Evaluation of Sample Size Formulae for Developing Adaptive Treatment Strategies Using a SMART Design.

Alena I. Scott* & Janet A. Levy† & Susan A. Murphy‡

April 30, 2007

* The Institute for Social Research, University of Michigan. Contact: ixmocane@umich.edu
† Center for Clinical Trials Network, National Institute on Drug Abuse. Contact: jlevy@nida.nih.gov
‡ Department of Statistics and The Institute for Social Research, University of Michigan. Contact: samurphy@umich.edu

1. INTRODUCTION

A specialized experimental design called sequential multiple assignment randomized trials (SMART) has been developed to support the investigation of a sequence of treatments in a principled way. In Scott et al[1], we presented statistical methodology and an evaluation of that methodology for the design and analysis of simple SMART trials. We also introduced and evaluated two new methods for sizing SMART trials. In this report, we present the simulation designs and the programs used to evaluate the sample size formulae for the SMART design presented in Scott et al[1].

2. THE SMART DESIGN AND RELATED RESEARCH QUESTIONS

We assume that we have data from a SMART design modeled in the following way: there are two options for the initial treatment followed by two treatment options for non-responders and one treatment option for responders. A representation of this design is presented in Figure 1. Note that this design is balanced; that is, the two treatment options for non-responders are the same regardless of initial treatment.

We use the following notation. A_1 is the indicator (0 or 1) for the initial treatment, R denotes the response to the initial treatment (non-response = 1 and response = 0), A_2 is the treatment indicator (0 or 1) for non-responders, and Y denotes a continuous final outcome. We use the convention of designating A_2=0 for responders. We further assume the patients are randomized equally to the two treatment options at each level; that is, $Pr\{A_1=1\}= Pr\{A_1=0=0.5$ and $Pr\{A_2=1|R=1, A_1=j\}= Pr\{A_2=0|R=1, A_1=j\}=0.5, j\in\{0,1\}$.

A summary of the four research questions we focus on answering with this SMART design is presented in Table 1; this is the generic version of Table 2 in Scott et al[1]. Note that Analyses 1 and 2 concern the components of an adaptive treatment strategy, and Analyses 3 and 4 concern strategies as a whole.

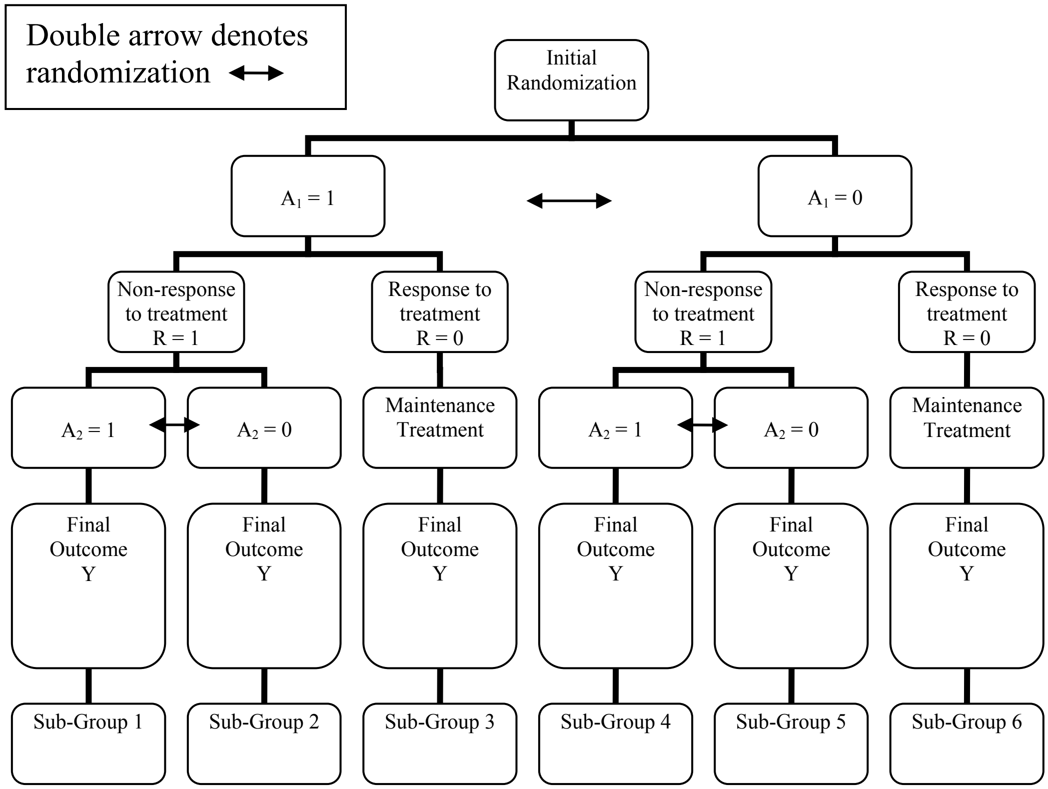

Figure 1. A SMART design to develop adaptive treatment strategies.

Table 1. Four research questions of interest
to guide the development of adaptive treatment strategies

Analysis	Research Question	Null Hypothesis
Two analyses that concern components of adaptive treatment strategies		
1	What is the effect of initial treatment assignment on long term outcome given specified treatments provided in the interim?	The mean long term outcome for all patients assigned to A_1=1 initially will be equal to the mean long term outcome of all patients assigned to A_1=0.
2	Considering only patients who did not respond to the initial treatment, what is the best subsequent treatment?	Considering only patients who did not respond to the initial treatment, the average long term outcome for those provided subsequently with A_2=1 will be the same as the long term average of those provided with A_1=0.
Two analyses that concern entire adaptive treatment strategies		
3	What is the difference in long term outcomes between two strategies that have different initial treatments, e.g. A_1=1, A_2=1 vs. A_1=0, A_2=0?	The mean long term outcome for all those given strategy A_1=1, A_2=1 will be equal to the long term mean outcome of all given treatment strategy A_1=0, A_2=0.
4	Which treatment strategy produces the best outcome?	This is an estimation problem not a hypothesis testing problem.

3. TEST STATISTICS AND SAMPLE SIZE FORMULAE

In this section, we present the test statistics and sample size formulae for the four different types of research questions summarized in Table 1. Without loss of generality, in Analysis 3, for comparing two strategies, (a_1, a_2) and (b_1, b_2), that have different initial treatments, we let $a_1=1$ and $b_1=0$.

3.1 Statistics for Different Analyses

The test statistics for Analyses 1-3 are presented in Table 2; the method for performing Analysis 4 is also given in Table 2. Note that Analyses 1-3 are hypotheses tests and that Analysis 4 is not a hypothesis test. The test statistics for Analyses 1 and 2 are the standard test statistics for a two group comparison with large samples[2] and are not unique to the SMART design. The estimator of a strategy mean, used in both Analysis 3 and Analysis 4, as well as the test statistic for Analysis 3 are given in Murphy[3]. In large samples, the test statistics used in Table 2 are normally distributed (with mean zero under the null hypothesis of no effect).

Table 2. Test statistics for each of the possible hypotheses

Type of Analysis	Test Statistic
1[(1)]	$Z = \dfrac{\left(\overline{Y}_{A1=1} - \overline{Y}_{A1=0}\right)}{\sqrt{\dfrac{S^2_{A1=1}}{N_{A1=1}} + \dfrac{S^2_{A1=0}}{N_{A1=0}}}}$ where $N_{A1=i}$ denotes the number of subjects who received i as the initial treatment
2[(1)]	$Z = \dfrac{\left(\overline{Y}_{R=1, A2=1} - \overline{Y}_{R=1, A2=0}\right)}{\sqrt{\dfrac{S^2_{R=1, A2=1}}{N_{R=1, A2=1}} + \dfrac{S^2_{R=1, A2=0}}{N_{R=1, A2=0}}}}$ where $N_{R=1, A2=i}$ denotes the number of non-responders who received i as the second treatment
3[(2)]	$Z = \dfrac{\sqrt{N}\left(\hat{\mu}_{A1=1, A2=a2} - \hat{\mu}_{A1=0, A2=b2}\right)}{\sqrt{\hat{\tau}^2_{A1=1, A2=a2} + \hat{\tau}^2_{A1=0, A2=b2}}}$ where N is the ***total*** number of subjects, and a2 and b2 are the second treatments in the two prespecified strategies being compared.
4	Choose largest of $\hat{\mu}_{A1=1, A2=1}, \hat{\mu}_{A1=0, A2=1}, \hat{\mu}_{A1=1, A2=0}, \hat{\mu}_{A1=0, A2=0}$

(1) $\overline{Y}$ and S^2 the sample mean and the sample variance; the subscript on *N* denotes the group of subjects

(2) See Table 3 for a definition of $\hat{\mu}$ and $\hat{\tau}^2$.

Table 3. Estimators for strategy means and
estimators for variance of estimator of strategy means.
Data for i^{th} patient is of the form $(A_{1i}, R_i, A_{2i}, Y_i)$,
where A_{1i}, R_i, A_{2i}, and Y_i are defined as in Section 4, and N is the *total* sample size.

Strategy sequence (a_1, a_2)	$\hat{\mu}_{A1=a1, A2=a2}$**: Estimator for strategy mean**
(1, 1)	$\dfrac{\frac{1}{N}\sum_{i=1}^{N}\frac{Y_i * A_{1i} * (R_i * A_{2i} + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}{\frac{1}{N}\sum_{i=1}^{N}\frac{A_{1i} * (R_i * A_{2i} + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}$
(1, 0)	$\dfrac{\frac{1}{N}\sum_{i=1}^{N}\frac{Y_i * A_{1i} * (R_i * (1 - A_{2i}) + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}{\frac{1}{N}\sum_{i=1}^{N}\frac{A_{1i} * (R_i * (1 - A_{2i}) + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}$
(0, 1)	$\dfrac{\frac{1}{N}\sum_{i=1}^{N}\frac{Y_i * (1 - A_{1i}) * (R_i * A_{2i} + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}{\frac{1}{N}\sum_{i=1}^{N}\frac{(1 - A_{1i}) * (R_i * A_{2i} + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}$
(0, 0)	$\dfrac{\frac{1}{N}\sum_{i=1}^{N}\frac{Y_i * (1 - A_{1i}) * (R_i * (1 - A_{2i}) + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}{\frac{1}{N}\sum_{i=1}^{N}\frac{(1 - A_{1i}) * (R_i * (1 - A_{2i}) + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}}$
Strategy sequence (a_1, a_2)	$\hat{\tau}^2_{A1=a1, A2=a2}$**: N*Estimator for variance of above estimator of strategy mean**
(1, 1)	$\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(Y_i - \hat{\mu}_{11}) * A_{1i} * (R_i * A_{2i} + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}\right)^2$
(1, 0)	$\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(Y_i - \hat{\mu}_{10}) * A_{1i} * (R_i * (1 - A_{2i}) + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}\right)^2$
(0, 1)	$\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(Y_i - \hat{\mu}_{01}) * (1 - A_{1i}) * (R_i * A_{2i} + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}\right)^2$
(0, 0)	$\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(Y_i - \hat{\mu}_{00}) * (1 - A_{1i}) * (R_i * (1 - A_{2i}) + (1 - R_i))}{0.5 * (R_i * 0.5 + (1 - R_i))}\right)^2$

3.2 Sample Size Calculations

All sample size formulae assume a two-tailed z-test. Let α be the desired size of the hypothesis test and 1–β denote the power of the test; let $z_{\alpha/2}$ be the standard normal (1–α/2) percentile. Approximate normality of the test statistic is assumed throughout.

In order to calculate the sample size, one must specify the desired detectable standardized effect size, denoted here by δ. We use the definition for standardized effect size found in Cohen[4]: the standardized effect size between two groups is the difference between the means of the two groups divided by the square root of the pooled variance, which is the square root of the average of the variances of the two groups being compared. All of the sample size formulae make the working assumption that the variances of the two groups under consideration are equal. The definition of the variance changes with the analysis under consideration; we will explicitly define the variance assumption as we present each sample size formulae. In Table 4, we summarize the standardized effect sizes for the various analyses we are considering.

Table 4. Standardized effect sizes for the four analyses in Table 3

Analysis	Formula for Standardized Effect Size δ
1	$\delta = \frac{E[Y \mid A_1 = 1] - E[Y \mid A_1 = 0]}{\sqrt{\frac{Var[Y \mid A_1 = 1] + Var[Y \mid A_1 = 0]}{2}}}$
2	$\delta = \frac{E[Y \mid R = 1, A_2 = 1] - E[Y \mid R = 1, A_2 = 0]}{\sqrt{\frac{Var[Y \mid R = 1, A_2 = 1] + Var[Y \mid R = 1, A_2 = 0]}{2}}}$
3	$\delta = \frac{E[Y \mid A_1 = 1, A_2 = a2] - E[Y \mid A_1 = 0, A_2 = b2]}{\sqrt{\frac{Var[Y \mid A_1 = 1, A_2 = a2] + Var[Y \mid A_1 = 0, A_2 = b2]}{2}}}$ where a2 and b2 are the second components in the two prespecified strategies being compared.
4	$\delta = \frac{E[Y \mid A_1 = a1, A_2 = a2] - E[Y \mid A_1 = b1, A_2 = b2]}{\sqrt{\frac{Var[Y \mid A_1 = a1, A_2 = a2] + Var[Y \mid A_1 = b1, A_2 = b2]}{2}}}$ where (a_1, a_2) = strategy with the highest mean outcome, (b1, b2) = strategy with the next highest mean outcome.

The sample size formulae for each of the analyses in Table 1 are summarized in Table 5. Each formula makes certain working assumptions which are presented below. The working assumptions are only used to size the SMART design and are not used to analyze the data from the trial.

Working assumptions for sample size formula N_1:

1. The variance of outcome Y given the first treatment $A_1=1$ is equal to the variance of outcome Y given the first treatment $A_1=0$; i.e. $\sigma^2 = Var[Y|A_1=1] = Var[Y|A_1=0]$.

Working assumptions for sample size formula N_2:

1. The variance of outcome Y for non-responders who were given second treatment $A_2=1$ is equal to the variance of outcome Y for non-responders who were given second treatment $A_2=0$; i.e. $\sigma^2 = Var[Y|R=1, A_2=1] = Var[Y|R=1, A_2=0]$.
2. The intermediate non-response rates are equal; that is, that the probability of non-response for a patient given initial treatment $A_1=1$ is the same as the probability of non-response for a patient given initial treatment $A_1=0$. We will denote this identical non-response rate by p.

Working assumptions for sample size formula N_{3a} (sample size varies by non-response rate):

1. The variance of outcome Y given treatment strategy ($A_1=1$, $A_2=a_2$) is equal to the variance of outcome Y given treatment strategy ($A_1=0$, $A_2=b_2$); i.e. $\sigma^2 = Var[Y|A_1=1, A_2=a_2] = Var[Y|A_1=0, A_2=b_2]$.
2. The variability of the outcome Y around the strategy mean ($A_1=1$, $A_2=a_2$), among either responders or non-responders, is less than the variance of the strategy mean and similarly for strategy ($A_1=0$, $A_2=b_2$).
3. The intermediate non-response rates are equal; that is, $p = Pr\{R=1|A_1=1\} = Pr\{R=1|A_1=0\}$.

Working assumptions for sample size formula N_{3b} (sample size is invariant to the non-response rate):

1. The variance of outcome Y given treatment strategy ($A_1=1$, $A_2=a_2$) is equal to the variance of outcome Y given treatment strategy ($A_1=0$, $A_2=b_2$); i.e. $\sigma^2 = Var[Y|A_1=1, A_2=a_2] = Var[Y|A_1=0, A_2=b_2]$.
2. The sample size formulae use the working assumption that the intermediate non-response rates are both equal to 1; that is, $p = 1$.

Working assumptions for sample size calculation N_4:

1. The marginal variances of the final outcome given the strategy are all equal and we denote this variance by σ^2. This means that, $\sigma^2 = Var[Y|A_1=a_1, A_2=a_2]$ for all (a_1, a_2) in {(1,1), (1,0), (0,1), (0,0)}
2. The sample sizes will be large enough so that $\hat{\mu}_{(a1,\, a2)}$ is approximately normally distributed.
3. The correlation between the final outcome Y given treatment strategy (1, 1) and Y given treatment strategy (1, 0) is the same as the correlation between Y given treatment strategy (0, 1) and Y given treatment strategy (0, 0).

Table 5. Sample size formulae for the four analyses of interest.

Analysis	Formula for Standardized Effect Size δ
1	$N_1 = 2*2*(z_{\alpha/2}+z_\beta)^2*(1/\delta)^2$
2	$N_2 = 2*2*(z_{\alpha/2}+z_\beta)^2*(1/\delta)^2/p$
3	$N_{3a} = 2*(z_{\alpha/2}+z_\beta)^2*(2*(2*p+1*(1-p)))*(1/\delta)^2$
	$N_{3b} = 2*(z_{\alpha/2}+z_\beta)^2*4*(1/\delta)^2$
4	Algorithm[(1)]

[(1)] See Appendix for Matlab code; see Scott et al[1] for more details.

4. SIMULATION DESIGN FOR THE EVALUATION OF SAMPLE SIZE FORMULAE

In this section, we present the method for designing the simulations to evaluate the sample size formulae presented in Section 3.2. Since the sample size formulae for Analyses 1 and 2 are standard formulae, we focus on evaluating the newly developed sample size formulae for Analyses 3 and 4. For each analysis, we present simulation parameters for generating data that follows the working assumptions; these datasets will be used in order to evaluate the accuracy of the sample size formulae for Analyses 3 and 4 (i.e. to see if we in fact achieve the desired power).

We also present simulation parameters for generating data that tests the robustness of the sample size formulae for Analyses 3 and 4. To test the robustness of a given formula, we calculate a sample size given by the relevant formula in Section 4.2, and then simulate data sets of this sample size that do not satisfy the working assumptions in one of the following ways:

- the intermediate non-response rates to first level treatments are unequal, i.e. $Pr\{R=1|A_1=1\} \neq Pr\{R=1|A_1=0\}$,
- the variances relevant to the analysis of interest are unequal,
- the distribution of the final outcome, Y, is right skewed (thus for a given sample size, the test statistic is more likely to have a non-normal distribution).

The sample sizes used for the simulations were chosen to give a power level of 0.90 and a type I error of 0.05 when one of Analyses 1-3 is used to size the trial, and a 0.90 probability of choosing the best strategy for Analysis 4 when it is used to size the trial; these sample sizes are shown in Table 6.

We sized the studies to detect a prespecified standardized effect size of 0.2 or 0.5. We simulated data with intermediate non-response rates of 0.5, 0.7 and 0.9 and with mean outcomes for the responders usually higher than those for non-responders.

Table 6. Sample Sizes Used for Simulations(1)(2)
All entries are for ***total*** sample size

Effect Size δ	Non-response rate(3) p	Analysis # 1	Analysis #2	Analysis #3 (sample size varies by p)	Analysis #3 (sample size invariant to p)	Analysis #4
δ = 0.20						
	p = 0.5	1056	2112	1584	2112	608
	p = 0.7	1056	1509	1796	2112	608
	p = 0.9	1056	1174	2007	2112	608
δ = 0.50						
	p = 0.5	169	338	254	338	97
	p = 0.7	169	241	287	338	97
	p = 0.9	169	188	321	338	97

(1) All entries assume each statistical test is two tailed of size $\alpha = 0.05$ and power $1-\beta = 0.90$; size is not required for Analysis 4 since it is not a hypothesis test.
(2) Analysis 4 is not a hypothesis test; we choose the sample size so that the probability that we choose the best treatment, given such a "best" treatment exists (i.e. given that there is a treatment strategy that has a higher mean outcome than the rest) is $1-\beta$.
(3) In each formula, non-response rates are assumed to be equal, i.e. $p=Pr\{R=1|A_1=1\}=Pr\{R=1|A_1=0\}$.

For Analysis 1-3, power is estimated by the proportion of times out of 1000 simulations that the null hypothesis is correctly rejected; for Analysis 4, the probability of choosing the best strategy is estimated by the proportion of times out 1000 simulations that the correct strategy with the highest mean is chosen.

For Analysis 3, we need to specify the strategies of interest, and for the purposes of these simulations, we will compare strategies (A_1=1, A_2=1) and (A_1=0, A_2=0). For the simulations to evaluate the robustness of the sample size calculation for Analysis 4, we choose (A_1=1, A_2=1) to always have the highest mean outcome and generate the data according to two different "patterns": 1) the strategy means are all different and 2) the mean outcomes of the other three strategies besides (A_1=1, A_2=1) are all equal. In the second pattern, it is more difficult to detect the "best" strategy because the highest mean must be distinguished from *all* the rest, which are all the "next highest", instead of just *one* next highest mean.

4.1 How the Data Was Simulated

Each simulated data set of size N consists of N vectors of the form (A_1, R, A_2, Y). The entries in each vector are generated in the following way:

1. Generate A_1 from a Bernoulli distribution with mean 0.5; that is, A_1=0 with probability 0.5 and A_1=1 with probability 0.5. This step represents the randomization between the two options for initial treatment.
2. Generate a response R to treatment A_1 from a Bernoulli distribution with mean p_{a1}, where $p_{a1}=Pr\{R=1|A_1=a_1\}$, i.e. the non-response rate for treatment $A_1=a_1$.
3. Next, generate the second treatment for non-responders from a Bernoulli distribution with mean 0.5; again, this step represents the randomization between the two options for the second treatment. If R=0, then we use the convention A_2 = 0 (see Table 7).
4. Finally, generate the final outcome Y given the history A_1, R, A_2. We assume that this final outcome given the past, $Y|A_1, R, A_2$, is normally distributed with mean $E[Y|A_1, R, A_2]$ and variance $Var[Y|A_1, R, A_2]$. Table 7 shows that there are six possible histories, and we must specify a mean and variance for each of these groups.

Table 7. Dummy coding to reflect treatments potentially tested by the SMART design

Sub-Group Number	**Initial Treatment A_1**	**Response to Initial Treatment R** **0 = Response** **1 = Non-response**	**Second Treatment for Non-responders A_2**
1	1	1	1
2	1	1	0
3	1	0	0
4	0	1	1
5	0	1	0
6	0	0	0

In summary, the model for generating a dataset is:

1. $A_1 \sim Bern(0.5)$
2. $R|A_1 \sim Bern(p_{A1})$, where
   a. $p_1 = Pr\{R=1|A_1=1\}$, the non-response rate when treatment A_1=1
   b. $p_0 = Pr\{R=1|A_1=0\}$, the non-response rate when treatment A_1=0
3. $A_2|R=1 \sim Bern(0.5)$; $A_2|R=0$ is coded as 0.
4. $Y|A_1, R, A_2 \sim N(E[Y|A_1, R, A_2], Var[Y|A_1, R, A_2])$, with notation
   a. $\nu_{a1, r, a2} = E[Y|A_1=a_1, R=r, A_2=a_2]$ and
   b. $\varsigma^2_{a1, r, a2} = Var[Y|A_1=a_1, R=r, A_2=a_2]$,

and the parameters that must be specified to simulate a data set are:

- the two non-response rates, p_0 and p_1,
- the six means for the final responses $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$, and
- the six variances for the final responses $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$.

For the simulations we performed, in all cases except for the simulations which challenge the equal variance assumption, we set $\sigma^2 = 100$.

4.2 Simulation Parameters vs. Parameters for the Effect Size of Interest

One detail to address is the relationship between the means and variances specified in the simulation design and the means and variances required for the specified effect size for the particular analysis of interest (Table 4). The variance that is required to calculate the sample size for a particular effect size is different than the ones that are specified in the simulations. Therefore, we must define the relationship between the means and variances in the simulation design (the conditional variances) and the means and variances in the specified effect size for the analysis of interest (the marginal variances).

Here we derive the means and variances required for each analysis in terms of the conditional means and variances specified in the simulation model.

We will use the following representation for the conditional mean:
$E[Y|A_1=a_1, R=r, A_2=a_2] = \gamma_1 + \gamma_2 A_1 + \gamma_3 R + \gamma_4 A_1 R + \gamma_5 R A_2 + \gamma_6 A_1 R A_2$. For reference, here are the means for the six groups in Table 7 in terms of these "γ"s:

- $E[Y|A_1=1, R=1, A_2=1] = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5 + \gamma_6$
- $E[Y|A_1=1, R=1, A_2=0] = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$
- $E[Y|A_1=1, R=0, A_2=0] = \gamma_1 + \gamma_2$
- $E[Y|A_1=0, R=1, A_2=1] = \gamma_1 + \gamma_3 + \gamma_4 + \gamma_5$
- $E[Y|A_1=0, R=1, A_2=0] = \gamma_1 + \gamma_3$
- $E[Y|A_1=0, R=0, A_2=0] = \gamma_1$.

Throughout the derivations, note that since A_1, A_2, and R are coded as binary variables, then $A_1^2=A_1$, $A_2^2=A_2$, and $R^2=R$.

Means and Variances for Analysis 1

<u>Goal</u>: Find a formula for $E[Y|A_1=a_1]$ and $Var[Y|A_1=a_1]$ in terms of $E[Y|A_1=a_1, R=r, A_2=a_2]$ and $Var[Y|A_1=a_1, R=r, A_2=a_2]$.

<u>Derivation of $E[Y|A_1=a_1]$</u>:

$$\begin{aligned} E[Y|A_1=a_1] &= E[\, E[Y|A_1, R, A_2]\, |A_1=a_1], \quad \text{(Law of Total Expectation)} \\ &= E[\gamma_1 + \gamma_2 A_1 + \gamma_3 R + \gamma_4 A_1 R + \gamma_5 R A_2 + \gamma_6 A_1 R A_2 \,|A_1=a_1] \\ &= \gamma_1 + \gamma_2 a_1 + (\gamma_3 + \gamma_4 a_1) E[R|A_1=a_1] + (\gamma_5 + \gamma_6 a_1) E[RA_2|A_1=a_1]. \end{aligned}$$

Now,

$$\begin{aligned} E[RA_2|A_1=a_1] &= 0*0*Pr\{R=0, A_2=0|A_1=a_1\} + 1*0*Pr\{R=1, A_2=0|A_1=a_1\} \\ &\quad + 1*1*Pr\{R=1, A_2=1|A_1=a_1\} \\ &= 1*1*Pr\{R=1, A_2=1|A_1=a_1\} \\ &= (0.5)Pr\{R=1|A_1=a_1\}. \end{aligned}$$

Also, note that $E[R|A_1=a_1]=p_{a1}$, the non-response rate for those given $A_1=a_1$. Therefore, we have

$$E[Y|A_1=a_1] = \gamma_1 + \gamma_2 a_1 + p_{a1}(\gamma_3 + \gamma_4 a_1 + 0.5\gamma_5 + 0.5\gamma_6 a_1).$$

Derivation of $Var[Y|A_1=a_1]$:

Note that $Var[Y|A_1=a_1] = E[Y^2|A_1=a_1] - E[Y|A_1=a_1]^2$.

Using the result for $E[Y|A_1=a_1]$ above:

$$\begin{aligned} E[Y|A_1=a_1]^2 &= (\gamma_1 + \gamma_2 a_1 + p_{a1}(\gamma_3 + \gamma_4 a_1 + 0.5\gamma_5 + 0.5\gamma_6 a_1))^2 \\ &= (\gamma_1 + \gamma_2 a_1)^2 + 2p_{a1}(\gamma_1 + \gamma_2 a_1)(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1)) \\ &\quad + p_{a1}{}^2(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1))^2. \end{aligned}$$

Next,

$$\begin{aligned} E[Y^2|A_1=a_1] &= E[\, E[Y^2|A_1, R, A_2] \,|A_1=a_1], \quad \text{(Law of Total Expectation)} \\ &= E[\, Var[Y|A_1, R, A_2] + E[Y|A_1, R, A_2]^2 \,|A_1=a_1] \\ &= E[\, Var[Y|A_1, R, A_2] \,|A_1=a_1] \\ &\quad + E[(\gamma_1 + \gamma_2 A_1)^2 + 2R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 A_1) \\ &\qquad + R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|A_1=a_1] \\ &= E[\, Var[Y|A_1, R, A_2] \,|A_1=a_1] + (\gamma_1 + \gamma_2 a_1)^2 \\ &\quad + 2E[R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 a_1)|A_1=a_1] \\ &\quad + E[R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|A_1=a_1]. \end{aligned}$$

Using basic properties of conditional expectation, we have the following.

$$\begin{aligned} &E[R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 A_1)|A_1=a_1] \\ &\quad = E[R|A_1=a_1]\, E[(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2) \,|A_1=a_1]\, E[(\gamma_1 + \gamma_2 A_1) \,|A_1=a_1] \\ &\quad = p_{a1}\, (\gamma_3 + \gamma_4 a_1 + \gamma_5 E[A_2] + \gamma_6 a_1 E[A_2])\, (\gamma_1 + \gamma_2\, a_1) \\ &\quad = p_{a1}\, (\gamma_1 + \gamma_2 a_1)(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1)) \end{aligned}$$

$$\begin{aligned} &E[R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|A_1=a_1] \\ &\quad = E[R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|A_1=a_1] \\ &\quad = E[R|A_1=a_1]\, E[(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|A_1=a_1] \\ &\quad = p_{a1}\, E[(\gamma_3 + \gamma_4 A_1)^2 + 2(\gamma_3 + \gamma_4 A_1)(\gamma_5 A_2 + \gamma_6 A_1 A_2) + (\gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|A_1=a_1] \\ &\quad = p_{a1}\, ((\gamma_3 + \gamma_4 a_1)^2 + 2(\gamma_3 + \gamma_4 A_1)(\gamma_5 + \gamma_6 A_1)E[A_2|A_1=a_1] \\ &\qquad + (\gamma_5 + \gamma_6 A_1)^2\, E[A_2|A_1=a_1]) \\ &\quad = p_{a1}\, ((\gamma_3 + \gamma_4 a_1)^2 + 2(\gamma_3 + \gamma_4 A_1)(\gamma_5 + \gamma_6 A_1)E[A_2] + (\gamma_5 + \gamma_6 A_1)^2\, E[A_2]) \\ &\quad = p_{a1}\, ((\gamma_3 + \gamma_4 a_1)^2 + 2(\gamma_3 + \gamma_4 a_1)(\gamma_5 + \gamma_6 a_1)(0.5) + (\gamma_5 + \gamma_6 a_1)^2(0.5)) \end{aligned}$$

Plugging back into $Var[Y|A_1=a_1] = E[Y^2|A_1=a_1] - E[Y|A_1=a_1]^2$ we have the following:

$$\begin{aligned} Var[Y|A_1=a_1] &= E[\, Var[Y|A_1, R, A_2] \,|A_1=a_1] + (\gamma_1 + \gamma_2 a_1)^2 \\ &\quad + 2p_{a1}(\gamma_1 + \gamma_2 a_1)(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1)) \end{aligned}$$

$+ p_{a1}((\gamma_3 + \gamma_4 a_1)^2 + (\gamma_3 + \gamma_4 a_1)(\gamma_5 + \gamma_6 a_1) + (\gamma_5 + \gamma_6 a_1)^2 (0.5))$
$- (\gamma_1 + \gamma_2 a_1)^2 - 2p_{a1}(\gamma_1 + \gamma_2 a_1)(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1))$
$- p_{a1}^2(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1))^2.$

Cancelling out terms gives us

$\mathrm{Var}[Y|A_1=a_1] = E[\ \mathrm{Var}[Y|A_1, R, A_2]\ |A_1=a_1] +$
$+ p_{a1}((\gamma_3 + \gamma_4 a_1)^2 + (\gamma_3 + \gamma_4 a_1)(\gamma_5 + \gamma_6 a_1) + (\gamma_5 + \gamma_6 a_1)^2 (0.5))$
$- p_{a1}^2(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1))^2.$

Using algebra to simplify, we get

$$\mathrm{Var}[Y|A_1=a_1] = E[\ \mathrm{Var}[Y|A_1, R, A_2]\ |A_1=a_1] + p_{a1}(1-p_{a1})(\gamma_3 + \gamma_4 a_1 + 0.5(\gamma_5 + \gamma_6 a_1))^2 + p_{a1}(0.5(\gamma_5 + \gamma_6 a_1))^2.$$

Means and Variances for Analysis 2

Goal: Find a formula for $E[Y|R=1, A_2=a_2]$ and $\mathrm{Var}[Y|R=1, A_2=a_2]$ in terms of $E[Y|A_1=a_1, R=r, A_2=a_2]$ and $\mathrm{Var}[Y|A_1=a_1, R=r, A_2=a_2]$.

Derivation of $E[Y|R=1, A_2=a_2]$:

$$\begin{aligned} E[Y|R=1, A_2=a_2] &= E[\ E[Y|A_1, R, A_2]\ |R=1, A_2=a_2], \quad \text{(Law of Total Expectation)} \\ &= E[\gamma_1 + \gamma_2 A_1 + \gamma_3 R + \gamma_4 A_1 R + \gamma_5 R A_2 + \gamma_6 A_1 R A_2\ |R=1, A_2=a_2] \\ &= \gamma_1 + \gamma_3 + (\gamma_2 + \gamma_4)E[A_1|R=1] + \gamma_5 a_2 + \gamma_6\ a_2 E[A_1|\ R=1] \end{aligned}$$

Using Bayes Rule,

$$\begin{aligned} \Pr\{A_1=a_1|R=1\} &= \frac{\Pr\{R=1|A_1=a_1\} * \Pr\{A_1=a_1\}}{\sum_{i\in\{0,1\}} \Pr\{R=1|A_1=i\} * \Pr\{A_1=i\}} \\ &= \frac{\Pr\{R=1|A_1=a_1\}(0.5)}{(\Pr\{R=1|A_1=0\} + \Pr\{R=1|A_1=1\})(0.5)} \\ &= \frac{\Pr\{R=1|A_1=a_1\}}{\Pr\{R=1|A_1=0\} + \Pr\{R=1|A_1=1\}}. \end{aligned}$$

Using the above formula for $\Pr\{A_1=a_1|R=1\}$,

$$\begin{aligned} E[A_1|R=1] &= 0*\Pr\{A_1=0|R=1\} + 1*\Pr\{A_1=1|R=1\} \\ &= \frac{\Pr\{R=1|A_1=1\}}{\Pr\{R=1|A_1=0\} + \Pr\{R=1|A_1=1\}} \\ &= \frac{p_1}{p_0 + p_1}. \end{aligned}$$

Therefore, we have

$$E[Y|R=1, A_2=a_2] = \gamma_1 + \gamma_3 + \gamma_5 a_2 + \frac{p_1}{p_0 + p_1}(\gamma_2 + \gamma_4 + \gamma_6 a_2).$$

<u>Derivation of Var[Y|R=1, A_2=a_2]</u>:

$Var[Y|R=1, A_2=a_2] = E[Y^2|R=1, A_2=a_2] - E[Y|R=1, A_2=a_2]^2$.

Using the result for $E[Y|R=1, A_2=a2]$ above:

$$E[Y|R=1, A_2=a_2]^2 = (\gamma_1 + \gamma_3 + \gamma_5 a_2)^2 + \frac{p_1}{p_0 + p_1} 2(\gamma_2 + \gamma_4 + \gamma_6 a_2)(\gamma_1 + \gamma_3 + \gamma_5 a_2) + \left(\frac{p_1}{p_0 + p_1}\right)^2 (\gamma_2 + \gamma_4 + \gamma_6 a_2)^2$$

Next,

$$\begin{aligned} E[Y^2|R=1, A_2=a_2] &= E[\, E[Y^2|A_1, R, A_2] \,|R=1, A_2=a_2], \text{ (Law of Total Expectation)} \\ &= E[\, Var[Y|A_1, R, A_2] + E[Y|A_1, R, A_2]^2 \,|R=1, A_2=a_2] \\ &= E[\, Var[Y|A_1, R, A_2] \,|R=1, A_2=a_2] \\ &\quad + E[(\gamma_1 + \gamma_2 A_1)^2 + 2R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 A_1) \\ &\qquad + R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|R=1, A_2=a_2] \\ &= E[\, Var[Y|A_1, R, A_2] \,|R=1, A_2=a_2] \\ &\quad + E[(\gamma_1 + \gamma_2 A_1)^2|R=1, A_2=a_2] \\ &\quad + 2E[R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 A_1) \,|R=1, A_2=a_2] \\ &\quad + E[R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|R=1, A_2=a_2] \end{aligned}$$

Now, we solve each of the components in the second part of $E[Y^2|R=1, A_2=a_2]$.

$$\begin{aligned} E[(\gamma_1 + \gamma_2 A_1)^2 \,|R=1, A_2=a_2] &= E[\gamma_1^2 + 2\gamma_1\gamma_2 A_1 + \gamma_2^2 A_1^2 \,|R=1] \\ &= \gamma_1^2 + 2\gamma_1\gamma_2 E[A_1|R=1] + \gamma_2^2 E[A_1|R=1] \\ &= \gamma_1^2 + (2\gamma_1\gamma_2 + \gamma_2^2)\frac{p_1}{p_0 + p_1} \end{aligned}$$

Recalling that $A_1^2=A_1$,

$$\begin{aligned} &E[R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 A_1) \,|R=1, A_2=a_2] \\ &= E[R(\gamma_3(\gamma_1 + \gamma_2 A_1) + \gamma_4 A_1(\gamma_1 + \gamma_2) + \gamma_5 A_2(\gamma_1 + \gamma_2 A_1) + \gamma_6 A_1 A_2(\gamma_1 + \gamma_2))|R=1, A_2=a_2] \\ &= \gamma_3(\gamma_1 + \gamma_2\frac{p_1}{p_0 + p_1}) + \gamma_4(\gamma_1 + \gamma_2)\frac{p_1}{p_0 + p_1} + \gamma_5 a_2(\gamma_1 + \gamma_2\frac{p_1}{p_0 + p_1}) + \gamma_6 a_2\frac{p_1}{p_0 + p_1}(\gamma_1 + \gamma_2) \\ &= (\gamma_3 + \gamma_5 a_2)(\gamma_1 + \gamma_2\frac{p_1}{p_0 + p_1}) + (\gamma_4 + \gamma_6 a_2)(\gamma_1 + \gamma_2)\frac{p_1}{p_0 + p_1}. \end{aligned}$$

$E[R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2 \,|R=1, A_2=a_2]$
$= E[R((\gamma_3 + \gamma_4 A_1)^2 + 2(\gamma_3 + \gamma_4 A_1)(\gamma_5 A_2 + \gamma_6 A_1 A_2) + (\gamma_5 A_2 + \gamma_6 A_1 A_2)^2) \,|R=1, A_2=a_2]$
$= E[R(\gamma_3 + \gamma_4 A_1)^2 + 2RA_2(\gamma_3 + \gamma_4 A_1)(\gamma_5 + \gamma_6 A_1) + R\, A_2\, (\gamma_5 + \gamma_6 A_1)^2 \,|R=1, A_2=a_2]$

$$= \gamma_3^2 + (2\gamma_3\gamma_4 + \gamma_4^2)\frac{p_1}{p_0 + p_1} + 2a_2\gamma_3\gamma_5 + 2a_2(\gamma_3\gamma_5 + \gamma_5\gamma_4 + \gamma_4\gamma_6)\frac{p_1}{p_0 + p_1}$$

$$+ a_2\gamma_5^2 + a_2(2\gamma_5\gamma_6 + \gamma_6^2)\frac{p_1}{p_0 + p_1}.$$

Substituting back into $Var[Y|R=1, A_2=a_2] = E[Y^2|R=1, A_2=a_2] - E[Y|R=1, A_2=a_2]^2$ and simplifying, we get

$$Var[Y|R=1, A_2=a_2] = E[\, Var[Y|A_1, R, A_2] \,|R=1, A_2=a_2] + (\gamma_2 + \gamma_4 + \gamma_6 a_2)^2 \frac{p_0 p_1}{p_0 + p_1}.$$

Means and Variances for Analyses 3 and 4

Goal: Find a formula for $E[Y|A_1=a_1, A_2=a_2]$ and $Var[Y|A_1=a_1, A_2=a_2]$ in terms of $E[Y|A_1=a_1, R=r, A_2=a_2]$ and $Var[Y|A_1=a_1, R=r, A_2=a_2]$.

Derivation of $E[Y|A_1=a1, A_2=a_2]$:

$$\begin{aligned} E[Y|A_1=a_1, A_2=a_2] &= E[\, E[Y|A_1, R, A_2] \,|A_1=a_1, A_2=a_2], \\ &= E[\gamma_1 + \gamma_2 A_1 + \gamma_3 R + \gamma_4 A_1 R + \gamma_5 R A_2 + \gamma_6 A_1 R A_2 \,|A_1=a_1, A_2=a_2] \\ &= \gamma_1 + \gamma_2 a_1 + E[R|A_1=a_1]\,(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2) \end{aligned}$$

Therefore, we have

$$E[Y|A_1=a_1, A_2=a_2] = \gamma_1 + \gamma_2 a_1 + p_{a1}(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2).$$

Derivation of $Var[Y|A_1=a_1, A_2=a_2]$:

$$Var[Y|A_1=a_1, A_2=a_2] = E[Y^2|A_1=a_1, A_2=a_2] - E[Y|A_1=a_1, A_2=a_2]^2.$$

Using the result for $E[Y|A_1=a_1, A_2=a_2]$ above:

$$\begin{aligned} E[Y|A_1=a_1, A_2=a_2]^2 &= (\gamma_1 + \gamma_2 a_1 + p_{a1}(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2))^2 \\ &= (\gamma_1 + \gamma_2 a1)^2 + 2p_{a1}(\gamma_1 + \gamma_2 a_1)\,(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2) \\ &\quad + p_{a1}^2(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2)^2 \end{aligned}$$

Next,

$$
\begin{aligned}
E[Y^2|A_2=a_1, A_2=a_2] &= E[\ E[Y^2|A_1, R, A_2]\ |A_2=a_1, A_2=a_2] \\
&= E[\ Var[Y|A_1, R, A_2] + E[Y|A_1, R, A_2]^2\ |A_2=a_1, A_2=a_2] \\
&= E[\ Var[Y|A_1, R, A_2]\ |A_2=a_1, A_2=a_2] \\
&\quad + E[(\gamma_1 + \gamma_2 A_1)^2 + 2R(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)(\gamma_1 + \gamma_2 A_1) \\
&\qquad + R^2(\gamma_3 + \gamma_4 A_1 + \gamma_5 A_2 + \gamma_6 A_1 A_2)^2\ |A_2=a_1, A_2=a_2] \\
&= E[\ Var[Y|A_1, R, A_2]\ |A_2=a_1, A_2=a_2] \\
&\quad + (\gamma_1 + \gamma_2 a_1)^2 \\
&\quad + 2p_{a1}(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2)(\gamma_1 + \gamma_2 a_1) \\
&\quad + p_{a1}(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2)^2
\end{aligned}
$$

Plugging back into $Var[Y|A_1=a_1, A_2=a_2] = E[Y^2|A_1=a_1, A_2=a_2] - E[Y|A_1=a_1, A_2=a_2]^2$ and simplifying gives

$$
\begin{aligned}
Var[Y|A_1=a_1, A_2=a_2] &= E[\ Var[Y|A1, R, A2]\ |A_1=a_1, A_2=a_2] \\
&\quad + p_{a1}(1-p_{a1})(\gamma_3 + \gamma_4 a_1 + \gamma_5 a_2 + \gamma_6 a_1 a_2)^2.
\end{aligned}
$$

4.3 Parameter Values for Data that Follows the Working Assumptions

Now, for each of the four analyses, we present the parameters which give data sets that conform to the working assumptions presented in Section 3.2

Analysis 1

The only working assumption to satisfy for N_1 is that the variance of outcome Y given the first treatment $A_1=1$ is equal to the variance of outcome Y given the first treatment $A_1=0$; in other words, $\sigma^2 = Var[Y|A_1=1] = Var[Y|A_1=0]$. Since the formula does not depend on an intermediate non-response rate, for the sake of simplicity, we let $p_0 = p_1$, and denote the common rate by p. Without loss of generality, we choose $A_1=1$ to have the larger mean outcome. We let $\sigma^2 = 100$. The actual values we chose for the simulations that follow the working assumptions for the sample size formula for Analysis 1 are summarized in Tables 8a and 8b.

Table 8a. Simulation parameters for Analysis 1 data
that follows the working assumptions for sample size formula N_1

Scenario	Effect size	Non-response rate $p_0 = p_1$	$E[Y\|A_1=a_1, R=r, A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1 R=r, A_2=a_2]^{(2)}$ $\varsigma^2_{a1,r,a2}$
1	0.2	0.5	{10.25, 3.25, 15.25, 7, 5, 12}	{90, 90, 61.625, 95, 95, 86}
2	0.2	0.7	{10.75, 3.75, 15.75, 7, 5, 12}	{80, 80, 67.5083, 95, 95, 84.1333}
3	0.2	0.9	{11.25, 4.25, 16.25, 7, 5, 12}	{82, 82, 86.725, 96, 96, 94.6}
4	0.5	0.5	{13.25, 6.25, 18.25, 7, 5, 12}	{80, 80, 71.625, 94, 94, 87}
5	0.5	0.7	{13.75, 6.75, 18.75, 7, 5, 12}	{80, 80, 67.5083, 90, 90, 95.8}
6	0.5	0.9	{14.25, 7.25, 19.25, 7, 5, 12}	{82, 82, 86.725, 96, 96, 94.6}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$

Table 8b. Related values for simulation parameters in Table 8a

Scenario	Effect size	Non-response rate $p_0 = p_1$	Corresponding γ values	$E[Y\|A_1=a_1, A_2=a_2]^{(1),(2)}$ $\mu_{a1,a2}$
1	0.2	0.5	{12, 3.25, -7, -5, 2, 5}	{12.75, 9.25, 9.5, 8.5}
2	0.2	0.7	{12, 3.75, -7, -5, 2, 5}	{12.25, 7.35, 8.5, 7.1}
3	0.2	0.9	{12, 4.25, -7, -5, 2, 5}	{11.75, 5.45, 7.5, 5.7}
4	0.5	0.5	{12, 6.25, -7, -5, 2, 5}	{15.75, 12.25, 9.5, 8.5}
5	0.5	0.7	{12, 6.75, -7, -5, 2, 5}	{15.25, 10.35, 8.5, 7.1}
6	0.5	0.9	{12, 7.25, -7, -5, 2, 5}	{14.75, 8.45, 7.5, 5.7}

(1) Order in which γ parameters are listed: $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6\}$
(2) Given for informational purposes, not for generative model
(3) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

Analysis 2

The working assumptions to satisfy for N_2 are

- the variance of outcome Y for non-responders given second treatment $A_2=1$ is equal to the variance of outcome Y for non-responders given $A_2=0$, denoted $\sigma^2 = Var[Y|R=1, A_2=1] = Var[Y|R=1, A_2=0]$, and
- the intermediate non-response rate, p_0 and p_1 are equal, denoted by p.

Without loss of generality, we choose $A_2=1$ to have the larger mean outcome for non-responders. We let $\sigma^2 = 100$. The actual values we chose for the simulating data that follow the working assumptions for the sample size formula for Analysis 2 are summarized in Tables 9a and 9b.

Table 9a. Simulation parameters for Analysis 2 data
that follows the working assumptions for sample size formula N_1

Scenario	Effect size	Non-response rate $p_0 = p_1$	$E[Y\|A_1=a_1, R=r, A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1\ R=r, A_2=a_2]^{(2)}$ $\varsigma^2_{a1,r,a2}$
1	0.2	0.5	{6.25, 3.25, 15.25, 6, 5, 12}	{99.9688, 98.4688, 100, 100, 100, 100}
2	0.2	0.7	{6.25, 3.25, 15.25, 6, 5, 12}	{99.9563, 97.8562, 100, 100, 100, 100}
3	0.2	0.9	{6.25, 3.25, 15.25, 6, 5, 12}	{99.9437, 97.2437, 100, 100, 100, 100}
4	0.5	0.5	{11.25, 3.25, 15.25, 7, 5, 12}	{91.9688, 99.4688, 99, 99, 99, 99}
5	0.5	0.7	{11.25, 3.25, 15.25, 7, 5, 12}	{89.3563, 99.8562, 98, 98, 98, 98}
6	0.5	0.9	{11.25, 3.25, 15.25, 7, 5, 12}	{93.7438, 98.2437, 90, 90, 99, 99}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$

Table 9b. Related values for simulation parameters in Table 9a

Scenario	Effect size	Non-response rate $p_0 = p_1$	Corresponding γ values	$E[Y\|A_1=a_1, A_2=a_2]^{(1),(2)}$ $\mu_{a1,a2}$
1	0.2	0.5	{12, 3.25, -7, -5, 2, 5}	{10.75, 9.25, 9, 8.5}
2	0.2	0.7	{12, 3.25, -7, -5, 2, 5}	{8.95, 6.85, 7.8, 7.1}
3	0.2	0.9	{12, 3.25, -7, -5, 2, 5}	{7.15, 4.45, 6.6, 5.7}
4	0.5	0.5	{12, 3.25, -7, -5, 2, 6}	{13.25, 9.25, 9.5, 8.5}
5	0.5	0.7	{12, 3.25, -7, -5, 2, 6}	{12.45, 6.85, 8.5, 7.1}
6	0.5	0.9	{12, 3.25, -7, -5, 2, 6}	{11.65, 4.45, 7.5, 5.7}

(1) Order in which γ parameters are listed: $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6\}$
(2) Given for informational purposes, not for generative model
(3) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

Analysis 3

As noted previously, to generate data for Analysis 3, we need to specify the strategies of interest, and we will compare strategies (A_1=1, A_2=1) and (A_1=0, A_2=0).

We generated the data according to the working assumptions in common for N_{3a} and N_{3b} which are

- the variance of outcome Y given treatment (A_1=1, A_2=1) and (A_1=0, A_2=0), denoted $\sigma^2 = Var[Y|A_1=1, A_2=1] = Var[Y|A_1=0, A_2=0]$, and
- the intermediate non-response rate, p_0 and p_1 are equal, denoted by p.

Additional assumptions were ignored. Note that when we generate a data set of size N_{3b} using the parameters below, we are always violating working assumption that $p_1 = p_0 = 1$.

Without loss of generality, we choose the strategy (A_1=1, A_2=1) to have the larger mean outcome than (A_1=0, A_2=0). We let $\sigma^2 = 100$. The actual values we chose for the simulations that follow the common working assumptions for the sample size formulae for Analysis 3 are summarized in Tables 10a and 10b.

Table 10a. Simulation parameters for Analysis 3 data that follows the common working assumptions for sample size formulae N_{3a} and N_{3b}

Scenario	Effect size	Non-response rate $p_0 = p_1$	$E[Y\|A_1=a_1, R=r, A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1\ R=r, A_2=a_2]^{(2)}$ $\varsigma^2_{a1,r,a2}$
1	0.2	0.5	{6.5, 1.5, 14.5, 7, 5, 12}	{99, 46.5, 69, 95, 83, 92.5}
2	0.2	0.7	{6.7, 1.7, 14.7, 7, 5, 12}	{95, 63.5, 66.8667, 95, 87.8, 94.1667}
3	0.2	0.9	{6.9, 1.9, 14.9, 7, 5, 12}	{95, 84.5, 87.4, 98, 95.6, 95.5}
4	0.5	0.5	{9.5, 4.5, 17.5, 7, 5, 12}	{98, 45.5, 70, 97, 85, 90.5}
5	0.5	0.7	{9.7, 4.7, 17.7, 7, 5, 12}	{94, 62.5, 69.2, 97, 89.8, 89.5}
6	0.5	0.9	{9.9, 4.9, 17.9, 7, 5, 12}	{94, 83.5, 96.4, 98, 95.6, 95.5}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$

Table 10b. Related values for simulation parameters in Table 10a

Scenario	Effect size	Non-response rate $p_0 = p_1$	Corresponding γ values	$E[Y\|A_1=a_1, A_2=a_2]^{(1),(2)}$ $\mu_{a1,a2}$
1	0.2	0.5	{12, 2.5, -7, -6, 2, 3}	{10.5, 8, 9.5, 8.5}
2	0.2	0.7	{12 2.7 -7 -6 2 3}	{9.1, 5.6, 8.5, 7.1}
3	0.2	0.9	{12 2.9 -7 -6 2 3}	{7.7, 3.2, 7.5, 5.7}
4	0.5	0.5	{12 5.5 -7 -6 2 3}	{13.5, 11, 9.5, 8.5}
5	0.5	0.7	{12 5.7 -7 -6 2 3}	{12.1, 8.6, 8.5, 7.1}
6	0.5	0.9	{12 5.9 -7 -6 2 3}	{10.7, 6.2, 7.5, 5.7}

(1) Order in which γ parameters are listed: $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6\}$
(2) Given for informational purposes, not for generative model
(3) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

Analysis 4

For the simulations to evaluate the robustness of the sample size calculation for Analysis 4, we choose (A_1=1, A_2=1) to always have the highest mean outcome and generate the data according to two different "patterns": 1) the strategy means are all different and 2) the mean outcomes of the other three strategies besides (A_1=1, A_2=1) are all equal. In the second pattern, it is more difficult to detect the "best" strategy because the highest mean must be distinguished from *all* the rest, which are all the "next highest", instead of just *one* next highest mean. In both patterns, we set the variance for all strategies (a_1, a_2) to $Var[Y|A_1=a_1, A_2=a_2] = 100$.

Table 11a. Simulation parameters for Analysis 4 data
that follows the working assumptions for sample size formula N_1

Scenario	Effect size	Non-response rate $p_0 = p_1$	$E[Y\|A_1=a_1=r, A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1\ R=r, A_2=a_2]^{(2)}$ $\varsigma^2_{a1,r,a2}$
Pattern 1: the mean for (1,1) is the highest mean, the other three are allowed to vary				
1	0.2	0.5	{7.5, 2.5, 15.5, 7, 5, 12}	{97, 44.5, 71, 95, 83, 92.5}
2	0.2	0.7	{8.1, 3.1, 16.1, 7, 5, 12}	{97, 65.5, 62.2, 95, 87.8, 94.1667}
3	0.2	0.9	{8.7, 3.7, 16.7, 7, 5, 12}	{97, 86.5, 69.4, 98, 95.6, 95.5}
4	0.5	0.5	{13.5, 2.5, 15.5, 7, 5, 12}	{100, 17.5, 98, 97, 85, 90.5}
5	0.5	0.7	{12.75, 2.5, 15.5, 7, 5, 12}	{100, 51.5687, 94.7062, 97, 89.8, 89.5}
6	0.5	0.9	{12.25, 2.5, 15.5, 7, 5, 12}	{100, 84.1563, 90.4938, 97, 94.6, 104.5}
Pattern 2: the mean for (1,1) is the highest mean and the other three means are all equal				
7	0.2	0.5	{9, 5, 12, 5, 5, 12}	{97, 77, 98.5, 95, 95, 80.5}
8	0.2	0.7	{7.85, 5, 12, 5, 5, 12}	{97, 87.4668, 94.9443, 95, 95, 77.3667}
9	0.2	0.9	{7.2, 5, 12, 5, 5, 12}	{98, 95.404, 97.264, 97, 97, 82.9}
10	0.5	0.5	{15, 5, 12, 5, 5, 12}	{100, 80, 95.5, 97, 97, 78.5}
11	0.5	0.7	{12.15, 5, 12, 5, 5, 12}	{100, 85.3067, 99.9843, 97, 97, 72.7}
12	0.5	0.9	{10.6, 5, 12, 5, 5, 12}	{100, 95.296, 98.236, 97, 97, 82.9}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$

Table 11b. Related values for simulation parameters in Table 11a

Scenario	Effect size	Non-response rate $p_0 = p_1$	Corresponding γ values	$E[Y\|A_1=a_1, A2=a_2]^{(1), (2)}$ $\mu_{a1,a2}$
Pattern 1: the mean for (1,1) is the highest mean, the other three are allowed to vary				
1	0.2	0.5	{12, 3.5, -7, -6, 2, 3}	{11.5, 9, 9.5, 8.5}
2	0.2	0.7	{12, 4.1, -7, -6, 2, 3}	{10.5, 7, 8.5, 7.1}
3	0.2	0.9	{12, 4.7, -7, -6, 2, 3}	{9.5, 5, 7.5, 5.7}
4	0.5	0.5	{12, 3.5, -7, -6, 2, 9}	{14.5, 9, 9.5, 8.5}
5	0.5	0.7	{12, 3.5, -7, -6, 2, 8.25}	{13.575, 6.4, 8.5, 7.1}
6	0.5	0.9	{12, 3.5, -7, -6, 2, 7.75}	{12.575, 3.8, 7.5, 5.7}
Pattern 2: the mean for (1,1) is the highest mean and the other three means are all equal				
7	0.2	0.5	{12, 0, -7, 0, 0, 4}	{10.5, 8.5, 8.5, 8.5}
8	0.2	0.7	{12, 0, -7, 0, 0, 2.85}	{9.095, 7.1, 7.1, 7.1}
9	0.2	0.9	{12, 0, -7, 0, 0, 2.2}	{7.68, 5.7, 5.7, 5.7}
10	0.5	0.5	{12, 0, -7, 0, 0, 10}	{13.5, 8.5, 8.5, 8.5}
11	0.5	0.7	{12, 0, -7, 0, 0, 7.15}	{12.105, 7.1, 7.1, 7.1}
12	0.5	0.9	{12, 0, -7, 0, 0, 5.6}	{10.74, 5.7, 5.7, 5.7}

(1) Order in which γ parameters are listed: $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6\}$
(2) Given for informational purposes, not for generative model
(3) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

4.4 Parameter Values for Data that Challenges the Working Assumptions of Equal Non-Response Rate

Since the sample size formulae for Analyses 1 and 2 are standard formulae, and we only present simulation designs for testing the robustness of the newly developed sample size formulae for Analyses 3 and 4. When we challenge the non-response rate equality assumption, we calculate the sample size formula for a particular non-response rate p, then generate the data with non-response rates $Pr\{R=1|A_1=1\} = p-0.05$ and $Pr\{R=1|A_1=0\} = p+0.05$.

Analysis3

Note that when we generate a data set of size N_{3b} using the parameters below, we are violating both N_{3b} working assumptions that $p_1 = p_0 = p$ and $p = 1$.

Table 12a. Simulation parameters for Analysis 3 data that violate the working assumption of equal non-response rate for sample size formula N_{3a} and N_{3b}

Scenario	Effect size	Non-response rate $\{p_0, p_1\}$	$E[Y\|A_1=a_1=r,A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1 R=r, A_2=a_2]^{(2)}$ $\varsigma^2_{a1,r,a2}$
1	0.2	{0.55, 0.45}	{5.75, 0.75, 13.75, 7, 5, 12}	{99, 51.75, 77.2545, 90, 79.2, 98.4722}
2	0.2	{0.75, 0.65}	{5.95, 0.95, 13.95, 7, 5, 12}	{97, 70.75, 75.8571, 95, 89, 96.25}
3	0.2	{0.95, 0.85}	{6.15, 1.15, 14.15, 7, 5, 12}	{97, 91.75, 98.8667, 99, 97.8, 95.25}
4	0.5	{0.55, 0.45}	{8.75, 3.75, 16.75, 7, 5, 12}	{98, 50.75, 78.0727, 95, 84.2, 92.3611}
5	0.5	{0.75, 0.65}	{8.25, 3.25, 16.25, 7, 5, 12}	{98, 82.25, 85.8857, 97, 93.4, 95.75}
6	0.5	{0.95, 0.85}	{9.15, 4.15, 17.15, 7, 5, 12}	{97, 91.75, 98.8667, 99, 97.8, 95.25}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$

Table 12b. Related values for simulation parameters in Table 12a

Scenario	Effect size	Non-response rate $\{p_0, p_1\}$	Corresponding γ values	$E[Y\|A_1=a_1, A_2=a_2]^{(1), (2)}$ $\mu_{a1,a2}$
1	0.2	{0.55, 0.45}	{12, 1.75, -7, -6, 2, 3}	{10.15, 7.9, 9.25, 8.15}
2	0.2	{0.75, 0.65}	{12, 1.95, -7, -6, 2, 3}	{8.75, 5.5, 8.25, 6.75}
3	0.2	{0.95, 0.85}	{12, 2.15, -7, -6, 2 3}	{7.35, 3.1, 7.25, 5.35}
4	0.5	{0.55, 0.45}	{12, 4.75, -7, -6, 2, 3}	{13.15, 10.9, 9.25, 8.15}
5	0.5	{0.75, 0.65}	{12, 4.25, -7, -6, 2, 3}	{11.05, 7.8, 7.75, 6.05}
6	0.5	{0.95, 0.85}	{12, 5.15, -7, -6, 2, 3}	{10.35, 6.1, 7.25, 5.35}

(1) Given for informational purposes, not for generative model
(2) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

Analysis4

Table 13a. Simulation parameters for Analysis 4 data
that violate working assumption of equal non-response rate
for sample size formula N_4

Scenario	Effect size	Non-response rate $\{p_0,\ p_1\}$	$E[Y\|A_1=a_1 =r,A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1\ R=r,\ A_2=a_2]^{(2)}$ $\varsigma^2_{a1,r,a2}$
Pattern 1: the mean for (1,1) is the highest mean, the other three are allowed to vary				
1	0.2	{0.55, 0.45}	{7.4, 1.4, 14.4, 7, 5, 12}	{100, 34, 77.95, 98, 87.2, 88.6944}
2	0.2	{0.75, 0.65}	{7.8, 1.8, 14.8, 7, 5, 12}	{100, 58, 68.15, 98, 92, 87.25}
3	0.2	{0.95, 0.85}	{8.2, 2.2, 15.2, 7, 5, 12}	{100, 82, 58.35, 98, 96.8, 114.25}
4	0.5	{0.55, 0.45}	{16, 5, 14, 2, 5, 13}	{100, 57.65, 98.2, 70, 95.65, 70.1167}
5	0.5	{0.75, 0.65}	{11, 3, 14, 2, 5, 13}	{100, 60.8, 94.15, 70, 84.25, 99.25}
6	0.5	{0.95, 0.85}	{14, 8, 14, 6, 9, 13}	{100, 94.6, 100, 98, 99.65, 91.45}
Pattern 2: the mean for (1,1) is the highest mean and the other three means are all equal				
7	0.2	{0.55, 0.45}	{7.895, 3.445, 12, 5, 5, 12}	{100, 69.0147, 92.4170, 98, 98, 75.4944}
8	0.2	{0.75, 0.65}	{6.995, 3.92, 12, 5, 5, 12}	{100, 85.9173, 83.7175, 98, 98, 69.25}
9	0.2	{0.95, 0.85}	{6.53, 4.175, 12, 5, 5, 12}	{100, 95.3035, 74.5672, 98, 98, 91.45}
10	0.5	{0.55, 0.45}	{14.3, 3.2, 13, 5, 5, 13}	{110, 58.1075, 91.0577, 98, 98, 67.2444}
11	0.5	{0.75, 0.65}	{11.5, 3.8, 13, 5, 5, 13}	{110, 81.1635, 79.9661, 98, 98, 58}
12	0.5	{0.95, 0.85}	{9.925, 4.05, 13, 5, 5, 13}	{100, 89.4030, 91.9627, 98, 98, 77.2}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1},\ \nu_{1,1,0},\ \nu_{1,0,0},\ \nu_{0,1,1},\ \nu_{0,1,0},\ \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1},\ \varsigma^2_{1,1,0},\ \varsigma^2_{1,0,0},\ \varsigma^2_{0,1,1},\ \varsigma^2_{0,1,0},\ \varsigma^2_{0,0,0}\}$

Table 13b. Related values for simulation parameters in Table13a

Scenario	Effect size	Non-response rate $\{p_0,\ p_1\}$	Corresponding γ values	$E[Y\|A_1=a_1,\ A_2=a_2]^{(1),(2)}$ $\mu_{a1,a2}$
Pattern 1: the mean for (1,1) is the highest mean, the other three are allowed to vary				
1	0.2	{0.55, 0.45}	{12, 2.4, -7, -6, 2, 4}	{11.25, 8.55, 9.25, 8.15}
2	0.2	{0.75, 0.65}	{12, 2.8, -7, -6, 2, 4}	{10.25, 6.35, 8.25, 6.75}
3	0.2	{0.95, 0.85}	{12, 3.2, -7, -6, 2, 4}	{9.25, 4.15, 7.25, 5.35}
4	0.5	{0.55, 0.45}	{13, 1, -8, -1, -3, 14}	{14.9, 9.95, 6.95, 8.6}
5	0.5	{0.75, 0.65}	{13, 1, -8, -3, -3, 11}	{12.05, 6.85, 4.75, 7}
6	0.5	{0.95, 0.85}	{13, 1, -4, -2, -3, 9}	{14, 8.9, 6.35, 9.2}
Pattern 2: the mean for (1,1) is the highest mean and the other three means are all equal				
7	0.2	{0.55, 0.45}	{12, 0, -7, -1.555, 0, 4.45}	{10.1528, 8.1502, 8.15, 8.15}
8	0.2	{0.75, 0.65}	{12, 0, -7, -1.08, 0, 3.075}	{8.7468, 6.7480, 6.75, 6.75}
9	0.2	{0.95, 0.85}	{12, 0, -7, -.825, 0, 2.355}	{7.3505, 5.3487, 5.35, 5.35}
10	0.5	{0.55, 0.45}	{13, 0, -8, -1.8, 0, 11.1}	{13.585, 8.59, 8.6, 8.6}
11	0.5	{0.75, 0.65}	{13, 0, -8, -1.2, 0, 7.7}	{12.025, 7.02, 7, 7}
12	0.5	{0.95, 0.85}	{13, 0, -8, -.95, 0, 5.875}	{10.3863, 5.3925, 5.4, 5.4}

(1) Given for informational purposes, not for generative model
(2) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1},\ \mu_{1,0},\ \mu_{1,0},\ \mu_{0,0}\}$

4.4 Parameter Values for Data that Challenges the Working Assumptions of Equal Variance

Again, we only present simulation designs for evaluating the newly developed sample size formulae for Analyses 3 and 4. In challenging the equal variance assumption, we set one of the variances at 81% of the other variance.

Analysis3

For these simulations, we set $Var[Y|A_1=1, A_2=1] = 100$ and $Var[Y|A_1=0, A_2=0] = 81$. Additionally, we assume $Var[Y|A_1=1, A_2=1] = Var[Y|A_1=1, A_2=0]$ and $Var[Y|A_1=0, A_2=0] = Var[Y|A_1=0, A_2=1]$.

Table 14a. Simulation parameters for Analysis 3 data that violate working assumption of equal variance for sample size formula N_{3a} and N_{3b}

Scenario	Effect size	Non-response rate $p_0 = p_1$	$E[Y\|A_1=a_1 =r,A_2=a_2]^{(1)}$ $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1\ R=r, A_2=a_2]^{(2)(3)}$ $\varsigma^2_{a1,r,a2}$ chosen so that $Var[Y\|A_1=1, A_2=1] = 100$, $Var[Y\|A_1=0, A_2=0] = 81$
1	0.2	0.5	{6.403, 1.403, 14.403, 7, 5, 12}	{97, 44.5, 71, 75, 63, 74.5}
2	0.2	0.7	{6.603, 1.603, 14.603, 7, 5, 12}	{95, 63.5, 66.8667, 75, 67.8, 77.5}
3	0.2	0.9	{6.803, 1.803, 14.803, 7, 5, 12}	{95, 84.5, 87.4, 80, 77.6, 67.5}
4	0.5	0.5	{9.257, 4.257, 17.257, 7, 5, 12}	{98, 45.5, 70, 70, 58, 79.5}
5	0.5	0.7	{9.457, 4.457, 17.457, 7, 5, 12}	{98, 66.5, 59.8667, 70, 62.8, 89.1667}
6	0.5	0.9	{9.657, 4.657, 17.657, 7, 5, 12}	{98, 87.5, 60.4, 80, 77.6, 67.5}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$
(3) Additionally, we assume $Var[Y|A_1=1, A_2=1] = Var[Y|A_1=1, A_2=0]$ and $Var[Y|A_1=0, A_2=0] = Var[Y|A_1=0, A_2=1]$.

Table 14b. Related values for simulation parameters in Table 14a

Scenario	Effect size	Non-response rate $p_0 = p_1$	Corresponding γ values	$E[Y\|A_1=a_1, A_2=a_2]^{(1),(2)}$ $\mu_{a1,a2}$
1	0.2	0.5	{12, 2.403, -7, -6, 2, 3}	{10.403, 7.903, 9.5, 8.5}
2	0.2	0.7	{12, 2.603, -7, -6, 2, 3}	{9.003, 5.503, 8.5, 7.1}
3	0.2	0.9	{12, 2.803, -7, -6, 2, 3}	{7.6030, 3.103, 7.5, 5.7}
4	0.5	0.5	{12, 5.257, -7, -6, 2, 3}	{13.257, 10.757, 9.5, 8.5}
5	0.5	0.7	{12, 5.457, -7, -6, 2, 3}	{11.857, 8.357, 8.5, 7.1}
6	0.5	0.9	{12, 5.657, -7, -6, 2, 3}	{10.457, 5.957, 7.5, 5.7}

(1) Given for informational purposes, not for generative model
(2) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

Analysis4

For these simulations, we set $Var[Y|A_1=1, A_2=1] = 100$ and for next best strategy (0, 1), $Var[Y|A_1=0, A_2=1] = 81$. Additionally, for Pattern 1, we assume $Var[Y|A_1=1, A_2=1] = Var[Y|A_1=1, A_2=0]$ and $Var[Y|A_1=0, A_2=1] = Var[Y|A_1=0, A_2=0]$. For Pattern 2, since the means for all the strategies besides (1, 1) are equal, we assume $Var[Y|A_1=1, A_2=0] = Var[Y|A_1=0, A_2=1] = Var[Y|A_1=0, A_2=0] = 81$.

Table 15a. Simulation parameters for Analysis 4 data
that violate working assumption of equal variance for sample size formula N_4

Scenario	Effect size	Non-response rate $p_0 = p_1$	$E[Y\|A_1=a_1 =r,A_2=a_2]$ (1) $\nu_{a1,r,a2}$	$Var[Y\|A_1=a_1\ R=r, A_2=a_2]$ (2) $\varsigma^2_{a1,r,a2}$
Pattern 1 (3): the mean for (1,1) is the highest mean, (0,1) has the next highest mean, the other two are allowed to vary below the mean of (0,1)				
1	0.2	0.5	{7.5, 2.5, 15.5, 7, 5, 12}	{98, 45.5, 70, 80, 68, 69.5}
2	0.2	0.7	{8.1, 3.1, 16.1, 7, 5, 12}	{90, 58.5, 78.5333, 80, 72.8, 65.8333}
3	0.2	0.9	{8.7, 3.7, 16.7, 7, 5, 12}	{95, 84.5, 87.4, 79, 76.6, 76.5}
4	0.5	0.5	{13.5, 2.5, 15.5, 7, 5, 12}	{120, 37.5, 78, 76, 64, 73.5}
5	0.5	0.7	{12.75, 2.5, 15.5, 7, 5, 12}	{100, 51.5687, 94.7062, 76, 68.8, 75.1667}
6	0.5	0.9	{12.25, 2.5, 15.5, 7, 5, 12}	{100, 84.1563, 90.4938, 76, 73.6, 103.5}
Pattern 2 (4): the mean for (1,1) is the highest mean and the other three means are all equal				
7	0.2	0.5	{9, 5, 12, 5, 5, 12}	{120, 62, 75.5, 76, 76, 61.5}
8	0.2	0.7	{7.85, 5, 12, 5, 5, 12}	{100, 63.3239, 87.9442, 75, 75, 60.7}
9	0.2	0.9	{7.2, 5, 12, 5, 5, 12}	{100, 76.2929, 79.2640, 75, 75, 90.9}
10	0.5	0.5	{15, 5, 12, 5, 5, 12}	{100, 42, 95.5, 75, 75, 62.5}
11	0.5	0.7	{12.15, 5, 12, 5, 5, 12}	{100, 58.1639, 99.9843, 75, 75, 60.7}
12	0.5	0.9	{10.6, 5, 12, 5, 5, 12}	{100, 74.1849, 98.2360, 75, 75, 90.9}

(1) Order in which parameters $\nu_{a1,r,a2}$ are listed: $\{\nu_{1,1,1}, \nu_{1,1,0}, \nu_{1,0,0}, \nu_{0,1,1}, \nu_{0,1,0}, \nu_{0,0,0}\}$
(2) Order in which parameters $\varsigma^2_{a1,r,a2}$ are listed: $\{\varsigma^2_{1,1,1}, \varsigma^2_{1,1,0}, \varsigma^2_{1,0,0}, \varsigma^2_{0,1,1}, \varsigma^2_{0,1,0}, \varsigma^2_{0,0,0}\}$
(3) Parameters $\varsigma^2_{a1,r,a2}$ are chosen so that $Var[Y|A_1=1, A_2=1] = Var[Y|A_1=1, A_2=0] = 100$, $Var[Y|A_1=0, A_2=1] = Var[Y|A_1=0, A_2=0] = 81$
(4) Parameters $\varsigma^2_{a1,r,a2}$ are chosen so that $Var[Y|A_1=1, A_2=1] = 100$, and $Var[Y|A_1=1, A_2=0] = Var[Y|A_1=0, A_2=1] = Var[Y|A_1=0, A_2=0] = 81$

Table 15b. Related values for simulation parameters in Table 15a

Scenario	Effect size	Non-response rate $p_0 = p_1$	Corresponding γ values	$E[Y\|A_1{=}a_1, A_2{=}a_2]^{(1),(2)}$ $\mu_{a1,a2}$
Pattern 1: the mean for (1,1) is the highest mean, (0,1) has the next highest mean, the other two are allowed to vary below the mean of (0,1)				
1	0.2	0.5	{12, 3.5, -7, -6, 2, 3}	{11.5, 9, 9.5, 8.5}
2	0.2	0.7	{12, 4.1, -7, -6, 2, 3}	{10.5, 7, 8.5, 7.1}
3	0.2	0.9	{12, 4.7, -7, -6, 2, 3}	{9.5, 5, 7.5, 5.7}
4	0.5	0.5	{12, 3.5, -7, -6, 2, 9}	{14.5, 9, 9.5, 8.5}
5	0.5	0.7	{12, 3.5, -7, -6, 2, 8.25}	{13.575, 6.4, 8.5, 7.1}
6	0.5	0.9	{12, 3.5, -7, -6, 2, 7.75}	{12.575, 3.8, 7.5, 5.7}
Pattern 2: the mean for (1,1) is the highest mean and the other three means are all equal				
7	0.2	0.5	{12, 0, -7, 0, 0, 4}	{10.5, 8.5, 8.5, 8.5}
8	0.2	0.7	{12, 0, -7, 0, 0, 2.85}	{9.0950, 7.1, 7.1, 7.1}
9	0.2	0.9	{12, 0, -7, 0, 0, 2.2}	{7.6800, 5.7, 5.7, 5.7}
10	0.5	0.5	{12, 0, -7, 0, 0, 10}	{13.5, 8.5, 8.5, 8.5}
11	0.5	0.7	{12, 0, -7, 0, 0, 7.15}	{12.105, 7.1, 7.1, 7.1}
12	0.5	0.9	{12, 0, -7, 0, 0, 5.6}	{10.74, 5.7, 5.7, 5.7}

(1) Given for informational purposes, not for generative model
(2) Order in which parameters $\mu_{a1,a2}$ are listed: $\{\mu_{1,1}, \mu_{1,0}, \mu_{1,0}, \mu_{0,0}\}$

4.5 Parameter Values for Data that Challenges the Working Assumptions of Normally Distributed Final Outcome.

Recall that in Section 4.1, we outlined the method for generating the data. To challenge the normality assumption, instead of generating the final outcome given a particular history (A_1, R, A_2) by a normal distribution with mean $E[Y|A_1, R, A_2]$ and variance $Var[Y|A_1, R, A_2]$, we generate from a Gamma distribution with the same mean and variance. That is, instead of generating

$$Y|A_1, R, A_2 \sim N(E[Y|A_1, R, A_2], Var[Y|A_1, R, A_2])$$

we generate $Y|A_1, R, A_2 \sim Gamma(a, b)$ where

$$a = \frac{E[Y \mid A_1, R, A_2]^2}{Var[Y \mid A_1, R, A_2]} \text{ and } b = \frac{Var[Y \mid A_1, R, A_2]}{E[Y \mid A_1, R, A_2]}.$$

The mean of this gamma distribution is a*b and the variance is a*b^2. The skewness of this distribution is calculated by 2*sqrt(1/a), in other words:

$$\text{skewness} = 2 * \sqrt{\frac{Var[Y \mid A_1, R, A_2]}{E[Y \mid A_1, R, A_2]^2}}.$$

Use the values for $E[Y|A_1, R, A_2]$ and $Var[Y|A_1, R, A_2]$ specified in Tables 10a and 11a, but generate the final outcome from a Gamma distribution as outlined here.

5. RESULTS OF SIMULATIONS FOR EVALUATING THE SAMPLE SIZE FORMULAE

We present the simulation results in their entirety, which we could not do in Scott et al[1] due to space constraints. Tables 16a and 16b provide the results of the simulations designed to evaluate the sample size formulas for analyses 3 and 4 respectively. We also examined the ability to perform Analysis 4 when we size for one of the other Analyses, in settings which follow the working assumptions and those which challenge them. Tables Table 17a-c show the results for detecting the best strategy, i.e. Analysis 4, when it is not used in sizing the trial.

Table 16a. Investigation of Sample Size Assumption Violations for Analysis 3 comparing strategies (1,1) and (0,0);
The power to reject the null hypothesis for Analysis 3 is shown when sample size is calculated to reject the null hypothesis for Analysis 3 with power of 0.90 and type I error of 0.05 (two-tailed)

Simulation Parameters				Simulation Results (power)			
Effect size	**Non-response rate (Default)**	**Sample size formula**	**Total sample size**	**Default working assumptions are correct**	**Non-equal non-response rates (1)**	**Non-equal variance(2)**	**Non-normal outcome Y(3)**
0.2	0.5	N_{3a}	1584	0.893	0.902	0.900	0.882
0.2	0.7	N_{3a}	1796	0.922*	0.884	0.892	0.896
0.2	0.9	N_{3a}	2007	0.882	0.910	0.916	0.877*
0.5	0.5	N_{3a}	254	0.896	0.864*	0.920	0.851*
0.5	0.7	N_{3a}	287	0.851*	0.836*	0.872*	0.891
0.5	0.9	N_{3a}	321	0.926*	0.886	0.880	0.898
0.2	0.5	N_{3b}	2112	0.950*	0.958*	0.954*	0.974*
0.2	0.7	N_{3b}	2112	0.960*	0.943*	0.927*	0.945*
0.2	0.9	N_{3b}	2112	0.903	0.934*	0.931*	0.898
0.5	0.5	N_{3b}	338	0.973*	0.938*	0.971*	0.916
0.5	0.7	N_{3b}	338	0.904	0.888	0.939*	0.917
0.5	0.9	N_{3b}	338	0.937*	0.890	0.889	0.922*

(1) $Pr\{R=1| A_1=1, A_2=1\} = p-0.05$ and $Pr\{R=1| A_1=0, A_2=0\} = p+0.05$, where p is the "default" non-response rate.
(2) $Var[Y| A_1=0, A_2=0] = .81*Var[Y| A_1=1, A_2=1]$
(3) The final outcome comes from a gamma distribution.
* The 95% confidence interval for this proportion does not contain 0.90

Table 16b. Investigation of Sample Size Violations for Analysis 4; Probability[1] to detect the correct "best" strategy when the sample size is calculated to detect the correct maximum strategy mean 90% of the time.

Simulation Parameters				**Simulation Results (probability)**			
Effect size	**Non-response rate (Default)**	**Pattern[2]**	**Sample size[3]**	**Default working assumptions are correct**	**Non-equal non-response rates [4]**	**Non-equal variance [5]**	**Non-normal outcome Y [6]**
0.2	0.5	1	608	0.966*	0.984*	0.965*	0.972*
0.2	0.7	1	608	0.972*	0.975*	0.975*	0.979*
0.2	0.9	1	608	0.962*	0.969*	0.964*	0.962*
0.5	0.5	1	97	0.980*	0.985*	0.966*	0.956*
0.5	0.7	1	97	0.961*	0.974*	0.969*	0.972*
0.5	0.9	1	97	0.960*	0.919*	0.976*	0.947*
0.2	0.5	2	608	0.964*	0.953*	0.952*	0.944*
0.2	0.7	2	608	0.926*	0.920*	0.957*	0.937*
0.2	0.9	2	608	0.905	0.929*	0.922*	0.923*
0.5	0.5	2	97	0.922*	0.974*	0.976*	0.948*
0.5	0.7	2	97	0.933*	0.901	0.951*	0.913
0.5	0.9	2	97	0.893	0.917	0.927*	0.885

(1) Probability calculated as the percentage of 1000 simulations on which correct strategy mean was selected as the maximum.
(2) 1 refers to the pattern of strategy means such that all are different, but that for (1,1) is always the highest. 2 refers to the pattern of strategy means such that the mean for (1,1) is higher than the other three and the other three are all equal.
(3) Calculated to detect the correct maximum strategy mean 90% of the time when the sample size assumptions hold.
(4) $Pr\{R=1|A_1=1, A_2=1\} = p-0.05$ and $Pr\{R=1|A_1=a1, A_2=a2\} = p+0.05$, where p is the "default" non-response rate and (a1, a2) is the strategy with the next highest mean.
(5) $Var[Y|A_1=1, A_2=1] = .81*Var[Y|A_1=a1, A_2=a2]$, where (a1, a2) is the strategy with the next highest mean.
(6) The final outcome comes from a gamma distribution
* The 95% confidence interval for this proportion does not contain 0.90.

Table 17a. The probability[1] of choosing the correct strategy for Analysis 4 when sample size is calculated to reject the null hypothesis for Analysis 1 (for a two-tailed test with power of 0.90 and type I error of 0.05)

Simulation Parameters			**Simulation Results**		
Effect size for Analysis 1	**Non-response Rate**	**Sample size**	**Analysis 1 (power)**	**Analysis 4 (probability[1])**	**Effect size for Analysis 4**
0.2	0.5	1056	0.880	1.000	0.325
0.2	0.7	1056	0.896	1.000	0.375
0.2	0.9	1056	0.904	1.000	0.425
0.5	0.5	169	0.934	0.987	0.350
0.5	0.7	169	0.910	0.998	0.490
0.5	0.9	169	0.920	0.998	0.630

(1) Probability calculated as the percentage of 1000 simulations on which correct strategy mean was selected as the maximum.

Table 17b. The probability[1] of choosing the correct strategy for Analysis 4 when sample size is calculated to reject the null hypothesis for Analysis 2 (for a two-tailed test with power of 0.90 and type I error of 0.05)

Simulation Parameters			**Simulation Results**		
Effect size for Analysis 2	**Non-response Rate**	**Sample size**	**Analysis 2 (power)**	**Analysis 4 (probability[1])**	**Effect size for Analysis 4**
0.2	0.5	2112	0.906	0.999	0.133
0.2	0.7	1509	0.897	0.956	0.109
0.2	0.9	1174	0.895	0.716	0.054
0.5	0.5	338	0.895	0.997	0.372
0.5	0.7	241	0.913	0.993	0.397
0.5	0.9	188	0.901	0.978	0.420

[1] Probability calculated as the percentage of 1000 simulations on which correct strategy mean was selected as the maximum.

Table 17c. The probability[1] of choosing the correct strategy for Analysis 4 when sample size is calculated to reject the null hypothesis for Analysis 3 (for a two-tailed test with power of 0.90 and type I error of 0.05)

Simulation Parameters				**Simulation Results**		
Effect size for Analysis 3	**Non-response rate**	**Sample size formula**	**Sample size**	**Analysis 3 (power)**	**Analysis 4 (probability[1])**	**Effect size for Analysis 4**
0.2	0.5	N_{3a}	1584	0.893	0.939	0.10
0.2	0.7	N_{3a}	1796	0.922	0.839	0.06
0.2	0.9	N_{3a}	2007	0.882	0.614	0.02
0.5	0.5	N_{3a}	254	0.896	0.976	0.25
0.5	0.7	N_{3a}	287	0.851	0.990	0.35
0.5	0.9	N_{3a}	321	0.926	0.978	0.32
0.2	0.5	N_{3b}	2112	0.950	0.953	0.10
0.2	0.7	N_{3b}	2112	0.960	0.878	0.06
0.2	0.9	N_{3b}	2112	0.903	0.613	0.02
0.5	0.5	N_{3b}	338	0.973	0.989	0.25
0.5	0.7	N_{3b}	338	0.904	0.990	0.35
0.5	0.9	N_{3b}	338	0.937	0.985	0.32

[1] Probability calculated as the percentage of 1000 simulations on which correct strategy mean was selected as the maximum.

APPENDIX

<u>Matlab Code to Calculate Sample Size for Analysis 4</u>

```
function N = samplesize4(maxn, sigma2, delta, conf)
%
% DESCRIPTION:
% This formula calculates the sample size for sizing a
% SMART trial to select the strategy with the highest
% mean outcome with some specified probability.
% We assume a SMART design with two decision points:
%   1) two initial treatments (A1 in {0,1})
%   2) two treatments (A2 in {0,1}) for
%      non-responders (R = 1) and maintenance treatment
%      for responders (R = 0).
% We make the following assumptions:
% - The marginal variances of the final outcome given the
%   strategy are all equal and we denote this variance by
%   sigma2.  This means that,
%   sigma2 = Var[Y|A1=a1, A2=a2] for all (a1, a2) in
%       {(1,1), (1,0), (0,1), (0,0)}
% - The final outcome Y is normally distributed.
% - The sample sizes will be large enough so that the
%   estimator for the mean is approximately normally
%   distributed.
% - The correlation between
%   * the final outcome Y given treatment strategy (1,1)
%     and Y given treatment strategy (1,0) is the same as
%   * the correlation between Y given treatment strategy
%     (0,1) and Y given treatment strategy (0,0);
%   we denote this identical correlation by r.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% INPUT
% maxn: the maximum sample size the user has available,
%       might be constrained by cost
% sigma2: variance of the final outcome given a strategy
% delta: standardized effect size the user desires to be
%        able to detect
% conf: the probability that the strategy estimated to have
%       the largest mean does indeed have the largest mean
%
% OUTPUT
% N: sample size
%
% EXAMPLE
```

```
% samplesize4(3000, 100, .2, .90);
%
%
%

if nargin < 4
   conf=.9;
end

x0=[1,maxn]';
options = optimset('Display','off','TolFun', 1,
'TolX',.01);

% find u (i.e. sample size) such that p0 = conf
[x,fval,exitflag] = fzero(@(u) prob(u, sigma2,delta,conf),
x0, options)

N=ceil(x) %N is the sample size

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [p0] = prob(n, sigma2, delta, conf)

% function used in ‘samplesize4’

niter=20000;
% we assume the variance is constant across the treatment
% strategies;
% s is the variance of the *estimator* of the strategy
% means:
s=sigma2.*4;
% only making one treatment strategy have an effect, this
% would be the hardest to detect:
one1=[delta.*sqrt(sigma2).*ones(niter,1),zeros(niter,3)];

for i=1:100

    % full range of possibilities for the correlation:
    r=(i-1)./100;
    % Correlation comes from people the treatment share,
    % directly related to the response rate.
    % Structure of Sigma comes from sharing responders and
    % treatments:
    Sigma=[s,r*s,0,0;r*s,s,0,0;0,0,s,r*s;0,0,r*s,s]./n;
        try % to make sure the covariance matrix is valid:
```

```
            R=chol(Sigma);
         catch
            i
            break;
         end;
     % here is the normality assumption:
     R=mvnrnd(one1,Sigma);

     % how often out of 20000 is the first larger than the
     % remaining three, this indicates probability (somewhat
     % analagous to power for hypothesis test) of finding
     % the best strategy:
     test=(R(:,1)> max(R(:,2:4),[],2));
     order(i)=mean(test);

end;

% take the one with the worst power, use that to find the
% sample size:
[p0,j]=min(order');
p0=p0-conf;
```

REFERENCES

1. Scott AI, Levy J, Murphy SA. Statistical Methodology for a SMART Design in the Development of Adaptive Treatment Strategies. American Psychopathological Association. In Press 2007.
2. Hoel P. Introduction to Mathematical Statistics. 5th ed. New York: John Wiley and Sons; 1984.
3. Murphy SA. An Experimental Design for the Development of Adaptive Treatment Strategies. *Statistics in Medicine*. 2005; **24**:1455-1481.
4. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1988.