Effect Size Estimates for the ESCAPE Trial
Proportional Odds Regression Versus Other Statistical Methods

Tolulope T. Sajobi, PhD; Yukun Zhang, MSc; Bijoy K. Menon, MD, MSc; Mayank Goyal, MD; Andrew M. Demchuk, MD; Joseph P. Broderick, MD; Michael D. Hill, MD, MSc

Background and Purpose—Ordinal outcomes, such as modified Rankin Scale (mRS), are the standard primary end points in acute stroke trials. Regression models for assessing treatment efficacy after adjusting for baseline covariates have been developed for continuous, binary, or ordinal end points. There has been no consensus on the best choice of method for analyzing these data.

Methods—We compared several regression models for assessing treatment efficacy in acute stroke trials using existing data sets from the Interventional Management of Stroke-III and Prolyse in Acute Cerebral Thromboembolism II (PROACT-2) trials. Patients with baseline non–contrast computed tomographic Alberta Stroke Program Early CT Score (ASPECTS) > 5, baseline computed tomographic angiography, or conventional angiogram showing an intracranial internal carotid artery or middle cerebral artery trunk (M-1) occlusion, adequate collateral circulation shown on computed tomographic angiography, and treatment times of non–contrast computed tomographic to groin puncture of ≤90 minutes, were included. Monte Carlo techniques were used to compare the statistical power of these regression models under a variety of simulated data analytic scenarios.

Results—Binary logistic regression showed greater power when the treatment is predicted to show evidence of benefit on one end of the mRS with no other gains across other levels of the scale. Proportional odds regression showed greater power when the treatment is predicted to show evidence of improvement on both ends of the mRS.

Conclusions—The mRS distribution for both treatment and control groups influences the power of the investigated statistical models to assess treatment efficacy. A careful evaluation of the expected outcome distribution across the mRS scale is required to determine the best choice of primary analysis. (Stroke. 2015;46:1800-1805. DOI: 10.1161/STROKEAHA.115.009328.)

Key Words: analysis ■ angiography ■ carotid artery, internal ■ middle cerebral artery

Stroke trials have typically reported outcomes on ordinal scales such as the modified Rankin Scale (mRS) and Barthel Index. Existing statistical methods for analyzing randomized control trials data have been developed for continuous or binary end points; ordinal scales have been dichotomized and binary logistic regression used for analysis. However, the dichotomous analysis of the ordinal outcomes data may be statistically less powerful.

Although several models for continuous data have been applied to ordinal outcomes, the assumption of normality may not be tenable when ordinal data are treated as continuous. There are limited statistical methods specifically developed for planning power calculation and analyzing ordinal end points. These methods include Cochran–Mantel–Haenszel tests, Mann–Whitney test, Wilcoxon rank tests, and proportional odds regression.1 The Cochran–Mantel–Haenszel test, Mann–Whitney test, and Wilcoxon rank tests can be used for comparisons between 2 treatment groups, but they cannot provide adjusted treatment effects. The proportional odds regression model assesses the change across the full ordinal scale and the output is a common odds ratio that can be interpreted as the odds of shifting from 1 category to the next. It is appealing...
because of the ability to adjust for covariates and the possibility of analyzing smaller samples. This model relies on the assumption of proportional change (shift) across the entire ordinal scale and this assumption can be tested. This model may mask the magnitude and direction of true treatment effects when the assumption of proportional odds is violated. Furthermore, the use of the proportional odds model for the primary analysis may not necessarily result in a smaller sample size requirement.

Howard et al. proposed a distribution free method of analysis. Another alternative is to use a sliding-dichotomy approach where the expected outcome, a treatment response, is varied according to a baseline measure of clinical severity. The theoretical concern with the mRS, that it is a categorical but not necessarily an interval scale, has recently been shown on empirical grounds to be moot; the mRS can be considered an interval scale, and there has been a growing consensus that the proportional odds model (shift analysis) should be favored to model the treatment response on the mRS in the analysis of acute stroke trials.

In the design of the Endovascular Treatment for Small Core and Anterior Circulation Proximal Occlusion With Emphasis on Minimizing CT to Recanalization Times (ESCAPE) trial, which focused on image-based patient selection for fast endovascular therapy, we were particularly interested in analytic approaches for endovascular ischemic stroke treatment. Using modern imaging and time criteria, we investigated past randomized control trials of endovascular stroke therapy to estimate an expected treatment effect. Using this paradigm, we conducted simulation studies and examined a variety of statistical regression models for estimating the treatment effects with ordinal clinical end points.

Methods

We used existing data from the Interventional Management of Stroke (IMS)-III and Prolyse in Acute Cerebral Thromboembolism II (PROACT-2) trials to examine different statistical methods for estimating treatment effect and expected outcome rates among patients who received endovascular treatments compared with patients in the control group. These 2 trials were selected because data on baseline imaging, baseline occlusion location, and fast treatment times are available. The IMS-III study was an international, randomized, open-label phase 3 clinical trial with blinded outcome assessment conducted from August 2006 to April 2012 that was stopped early for futility. Subjects with moderate-to-severe acute ischemic stroke who had received intravenous tissue-type plasminogen activator (tPA) within 3 hours after symptom onset were randomized in a 2:1 ratio to endovascular treatment or continued intravenous tPA. The primary goal of this trial was to test the efficacy of the intravenous tPA followed by protocol-approved endovascular treatment as compared with standard intravenous tPA. The primary outcome of this study was neurological disability as measured by the mRS 0 to 2 at 90 days measured by assessors blinded to treatment allocation. PROACT-2 was an international, randomized, controlled, multicenter, blinded clinical trial conducted between February 1996 and August 1998. Subjects with acute ischemic stroke of <6 hours of duration caused by angiographically proven occlusion of the middle cerebral artery and without hemorrhage or major early infarction signs on computed tomographic (CT) scan were randomized in a 2:1 ratio to intra-arterial recombinant prourokinase plus unfractionated heparin versus unfractionated heparin control. The objective of the study was to determine the clinical efficacy and safety of intra-arterial recombinant prourokinase infusion. The primary outcome was neurological disability as measured by the mRS 0 to 2 at 90 days, measured by assessors blinded to treatment allocation.

Data From IMS-III and PROACT-2

We selected patients with baseline non–contrast CT Alberta Stroke Program Early CT Score (ASPECTS) >5, baseline CT angiography or angiogram showing an intracranial ICA or M1-segment middle cerebral artery occlusion, adequate or good collateral circulation shown on CT angiography (IMS-III only), and treatment times of non–contrast CT to groin puncture of 90 minutes (IMS-III only). For IMS-III, the efficacy of the endovascular treatments was compared using proportional odds regression, binary logistic regression, multiple linear regression, and robust regression after adjusting for age, sex, stroke severity as measured by the National Institutes of Health Stroke Scale. For PROACT-2, these 4 regression models were adjusted for age, sex, baseline stroke severity as measured by the National Institutes of Health Stroke Scale, and time to treatment (defined as onset-to-randomization time which occurred immediately on angiographic identification of a relevant occlusion). For the proportional odds regression, the assumption of proportionality was tested using the Score test. Binary logistic regression was conducted by estimating the association between the type of treatment received and the proportion of patients who reported good outcomes (mRS=0–2) at 90 days. The estimates of the regression coefficients and the corresponding odds ratio (where appropriate) for the treatment effect along with the corresponding 95% confidence intervals were reported. Statistical significance was assessed as α=0.05.

Independent Simulation Study

A Monte Carlo study was conducted to evaluate the statistical power of several regression methods for testing differences in mRS scores of patients randomized to treatment or control under a variety of data analytic conditions. This included (1) proportional odds regression, (2) binary logistic regression, (3) ordinary least squares regression, and (4) robust regression. In the latter 2 linear regression methods, we make the assumption that the mRS scale can be treated as a normally distributed continuous variable. The simulation conditions investigated included (1) total sample size (n=80, 160, 240, and 400), (2) ratio of group sizes (equal [N_T=N_C] and unequal group sizes [N_T=3N_C]), (3) outcome distribution on the mRS, and (4) violation of distributional assumptions.

Distributions I and II describe the scenario in which the proportion of patients who reported good outcomes (mRS=0, 1, and 2) was significantly higher in the treatment arm than in the controls, and the proportion of patients who reported worse outcomes (mRS=5, 6) were substantially lower in the treatment arm than in the control arm. In distribution III, the proportion of patients who reported worse outcomes was significantly lower in the treatment group than in the controls, but there was no substantial difference in the conditional probabilities for the remaining mRS levels. In distribution IV, the proportion of patients who reported good outcomes was higher in the treatment group than in control groups, but there was no substantial difference in the conditional probabilities for the remaining mRS levels (Table 1).

Ordinal data were generated from a proportional odds regression model as a function of the type of treatment received (treatment versus control), age, and sex (women versus men). When the proportional odds assumption was satisfied (distributions I and II), the data were generated from a proportional odds regression model

$$\log \frac{Pr(Y \leq k \mid X)}{Pr(Y > k \mid X)} = \alpha_k + \beta_1 \cdot \text{trt} + \beta_2 \cdot \text{sex} + \beta_3 \cdot \text{age} + \epsilon_k,$$

where $Pr(Y \leq k \mid X)$ is the probability that an individual’s score on the mRS is less than or equal to k (k=0,1,2,...,5), $\alpha_k$ represents kth intercepts for the proportional odds model, and $(\beta_1, \beta_2, \beta_3)$ represents the vector of corresponding regression coefficients associated with the model covariates. The conditional probabilities derived from the proportional odds model were then used to generate the ordinal data. When the assumption of proportionality was violated, the simulation data based on the distributions were generated using 2 approaches. First, for distributions I and II, the ordinal data were generated from a mis-specified proportional odds logistic regression.
log \[ \frac{\Pr(Y \leq k | X)}{\Pr(Y > k | X)} \] = \alpha_k + \beta_k \text{trt} + \beta_k \text{sex} + (\delta - (k - 1)\beta_k \text{age} + \epsilon 

where \( k=0, \ldots, 5 \) and \( \alpha_k, \beta_k, \delta \), and \( \beta_k \) are as defined in Equation 1.

Second, the conditional probabilities for distributions III and IV were arbitrarily chosen in such a way that the assumption of proportional odds was violated. Description of the simulation conditions is available in Table I in the online-only Data Supplement.

Statistical power of the analytic approaches was considered as the probability to detect a statistical significant treatment effect. For the binary logistic regression, the mRS scale is dichotomized as [0–2] versus [3–6]. In all, 64 simulation conditions were investigated with 1000 replications for each combination of simulation parameters. All analyses were conducted using R software (R Development Core Team, 2013).

### Table 1. Selected mRS Outcome Distributions (%) for Treatment and Control Groups for Simulation Data Generation (1000 Replications)

<table>
<thead>
<tr>
<th>Proportional Odds Assumption</th>
<th>Distribution</th>
<th>Group</th>
<th>mRS Levels</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfied</td>
<td>I</td>
<td>Control</td>
<td></td>
<td>6.8</td>
<td>14.3</td>
<td>9.5</td>
<td>16.4</td>
<td>19.4</td>
<td>20.4</td>
<td>13.2</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td>10.5</td>
<td>19.5</td>
<td>11.4</td>
<td>17.3</td>
<td>17.3</td>
<td>15.3</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>Control</td>
<td></td>
<td>5.5</td>
<td>12.1</td>
<td>8.5</td>
<td>15.5</td>
<td>19.7</td>
<td>22.7</td>
<td>16.0</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td>10.5</td>
<td>19.5</td>
<td>11.4</td>
<td>17.3</td>
<td>17.3</td>
<td>15.3</td>
<td>8.7</td>
</tr>
<tr>
<td>Violated</td>
<td>I</td>
<td>Control</td>
<td></td>
<td>6.8</td>
<td>20.4</td>
<td>19.0</td>
<td>24.5</td>
<td>17.5</td>
<td>9.0</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td>10.5</td>
<td>26.9</td>
<td>20.5</td>
<td>21.5</td>
<td>12.9</td>
<td>5.9</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>Control</td>
<td></td>
<td>5.5</td>
<td>17.5</td>
<td>17.7</td>
<td>25.2</td>
<td>19.8</td>
<td>10.8</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td>10.5</td>
<td>26.9</td>
<td>20.5</td>
<td>21.5</td>
<td>12.9</td>
<td>5.9</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>Control</td>
<td></td>
<td>12.0</td>
<td>11.0</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td>14.0</td>
<td>15.0</td>
<td>17.0</td>
<td>17.0</td>
<td>17.0</td>
<td>17.0</td>
<td>17.0</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>Control</td>
<td></td>
<td>10.0</td>
<td>10.0</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
<td>17.0</td>
<td>18.0</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td>16.0</td>
<td>16.0</td>
<td>13.0</td>
<td>13.0</td>
<td>13.0</td>
<td>14.0</td>
<td>15.0</td>
</tr>
</tbody>
</table>

Distributions I and II: The proportion of patients who reported good outcomes (mRS=0, 1) was significantly higher in the treatment arm than in the controls, but the proportion of patients reported worse outcomes (mRS=5, 6) were substantially lower in the treatment arm than in the control arm; III: the proportion of patients who reported worse outcomes was significantly lower in the treatment group than in the controls, but there was no substantial difference in the conditional probabilities for the remaining mRS levels; IV: the proportion of patients who reported good outcomes was higher in the treatment group than in the control groups, but there was no substantial difference in the conditional probabilities for the remaining mRS levels. mRS indicates modified Rankin Scale.

### Results

The Figure shows the distribution of patients across the mRS for treatment and control groups for both data sets. From IMS-III (N=656), approximately half of the patients enrolled had baseline CT angiography and of these 52 met our inclusion criteria. From PROACT-2 (N=180), 86 met our inclusion criteria. In the IMS-III trial, the magnitude of the treatment effect varied by type of regression model adopted (Tables 2 and 3).

For proportional odds regression, the score test revealed that the assumption of proportional odds is satisfied across all the categories of mRS for each covariate in the model \( P<0.05 \). The proportional odds regression analysis showed an adjusted

![Figure](http://stroke.ahajournals.org/)
common odds ratio of 6.1, indicating that the odds of a 1-point improvement (shift along the mRS scale) is 6.1× greater in favor of treatment. Binary logistic regression showed an odds ratio of 1.4 in favor of endovascular treatment (Tables 2 and 3). Regression models that treat mRS scores as continuous data (ie, multiple linear regression and robust regression) revealed that patients who received the endovascular treatment are likely to have lower scores on the mRS (functional improvement) than patients who received tPA alone.

For PROACT-2, no collateral status or time criteria were applied as these data were unavailable. Furthermore, the identification of occlusions was based on angiography and not CT angiography. A comparison of the regression models showed that there is no evidence of statistically significant association between the type of treatment received and improvement on mRS (Tables 2 and 3). However, age, time to treatment, stroke severity, as measured by the National Institutes of Health Stroke Scale score, were statistically significant predictors of mRS score at 90 days.

Simulation Study Results
The statistical power of each procedure varied by total sample size and outcome distribution across the mRS levels (Table 4). The proportional odds regression was at least 10% more powerful than the binary logistic regression when the proportion of patients with good outcomes and bad outcomes are higher and lower, respectively, in the treatment group than the control groups (ie, distributions I or II). This was also true when the proportional odds assumption was violated.

In contrast, binary logistic regression had the highest power to detect treatment differences among the investigated models.
when the proportion of patients who reported worse outcomes was lower in the treatment group than in control groups (distribution III), or when the proportion of patients who reported good outcomes is higher in the treatment group than in the control group (distribution IV), whereas there were no differences in the conditional probabilities for the remaining mRS levels.

Table 5 describes the average power rates for the regression procedures for equal and unequal group size conditions. The effects of group size ratio on the statistical power of proportional odds regression varied across the distribution of the patients on the mRS level. The average power rates for proportional odds regression were higher under unequal group size conditions than the power rates under equal group size conditions, when the data were sampled from distribution I and II. This procedure was more powerful under equal group size conditions than unequal group size conditions when the proportion of patients who reported good (bad) outcomes was higher (lower) in the treatment group than in control groups, whereas there were no differences in the conditional probabilities for the remaining mRS levels (distributions III or IV). In contrast, the other regression procedures were more powerful under equal group sizes than under unequal group size conditions.

Finally, the multiple linear regression and robust regression were equally as powerful as the proportional odds regression when the proportional odds assumption is satisfied (Table 4). However, when the proportional odds assumption was violated, the power rates for these procedures varied by the distributions patients across the mRS levels and group size conditions (Table 5). The ideal conditions for maximum power for each analytic technique are summarized in Table 6.

**Discussion**

Our analysis confirms that statistical power for ordinal scale analysis varies by mRS outcome distribution. The proportional odds and ordinary least squares regression models are most powerful when the proportion of patients who have good outcomes is higher, and poor outcomes lower in the treatment group than in the control group. The binary logistic regression model is most powerful when the proportion of patients who had good (or bad) outcomes are higher (or lower) in the I group than the other group, while simultaneously there are no other significant difference in the proportion of patients in either groups across other levels of the mRS scale (Table 6). Classically in stroke thrombolytic trials, although there has been demonstrable benefit in good outcome (mRS 0–1 or mRS 0–2), there has not been a similar reduction in major disability or death (mRS 5–6)—an ideal circumstance for binary logistic regression. Part of the art of trial design is to accurately predict the distribution of outcomes.

Although proportional odds regression has been interpreted as ideal because of a smaller sample size requirement and decreased costs and time to complete a trial, this interpretation is only valid if the outcome distribution is correct. The assumption of proportional odds may not be satisfied in practice; the distribution of outcomes may not neatly show the same direction of effect in every category. Furthermore, there are no good methods for estimating interim analysis efficacy or futility boundaries using proportional odds analysis. Where interim analyses for futility or efficacy are desired and the proportional odds approach is the primary outcome, there is an urgent need to develop techniques that allow for appropriate statistical stopping boundaries to be calculated.

Models that treat the ordinal outcome as continuous are equally as powerful as the proportional odds regression when the distribution of outcomes adheres to the proportional odds assumption. This finding is consistent with previous research, which suggests that ordinal outcome scales, for which the
underlying latent construct can be assumed to be continuous, can be analyzed using regression models for continuous data. It would be reasonable to use continuous methods for assessing the mRS in future stroke trials where intervention is predicted to change the expected distribution of scores across the entire scale. This would be a relatively new approach, to use linear regression with the mRS, in the stroke literature. We caution that the robust regression method, which trims observations at either tails of the distributions, is not recommended for analyzing mRS as it may result in loss of clinically relevant information at both ends of the scale leading to bias in the detection of the intended treatment effects.

Finally, an important issue for any analytic approach is the clinical interpretability of the outcome measure. Models that quantify treatment effects as averages on the ordinal scale may be difficult to interpret meaningfully at the bedside. For example, a score of 2.37 on the mRS does not have direct meaning to the individual patient, whereas an integer score does. Similarly, the odds ratio, and the common odds ratio particularly, is not as easy to translate to the bedside as a relative chance of benefit. The risk ratio is directly related to the absolute risk difference and intuitively easier to comprehend. An important limitation of our analysis is that we chose subsets of patients from past randomized controlled trials. There is a risk that our chosen group of interventionally treated patients compared with all controls results in a bias in favor of intervention. We suspect, but do not know, that this bias is relatively small because subsequent randomized controlled trials using the paradigm of imaging selection and fast treatment protocols have shown a similarly large treatment effect.

In conclusion, a careful evaluation of the expected outcome distribution across the mRS scale is required to determine the best choice of primary analysis before trial initiation.

Acknowledgments
The article was written by Drs Sajobi, Menon, and Hill. Statistical simulations were conducted by Y. Zhang and Dr Sajobi. All authors made substantive revisions and approved the final version.

Sources of Funding
The ESCAPE trial was funded by a consortium with grants to the University of Calgary from Covidien, the University of Calgary (Hotchkiss Brain Institute, the Department of Clinical Neurosciences and Calgary Stroke Program, and the Department of Radiology), Alberta Innovates—Health Solutions, the Heart and Stroke Foundation of Canada, and Alberta Health Services.

Table 6. Summary of Ideal Analytic Technique to Maximize Power

<table>
<thead>
<tr>
<th>Expected Distribution of Outcomes on the mRS (Intervention vs Control)</th>
<th>Randomization Ratio</th>
<th>Proportional Odds Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportional odds logistic regression</td>
<td>More good outcomes and less bad outcomes across the scale</td>
<td>n:1</td>
</tr>
<tr>
<td>Binary logistic regression</td>
<td>More good outcomes but no difference in bad outcomes or more bad outcomes but no difference in good outcomes</td>
<td>1:1</td>
</tr>
<tr>
<td>Ordinary least squares regression</td>
<td>More good outcomes and less bad outcomes across the scale</td>
<td>1:1</td>
</tr>
<tr>
<td>Robust linear regression</td>
<td>Not recommended because of truncation</td>
<td>...</td>
</tr>
</tbody>
</table>

mRS indicates modified Rankin Scale; and NA, not applicable.

Disclosures
Dr Broderick provided research monies to the Department of Neurology from Genentech for PRISMS (Phase IIIIB, Double-Blind, Multicenter Study to Evaluate the Efficacy and Safety of Alteplase in Patients With Mild Stroke: Rapidly Improving Symptoms and Neurologic Deficits) Trial; travel to Australian stroke conference paid for by Boehringer Ingelheim. Study medication from Genentech for Interventional Management of Stroke (IMS)-III trial and study catheters supplied during Protocol Versions 1 to 3 by Concentric Inc, EKOS Corp, and Cordis Neurovascular. He was the PI for the IMS-III trial funded by the National Institute of Neurological Disorders and Stroke (National Institutes of Health). The other authors report no conflicts.

References
Effect Size Estimates for the ESCAPE Trial: Proportional Odds Regression Versus Other Statistical Methods
Tolulope T. Sajobi, Yukun Zhang, Bijoy K. Menon, Mayank Goyal, Andrew M. Demchuk, Joseph P. Broderick and Michael D. Hill

Stroke. 2015;46:1800-1805; originally published online May 28, 2015;
doi: 10.1161/STROKEAHA.115.009328

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/46/7/1800

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2015/06/22/STROKEAHA.115.009328.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/
Supplemental Material

Sajobi TT et al. Effect Size Estimates for the ESCAPE trial: Proportional Odds Regression versus other statistical methods

**Supplementary Table I: Description of Simulation Parameters for the Simulation Study**

<table>
<thead>
<tr>
<th>Simulation Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sample Size (N)</td>
<td>80, 160, 240, 400</td>
</tr>
<tr>
<td>Group Sizes</td>
<td>(40, 40), (20, 60) for $N = 80$. Similar patterns for $N = 160, 240, 400$</td>
</tr>
<tr>
<td>Group Size Ratio</td>
<td>1:1, 3:1</td>
</tr>
<tr>
<td>Outcome Distributions*</td>
<td>I, II, III, IV</td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td></td>
</tr>
<tr>
<td>Intercept ($\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$)</td>
<td>(0.5, 1.8, 2.3, 3, 3.8, 5)</td>
</tr>
<tr>
<td>Treatment ($\beta_1$)</td>
<td>0.22, 0.47, 0.69, 1.01, 1.61</td>
</tr>
<tr>
<td>Sex ($\beta_2$)</td>
<td>-0.04</td>
</tr>
<tr>
<td>Age ($\beta_3$)</td>
<td>0.02</td>
</tr>
<tr>
<td>Treatment Effect Size ($e^{\beta_1}$)</td>
<td>1.25, 1.6, 2.0, 3.0, 5.0</td>
</tr>
</tbody>
</table>

*See Table 1 for a description of the conditional probabilities for the mRS outcomes*