Title Page

Title: Cluster Designs to Assess the Prevalence of Acute Malnutrition by LQAS: A Validation Study by Computer Simulation

Authors: Casey Olives[*], Marcello Pagano[*], Megan Deitchler[†], Bethany Hedt[*], Kari Egge[§], Joseph J Valadez[‡]

Corresponding author: Casey Olives

Harvard School of Public Health, Department of Biostatistics,

655 Huntington Avenue, SPH2, 4th Floor

Boston, Massachusetts 02115

colives@hsph.harvard.edu

(612) 501-7947

Sources of support: Financial support for this project was provided to Casey Olives and the Academy for Educational Development (AED), Food and Nutrition Technical Assistance (FANTA) Project by the Office of Health, Infectious Diseases and Nutrition of the Bureau for Global Health at the US Agency for International Development (USAID), under terms of Cooperative Agreement No HRN-A-00-98-00046-00. Marcello Pagano and Bethany Hedt received support from the National Institutes of Health (NIH) grants R01EB006195 and T32AI07358.

Word count for manuscript: 3905 (not including text for table placement)

Word count for abstract: 136 words

Number of tables and figures: 6 tables

[*] Harvard School of Public Health, Department of Biostatistics, Boston, MA, USA; [†] Academy for Educational Development, FANTA Project, Washington DC, USA; [§] Independent Consultant; [‡] The World Bank, Washington DC, USA

Number of appendices: 1 appendix

ABSTRACT

Traditional Lot Quality Assurance Sampling (LQAS) methods require random sampling to guarantee valid results. However, cluster sampling has been proposed to reduce the number of random starting points. This study uses simulations to examine the classification error of two such designs, a 67x3 (67 clusters of 3 observations) and a 33x6 (33 clusters of 6 observations) sampling scheme to assess the prevalence of Global Acute Malnutrition (GAM). Further, we explore the use of a 67x3 sequential sampling scheme for LQAS classification of GAM prevalence. Results indicate that for independent clusters with moderate intra-cluster correlation for the GAM outcome, the three sampling designs maintain approximate validity for LQAS analysis. Sequential sampling can substantially reduce the average sample size required for data collection. The presence of inter-cluster correlation can impact dramatically the classification error associated with LQAS analysis.

Key Words: Acute malnutrition, emergency, lot quality assurance sampling, LQAS, sequential sampling, wasting

INTRODUCTION

In the last twenty years, development organizations working in international health have increasingly adopted Lot Quality Assurance Sampling (LQAS) to assess health care parameters. Nearly all of the 805 studies identified in a recent review of LQAS implemented between January 1984 and December 2004 employed traditional LQAS sampling methods (Robertson 2006), in which a Simple Random Sample (SRS) is used for data collection. The exceptions are studies in which a two-stage LQAS design was combined with cluster sampling to assess neonatal tetanus eradication (WHO 2001, WHO 2002, WHO 2004), and a study in which small clusters instead of a SRS were used to assess the prevalence of Global Acute Malnutrition (GAM) by LQAS analysis methods (Deitchler 2007, Deitchler 2008).

In the international health setting, small sample sizes (e.g., n=19) have often been used for LQAS assessment of service provision indicators (Valadez 2003). The small samples sizes have meant that LQAS has been feasible for use by local managers (Valadez 1991). However, use of LQAS for assessment of anthropometric indicators requires large sample sizes due to the increased precision needed for hypothesis testing. To use SRS with large sample sizes means an increase in time and cost, as data collection for each observation in the sample can require travel to a different site. Sampling observations in batches, or clusters, is an alternative method which reduces the number of site visits needed to complete data collection. However, if the observations within each cluster are highly correlated with respect to the outcome being assessed, cluster sampling leads to increased misclassification with the LQAS analysis method. On the other hand,

cluster sampling could be a viable option if it does not undermine the validity of the independence assumption for hypothesis testing, as required by LQAS.

Deitchler *et al* field tested both a 67x3 and 33x6 cluster design (67 clusters of size 3 and 33 clusters of size 6, respectively) for LQAS assessment of GAM prevalence in the Siraro woreda of Ethiopia in 2003 and in the Administrative Units of Fur Baranga and Habila in West Darfur in 2008 (Deitchler 2007, Deitchler 2008). The use of a 67x3 sequential sampling design was also investigated in the Ethiopia study. In comparison to the 67x3 and 33x6 design, the sequential design allowed for a reduction in the total sample size required to assess the prevalence of GAM by LQAS analysis methods (Deitchler 2007). Similar sequential designs have been used for categorizing HIV drug resistance (Bennett 2006). However, those designs relied on SRS for validity.

The current study uses computer simulations to assess the validity of the small cluster approach used to assess GAM prevalence (Deitchler 2007). The principal sampling strategy uses a cluster model to minimize the number of random sites to visit. We focus on a 67x3 and a 33x6 cluster design as these were the designs tested in Ethiopia and Sudan. Additionally, we develop and investigate a second strategy which applies a sequential sampling scheme to the 67x3 cluster design. Here, we use more robust statistical assumptions for the sequential design than had been applied to the work in Ethiopia, in order to improve the design.

METHODS

Traditional LQAS Methods

LQAS inference uses the binomial approximation to the hypergeometric distribution to test whether the prevalence of the health parameter of interest is exhibited

at a proportion greater than or equal to some pre-specified threshold P_o. This is equivalent to the hypothesis test:

$$H_o: P \geq P_o \text{ vs. } H_a: P < P_o$$

where P is the true prevalence in the population and P_o, the upper threshold, is the prevalence level the data are tested against. In the case of GAM, P_o represents an unacceptable level of acute malnutrition in the population. It is chosen to reflect the prevalence at which a population would be considered a priority for humanitarian intervention. The null hypothesis is rejected if the number of individuals in the sample exhibiting GAM, *s,* is less than or equal to an *a priori* defined critical, *d* ($s \leq d$). This critical value is often referred to as the decision rule in LQAS literature (Valadez 1991). In addition, LQAS requires that we define a lower threshold, P_a. The lower threshold reflects the GAM prevalence at which the population would not be considered a priority intervention.

As with any hypothesis test, there is an alpha (α) and beta (β) error associated with LQAS. The alpha error is the highest probability that the null hypothesis is incorrectly rejected. In the case of GAM, this would mean concluding that the assessment area does not have a high level of acute malnutrition when in fact it does. This probability is controlled for at the upper threshold.

$$\alpha \geq \sum_{i=0}^{d} \frac{n!}{(n-i)!i!} P_o^i Q_o^{n-i}, \text{ where } Q_o = 1 - P_o .$$

The beta error is the highest probability that we incorrectly fail to reject the null hypothesis. This would mean concluding that the assessment area does have a high level

of acute malnutrition when in fact it does not. The beta error is controlled for at the lower threshold.

$$\beta \geq \sum_{i=d+1}^{n} \frac{n!}{(n-i)!i!} P_a^i Q_a^{n-i}, \text{ where } Q_a = 1 - P_a .$$

The critical value is chosen to approximate the desired alpha and beta given the upper and lower thresholds, and the sample size. In practice, it is difficult to attain the alpha and beta errors exactly due to the discrete nature of the binomial distribution. Further, more than one critical value can achieve the specified constraints. The actual error probabilities for a specific sample size, and upper and lower thresholds therefore depend on the critical value, *d*, chosen.

In this study, we investigate the upper and lower thresholds field tested in Ethiopia and Sudan (Deitchler 2007, Deitchler 2008). Three couplets (i.e., upper/lower threshold pairs) are investigated: the upper thresholds of 10%, 15%, and 20%, and the respective lower thresholds of 5%, 10%, and 15%. The 10%/5% and 15%/10% couplet are of primary concern as these are the benchmarks most commonly used by humanitarian agencies to assess GAM (FAO/FSAU 2006). The 20%/15% couplet is of secondary consideration as GAM prevalence above 20% is fairly rare, even in emergency settings (Medecins Sans Frontieres 1995).

For each upper and lower threshold couplet, we determined the critical value subject to the constraints of an alpha error of approximately 0.10 and a beta error of approximately 0.20 for samples of sizes 198 (33x6) and 201 (67x3). Table 1 gives the sample size, critical value, and associated alpha and beta errors for each upper and lower threshold couplet when a traditional SRS is used for data collection. For the 10%/5% couplet, a critical value of 13 meets the constraints of alpha $\leq$ 0.10 and beta $\leq$ 0.20 for

both sample sizes. For the 15%/10% couplet, the desired error limits are approximately maintained for a critical value of 23. For the 20%/15% couplet, no critical value attains or closely approximates the desired alpha and beta constraints for samples of size 198 and 201. The critical value 33 minimizes the total error for a sample of size 198 and the critical value 34 minimizes the total error for a sample of size 201. We chose to use the critical value 33 for this couplet, with a corresponding alpha of 0.138 and beta of 0.221.

[Table 1 here]

LQAS Methods for Sequential Cluster Designs

In this section we investigate a sequential LQAS cluster design to test the same three null hypotheses as above. The sequential cluster design differs from traditional LQAS as a decision can be made to reject or accept the null hypothesis after each individual cluster is observed. In a k x m sequential sampling design, there are at most k stages of sampling. At each stage, m sampling elements are observed for a maximum of n possible observations. At the i^{th} stage of sampling, we define an acceptance rule, a_i, a rejection rule, r_i, and the cumulative number of outcomes, s_i (in our application, an outcome is a child exhibiting GAM). If $s_i \geq a_i$, then we conclude that the prevalence of GAM is greater than or equal to P_o, and sampling stops. Likewise, if $s_i \leq r_i$, then we conclude the prevalence is less than P_o, and sampling stops. Otherwise, if $r_i < s_i < a_i$, sampling proceeds to the next stage. If no decision is made by the time the final (k^{th}) stage of sampling is reached, then a decision is made to reject if $s_n \leq \frac{a_n + r_n}{2}$ and to accept if $s_n > \frac{a_n + r_n}{2}$.

Wald outlined the calculation of LQAS critical values at each stage of a sequential design applied to observations selected by SRS (Wald 1947). These critical values are linear in the individual observations. We adapt this theory to accommodate clusters of size m (m>1), under the assumption that observations within each cluster are independent. Namely, define

$$r_i = \frac{\log\frac{1-\beta}{\alpha} + m \cdot i \log\frac{Q_0}{Q_a}}{\log\frac{P_a Q_0}{P_0 Q_a}} \text{ and } a_i = \frac{\log\frac{\beta}{1-\alpha} + m \cdot i \log\frac{Q_0}{Q_a}}{\log\frac{P_a Q_0}{P_0 Q_a}},$$

where alpha and beta refer to the target classification errors. These critical values are linear in the sampling stage, and thus reflect a cluster sampling design.

One of the benefits of sequential designs is the potential for reduction of the overall sample size required for data collection. With respect to the outcome of acute malnutrition, a reduction in sample size could lead to a more rapid response to an emergency situation. The average sample number (ASN), or the average number of clusters sampled to reject or accept the null hypothesis, characterizes this reduction. The average sample size is equal to the number of sampling elements per cluster times the ASN (m x ASN) and is given by the following formula

$$ASN = \left\lceil \sum_{i=1}^{k} \mathrm{i}\{\Pr(\text{Reject at stage i}) + \Pr(\text{Accept at stage i})\} \right\rceil,$$

where $f(x) = \lceil(\cdot)\rceil$ is the next largest integer function (Aroian 1976).

The Wald critical values rely on the assumption that the number of possible observations is unbounded. However, in virtually all applications, this is not the case. When the number of possible observations is bounded, the design is said to be truncated. The use of Wald critical values in truncated sequential designs does not generally yield

the appropriate alpha and beta (Wald 1947). Aroian suggests treating a sequential sample as a random walk to directly calculate the classification error for a truncated design (Aroian 1965, 1976). We used Aroian's direct method to calculate the true classification error for a range of sequential designs varied over the parameter space of α and β to arrive within the desired targets of classification error.

Here we investigate a 67x3 sequential sampling design with application to the three upper/lower threshold couplets of interest. In terms of the above notation, k=67, m=3, n=201, and the upper bound for the ASN is 67. For each upper and lower threshold couplet, we determine the acceptance and rejection rules using Wald theory. We calculated critical values for a range of alpha and beta errors around the target levels of 0.10 and 0.20, respectively. The final critical values chosen are those that yield the true alpha and beta nearest to the desired levels as calculated using the direct method. For both the 15%/10% and 20%/15% couplets, we were unable to find a design that yielded the desired alpha and beta targets. For these couplets we selected the design that jointly minimized the alpha and beta errors. For the 10%/5% couplet we expect an alpha of 0.10 and a beta of 0.16. For the 15%/10% couplet, we expect an alpha of 0.10 and a beta of 0.24. And for the 20%/15% couplet, we expect an alpha of 0.17 and a beta of 0.22. The critical values for each couplet are given in Table 2.

[Table 2 here]

Simulation Validation of Cluster Designs for LQAS Analysis

One key assumption in LQAS theory is that a SRS is used for data collection of binary outcomes (Hoshaw-Woodard 2001, Valadez 1991). Cluster sampling often results in an intra-cluster correlation (correlation between subjects within the same cluster with

respect to the outcome of interest). For the cluster designs of concern here, intra-cluster correlation could result from within household correlation (i.e., correlation of GAM among multiple children sampled in one household) or as correlation of GAM among multiple households sampled within the same cluster (Deitchler 2007). Inter-cluster correlation (correlation between subjects in different clusters) is also possible although this is likely to be minimal for acute malnutrition, and can be assumed to be less than or equal to the intra-cluster correlation (Fenn 2004, Reed 2000). Validation of the 67x3, 33x6 and sequential cluster design requires assessing the effect of these potential correlations on the alpha and beta errors associated with LQAS hypothesis testing.

For the cluster sampling techniques investigated here, we assume that intra-cluster correlation is homogeneous and non-negative. Inter-cluster correlation is also assumed homogeneous and non-negative, and less than or equal to the intra-cluster correlation. This study confines the investigation to the inter- and intra-cluster correlations of 0.00, 0.05, 0.10, 0.15, 0.20, and 0.25, because these provide a broad set of acceptable alternatives. Kalton's work on cluster sampling suggests that intra-cluster correlation is usually less than 0.15 for most indicators (Kalton 1983). The well-documented multiple causes of malnutrition along with the age dependence vulnerability of children to acute malnutrition (Shrimpton 2001, UNICEF 1990) further suggest a low intra-cluster correlation to be likely. Moreover, a review of Demographic and Health Surveys conducted in 46 developing countries reported intra-cluster correlations of less than 0.10 for acute malnutrition in 90 percent of the countries studied (Fenn 2004) and intra-cluster correlations of less than 0.05 for acute malnutrition in field applications of the 33x6 and 67x3 designs in Sudan (Deitchler 2008). With these considerations in mind, we

anticipate intra-cluster correlations using the three cluster sampling schemes used here to be less than 0.05 in most field settings. Intra-cluster correlation levels equal to and above 0.05 for GAM, while unlikely, are investigated in this study in order to understand the effect of unusually high levels of intra-cluster correlation on LQAS classification error for these designs.

Simulation Methods

To reproduce the correlation structure arising from the 67x3 and 33x6 sampling schemes and the 67x3 sequential sampling scheme, it is necessary to generate correlated binary vectors **D** such that $\mathbf{D} \sim (\mathbf{P}, \Sigma)$ where **P** is the $n \times 1$ mean vector of P's and **Σ** is the $n \times n$ variance-covariance matrix describing the correlation structure. For each couplet, samples of size 201 and 198 were generated under the various inter- and intra-cluster correlation constraints. This procedure was repeated 10,000 times for each couplet and inter/intra-cluster correlation pair for each design. All simulations were performed using the statistical package R v.2.6.0. The simulation methodology is described in detail in the Appendix.

RESULTS

Cluster Sampling Strategy: The 67x3 and 33x6 Designs

Tables 3-5 contain the results of the simulations for the 67x3 and 33x6 designs along with estimated standard error. As expected, those simulations with an inter-cluster and intra-cluster correlation equal to zero for GAM demonstrate alpha and beta errors approximately equal to the binomial alpha and beta errors presented in Table 1, as this situation corresponds to a SRS.

In the correlated samples, the least effect on alpha and beta error occurs when the inter-cluster correlation equals zero. For example, in the case of the 67x3 design, if the inter-cluster correlation is equal to zero and the intra-cluster correlation is less than or equal to 0.25, the 10%/5% couplet maintains the desired error limits of alpha ≤0.10 and beta ≤0.20 (Table 3). With intra-cluster correlations less than 0.10 the 15%/10% couplet performs approximately within the desired error limits (Table 4). Although the 20%/15% couplet has errors slightly above the desired limits at this correlation level, these were expected from the outset as the targets were untenable under SRS (Tables 5).

In the case of the 33x6 design, assuming an inter-cluster correlation equal to zero, the 10%/5% couplet conforms to the desired error limits of alpha ≤ 0.10 and beta ≤ 0.20 for intra-cluster correlations up to 0.10 (Table 3); the 15%/10 couplet conforms approximately to the desired error limits when the intra-cluster correlation equals zero, and, as expected, the 20%/15% couplet does not attain the desired performance (Table 4, Table 5).

In cases where both the inter- and intra-cluster correlation are greater than zero, there is a substantial increase in the alpha error for both the 67x3 and 33x6 LQAS designs; although the beta error is less affected. This result suggests that when inter-cluster correlation is greater than zero, larger samples may be required to attain the desired alpha and beta levels. Upon use of random methods for selection of clusters to sample, it is, however, reasonable to assume an inter-cluster correlation equal to zero for LQAS assessment of GAM prevalence with the 67x3 or 33x6 design.

[Table 3-5 here]

Sequential Sampling Strategy: The 67x3 Sequential Design

Table 6 shows the simulation results for the 67x3 sequential design. As expected, when inter- and intra-cluster correlations are equal to zero, the results closely approximate the alpha and beta errors calculated under SRS. Additionally, the least effect on the alpha and beta errors occurs in simulations where the inter-cluster correlation is equal to zero. Assuming an inter-cluster correlation equal to zero and an intra-cluster correlation as high as 0.25, the alpha error is ≤ 0.16 and the beta error is ≤ 0.25 for the 10%/5% couplet. For the 15%/10% couplet, the alpha and beta errors are ≤ 0.14 and ≤ 0.30, respectively. The errors for the 20%/15% couplet are slightly higher with the alpha error ≤ 0.211 and the beta error ≤ 0.284.

For all simulated sequential samples, the ASN is substantially less than the maximum of 67. For the 10%/5% couplet, the maximum ASN is approximately 23 under the null and 34 under the alternative (n=69 and n=102, respectively). For the 15%/10% couplet, the maximum ASN is approximately 35 under the null and 50 under the alternative (n=105 and n=150, respectively), and for the 20%/15% couplet, the maximum ASN is 40 under the null and 47 under the alternative (n=120 and n=141, respectively). This result suggests that the 67x3 sequential design could be utilized to decrease the total number of clusters sampled, and thus the overall sample size required for data collection. A slightly elevated level of misclassification, beyond alpha ≤ 0.10 and beta ≤ 0.20, would need to be acceptable for the 15%10% and 20%/15% couplets, but in cases where uncorrelated clusters and a low intra-cluster correlation can be assumed for GAM, the design may be appropriate to use.

[Tables 6 here]

DISCUSSION

This study uses computer simulations to assess three cluster sampling schemes that were field-tested in Ethiopia to assess the prevalence of GAM by LQAS analysis methods (Deitchler 2007). The simulation results show the 67x3 and 33x6 cluster designs conform to the desired error limits of alpha ≤ 0.10 and beta ≤ 0.20 for the 10%/5% and 15%/10% couplet at numerous intra-cluster correlation levels when the inter-cluster correlation is equal to zero. It stands to reason that the 67x3 design conforms to the desired alpha and beta limits at higher intra-cluster correlation levels than the 33x6 design for both the 10%/5% and 15%/10% couplet. For the 10%/5% couplet, the 67x3 design maintains the desired error limits when the inter-cluster correlation is zero and the intra-cluster correlation is as high as 0.25. For the 15%/10% couplet, the 67x3 design maintains alpha and beta approximately equal to 0.10 and 0.20 when the inter-cluster correlation is equal to zero and the intra-cluster correlation is less than 0.10. Therefore, when clusters can be assumed independent and correlation within the clusters can be assumed to be less than 0.10, the 67x3 design can be an effective method to reduce the number of sites that would otherwise need to be visited by a SRS of the same size. In cases where the clusters can be assumed independent and correlation within the clusters less than 0.15, the 33x6 design can also be an effective method for assessment of GAM, allowing for LQAS inference within the desired error limits for the 10%/5% couplet. In order to maintain the same error limits for the 15%/10% couplet with the 33x6 design, there can be no intra-cluster correlation. Intuitively, we expect the 67x3 design to perform within the desired error limits at higher levels of intra-cluster correlation than the 33x6 design, as smaller clusters would suffer less from intra-cluster correlation.

The simulation results for the 67x3 sequential design indicate a potential time advantage over the 67x3 and 33x6 cluster designs because the total sample required for data collection are likely to be smaller. However, notwithstanding two exceptions, the simulation results indicate that the alpha and beta errors for all inter- and intra-correlation levels, for each threshold couplet, exceed the desired alpha and beta limits of 0.10 and 0.20, respectively. Use of the sequential design with these maximal sample sizeswould therefore be recommended only when it is acceptable to deviate slightly from the above stated limits of alpha and beta.

The results of this simulation study demonstrate that information about the intra-cluster correlation of GAM is needed in order to reliably use the 67x3, 33x6, and 67x3 sequential sampling designs for LQAS assessment of the prevalence of GAM. Fenn's review of Demographic and Health Surveys suggests that most field settings will have an acute malnutrition intra-cluster correlation <0.10 (Fenn 2004) while Deitchler's field application of the 67x3 and 33x6 designs in Sudan suggests an intra-cluster correlation <0.05 to be likely (Deitchler 2008) . These studies provide useful information about the plausible upper limit of intra cluster correlation for acute malnutrition. However, investigators rarely know in advance the exact intra-cluster correlation that exists in a field setting where a malnutrition assessment will be conducted. Until there is more clarity about the conditions in which the upper levels of 0.05-0.10 intra-cluster correlation of GAM would be expected, or possibly exceeded, investigators desiring strict adherence to the stated LQAS error limits of alpha ≤ 0.10 and beta ≤ 0.20 may prefer to err on the side of caution by using the better performing 67x3 design, while investigators who require data rapidly may prefer instead to use the 67x3 sequential design. Finally,

those investigators seeking a balance between limited classification error and potential expediency of data collection may find the 33x6 design to best meet their data requirements.

The results of this study support use of the cluster designs used in Ethiopia and Sudan (Deitchler 2007, Deitchler 2008) for detecting threshold levels of GAM by LQAS analysis methods. Further, the findings from this study provide useful information to investigators who need to decide which design (i.e., 67x3, 33x6, or 67x3 sequential design) best suits their analytic needs, with respect to expediency of data collection, and desired limits of classification error. The cluster sampling schemes analyzed here offer both time-efficient and statistically valid alternatives to the conventional methodology for assessment of acute malnutrition in emergency settings.

ACKNOWLEDGMENTS

The authors are grateful to Stacy Hoshaw-Woodard who conducted the preliminary simulations, which led to the development of the 33x6, 67x3, and sequential design in 2003. The authors acknowledge the contributions Gilles Bergeron, Hedwig Deconinck, Soledad Fernandez, and Mary Hennigan made to earlier studies, in which the designs were field tested and those results reported. Gilles Bergeron and Nadra Franklin provided helpful comments on an earlier draft of this manuscript. Financial support for this project was provided to Casey Olives and the Academy for Educational Development (AED), Food and Nutrition Technical Assistance (FANTA) Project by the office of Health, Infectious Diseases and Nutrition of the Bureau for Global Health at the US Agency for International Development (USAID), under terms of Cooperative Agreement No HRN-A-00-98-00046-00. Marcello Pagano and Bethany Hedt received support from

the National Institutes of Health (NIH) grants R01EB006195 and T32AI07358. The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of USAID or NIH.

[Appendix]

Sample Size	10%/5%			15%/10%			20%/15%		
	d	Alpha	Beta	d	Alpha	Beta	d	Alpha	Beta
201	12	0.031	0.208	22	0.061	0.279	32	0.085	0.315
	13	0.054	0.134	23	0.091	0.209	33	0.117	0.250
	14	0.089	0.081	24	0.131	0.151	34	0.157	0.193
198	12	0.035	0.194	22	0.072	0.255	32	0.101	0.283
	13	0.062	0.123	23	0.106	0.188	33	0.138	0.221
	14	0.101	0.073	24	0.150	0.134	34	0.183	0.169

Table 1: LQAS Alpha and Beta Errors, and Critical values for Samples Sizes 198 and 201 for Three Upper and Lower Thresholds Assuming SRS.

Table 2: Rejection (*r*) and Acceptance (*a*) Rules for 67x3 Sequential Design for Three Upper and Lower Threshold Couplets Assuming Complete Independence.*

Stage	Couplet (P_o/P_a)					
	10%/5%		15%/10%		20%/15%	
	r	*a*	*r*	*a*	*r*	*a*
1	ND	3	ND	4	ND	5
2	ND	3	ND	5	ND	6
3	ND	3	ND	5	ND	7
4	ND	3	ND	5	ND	7
5	ND	4	ND	6	ND	8
6	ND	4	ND	6	ND	8
7	ND	4	ND	6	ND	9
8	ND	4	ND	7	ND	9
9	ND	4	ND	7	ND	10
10	ND	5	ND	8	ND	10
11	ND	5	ND	8	ND	11
12	ND	5	ND	8	0	11
13	ND	5	ND	9	0	12
14	ND	6	ND	9	1	12
15	0	6	ND	9	1	13
16	0	6	ND	10	2	13
17	0	6	0	10	2	14
18	0	6	0	11	3	14
19	1	7	1	11	3	15
20	1	7	1	11	4	15
21	1	7	1	12	4	16
22	1	7	2	12	5	16
23	1	8	2	12	6	17
24	2	8	2	13	6	17
25	2	8	3	13	7	18
26	2	8	3	14	7	19
27	2	8	4	14	8	19
28	2	9	4	14	8	20
29	3	9	4	15	9	20
30	3	9	5	15	9	21
31	3	9	5	15	10	21
32	3	9	5	16	10	22
33	4	10	6	16	11	22
34	4	10	6	16	11	23
35	4	10	6	17	12	23
36	4	10	7	17	12	24
37	4	11	7	18	13	24
38	5	11	8	18	13	25
39	5	11	8	18	14	25
40	5	11	8	19	14	26

41	5	11	9	19	15	26
42	6	12	9	19	15	27
43	6	12	9	20	16	27
44	6	12	10	20	16	28
45	6	12	10	21	17	28
46	6	13	11	21	18	29
47	7	13	11	21	18	30
48	7	13	11	22	19	30
49	7	13	12	22	19	31
50	7	13	12	22	20	31
51	7	14	12	23	20	32
52	8	14	13	23	21	32
53	8	14	13	24	21	33
54	8	14	14	24	22	33
55	8	14	14	24	22	34
56	9	15	14	25	23	34
57	9	15	15	25	23	35
58	9	15	15	25	24	35
59	9	15	15	26	24	36
60	9	16	16	26	25	36
61	10	16	16	26	25	37
62	10	16	16	27	26	37
63	10	16	17	27	26	38
64	10	16	17	28	27	38
65	11	17	18	28	27	39
66	11	17	18	28	28	39
67	14	15	23	24	34	35

* ND signifies that no decision is made and sampling continues.

Table 3: Alpha and Beta Errors for the 10%/5% Couplet w/ Varied Intra- and Inter-Cluster Correlation and *d*=13.

Correlation		67 x 3		33 x 6	
Inter	Intra	Alpha (s.e.)	Beta (s.e.)	Alpha (s.e.)	Beta (s.e.)
0.00	0.00	0.054 (0.002)	0.136 (0.003)	0.061 (0.002)	0.124 (0.003)
0.00	0.05	0.064 (0.002)	0.144 (0.004)	0.083 (0.003)	0.147 (0.004)
0.05		0.389 (0.005)	0.247 (0.004)	0.402 (0.005)	0.253 (0.004)
0.00	0.10	0.071 (0.003)	0.159 (0.004)	0.103 (0.003)	0.161 (0.004)
0.05		0.390 (0.005)	0.263 (0.004)	0.393 (0.005)	0.246 (0.004)
0.10		0.491 (0.005)	0.234 (0.004)	0.488 (0.005)	0.245 (0.004)
0.00	0.15	0.077 (0.003)	0.162 (0.004)	0.123 (0.003)	0.188 (0.004)
0.05		0.389 (0.005)	0.246 (0.004)	0.394 (0.005)	0.246 (0.004)
0.10		0.473 (0.005)	0.239 (0.004)	0.495 (0.005)	0.237 (0.004)
0.15		0.552 (0.005)	0.220 (0.004)	0.551 (0.005)	0.216 (0.004)
0.00	0.20	0.086 (0.003)	0.172 (0.004)	0.141 (0.003)	0.197 (0.004)
0.05		0.389 (0.005)	0.258 (0.004)	0.407 (0.005)	0.245 (0.004)
0.10		0.487 (0.005)	0.236 (0.004)	0.491 (0.005)	0.231 (0.004)
0.15		0.550 (0.005)	0.230 (0.004)	0.550 (0.005)	0.221 (0.004)
0.20		0.592 (0.005)	0.208 (0.004)	0.599 (0.005)	0.205 (0.004)
0.00	0.25	0.097 (0.003)	0.179 (0.004)	0.163 (0.004)	0.205 (0.004)
0.05		0.400 (0.005)	0.256 (0.004)	0.409 (0.005)	0.255 (0.004)
0.10		0.478 (0.005)	0.236 (0.004)	0.491 (0.005)	0.239 (0.004)
0.15		0.552 (0.005)	0.225 (0.004)	0.548 (0.005)	0.215 (0.004)
0.20		0.592 (0.005)	0.213 (0.004)	0.595 (0.005)	0.198 (0.004)
0.25		0.628 (0.005)	0.195 (0.004)	0.635 (0.005)	0.187 (0.004)

Table 4: Alpha and Beta Errors for the 15%/10% Couplet w/ Varied Intra- and Inter-Cluster Correlation and *d*=23.

Correlation		67 x 3		33 x 6	
Inter	Intra	Alpha (s.e.)	Beta (s.e.)	Alpha (s.e.)	Beta (s.e.)
0.00	0.00	0.095 (0.003)	0.211 (0.004)	0.107 (0.003)	0.185 (0.004)
0.00	0.05	0.096 (0.003)	0.215 (0.004)	0.137 (0.003)	0.210 (0.004)
0.05		0.407 (0.005)	0.317 (0.005)	0.420 (0.005)	0.310 (0.005)
0.00	0.10	0.111 (0.003)	0.220 (0.004)	0.155 (0.004)	0.230 (0.004)
0.05		0.410 (0.005)	0.324 (0.005)	0.426 (0.005)	0.314 (0.005)
0.10		0.479 (0.005)	0.312 (0.005)	0.488 (0.005)	0.304 (0.005)
0.00	0.15	0.115 (0.003)	0.230 (0.004)	0.173 (0.004)	0.241 (0.004)
0.05		0.393 (0.005)	0.327 (0.005)	0.418 (0.005)	0.317 (0.005)
0.10		0.489 (0.005)	0.315 (0.005)	0.497 (0.005)	0.307 (0.005)
0.15		0.538 (0.005)	0.301 (0.005)	0.525 (0.005)	0.288 (0.005)
0.00	0.20	0.135 (.003)	0.248 (0.005)	0.192 (0.004)	0.264 (0.004)
0.05		0.407 (0.005)	0.319 (0.005)	0.415 (0.005)	0.316 (0.005)
0.10		0.485 (0.005)	0.308 (0.005)	0.491 (0.005)	0.298 (0.005)
0.15		0.527 (0.005)	0.305 (0.005)	0.537 (0.005)	0.292 (0.005)
0.20		0.562 (0.005)	0.283 (0.005)	0.577 (0.005)	0.279 (0.004)
0.00	0.25	0.138 (0.003)	0.247 (0.004)	0.205 (0.004)	0.270 (0.004)
0.05		0.403 (0.005)	0.328 (0.005)	0.425 (0.005)	0.322 (0.005)
0.10		0.481 (0.005)	0.304 (0.005)	0.488 (0.005)	0.306 (0.005)
0.15		0.525 (0.005)	0.301 (0.005)	0.536 (0.005)	0.290 (0.005)
0.20		0.557 (0.005)	0.285 (0.005)	0.575 (0.005)	0.279 (0.004)
0.25		0.594 (0.005)	0.278 (0.004)	0.597 (0.005)	0.261 (0.004)

Table 5: Alpha and Beta Errors for the 20%/15% Couplet w/ Varied Intra- and Inter-Cluster Correlation and *d*=33.

Correlation		67 x 3		33 x 6	
Inter	Intra	Alpha (s.e.)	Beta (s.e.)	Alpha (s.e.)	Beta (s.e.)
0.00	0.00	0.118 (0.003)	0.248 (0.004)	0.135 (0.003)	0.227 (0.004)
0.00	0.05	0.129 (0.003)	0.256 (0.004)	0.165 (0.004)	0.244 (0.004)
0.05		0.401 (0.005)	0.361 (0.005)	0.423 (0.005)	0.342 (0.005)
0.00	0.10	0.138 (0.003)	0.266 (0.004)	0.190 (0.004)	0.262 (0.004)
0.05		0.409 (0.005)	0.357 (0.005)	0.422 (0.005)	0.343 (0.005)
0.10		0.475 (0.005)	0.358 (0.005)	0.481 (0.005)	0.339 (0.005)
0.00	0.15	0.154 (0.004)	0.276 (0.004)	0.210 (0.004)	0.281 (0.004)
0.05		0.415 (0.005)	0.366 (0.005)	0.425 (0.005)	0.344 (0.005)
0.10		0.474 (0.005)	0.350 (0.005)	0.487 (0.005)	0.338 (0.005)
0.15		0.510 (0.005)	0.339 (0.005)	0.525 (0.005)	0.336 (0.005)
0.00	0.20	0.159 (0.004)	0.274 (0.004)	0.223 (0.004)	0.283 (0.005)
0.05		0.418 (0.005)	0.364 (0.005)	0.421 (0.005)	0.352 (0.005)
0.10		0.481 (0.005)	0.348 (0.005)	0.484 (0.005)	0.334 (0.005)
0.15		0.511 (0.005)	0.328 (0.005)	0.527 (0.005)	0.334 (0.005)
0.20		0.547 (0.005)	0.329 (0.005)	0.544 (0.005)	0.323 (0.005)
0.00	0.25	0.168 (0.004)	0.282 (0.004)	0.239 (0.004)	0.291 (0.005)
0.05		0.408 (0.005)	0.365 (0.005)	0.435 (0.005)	0.353 (0.005)
0.10		0.481 (0.005)	0.353 (0.005)	0.477 (0.005)	0.353 (0.005)
0.15		0.511 (0.005)	0.340 (0.005)	0.523 (0.005)	0.337 (0.005)
0.20		0.540 (0.005)	0.335 (0.005)	0.553 (0.005)	0.322 (0.005)
0.25		0.561 (0.005)	0.314 (0.005)	0.574 (0.005)	0.313 (0.005)

Table 6: Alpha, Beta, and Average Sample Number (ASN) for Three Upper and Lower Threshold Couplets using the 67 x 3 Sequential Design for Varied Intra- and Inter- Cluster Correlation.

Correlation		10%/5%				15%/10%				20%/15%			
Inter	Intra	Alpha (s.e.)	Beta (s.e.)	ASN_o (s.e.)	ASN_a (s.e.)	Alpha (s.e.)	Beta (s.e.)	ASN_o (s.e.)	ASN_a (s.e.)	Alpha (se.)	Beta (s.e.)	ASN_o (s.e.)	ASN_a (s.e.)
0.00	0.00	0.095 (0.003)	0.160 (0.004)	22.933 (17.914)	33.048 (17.452)	0.087 (0.003)	0.240 (0.004)	34.273 (21.493)	49.094 (18.291)	0.150 (0.004)	0.217 (0.004)	39.707 (21.124)	46.955 (18.685)
0.00	0.05	0.103 (0.003)	0.178 (0.004)	22.567 (17.808)	31.543 (17.186)	0.096 (0.003)	0.259 (0.004)	34.015 (21.773)	47.494 (18.828)	0.162 (0.004)	0.235 (0.004)	38.436 (21.057)	45.529 (19.059)
0.05		0.401 (0.005)	0.251 (0.004)	22.576 (18.066)	23.309 (14.715)	0.392 (0.005)	0.335 (0.005)	30.509 (21.187)	32.414 (18.853)	0.435 (0.005)	0.345 (0.005)	30.348 (20.214)	30.596 (18.905)
0.00	0.10	0.125 (0.003)	0.197 (0.004)	22.519 (17.856)	30.109 (16.807)	0.107 (0.003)	0.268 (0.004)	33.95 (21.782)	46.158 (19.059)	0.182 (0.004)	0.247 (0.004)	37.939 (21.241)	44.01 (19.359)
0.05		0.401 (0.005)	0.257 (0.004)	21.946 (17.558)	22.831 (14.678)	0.396 (0.005)	0.330 (0.005)	30.305 (21.075)	32.322 (18.692)	0.427 (0.005)	0.335 (0.005)	29.858 (20.108)	30.289 (18.494)
0.10		0.486 (0.005)	0.245 (0.004)	20.37 (16.199)	20.75 (13.411)	0.479 (0.005)	0.308 (0.005)	27.638 (19.953)	28.16 (17.646)	0.485 (0.005)	0.345 (0.005)	25.809 (18.87)	25.617 (17.821)
0.00	0.15	0.135 (0.003)	0.206 (0.004)	21.747 (17.632)	28.99 (16.394)	0.119 (0.003)	0.290 (0.005)	33.104 (21.656)	44.276 (19.614)	0.186 (0.004)	0.258 (0.004)	36.302 (21.235)	42.393 (19.326)
0.05		0.394 (0.005)	0.257 (0.004)	21.381 (17.219)	22.3 (14.059)	0.398 (0.005)	0.329 (0.005)	29.672 (20.891)	31.612 (18.447)	0.429 (0.005)	0.344 (0.005)	29.245 (19.832)	29.591 (18.446)
0.10		0.493 (0.005)	0.239 (0.004)	20.32 (15.945)	20.182 (12.808)	0.470 (0.005)	0.314 (0.005)	27.101 (19.904)	27.32 (17.547)	0.495 (0.005)	0.349 (0.005)	25.779 (18.913)	25.253 (17.567)
0.15		0.561 (0.005)	0.227 (0.004)	19.201 (15.149)	19.046 (12.248)	0.517 (0.005)	0.296 (0.005)	25.143 (18.754)	25.399 (16.745)	0.527 (0.005)	0.340 (0.005)	23.265 (18.173)	23 (16.886)
0.00	0.20	0.149 (0.004)	0.214 (0.004)	21.406 (17.343)	28.003 (16.382)	0.131 (0.003)	0.298 (0.005)	32.52 (21.729)	43.029 (19.681)	0.199 (0.004)	0.269 (0.004)	35.989 (21.293)	41.115 (19.583)
0.05		0.396 (0.005)	0.259 (0.004)	20.641 (16.63)	21.945 (13.911)	0.397 (0.005)	0.343 (0.005)	28.926 (20.722)	30.674 (18.336)	0.425 (0.005)	0.350 (0.005)	28.693 (19.632)	29.18 (18.41)
0.10		0.494 (0.005)	0.248 (0.004)	19.927 (15.853)	19.948 (12.778)	0.475 (0.005)	0.319 (0.005)	26.894 (19.822)	26.982 (17.272)	0.489 (0.005)	0.342 (0.005)	25.352 (18.715)	25.06 (17.169)
0.15		0.553 (0.005)	0.225 (0.004)	19.04 (15.084)	18.989 (12.159)	0.516 (0.005)	0.304 (0.005)	24.959 (18.712)	24.75 (16.283)	0.519 (0.005)	0.327 (0.005)	22.867 (17.792)	22.871 (16.67)
0.20		0.590 (0.005)	0.210 (0.004)	18.432 (14.329)	18.034 (11.208)	0.554 (0.005)	0.281 (0.004)	23.452 (17.862)	23.621 (15.817)	0.553 (0.005)	0.319 (0.005)	21.601 (17.335)	21.155 (16.099)
0.00	0.25	0.158 (0.004)	0.243 (0.004)	20.872 (16.861)	26.806 (15.91)	0.140 (0.003)	0.300 (0.005)	31.899 (21.734)	41.276 (19.831)	0.211 (0.004)	0.284 (0.005)	34.936 (21.099)	39.666 (19.669)
0.05		0.400 (0.005)	0.261 (0.004)	20.294 (16.454)	21.574 (13.355)	0.396 (0.005)	0.346 (0.005)	28.724 (20.627)	30.247 (18.269)	0.428 (0.005)	0.352 (0.005)	27.96 (19.367)	28.537 (17.946)
0.10		0.479 (0.005)	0.247 (0.004)	19.054 (15.346)	19.646 (12.419)	0.469 (0.005)	0.320 (0.005)	26.375 (19.544)	27.095 (17.22)	0.479 (0.005)	0.351 (0.005)	24.808 (18.412)	24.777 (17.161)
0.15		0.545 (0.005)	0.231 (0.004)	18.609 (14.461)	18.752 (11.812)	0.516 (0.005)	0.299 (0.005)	24.36 (18.493)	24.741 (16.105)	0.528 (0.005)	0.333 (0.005)	22.457 (17.425)	22.523 (16.403)
0.20		0.593 (0.005)	0.208 (0.004)	18.097 (14.004)	17.829 (10.931)	0.557 (0.005)	0.287 (0.005)	23.245 (17.661)	23.094 (15.263)	0.546 (0.005)	0.322 (0.005)	20.935 (16.909)	20.758 (15.829)
0.25		0.622 (0.005)	0.192 (0.004)	17.634 (13.438)	17.479 (10.792)	0.594 (0.005)	0.268 (0.004)	22.23 (17.026)	22.177 (15.051)	0.575 (0.005)	0.315 (0.005)	19.744 (16.143)	19.707 (15.4)

REFERENCES

1. Aroian LA. Sequential analysis, direct method. *Technometrics.* 1968; **10**: 125-132.

2. Aroian LA. Applications of the direct method in sequential analysis (Com: V21 p 138). *Technometrics.* 1976; **18**: 301-306.

3. Bennett D, Smith A, McCormick L, et al. Categorization of transmitted HIV drug resistance using the WHO/CDC HIV drug resistance threshold survey method. Abstracts of the XV International Drug Resistance Workshop; June 13-17, 2006; Sitges, Spain. Abstract 103.

4. Deitchler M, Valadez JJ, Egge K, Fernandez S, Henningan M. A field test of three LQAS designs to assess the prevalence of acute malnutrition. *IJE* 2007 May; doi:10.1093/ije/dym092. Available at http://ije.oxfordjournals.org/papbyrecent.dtl

5. Deitchler M, Deconinck H, Bergeron G. Precision, time, and cost: a comparison of three sampling designs in an emergency setting. *Emerging Themes in Epidemiology.* 2008; 5:6.

6. Fenn B, Morris SS, Frost, C. Do childhood growth indicators in developing countries cluster? Implications for intervention strategies. *Public Health Nutrition.* 2004; **7**(7): 829-834.

7. Food and Agriculture Organization (FAO)/Food Security Analysis Unit (FSAU). Integrated food security and humanitarian phase classification: technical manual, version 1. Nairobi: FAO/FSAU Technical Series IV.11, 2006.

8. Hoshaw-Woodard S. Description and comparison of methods of cluster sampling and lot quality assurance sampling to assess immunization coverage. Geneva: World Health Organization; 2001. Available at http://www.who.int/vaccines-documents/DocsPDF01/www592.pdf

9. Hoshaw-Woodard S. The use of LQAS sampling to assess malnutrition. Ohio State University. Version II, Final Draft, 2003.

10. Kalton G. Introduction to survey sampling. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-035. Beverly Hills, CA: Sage, 1983.

11. Medecins Sans Frontieres (MSN). *Nutrition guidelines.* 1st Edition. Paris: MSF, 1995.

12. Reed JF. Eliminating bias in randomized cluster trials with correlated binomial outcomes. *Comput Methods Programs Biomed*. 2000; **61**(2): 119-123.

13. Robertson R, Valadez JJ. Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004. *Soc Sci Med* 2006 Sep; **63** (6): 1648-1660.

14. Shrimpton R, Victora CG, de Onis M, Lima RC, Blossner M, Clugston G. Worldwide timing of growth faltering : implications for nutritional interventions. Pediatrics. 2001 May; 107 (5): E75.

15. Spiegel PB, Salamca P, Maloney S, vander Veen A. Quality of Malnutrition assessment surveys collected during famine in Ethiopia. *J Am Med Ass* 2004; **292**: 613-18.

16. United Nations Children's Fund (UNICEF). Policy review paper E/ICEF/1990/1.6, JC 27/UNICEF-WHO/89.4: Strategy for improved nutrition of children and women in developing countries. New York, N.Y.: United Nations Children's Fund, 1990.

17. Valadez JJ, Weiss B, Leburg C, Davis R. Assessing community health programs: using LQAS for baseline surveys and regular monitoring. London: Teaching-aids at Low Cost, 2003.

18. Valadez JJ. *Assessing child survival programs in developing countries: testing lot quality assurance sampling*. Cambridge: Harvard University Press, 1991.

19. Wald A. *Sequential analysis.* Dover: Dover Publications, 1947.

20. World Health Organization. Assessment of elimination of neonatal tetanus in Zimbabwe, 2001. *Weekly Epidemiological Record* 2001; **76** (14), 101-108.

21. World Health Organization. Validation of neonatal tetanus elimination in Morocco, 2002. *Weekly Epidemiological Record* 2002; **77** (39), 325-328.

22. World Health Organization. Assessment of neonatal tetanus elimination in Eritrea. *Weekly Epidemiological Record* 2004; **79** (24), 222-227.

Appendix

This appendix details the methodology used for simulating binary random vectors with a correlation structure that reflects that which one might expect when applying the 67×3 and 33×6 LQAS cluster designs. There are few discrete probabilistic distributions which easily lend themselves to simulation of correlated binary observations. We outline a specific method for generating binary random vectors based on truncation of multivariate normal random variables.

Simulation

For a $k \times m$ cluster sample (k is the number of clusters and m is the size of each cluster for a total sample of size n), it is necessary to generate clusters with specific inter- and intra-cluster correlation subject to the constraint that the inter-cluster correlation is less than or equal to the intra-cluster correlation. Let $\boldsymbol{\tau_1} = \tau_1 \mathbf{1}\mathbf{1}^\top + (1-\tau_1)\boldsymbol{I}$ and $\boldsymbol{\tau_2} = \tau_2 \mathbf{1}\mathbf{1}^\top$, where $\mathbf{1}$ is an $m \times 1$ column vector of ones, and $\boldsymbol{I}$ is the $m \times m$ identity matrix. Then the desired correlation structure, $\mathbf{A}$, is a block diagonal matrix of dimension $n \times n$ with $\boldsymbol{\tau_1}$ on the diagonal blocks and $\boldsymbol{\tau_2}$ on the off diagonal blocks.

To achieve this structure for a binary random vector, first generate a realization, $\mathbf{Y}$, from the multivariate normal distribution of dimension n with mean equal to the zero vector and variance-covariance matrix equal to the above described correlation matrix, $\mathbf{A}$. Each component of the multivariate normal realization $\mathbf{Y}$ (Y_i, i=1,..., n) is marginally distributed as a normal random variable with mean zero and unit variance, and the correlation between any two components Y_i and Y_j is given by the $(i,j)^{th}$ entry of $\mathbf{A}$.

To attain the binary sample with the desired correlation structure, let

$$D_i = \begin{cases} 1 \text{ if } Y_i \leq \Phi^{-1}(P) \\ 0 \text{ otherwise, for i=1, 2,..., n,} \end{cases}$$

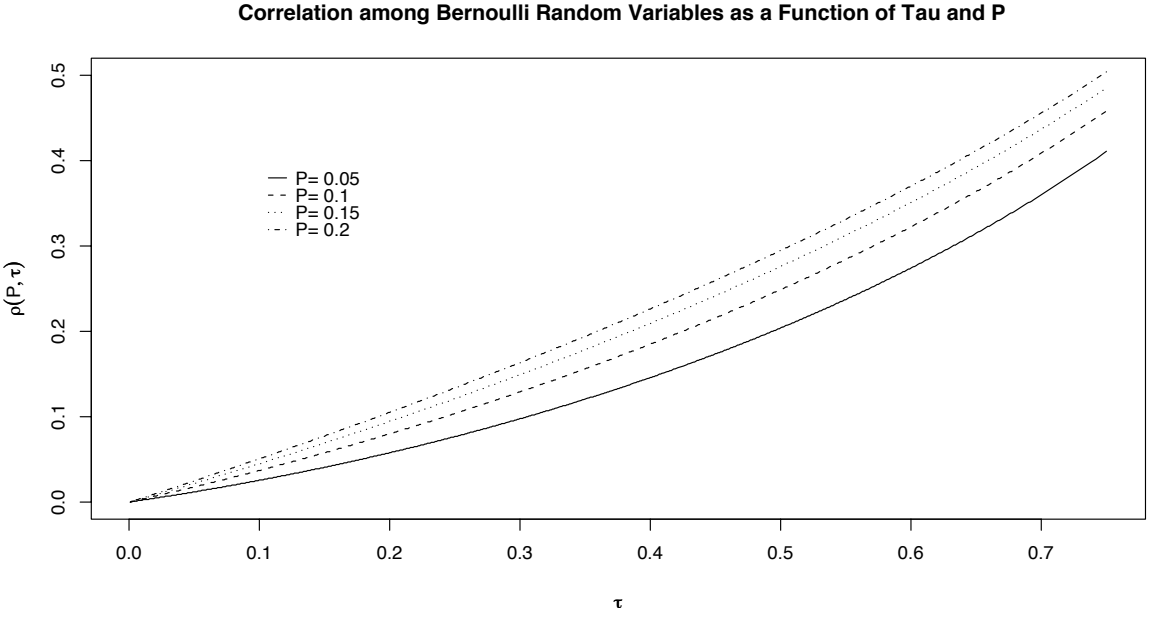
where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution and P is chosen to reflect the malnutrition prevalence ($P = P_o$ when simulating under the null hypothesis and $P = P_a$ when simulating under the alternative hypothesis). Then D_i is a bernoulli random variable with mean $P(Y_i \leq \Phi^{-1}(P)) = P$ for $i = 1, \ldots, n$ and $\mathbf{D}$= $(D_1, D_2, \ldots, D_n)$ is the resulting correlated sample of binary outcomes.

The correlation between any two of the resulting binary components D and D' is given by

$$Corr(D, D') = \frac{\int_{-\infty}^{y'} \int_{-\infty}^{y} f(y, y';\, \tau) dy dy' - P^2}{P(1-P)} \tag{1}$$

where $f(\cdot)$ is the probability density function of a bivariate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix given by $\boldsymbol{I}(1-\tau) + \mathbf{1}\mathbf{1}^\top\tau$, where $\boldsymbol{I}$ is the 2×2 identity matrix and $\mathbf{1}$ is a 2×1 vector of ones. The goal is to choose τ and P such that the resulting correlation is a specific value ρ. Define the function $\rho(P, \tau) \equiv Corr(D, D')$. Figure 1 plots $\rho(P, \tau)$ against τ for a range of P. Although no closed form solution to the double integral exists in (1), numerical integration yields highly precise approximations and can be implemented in many software packages. Here, numerical integration was performed using the `mvtnorm` library in `R` version 2.6.0.

Figure 1: Approximation of Bernoulli Correlation as a Function of Bivariate Normal Correlation and Prevalence



This approximation is used to simulate binary outcomes with correlation ρ and mean P. For example, to simulate binary outcomes with correlation $\rho = 0.05$ and mean $P = 0.10$ requires simulation in the multivariate normal with correlation $\tau = 0.131$. Table 7 outlines the values of τ used to simulate binary outcomes with correlation ρ and mean P in this study.

Table 7: Values of τ Needed to Simulate Binary Outcomes with Correlation ρ and mean P.

	Prevalence (P)			
Correlation (ρ)	0.05	0.10	0.15	0.20
0.05	0.177	0.131	0.110	0.098
0.10	0.305	0.242	0.210	0.191
0.15	0.407	0.339	0.301	0.277
0.20	0.493	0.424	0.385	0.359
0.25	0.568	0.501	0.462	0.435