Intent-to-Treat Analysis of Cluster Randomized Trials when Clusters Report Unidentifiable Outcome Proportions

Stacia M. DeSantis[1], **Ruosha Li**[1], **Yefei Zhang**[1], **Xueying Wang**[1], **Sally W. Vernon**[2], **Barbara C. Tilley**[1], **Gary Koch**[3]

[1]Department of Biostatistics and Data Science, The University of Texas Health Science Center at Houston, Houston, TX, USA.

[2]Department of Health Promotions and Behavioral Science, The University of Texas Health Science Center at Houston, Houston, TX, USA.

[3]Department of Biostatistics, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

Abstract

Background: Cluster randomized trials are designed to evaluate interventions at the cluster or group level. When clusters are randomized but some clusters report no or non-analyzable data, intent to treat analysis, the gold standard for the analysis of randomized controlled trials, can be compromised. This paper presents a very flexible statistical methodology for cluster randomized trials whose outcome is a cluster-level proportion (e.g., proportion from a cluster reporting an event) in the setting where clusters report non analyzable data (which in general could be due to non adherence, dropout, missingness, etc). The approach is motivated by a preciously-published stratified randomized controlled trial called, "The Randomized Recruitment Intervention Trial (RECRUIT)," designed to examine the effectiveness of a trust-based continuous quality improvement intervention aimed at increasing minority recruitment into clinical trials (ClinicalTrials.gov Identifier: NCT01911208).

Methods: The novel approach exploits the use of generalized estimating equations for cluster-level reports, such that all clusters randomized at baseline are able to be analyzed, and intervention effects are presented as risk ratios. Simulation studies under different outcome missingness scenarios and a variety of intra-cluster correlations are also conducted. A comparative analysis of the method with imputation and per protocol approaches for RECRUIT is presented.

Results: Simulation results show the novel approach produces unbiased and efficient estimates of the intervention effect that maintain the nominal type I error rate. Application to RECRUIT shows similar effect sizes when compared to the imputation and per protocol approach.

Corresponding author: Stacia M. DeSantis, University of Texas Health Science Center at Houston School of Public Health, Department of Biostatistics and Data Science, Houston, Texas, US. Stacia.M.DeSantis@uth.tmc.edu.

Conclusions: The paper demonstrate that an innovative bivariate generalized estimating equations framework allows one to implement an intent to treat analysis to obtain risk ratios or odds ratios, for a variety of cluster randomized designs.

Keywords

Cluster randomized trials; randomized controlled trials; missing data; clinical trials; intent to treat

Introduction

Randomized controlled trials are the gold standard of public health research. Typically, the unit of randomization and analysis is the individual. Cluster randomized trials are trials designed to evaluate interventions that operate at a cluster-level[1]. Such studies may manipulate the physical or social environment such that intervention cannot feasibly be delivered to individuals.[2–5] Examples include interventions delivered to schools, workplaces, or hospitals.[6–10] The unit of analysis for cluster randomized trials may be the individual or the cluster.

In order to uphold randomization and ensure unbiased (causal) effects are estimated in the primary outcome analysis, all randomized trials, should be analyzed on the principle of intent to treat. Under intent to treat, participants or clusters are analyzed as members of the treatment group to which they were randomized regardless of their adherence to, or whether they received, the intended treatment.[11–13] Thus it ignores nonadherence, protocol deviations such as randomization errors, withdrawal, dropout, and anything that happens after randomization, most of which are inevitable in human trials.[14] In the setting of cluster randomized trials, a design effect is typically included to model the correlation between individuals within a cluster.[15] However, applying the intent to treat principle to a cluster randomized can be more difficult because the issues of non-compliance, dropout, and missing data can be more complex. In fact, research has shown the intent to treat principle is more difficult to adhere to in cluster randomized trials, as loss of an entire cluster (versus say, one individual) both compromises inference and potentially decreases statistical power.[8,16–18] Researchers have presented remedies for this issue including randomizing clusters only when the first participant is included to prevent empty clusters (index case concept),[17] or for trials already in progress, using ad hoc missing data or propensity score methods to accommodate missing outcome data at the individual or cluster level. Others propose the use of Expectation-Maximization algorithm to obtain unbiased intent to treat-principle estimates.[15]

The current research is motivated by a stratified cluster randomized trial where 8 of 50 randomized clusters reported 0/0 proportion data. The Randomized Recruitment Intervention Trial (RECRUIT, ClinicalTrials.gov Identifier: NCT01911208) has been previously described in great detail.[10] Briefly, RECRUIT was the first multi-site randomized controlled trial to examine the effectiveness of a trust-based continuous quality improvement intervention aimed at health care workers, to increase minority recruitment into clinical trials. Four multi-site randomized controlled trials (parent trials) supported by 3 National Institutes of Health participated in RECRUIT. Fifty sites (i.e., clusters in this setting)

within the 4 parent trials were randomized, 24 to intervention, and 26 to no intervention. Sites, the unit of analysis, were matched within parent trial on site characteristics. Overall, 26 intervention and 24 control sites were enrolled. The primary outcome was the site-level (cluster-level) proportion minority enrollment (to produce 50 outcomes for analysis). A simple schematic of the design can be viewed in Figure 1 of the design paper.[10] Briefly, the study was powered to detect a 0.10 absolute difference in intervention versus control proportions of minorities recruited assuming a 2-sample test and intra-cluster (site) correlation of 0.10.[10] A generalized estimating equations (GEE) approach to model the proportion minority enrolled was pre-specified to account for clustering of people within a site; GEE provides consistent standard errors even when the correlation structure is incorrectly specified[19] and is appropriate for 50 or more clusters in a cluster randomized trial.[20]

An unforeseen issue that arose mid-trial was that 8 of 50 sites were unable to enroll any patients. This led to an analytic challenge since the proportion minority enrolled for those 8 sites returned non analyzable data. As omitting 8 randomized sites could compromise the final analysis, this paper proposes a remedy that enables the inclusion of all sites. This paper proposes and assesses via simulation, a remedy for analyzing the effect of the intervention on a cluster-level outcome measured as either a proportion or a count, when some clusters report non analyzable data, or do not report outcomes. Specifically, an exact statistical approach using GEEs for estimation is developed, tested via simulation, and applied to the RECRUIT study. The RECRUIT study is used as an example throughout the following development to demonstrate the flexibility of the approach to both straightforward, and more complex cluster randomized designs.

Methods

The relevant Institutional Review Board provided approval for the RECRUIT Study. The below formulation will be described in terms of the RECRUIT data example for ease of presentation, but is generalizeable to any cluster randomized trial where the outcome is a cluster-level proportion. Since the cluster in RECRUIT is a site, the term "site" is used henceforth in place of "cluster." Let D_i be the total number of enrolled patients in site i, and let X_{ij} denote a binary outcome, i.e., minority status (1 vs. 0) for the jth patient enrolled in site i, where $j = 1, 2, ..., D_i$. Therefore, $Y_i = \Sigma_{j=1}^{D_i} X_{ij}$ represents the total number of minority patients enrolled in site i. Let Z_{1i} be the intervention indicator, where $Z_{1i}=1$ represents intervention and $Z_{1i}=0$ represents control. When illustrating the below approach, assume only 2 parent trials, namely trial A and trial B (in order to simplify the notation). Simple cluster randomized trials will not have parent trials (the "strata" element) unique to RECRUIT. In the simpler setting, the following approach simplifies in an obvious manner. Define T_i as the indicator variable that equals 1 for trial A and 0 for trial B. More or no trial-specific indicator variables could be created, without loss of generality.

The sites with no participants enrolled provided an undefined value for the proportion of minorities enrolled (0/0) at that clinic. The following presents several possible approaches to the analysis in the presence of clusters that report no or undefined data; the last approach

derived is the most true intent to treat analysis, as it analyzes all clusters as they were randomized.

Method 1: A generalized estimating equations model for individual-level outcomes

The primary analysis plan assumed all clusters would report whether or not a person enrolled was a minority. The pre-specified analysis was therefore a binomial GEE with logit link function to model the individual-level outcome, X_{ij} and with trial as a covariate. Minority patients would be coded $X_{ij}=1$ while non-minority patients would be coded $X_{ij}=0$. The GEE with logit link would model the probability that $X_{ij} = 1$ adjusting for the intervention indicator Z_{1i}, and the parent trial indicator T_i (which is specific to RECRUIT and adds an additional layer of complexity not relevant to most cluster randomized trials). The effect of interest is that of the intervention on the proportion of minorities that enrolled; this corresponds to the coefficient, β_1 of Z_{1i} in the GEE regression model, such that $\exp(\beta_1)$ is the odds ratio for enrolling a minority in the intervention versus control sites.

Sites enrolling no individuals (i.e., $D_i = 0$) would be omitted from this individual-level analysis, thus 8 of the 50 randomized sites (6 control and 2 intervention) would be excluded, substantially reducing the effective number of clusters to 42. The GEE available case analysis would analyze only 42 clusters as opposed to 50 clusters.

To facilitate the pre-planned primary analysis while still including the 8 sites with $D_i = 0$, a feasible analytic strategy would be to impute 1 single non-minority patient for each site that enrolled 0 patients (denominator). Imputation of a non-minority is chosen since they represent the majority of participants and would be the most conservative approach to a modified analysis. One could easily impute the alternative as well, and conduct a sensitivity analysis. While this approach allows all sites that were randomized to provide outcome data, imputing participants that do not exist is not ideal. The parameter of interest would be β_1, where $\exp(\beta_1)$ is the OR as defined above. In the below, reference to the pre-planned GEE analysis with all available case data or with imputed data is termed, "available cases-GEE" and "Imputation-GEE," respectively.

Method 2: A generalized linear model for cluster-level outcomes

An alternative procedure is to consider modeling Y_i, the total number of minority patients for each site. One could analyze all available cases (42 sites) or use the same simple imputation strategy as above - impute a placeholder of 1 for the denominator (i.e., D_i) of the 8 sites with zero enrollment, and a 0 for the numerator, Y_i. A cluster-level generalized linear model (GLM) for count data would then be fit. To accommodate potential overdispersion in the counts, one may adopt the overdispersed Poisson model or negative binomial (NB) model. For illustration purpose, we present the overdispersed Poisson with offset, D_i, below. The model for the mean of the outcome conditional on the denominator (D_i), intervention (Z_{1i}), and trial (T_i), is,

$$\log\{E(Y_i \mid D_i, Z_{1i}, T_i)\} = \log(D_i) + \beta_0 + \beta_1 Z_{1i} + \beta_2 T_i, \tag{1}$$

where the D_i is the number of participants enrolled in site i. By including $\log(D_i)$ as an offset, the model targets the proportion minority, Y_i/D_i, the quantity of interest. Given D_i, the model for the mean is,

$$E\{(Y_i/D_i) \mid D_i, Z_{1i}, T_i\} = \exp(\beta_0) \times \exp(\beta_1 Z_{1i} + \beta_2 T_i). \tag{2}$$

The parameter of interest is still β_1, where $\exp(\beta_1)$ is the risk ratio (RR) to be discussed for enrolling a minority in the intervention versus control sites. Of course if the interaction between Z_{1i} and T_i was significant, then the RR would be stratified by trial and separately reported. For comparison purpose, it is sensible as with GEE, to consider the "available case" counterpart of the GLM, where we omit the sites with 0 enrollment (i.e., $D_i = 0$). As in the above, the methods are termed, "available cases-GLM" and "imputation-GLM."

Method 3: Intent to treat analysis for cluster level outcomes

Since none of the aforementioned approaches are ideal methods of analysis due to either imputation of or omission of randomized units, a novel intent to treat GEE approach was formulated based on a joint analysis of two correlated components: the numerator and the denominator of the outcome proportion. This simple approach treats the cluster-level outcome as a bivariate vector, representing a numerator and a denominator produced from each site (cluster). The formulation is used for convenience because as shown below, a simple reparameterization enables a cluster randomized intent to treat analysis, whereas any other obvious approach would require omission of sites, imputation, expectation maximization (EM), or otherwise.[15]

The trick to enabling intent to treat GEE analysis when some clusters do not provide analyzable proportion outcome data is by jointly modeling the numerator and denominator "components" such that sites that report no outcome, still enter the statistical model. The outcome is therefore now defined as a bivariate vector, $\mathbf{W}_i = (Y_i, D_i)$, where the components are as defined above - Y_i is the numerator count for cluster i (the number of minorities enrolled in cluster i) and D_i is the denominator count for cluster i (the total number of participants enrolled in cluster i). For illustration, first consider one-sample data. Suppose that the proportion of interest is p, and let μ_{num} and μ_{den} denote the marginal mean of Y and D respectively; then $\mu_{num} = \mu_{den} \times p$. Thus the ratio of the two means give the proportion of interest.

Specifically, consider the joint model of the mean of Y_i and D_i, where both count outcomes marginally follow a specific distribution for count data, such as the over-dispersed Poisson or negative binomial model. The following is the bivariate model parameterization: let (μ_{i1}, μ_{i2}) denote $(E(Y_i), E(D_i))$. First assume a typical cluster randomized trial where there are no strata (which is specific to RECRUIT). The intervention indicator is Z_{1i}; now include an indicator, Z_{2il}, where $Z_{2il} = 1$ for the numerator count ($l = 1$) and $Z_{2il} = 0$ for the denominator count ($l = 2$). The terms "num" and "den" are shorthand for numerator and denominator below. The mean model is parameterized as,

$$\mu_{il} = \mu_0 \times \exp\{Z_{1i}\beta_1 + Z_{2il}\beta_2 + Z_{1i}Z_{2il}\beta_3\}, l = 1,2 \, . \tag{3}$$

In this model, interaction between Z_{1i} and Z_{2il} are essential because as they will lead to the intent to treat intervention effect on the proportion. The mean counts under this model are given by,

$$
\begin{aligned}
\mu_{Intervention,\,num} &= \mu_0 \times \exp(\beta_1 + \beta_2 + \beta_3), \\
\mu_{Intervention,\,den} &= \mu_0 \times \exp(\beta_1), \\
\mu_{Control,\,num} &= \mu_0 \times \exp(\beta_2), \\
\mu_{Control,\,den} &= \mu_0 \, .
\end{aligned}
$$

From these formulae, the treatment effect on the proportion of interest for the cluster randomized trial is simply,

$$\frac{\mu_{Intervention,\,num} / \mu_{Intervention,\,den}}{\mu_{Control,\,num} / \mu_{Control,\,den}} = \frac{\exp(\beta_2 + \beta_3)}{\exp(\beta_2)} = \exp(\beta_3) \, . \tag{4}$$

Therefore, by modeling the cluster-level count outcomes, this model formulation allows for the estimation of the intervention effect on the proportion of interest.

For the RECRUIT stratified cluster randomized trial, assume 2 parent trials, trial A and trial B. Simply incorporate a parent trial indicator T_i, and the mean model is parameterized similarly,

$$\mu_{il} = \mu_0 \times \exp\{Z_{1i}\beta_1 + Z_{2il}\beta_2 + Z_{1i}Z_{2il}\beta_3 + T_i\beta_4 + Z_{1i}T_i\beta_5 + Z_{2il}T_i\beta_6\}, l = 1,2 \, . \tag{5}$$

The parameter of interest is still β_3, as shown below for parent trial A

$$
\begin{aligned}
\mu_{Intervention,\,Trial\,A,\,num} &= \mu_0 \times \exp(\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6), \\
\mu_{Intervention,\,Trial\,A,\,den} &= \mu_0 \times \exp(\beta_1 + \beta_4 + \beta_5), \\
\mu_{Control,\,Trial\,A,\,num} &= \mu_0 \times \exp(\beta_2 + \beta_4 + \beta_6), \\
\mu_{Control,\,Trial\,A,\,den} &= \mu_0 \times \exp(\beta_4) \, .
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\frac{\mu_{Intervention,\,Trial\,A,\,num} / \mu_{Intervention,\,Trial\,A,\,den}}{\mu_{Control,\,Trial\,A,\,num} / \mu_{Control,\,Trial\,A,\,den}} &= \frac{\exp(\beta_2 + \beta_3 + \beta_6)}{\exp(\beta_2 + \beta_6)} \\
&= \exp(\beta_3),
\end{aligned}
$$

where $\mu_{Intervention,\,Trial\,A,\,num} / \mu_{Intervention,\,Trial\,A,\,den}$ corresponds to the proportion of minority for the intervention arm for parent Trial A, and

$\mu_{Control, Trial A, num} / \mu_{Control, Trial A, den}$ corresponds to the proportion for the control arm for parent Trial A. Similar derivations for trial B yield,

$$\frac{\mu_{Intervention, Trial B, num} / \mu_{Intervention, Trial B, den}}{\mu_{control, Trial B, num} / \mu_{control, Trial B, den}} = \frac{\exp(\beta_2 + \beta_3)}{\exp(\beta_2)} = \exp(\beta_3).$$

It is clear from above that $\exp(\beta_3)$ represents the intent to treat intervention effect on the proportion minority enrolled (for all parent trials). In Equation (5), this effect is assumed constant across parent trials. If one suspects that parent trial moderates the intervention effect on the proportion, a three way interaction $Z_{1i}Z_{2il}T_i$ may be included and tested. This interaction was not part of the primary hypothesis and was not significant in the primarily analysis of RECRUIT. It would likely not be possible to detect this effect for most cluster randomized trials due to power limitations therein.[21,22]

To conduct the analysis under the model given by Equation (3) or (5), one adopts existing GEE software.[19] Specifically, an over-dispersed Poisson distribution or a negative-binomial model, with working independence correlation structure and log link may be specified. As the GEE model focuses on estimating the marginal means, it makes little assumptions about the joint distribution of Y and D. The model is therefore flexible for count data and robust to small to moderate deviations from model assumptions. One can decide whether to use either the overdispersed Poisson or NB in practice after consulting the large body of literature devoted to the deficiencies and benefits of either assumption.[23–26]

For all three GEE-based methods (i.e., available cases GEE, imputation GEE, intent to treat GEE) empirical standard error estimates are adopted. It is known that the empirical standard errors may be slightly biased downward, and the bias becomes more noticeable when the number of clusters is smaller than 50.[27] Researchers have discussed several bias correction methods for the empirical standard error estimates, which are implemented in the simulation study and application using the R package *geesmv*.[27]

Simulation Study

Simulation Setup

A simulation study is conducted to evaluate the performance of the intent to treat GEE approach versus the alternative approaches. The goal of the study is not necessarily to demonstrate exceptional performance, but to show near-equivalence between the methods such that the intent to treat GEE analysis may be used in place of the others in settings of either missing or non-analyzable outcome data. To parallel the RECRUIT design, two parent trials (trial A and B) with a total of 50 sites within those trials are assumed. The intervention indicator variable, Z_{1i}, and the parent trial indicator variable, T_i, are both generated from a Bernoulli distribution with probability 0.5, for $i = 1, 2, ..., 50$ sites. The total enrollment for each site (D_i) is a count variable with mean μ and variance $\mu(1 + 1/\phi)$, where ϕ concerns overdispersion (or deviation from standard Poisson assumptions). Mirroring observations from RECURIT, we set $\mu = 8$ and $\phi = 0.1$, which lead to approximately 13% sites with 0 enrollment. For a site, i, with D_i 1, correlated Bernoulli variables X_{ij} are generated with probability

$$p_{ij} = \begin{cases} (1 - Z_{1i})p_{11} + Z_{1i}p_{12} \text{ if } T_i = 0, \\ (1 - Z_{1i})p_{21} + Z_{1i}p_{22} \text{ if } T_i = 1, \end{cases}$$

for $j = 1, 2, ..., D_i$, where (p_{k1}, p_{k2}) stand for the case probabilities in the placebo group and the intervention group respectively for the k_{th} parent trial, i.e., $k = 1,2$. Clustered Bernoulli random variables are generated, X_{ij}, using the *rcbin* function in R package *ICCbin* for $j = 1,2,...,D_i$. The jth participant is assigned as a case (e.g., minority) if the corresponding Bernoulli random variable equals 1 and a non case (e.g., non minority) otherwise. Following this, Y_i is the total number of cases in site i. If a site, i, has zero enrollment, then $Y_i = D_i = 0$.

First consider a simpler scenario without a parent trial effect, setting $(p_{11},p_{12}) = (p_{21},p_{22}) = (p_1,p_2)$. Here, p_1 and p_2 stand for the case probability in the intervention and control arm respectively. Four different parameter values of (p_1,p_2) are considered by setting $(p_1,p_2) = (0.16,0.16)$, $(0.16,0.224)$, $(0.16,0.256)$, and $(0.16,0.32)$, respectively. The case probability of 0.16 is chosen to mimic the proportion of minorities in the control arm of RECRUIT; the magnitudes of the intervention effects correspond to RRs of 1.0, 1.4, 1.6 and 2.0, respectively.

Next, consider the scenario with a parent trial effect, with (p_{11},p_{12}) (p_{21},p_{22}). Four different parameter values of (p_{k1},p_{k2}) are considered by setting $(p_{11},p_{12}) = (0.12,0.12)$, $(0.12,0.168)$, $(0.12,0.192)$, $(0.12,0.24)$ for parent trial A and the corresponding $(p_{21},p_{22}) = (0.2,0.2)$, $(0.2,0.28)$, $(0.2,0.32)$, $(0.2,0.4)$ for parent trial B, respectively. These parameters correspond to RR, p_{k2}/p_{k1}, for the intervention effect of 1.0, 1.4, 1.6, and 2.0, respectively. For each setup, we considered the within group correlation $\rho = 0.01$, 0.025, 0.05, or 0.1 to mimic realistic degrees of intracluster correlation in clinical trials research.[8,28] This results in a total of 16 different simulation scenarios for which 10,000 data sets are generated for each scenario.

All simulated data are generated using the statistical package R. The simulated datasets are then analyzed using the function *"glm"* in R for the two GLM models and using the *"gee"* and *"geesmv"* package in R for the three GEE models.

Simulation Results

Table 1 displays the result from all five approaches presented in Section 2 under the typical cluster randomized trial scenario without a parent trial effect in terms of bias, relative bias, standard deviation of the estimates, average of estimated standard errors, relative standard errors and 95% coverage probability. Note that for the two approaches based on the GEE model with binomial distribution, namely the Imputation-GEE and the available cases-GEE, β corresponds to the log of the odds ratio. In comparison, for the intent to treat GEE approach and the two GLM approaches, β equals the log of the risk ratio. The true values for β are provided in the 2nd-3rd rows of Table 1.

Table 1 indicates that the intent to treat GEE approach has comparable bias under a variety of null and clinically meaningful non-null intervention effects, and correlation parameters (columns). The bias tends to be downward in scenarios without treatment effects, and

upwards in scenarios with positive treatment effects. However, the degree of bias is small; the relative bias is smaller than 3.2% under all settings. The relative bias increases slightly with the intracluster correlation coefficient for all methods, likely because that larger correlations correspond to a reduction in the effective sample size.[29] The standard deviation and average of estimated standard error reflect one another reasonably well, as evidenced by relative standard errors that are close to 1. The 95% coverage probabilities are near the nominal level for all but the two GLM approaches. Similarly, the type I error rates, which correspond to 1 minus the coverage probability when $\beta = 0$, are satisfactory for all methods except for the two GLM approaches. Interestingly, deletion of 0s in the available cases-GLM approach does not show evidence of loss of efficiency despite nearly 13% of the data being deleted; further, imputation shows no efficiency improvement over deletion.

Table 2 shows the results under the scenario with a parent trial effect. In the presence of strata (trial), the data generation scheme satisfies models (2) and (5) but not the binomial GEE model, as it is difficult if not impossible to generate data that satisfy the assumptions of all models simultaneously. Therefore, the two binomial GEE approaches are not implemented here to avoid unfair comparisons. The intent to treat GEE continues to perform very satisfactorily in terms of bias and average of estimated standard errors, with coverage rates that are close to the nominal level of 95%. By comparison, the GLM approaches gives coverage rates that are lower than the nominal level, likely because its standard error estimates tend to be smaller than the empirical counterparts.

The conclusion of the simulation study is that the performance of the intent to treat GEE approach is competitive when compared to alternative methods, and therefore could be used in place of imputation and available case analysis in order to include all clusters that were randomized.

Data Application

The approaches above are applied to the RECRUIT study using a sensitivity analytic approach, in complement to the analysis reported in the primary paper. From exploratory model fit criteria, it was determined a negative binomial was appropriate. As presented in the Methods Section, each model fit includes the 3 trial indicators to adjust for these design effects. The intent to treat GEE additionally includes the interactions between treatments and trial as necessary to produce the intent to treat intervention effect.

Table 3 presents the intervention effect, the standard error of the intervention effect, the 95% confidence interval, and associated p-values resulting from applying the 5 methods to RECRUIT. The estimated coefficients correspond to $\log(OR)$ in the top two rows and $\log(RR)$ in the bottom three rows. The intention to treat GEE approach results in a slightly smaller standard error than the two GLM approaches. Also of note is that all methods produce relatively similar, insignificant p-values, serving as a strong sensitivity analysis for the primary analysis of RECRUIT; inference and conclusions remain unchanged no matter which model is applied or whether OR/RR is the effect size of interest. Overall, the conclusion is there is no effect of the intervention on minority enrollment. The reported effects are all in the positive direction, indicating minority enrollment was increased, albeit

insignificantly, in the intervention versus no intervention clusters. It is worth mentioning that the novel analysis without the 8 sites led to smaller effect size of log risk ratio = 0.266, standard error= 0.254, thus the estimated effect size for 42 sites is smaller than that including all the sites, but inference remains unchanged.

Discussion

This paper presents a novel intent to treat approach to analyzing cluster randomized trials in the difficult setting where a binary cluster-level outcome is non analyzable. The simple and easy-to-implement approach decomposes the proportion outcome into numerator and denominator counts to facilitate an exact method such that all randomized units can be included in the analysis and rate ratios for the intervention effect can be produced. This is achieved by modeling the bivariate "count" vector via a generalized estimating equations approach for clustered count data, or any other statistical method that can accommodate the now-bivariate count outcome in the context of clustered data, (e.g., a generalized linear mixed model). Unlike previous work, the solution does not require ad hoc missing data methods such as imputation or expectation-maximization and is applicable to the very common setting when cluster-level proportions or counts are the outcome of interest.[15,17] On the other hand, simulation results suggest that the available case binomial GEE is also acceptable if the pre-specified analysis plan does not require the inclusion of all randomized units.

In cluster randomized trials where the group rather than individual is randomized, the the intent to treat principle is often challenging to implement because of the lack of statistical methods to handle empty clusters. Oftentimes, clusters are discarded from the analysis.[17] While methods of imputation have been proposed, there is currently no clear solution to the current problem in the literature.[15,17] The solution presented in this paper is viable, easy to implement, and as shown via the application in this paper, can be adapted to even more complex designed such as stratified cluster randomized trials. The approach could potentially be adaptable to other complex designs such as stepped wedge trials, where clusters are randomized to different sequences over time; more research into this would be needed as such trials have the additional complexity of potential confounding by time.[30,31] Further, if one wanted to obtain the intervention vs control comparison for the odds ratio (OR) rather than the RR, then the bivariate method would simply be applied to the minority and non-minority count (rather than the minority and total count). In this case, one obtains ORs instead of RRs under the NB distributional assumption.

There are limitations to the current approach, Primarily, it is only applicable when individual-level covariates are not important to the overall study hypotheses, as the method only accommodates cluster-level (and not individual-level) covariates. Further, as with any method, it should be applied in the context of a sensitivity analysis, for example, in comparison to per protocol analysis of the data. When inference is similar for per protocol, imputation, and intent to treat GEE approaches, the authors recommend intent to treat GEE be reported in conjunction with those analyses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors acknowledge the entire Randomized Recruitment Intervention Trial Study team.

References

1. Donner A, Birkett N and Buck C. Randomization by cluster: sample size requirements and analysis. American Journal of Epidemiology 1981; 114(6): 906–914. [PubMed: 7315838]

2. Cornfield J.Randomization by group: a formal analysis. American journal of epidemiology 1978; 108(2): 100–102. [PubMed: 707470]

3. Murray DM. Design and analysis of group-randomized trials: A review of recent developments. Annals of epidemiology 1997; 7(7): S69–S77.

4. Murray DM. Design and analysis of group-randomized trials, volume 29. Monographs in Epidemiology & B, 1998.

5. Donner A and Klar N. Design and analysis of cluster randomization trials. London:Arnold, 2000.

6. Glanz K, Kristal AR, Tilley BC et al. Psychosocial correlates of healthful diets among male auto workers. Cancer Epidemiology, biomarkers, and prevention 1998; 7: 119–126.

7. Tilley BC, Vernon SW, Myers R et al. The next step trial: impact of a worksite colorectal cancer screening promotion program. Preventive medicine 1999; 28(3): 276–283. [PubMed: 10072746]

8. Murray DM, Varnell SP and Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. American journal of public health 2004; 94(3): 423–432. [PubMed: 14998806]

9. Tilley BC, Mainous AGI, Elm JJ et al. A randomized recruitment intervention trial in parkinson's disease to increase participant diversity: early stopping for lack of efficacy. Clinical Trials 2012; 9: 188–197. [PubMed: 22496398]

10. Tilley BC, Mainous III AG, Smith DW et al. Design of a cluster-randomized minority recruitment trial: Recruit. Clinical Trials 2017; 14(3): 286–298. [PubMed: 28545336]

11. Lachin JM. Statistical considerations in the intent-to-treat principle. Controlled Clinical Trials 2000; 21: 167–189. [PubMed: 10822117]

12. Yelland LN, Sullivan TR, Voysey M et al. Applying the intention-to-treat principle in practice: Guidance on handling randomisation errors. Clinical Trials 2015; 12(4): 418–423. [PubMed: 26033877]

13. Livingston AH, and Lewis RJ. JAMA Guide to Statistics and Medicine. New York: McGraw Hill, 2020.

14. Gupta SK. Intention-to-treat concept: A review. Perspectives in Clinical Research 2011; 2(3): 109–112. [PubMed: 21897887]

15. Jo B, Asparouhov T and O MB. Intention-to-treat analysis in cluster randomized trials with noncompliance. Statistics in Medicine 2008; 27(27): 5565–5577. [PubMed: 18623608]

16. Puffer S, Torgerson D and Watson J. Evidence for risk of bias in cluster randomised trials: Review of recent trials published in three general medical journals. British Medical Journal 2003; 327: 785–789. [PubMed: 14525877]

17. Giraudeau B and Ravaud P. Preventing bias in cluster randomised trials. PLoS Medicine 2009; 6(5): e1000065.

18. Garrison MM and Mangione-Smith R. Cluster randomized trials for health care quality improvement research. Academic pediatrics 2013; 13(6): S31–S37. [PubMed: 24268082]

19. Liang KY and Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986; 73(1): 13–22.

20. Donner A.Sample size requirements for stratified cluster randomization designs. Statistics in medicine 1992; 11(6): 743–750. [PubMed: 1594813]

21. Heo M and Leon AC. Ssample size requirements to detect an intervention by time interaction in longitudinal cluster randomized clinical trials. Statistics in Medicine 2009; 28(6): 1017–1027. [PubMed: 19153969]

22. Heo M and Leon AC. Sample sizes required to detect two-way and three-way interactions involving slope differences in mixed-effects linear models. Journal of Biopharmaceutical Statistics 2010; 20(4): 787–802. [PubMed: 20496206]

23. Desantis S and Bandyopadhyay D. Hidden markov models for zero-inflated poisson counts with an application to substance use. Statistics in Medicine 2011; 30(14): 1678–1694. [PubMed: 21538455]

24. Bandyopadhyay D, Desantis S, Korte J et al. Some considerations for excess zeroes in substance abuse research. American Journal of Drug and Alcohol Abuse 2011; 37(5): 376–382. [PubMed: 21854280]

25. Van Hoef JM and Bovent PL. Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data. Ecology 2007; 88(11): 2766–2772. [PubMed: 18051645]

26. Zhu H, Luo S and DeSantis SM. Zero-inflated count models for longitudinal measurements with heterogeneous random effects. Statistical Methods in Medical Research 2017; 26(4): 1774–1786. [PubMed: 26113383]

27. Wang M, Kong L, Li Z et al. Covariance estimators for generalized estimating equations (gee) in longitudinal analysis with small samples. Statistics in medicine 2016; 35(10): 1706–1721. [PubMed: 26585756]

28. Wu S, Crespi CM and Wong WK. Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials. Contemporary clinical trials 2012; 33(5): 869–880. [PubMed: 22627076]

29. Killip S, Mahfoud Z and Pearch K. What is an intracluster correlation coefficient? crucial concepts for primary care researchers. Annals of Family Medicine 2004; 2(3): 204–208. [PubMed: 15209195]

30. Thompson J, Fielding K, Hargreaves J et al. The optimal design of stepped wedge trials with equal allocation to sequences and a comparison to other trial designs. Clinical Trials 2017; 14(6): 639–647. [PubMed: 28797179]

31. Hemming K, Taljaard M, McKenzie J et al. Reporting of stepped wedge cluster randomised trials: extension of the consort 2010 statement with explanation and elaboration. British Medical Journal 2018; 363: k1614.

Table 1.

Simulation results for the scenario with a trial effect.

		ρ = 0.01				ρ = 0.025				ρ = 0.05				ρ = 0.1			
	True β (log RR)	0	0.337	0.470	0.693	0	0.337	0.470	0.693	0	0.337	0.470	0.693	0	0.337	0.470	0.693
	True β (log RR)	0	0.400	0.560	0.847	0	0.400	0.560	0.847	0	0.400	0.560	0.847	0	0.400	0.560	0.847
Bias	Imputation-GEE	0.002	0.010	0.012	0.014	0.003	0.012	0.015	0.017	0.003	0.015	0.018	0.021	0.001	0.017	0.022	0.028
	available cases-GEE	−0.005	0.004	0.008	0.011	−0.004	0.007	0.010	0.014	−0.004	0.009	0.013	0.019	−0.007	0.012	0.017	0.026
	Imputation-GLM	0.003	0.010	0.012	0.015	0.003	0.012	0.015	0.018	0.003	0.014	0.017	0.021	0.002	0.017	0.020	0.026
	available cases-GLM	−0.003	0.004	0.007	0.009	−0.002	0.007	0.009	0.012	−0.002	0.009	0.011	0.015	−0.004	0.011	0.014	0.020
	intent to treat-GEE	−0.003	0.004	0.007	0.009	−0.002	0.007	0.009	0.012	−0.002	0.009	0.011	0.015	−0.004	0.011	0.014	0.020
Relative Bias	Imputation-GEE	-	0.024	0.021	0.015	-	0.029	0.025	0.018	-	0.036	0.030	0.024	-	0.042	0.037	0.031
	available cases-GEE	-	0.011	0.013	0.013	-	0.016	0.017	0.016	-	0.023	0.023	0.021	-	0.029	0.029	0.028
	Imputation-GLM	-	0.030	0.027	0.021	-	0.037	0.031	0.025	-	0.043	0.036	0.030	-	0.049	0.043	0.037
	available cases-GLM	-	0.013	0.014	0.013	-	0.020	0.019	0.017	-	0.026	0.024	0.022	-	0.032	0.031	0.029
	intent to treat-GEE	-	0.013	0.014	0.013	-	0.020	0.019	0.017	-	0.026	0.024	0.022	-	0.032	0.031	0.029
ASE	Imputation-GEE	0.323	0.304	0.298	0.290	0.350	0.331	0.325	0.317	0.391	0.370	0.364	0.355	0.459	0.437	0.430	0.420
	available cases-GEE	0.325	0.306	0.300	0.292	0.353	0.334	0.327	0.319	0.394	0.373	0.367	0.358	0.463	0.441	0.434	0.424
	Imputation-GLM	0.231	0.211	0.204	0.193	0.235	0.222	0.215	0.203	0.248	0.239	0.232	0.219	0.296	0.272	0.263	0.248
	available cases-GLM	0.243	0.221	0.213	0.201	0.248	0.233	0.225	0.212	0.263	0.253	0.244	0.230	0.314	0.288	0.277	0.262
	intent to treat-GEE	0.249	0.231	0.227	0.225	0.272	0.252	0.247	0.243	0.304	0.281	0.275	0.269	0.359	0.333	0.325	0.315
SD	Imputation-GEE	0.315	0.297	0.290	0.282	0.345	0.324	0.317	0.310	0.392	0.362	0.355	0.347	0.457	0.435	0.426	0.417
	available cases-GEE	0.316	0.297	0.291	0.283	0.346	0.326	0.318	0.311	0.394	0.364	0.357	0.349	0.459	0.435	0.428	0.419
	Imputation-GLM	0.261	0.239	0.231	0.221	0.285	0.261	0.253	0.242	0.319	0.292	0.283	0.271	0.382	0.351	0.339	0.325
	available cases-GLM	0.260	0.238	0.230	0.220	0.285	0.260	0.252	0.241	0.319	0.292	0.283	0.271	0.382	0.350	0.339	0.325
	intent to treat-GEE	0.247	0.228	0.224	0.222	0.270	0.249	0.244	0.240	0.301	0.278	0.272	0.266	0.356	0.330	0.322	0.312
RSE	Imputation-GEE	1.023	1.026	1.027	1.028	1.017	1.022	1.024	1.023	0.997	1.022	1.024	1.021	1.005	1.010	1.010	1.008
	available cases-GEE	1.027	1.029	1.031	1.031	1.020	1.025	1.027	1.025	1.001	1.026	1.027	1.025	1.009	1.014	1.014	1.011
	Imputation-GLM	0.884	0.883	0.881	0.872	0.823	0.851	0.849	0.839	0.778	0.820	0.818	0.807	0.775	0.775	0.775	0.763
	available cases-GLM	0.933	0.930	0.926	0.913	0.871	0.898	0.893	0.880	0.825	0.866	0.863	0.849	0.821	0.819	0.811	0.806

	ρ = 0.01				ρ = 0.025				ρ = 0.05				ρ = 0.1			
intent to treat-GEE	0.959	0.971	0.986	1.021	0.954	0.966	0.977	1.005	0.954	0.964	0.974	0.994	0.940	0.950	0.959	0.971
CP(%)																
Imputation-GEE	94.6	95.0	95.1	94.8	94.5	94.8	95.0	94.5	94.3	94.4	94.5	94.6	93.3	94.0	93.9	93.9
available cases-GEE	94.5	95.0	94.9	94.7	94.5	94.6	95.0	94.6	94.2	94.5	94.5	94.7	93.5	94.0	94.0	93.9
Imputation-GLM	92.7	92.8	92.7	92.5	91.5	91.6	91.4	91.1	89.8	89.9	90.0	89.4	87.9	88.1	87.9	87.5
available cases-GLM	94.1	94.1	94.0	93.8	93.0	93.0	93.0	92.7	91.8	91.9	91.8	91.2	89.9	90.3	89.9	89.3
intent to treat-GEE	94.0	94.6	95.1	95.7	94.0	94.5	94.7	95.3	94.1	94.4	94.5	95.1	93.8	94.1	94.3	94.6

Bias=mean estimate — true value, Relative bias= bias/true value, ASE=average of estimated standard errors, SD=standard deviation of the estimates, RSE = ASE/SD, CP=95% coverage probability, ρ = intracluster correlation coefficient, β = log RR of interest, RR=risk ratio, OR=odds ratio.

Table 2.

Simulation results for the scenario with a trial effect.

		ρ=0.01				ρ=0.025				ρ=0.05				ρ=0.1			
True β (log RR)		0	0.337	0.470	0.693	0	0.337	0.470	0.693	0	0.337	0.470	0.693	0	0.337	0.470	0.693
Bias	Imputation-GLM	0.004	0.008	0.010	0.014	0.005	0.010	0.012	0.016	0.004	0.011	0.014	0.018	0.003	0.013	0.018	0.023
	available cases-GLM	−0.002	0.003	0.005	0.008	−0.001	0.004	0.007	0.010	−0.002	0.005	0.008	0.013	−0.003	0.007	0.012	0.018
	intent to treat-GEE	−0.002	0.003	0.005	0.009	−0.002	0.004	0.007	0.011	−0.002	0.005	0.009	0.014	−0.003	0.007	0.013	0.018
Relative Bias	Imputation-GLM	-	0.025	0.022	0.020	-	0.029	0.026	0.023	-	0.032	0.029	0.026	-	0.039	0.038	0.033
	available cases-GLM	-	0.008	0.010	0.012	-	0.012	0.014	0.015	-	0.015	0.018	0.019	-	0.022	0.026	0.025
	intent to treat-GEE	-	0.008	0.011	0.012	-	0.012	0.015	0.016	-	0.015	0.018	0.020	-	0.022	0.027	0.027
ASE	Imputation-GLM	0.264	0.241	0.234	0.221	0.287	0.263	0.255	0.241	0.320	0.294	0.285	0.270	0.383	0.352	0.341	0.323
	available cases-GLM	0.264	0.241	0.233	0.221	0.287	0.263	0.255	0.241	0.320	0.294	0.285	0.269	0.383	0.352	0.341	0.323
	intent to treat-GEE	0.263	0.245	0.241	0.238	0.285	0.265	0.260	0.255	0.315	0.293	0.287	0.280	0.367	0.341	0.334	0.325
SD	Imputation-GLM	0.236	0.215	0.208	0.197	0.248	0.227	0.219	0.207	0.267	0.244	0.236	0.223	0.301	0.275	0.266	0.252
	available cases-GLM	0.249	0.226	0.219	0.206	0.262	0.239	0.230	0.217	0.282	0.257	0.249	0.234	0.319	0.291	0.282	0.266
	intent to treat-GEE	0.265	0.242	0.234	0.222	0.288	0.264	0.256	0.242	0.321	0.295	0.286	0.270	0.383	0.352	0.341	0.323
RSE	Imputation-GLM	0.891	0.891	0.890	0.888	0.863	0.861	0.859	0.858	0.832	0.828	0.825	0.826	0.785	0.782	0.781	0.778
	available cases-GLM	0.941	0.939	0.936	0.931	0.912	0.908	0.904	0.901	0.881	0.875	0.871	0.870	0.832	0.828	0.826	0.822
	intent to treat-GEE	0.993	1.012	1.029	1.073	0.988	1.003	1.016	1.056	0.982	0.992	1.004	1.038	0.958	0.969	0.980	1.004
CP(%)	Imputation-GLM	92.8	92.7	92.9	93.1	91.6	91.4	91.9	91.8	90.5	90.5	90.7	90.7	88.4	88.1	88.4	88.3
	available cases-GLM	94.1	94.2	94.2	94.3	93.3	93.3	93.6	93.3	92.3	92.3	92.5	92.3	90.6	90.3	90.3	90.4
	intent to treat-GEE	95.3	95.7	95.9	96.7	96.5	95.4	95.8	96.5	95.1	95.1	95.4	96.3	94.6	94.7	95.0	95.8

Bias=mean estimate — true value, Relative bias= bias/true value, ASE=average of estimated standard errors, SD=standard deviation of the estimates, RSE = ASE/SD, CP=95% coverage probability, ρ = intracluster correlation coefficient, β= log RR of interest, RR=risk ratio, OR=odds ratio

Table 3.

RECRUIT data analysis results for the intervention effect.

Model	Coefficient	SE	95%CI	p-value	Interaction
Imputation-GEE	0.379	0.312	−0.234, 0.991	0.226	No
available cases-GEE	0.239	0.316	−0.381, 0.858	0.450	No
Imputation-GLM	0.319	0.261	−0.192, 0.830	0.221	No
available cases-GLM	0.201	0.264	−0.316, 0.718	0.446	No
intent to treat-GEE	0.290	0.254	−0.207, 0.787	0.253	Yes

AC=available cases, GEE= generalized estimating equations, GLM = generalized linear model, SE = standard error (without correction), CI= confidence interval, Interaction = component interaction from Equation (3). Coefficient corresponds to $\log(OR)$ in the top 2 rows and $\log(RR)$ in the bottom 3 rows.