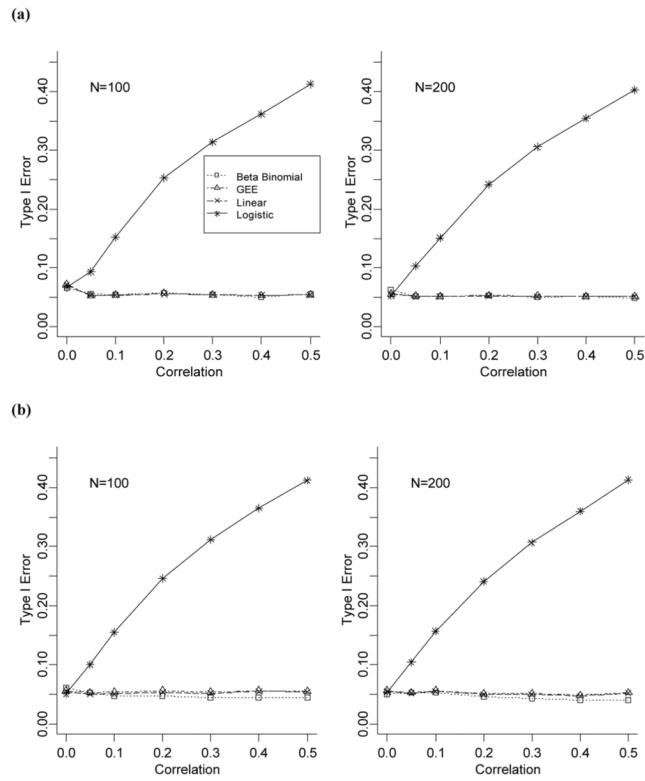


Figure 1. (a) Comparison of Type I error rates for data generated under a beta-binomial model with n fixed at 10 for all individuals. (b) Comparison of Type I error rates for data generated under a shared response model with n fixed at 10 for all individuals.

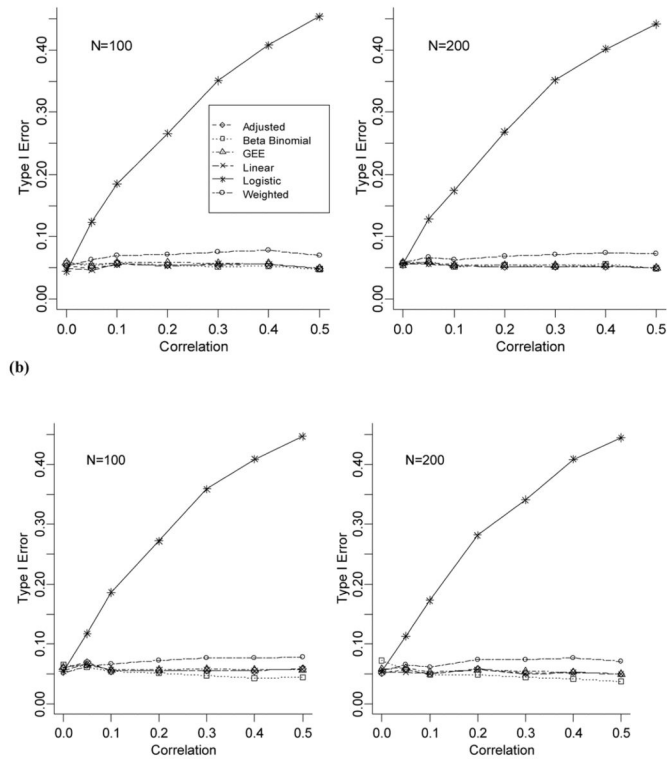


Figure 2. (a) Comparison of Type I error rates for data generated under a beta-binomial model with n_i varying for each individual. (b) Comparison of Type I error rates for data generated under a shared response model with n_i varying for each individual.

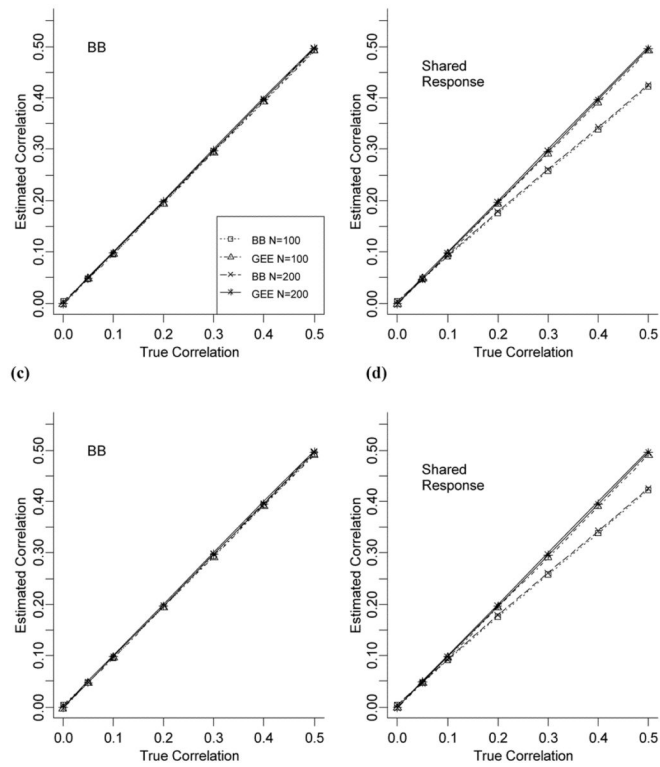


Figure 3. (a)-(b) Comparison of the true intra-individual correlation with the estimated intra-individual correlation for the GEE and BB models for data generated under both the beta-binomial and shared response models for fixed n . (c)-(d) Comparison of the true intra-individual correlation with the estimated intra-individual correlation for the GEE and BB models for data generated under both the beta-binomial and shared response models for random n .

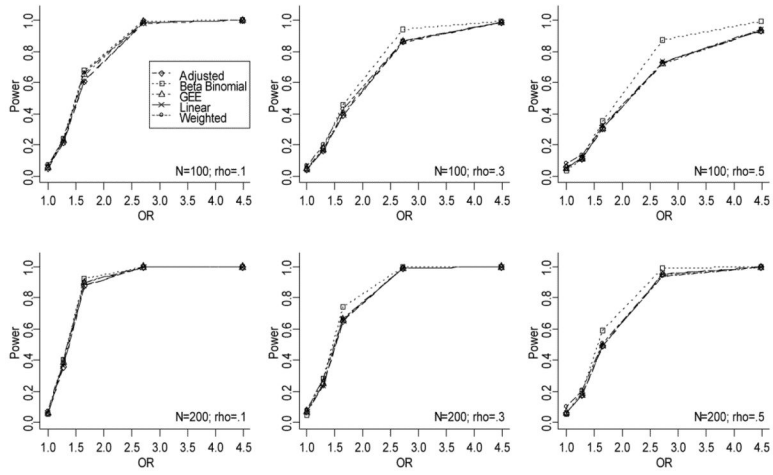


Figure 4. Power plots for data generated under a shared response model with random n .

Table 1

Analysis of the Example Data Set

Model#	Parameter	Parameter Estimate	SE Estimate	Z	p-value	Correlation Estimate	95% C.I.
Linear	DAR	0.05	0.05	1.10	0.271	---	(-0.04, 0.14)
Adjusted	DAR	0.05	0.05	1.08	0.281	---	(-0.04, 0.14)
Weighted	DAR	0.07	0.04	1.59	0.113	---	(-0.02, 0.15)
Naïve Logistic	InOR (OR)	0.42 (1.57)	0.16	2.77	0.006	---	(1.14, 2.16)
GEE Robust	InOR (OR)	0.36 (1.44)	0.29	1.24	0.216	0.18	(0.81, 2.56)
Beta Binomial	InOR (OR)	0.38 (1.47)	0.27	1.42	0.155	0.22	(0.86, 2.49)

DAR=difference in adherence rate, Rate in White - Rate in African American

Reference group is African American

Table 2
Empirical 95% Coverage for Data Generated under Shared Response Model with Random η_i .

β_1	Model	N=100			N=200		
		$\rho=0.1$	$\rho=0.3$	$\rho=0.5$	$\rho=0.1$	$\rho=0.3$	$\rho=0.5$
0	Linear	0.946	0.955	0.949	0.938	0.932	0.941
	Adjusted	0.947	0.956	0.943	0.941	0.931	0.940
	Weighted	0.928	0.930	0.915	0.927	0.920	0.901
	Naïve Logistic	0.814	0.662	0.565	0.802	0.638	0.552
	GEE Robust	0.940	0.948	0.948	0.942	0.928	0.936
	Beta Binomial	0.944	0.958	0.963	0.942	0.932	0.947
0.25	Linear	0.947	0.949	0.950	0.945	0.944	0.950
	Adjusted	0.949	0.944	0.952	0.946	0.943	0.949
	Weighted	0.934	0.925	0.936	0.929	0.921	0.926
	Naïve Logistic	0.825	0.632	0.555	0.822	0.652	0.558
	GEE Robust	0.946	0.938	0.951	0.939	0.948	0.948
	Beta Binomial	0.948	0.953	0.963	0.942	0.955	0.959
0.5	Linear	0.950	0.949	0.950	0.947	0.950	0.953
	Adjusted	0.949	0.949	0.947	0.946	0.951	0.953
	Weighted	0.937	0.930	0.932	0.934	0.933	0.932
	Naïve Logistic	0.823	0.653	0.552	0.818	0.662	0.549
	GEE Robust	0.948	0.943	0.940	0.944	0.951	0.951
	Beta Binomial	0.953	0.957	0.961	0.951	0.960	0.962
1	Linear	0.946	0.945	0.937	0.944	0.949	0.947
	Adjusted	0.947	0.945	0.936	0.945	0.949	0.948
	Weighted	0.935	0.929	0.927	0.934	0.927	0.927
	Naïve Logistic	0.827	0.655	0.543	0.822	0.655	0.562
	GEE Robust	0.938	0.939	0.940	0.944	0.946	0.946
	Beta Binomial	0.949	0.955	0.952	0.952	0.956	0.953
1.5	Linear	0.948	0.942	0.947	0.951	0.951	0.948
	Adjusted	0.948	0.942	0.946	0.952	0.952	0.948
	Weighted	0.932	0.925	0.930	0.939	0.928	0.931

β_1	Model	N=100			N=200		
		$\rho=0.1$	$\rho=0.3$	$\rho=0.5$	$\rho=0.1$	$\rho=0.3$	$\rho=0.5$
	Naïve Logistic	0.831	0.651	0.552	0.828	0.656	0.564
	GEE Robust	0.934	0.926	0.941	0.945	0.943	0.945
	Beta Binomial	0.951	0.956	0.955	0.953	0.948	0.949