Published in final edited form as:

Stroke. 2015 July ; 46(7): 1800–1805. doi:10.1161/STROKEAHA.115.009328.

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, and Michael D. Hill, MD MSc
Calgary Stroke Program, Department of Clinical Neurosciences, Hotchkiss Brain Institute, University of Calgary, Calgary, AB, CANADA

Abstract

Background and Purpose—Ordinal outcomes, such as modified Rankin scale (mRS), are the standard primary endpoints in acute stroke trials. Regression models for assessing treatment efficacy after adjusting for baseline covariates have been developed for continuous, binary, or ordinal endpoints. 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 datasets from IMS III and PROACT 2 trials. Patients with baseline non-contrast CT ASPECTS score > 5, baseline CT angiography (CTA) or conventional angiogram showing an intracranial internal carotid artery (ICA) or middle cerebral artery trunk (M-1) occlusion, adequate collateral circulation shown on CTA, and treatment times of non-contrast CT to groin puncture of 90 minutes or less, were included. Monte Carlo techniques were used to compare the statistical power of these regression models under a variety of simulated data analytic scenarios.

Correspondence to: Michael D. Hill, Professor, Calgary Stroke Program, Department of Clinical Neurosciences, Hotchkiss Brain Institute, University of Calgary, Cumming School of Medicine, Foothills Hospital, Rm 1242A, 1403 29th Street NW, Calgary, Alberta, T2N 2T9, CANADA, FAX 403 283 2270; michael.hill@ucalgary.ca.

Current affiliations:

Tolulope T. Sajobi – Departments of Community Health Sciences and Clinical Neurosciences, Hotchkiss Brain Institute, Faculty of Medicine, University of Calgary, Cumming School of Medicine, Calgary, AB CANADA

Yukun Zhang – Department of Clinical Neurosciences, Faculty of Medicine, University of Calgary, Cumming School of Medicine, Calgary, AB CANADA

Bijoy K Menon - Department of Clinical Neurosciences and Radiology, Hotchkiss Brain Institute, Faculty of Medicine, University of Calgary, Cumming School of Medicine, Calgary, AB CANADA

Mayank Goyal – Department of Radiology, Clinical Neurosciences, Faculty of Medicine, University of Calgary, Cumming School of Medicine, Calgary, AB CANADA

Andrew M Demchuk – Department of Clinical Neurosciences, Radiology, Hotchkiss Brain Institute, Faculty of Medicine, University of Calgary, Cumming School of Medicine, Calgary, AB CANADA

Joseph P Broderick – University of Cincinnati Neuroscience Institute, Department of Neurology and Rehabilitation Medicine, University of Cincinnati, Cincinnati, OH, USA

Michael D. Hill – Departments of Clinical Neurosciences, Medicine, Radiology and Community Health Sciences, Hotchkiss Brain Institute, Faculty of Medicine, University of Calgary, Cumming School of Medicine, Calgary, AB CANADA

Statement of Authorship: The paper was written by TTS, BKM, MDH. Statistical simulations were conducted by YZ and TTS. All authors made substantive revisions and approved the final version.

Disclosures:

Joseph Broderick: research monies to Department of Neurology from Genentech for PRISMS Trial; travel to Australian stroke conference paid for by Boehringer Ingelheim. Study medication from Genentech for IMS III Trial and study catheters supplied during Protocol Versions 1-3 by Concentric Inc, EKOS Corp, and Cordis Neurovascular. JB was the PI for the IMS-3 trial funded by NINDS (NIH).

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.

Keywords

ischemic stroke; analysis; proportional odds models; modified Rankin Scale; endovascular stroke treatment

Introduction

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 (RCTs) data have been developed for continuous or binary endpoints; 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 a number of 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 endpoints. 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 two 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 one 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 upon the assumption of proportional change (“shift”) across the entire ordinal scale and this assumption can be tested. This model may mask the magnitude and/or direction of true treatment effects when the assumption of proportional odds is violated. Further, 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.² 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.³ 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,⁴ 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 ESCAPE trial,^{5, 6} 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 RCTs 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 endpoints.

Methods

We used existing data from the IMS-III and PROACT-2 trials to examine different statistical methods for estimating treatment effect and expected outcome rates among patients who received endovascular treatments compared to patients in the control group.^{7, 8} These two 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 t-PA 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 t-PA followed by protocol-approved endovascular treatment as compared with standard intravenous t-PA. The primary outcome of this study was neurological disability as measured by the mRS 0-2 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 less than 6 hours' duration caused by angiographically proven occlusion of the MCA and without hemorrhage or major early infarction signs on computed tomographic scan were randomized in a 2:1 ratio to intra-arterial rproUK plus unfractionated heparin vs. unfractionated heparin control.⁸ The objective of the study was to determine the clinical efficacy and safety of intra-arterial (IA) recombinant prourokinase (r-proUK) infusion. The primary outcome was neurological disability as measured by the mRS 0-2 at 90 days, measured by assessors blinded to treatment allocation.

Data from IMS-III and from PROACT-2

We selected patients with baseline non-contrast CT ASPECTS score > 5, baseline CTA or angiogram showing an intracranial ICA or M1-segment MCA occlusion, adequate or good collateral circulation shown on CTA (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 (NIHSS). For PROACT-2, these four regression models were adjusted for age, sex, baseline stroke severity as measured by the NIHSS, and time to treatment (defined as onset-to-randomization time which occurred immediately upon angiographic identification of a relevant occlusion). For the proportional odds regression, the assumption of proportionality was tested using the Score

test.¹ 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/or the corresponding odds ratio (where appropriate) for the treatment effect along with the corresponding 95% confidence intervals were reported. Statistical significance was assessed as $\alpha = 0.05$.

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 (a) proportional odds regression, (b) binary logistic regression, (c) ordinary least squares regression, and (d) robust regression. In the latter two 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 (a) total sample size [$N = 80, 160, 240$, and 400], (b) ratio of group sizes [Equal ($N_1 = N_2$) and unequal group sizes ($N_1 = 3N_2$)], (c) outcome distribution on the mRS, and (d) violation of distributional assumptions.

Distributions I and II describe the scenario in which the proportion of patients who reported good outcomes (mRS = 0, 1, 2) was significantly higher in the treatment arm than the controls, and the proportion of patients who reported worse outcomes (mRS = 5, 6) were substantially lower in the treatment arm than the control arm. In distribution III, the proportion of patients who reported worse outcomes was significantly lower in the treatment group than 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 control groups, but there was no substantial difference in the conditional probabilities for the remaining mRS levels (See 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 (female vs. male). When the proportional odds assumption was satisfied (Distributions I and II), the data were generated from a proportional odds regression model

$$\log \left[\frac{Pr(Y \leq k|X)}{Pr(Y > k|X)} \right] = \alpha_k + \beta_1 Trt + \beta_2 Sex + \beta_3 Age + \varepsilon_i \quad (1)$$

where $P(Y \leq k|X)$ is the probability that an individual's score on the mRS scale is less than or equal to k ($k = 0, 1, \dots, 5$), α_k represents k th 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 two approaches. First, for distributions I and II, the ordinal data were generated from a mis-specified proportional odds logistic regression

$$\log \left[\frac{Pr(Y \leq k|X)}{Pr(Y > k|X)} \right] = \alpha_k + \beta_1 Trt + \beta_2 Sex + (\delta - (k - 1) \beta_3) Age + \varepsilon_i, \quad (2)$$

where $k = 0, \dots, 5$, and α_k , β_1 , β_2 , and β_3 are as defined in equation 1. Secondly, the conditional probabilities for distributions III and IV were arbitrarily chosen in such a way that the assumption of proportional odds was violated. See Supplementary Table I for a description of the simulation conditions.

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).

Results

Figure 1 shows the distribution of patients across the mRS for treatment and control groups for both datasets. 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 (See Table 2). 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 common odds ratio of 6.1, indicating that the odds of a 1 point improvement (“shift” along the mRS scale) is 6.1-times greater in favour of treatment. Binary logistic regression showed an OR = 1.4 in favour of endovascular treatment (Table 2). Regression models that treat mRS scores as continuous data (i.e., 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 t-PA alone.

For PROACT-2, no collateral status or time criteria were applied as these data were unavailable. Further, the identification of occlusions was based upon angiography and not CTA. 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 (Table 2). However, age, time to treatment, stroke severity, as measured by NIHSS 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 3). 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 (i.e., 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 control groups (distribution III), or when the proportion of patients who reported good outcomes is higher in the treatment group than the control group (distribution IV), while there were no differences in the conditional probabilities for the remaining mRS levels.

Table 4 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. Whereas, 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 control groups while 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 3). 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 4). The ideal conditions for maximum power for each analytic technique are summarized in Table 5.

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 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 one 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 5) Classically in stroke thrombolytic trials, while 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. Further, 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.⁹ 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, a very 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 to all controls results in a bias in favour 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.^{6, 10, 11}

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

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

None for this work which was done in planning for the ESCAPE trial.

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

References

1. Agresti, A. Categorical data analysis. Wiley; Hoboken, NJ: 2013. Models for Multinomial Response; p. 293-338.
2. Howard G, Waller JL, Voeks JH, Howard VJ, Jauch EC, Lees KR, et al. A simple, assumption-free, and clinically interpretable approach for analysis of modified rankin outcomes. *Stroke*. 2012; 43:664–669. [PubMed: 22343650]
3. Saver JL. Optimal end points for acute stroke therapy trials: Best ways to measure treatment effects of drugs and devices. *Stroke*. 2011; 42:2356–2362. [PubMed: 21719772]
4. Hong KS, Saver JL. Years of disability-adjusted life gained as a result of thrombolytic therapy for acute ischemic stroke. *Stroke*. 2010; 41:471–477. [PubMed: 20133917]
5. Demchuk AM, Goyal M, Menon BK, Eesa M, Ryckborst KJ, Kamal N, et al. Endovascular treatment for small core and anterior circulation proximal occlusion with emphasis on minimizing ct to recanalization times (escape) trial: Methodology. *Int J Stroke*. 2015; 10:429–438. [PubMed: 25546514]
6. Goyal M, Demchuk AM, Menon BK, Eesa M, Rempel JL, Thornton J, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med*. 2015; 372:1019–1030. [PubMed: 25671798]
7. Broderick JP, Palesch YY, Demchuk AM, Yeatts SD, Khatri P, Hill MD, et al. Endovascular therapy after intravenous t-pa versus t-pa alone for stroke. *The New England journal of medicine*. 2013; 368:893–903. [PubMed: 23390923]
8. Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, et al. Intra-arterial prourokinase for acute ischemic stroke. The proact ii study: A randomized controlled trial. *Prolyse in acute cerebral thromboembolism*. *JAMA*. 1999; 282:2003–2011. [PubMed: 10591382]
9. Walters SJ, Campbell MJ, Lall R. Design and analysis of trials with quality of life as an outcome: A practical guide. *J Biopharm Stat*. 2001; 11:155–176. [PubMed: 11725929]
10. Campbell BC, Mitchell PJ, Kleinig TJ, Dewey HM, Churilov L, Yassi N, et al. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med*. 2015; 372:1009–1018. [PubMed: 25671797]
11. Saver JL, Goyal M, Bonafe A, Diener HC, Levy EI, Pereira VM, et al. Stent-retriever thrombectomy after intravenous t-pa vs. T-pa alone in stroke. *N Engl J Med*. 2015 in press.

Figure 1. Percentage Distribution of Modified Rankin scores for IMS III and PROACT II Trials in patients with non-contrast CT ASPECTS >5, proximal occlusions and in IMS III patients, moderate to good baseline collaterals.

Table 1

Selected mRS Outcome Distributions (%) for Treatment and Control Groups for Simulation Data Generation (1000 replications)

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

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

Table 2

A Comparison of Odds Ratio, Regression Coefficients and 95% CI for Endovascular Treatment Effect Sizes

	Proportional odds regression		Binary Logistic Regression [@]	
	OR	95% CI	OR	95% CI
IMS-III (n= 52)				
Treatment (endovascular)	6.07 [*]	(2.01, 19.63)**	1.36 [*]	(1.06, 1.75)
Age	0.88	(0.31, 2.50)	0.94	(0.72, 1.24)
Sex (male)	1.51	(0.55, 4.19)	0.93	(0.72, 1.20)
Baseline NIHSS	0.14 [*]	(0.04, 0.43)	0.67 [*]	(0.51, 0.87)
PROACT-2 (n= 86)				
Treatment (endovascular)	1.75	(0.76, 4.13)	1.17	(0.98, 1.39)
Age	0.93 [*]	(0.90, 0.96)	0.98 [*]	(0.98, 0.99)
Sex (male)	0.80	(0.35, 1.81)	0.93	(0.79, 1.10)
Baseline NIHSS	0.86 [*]	(0.78, 0.94)	0.99	(0.97, 1.01)
Time to Treatment	0.42 [*]	(0.26, 0.67)	0.89 [*]	(0.81, 0.97)
	Multiple Linear Regression		Robust Regression	
	β	95% CI	β	95% CI
IMS-III (n= 52)				
Treatment (endovascular)	-1.45 [*]	(-2.46, -0.45)	-1.65 [*]	(-2.65, -0.66)
Age	0.48	(-0.56, 1.53)	0.31	(-0.74, 1.35)
Sex (male)	-0.28	(-1.30, 0.74)	-0.30	(-1.31, 0.72)
Baseline NIHSS	1.74 [*]	(0.69, 2.78)	1.92 [*]	(0.88, 2.96)
PROACT-2 (n=86)				
Treatment (endovascular)	-0.47	(-1.24, 0.30)	-0.40	(-1.16, 0.36)
Age	0.07 [*]	(0.04, 0.10)	0.07 [*]	(0.05, 0.10)
Sex (male)	0.16	(-0.57, 0.88)	0.15	(-0.57, 0.87)

NB: NIHSS = National Institutes of Health Stroke Scale

OR = Odds Ratio; 95%CI = 95% Confidence Interval

[@] The binary logistic regression uses mRS Of 0 -2 versus mRS of 3-6^{*} $p < 0.05$

Table 3

Average Power (%) for Regression Procedures by Distribution of Conditional Probabilities and Total Sample Size

Distribution	Sample Size	Proportional Odds	Binary LR	OLS	Robust
Proportional Odds Assumption Satisfied					
I	80	19.7	10.2	18.3	19.5
	160	34.8	20.8	33.9	34.8
	240	49.8	29.9	49.0	49.7
	400	69.7	45.0	69.6	69.4
II	80	35.2	15.5	33.0	35.5
	160	62.4	38.2	62.4	62.6
	240	80.2	54.1	79.6	79.6
	400	94.0	75.3	93.7	94.1
Proportional odds Assumption Violated					
I	80	19.4	12.0	18.4	20.5
	160	34.3	25.3	33.5	35.7
	240	48.9	34.9	48.2	50.2
	400	69.0	50.9	68.4	70.1
II	80	34.2	19.5	33.3	35.1
	160	61.1	41.3	60.4	62.7
	240	79.7	58.9	79.2	80.6
	400	94.0	78.3	93.7	93.9
III	80	18.3	16.8	17.4	19.2
	160	27.5	33.6	26.7	28.4
	240	39.3	51.1	38.4	39.3
	400	57.5	70.9	57.1	57.6
IV	80	16.1	24.0	15.4	16.9
	160	24.4	36.8	24.2	24.8
	240	31.8	50.9	32.3	32.6
	400	47.3	73.9	49.2	49.0

Note: Binary LR = Binary logistic regression; OLS = Ordinary least squares linear regression; Robust = Robust Regression

Table 4

Average Power Rates (%) for Regression Procedures by Distribution of Conditional Probabilities and Group Size Ratio for Treatment and Control Groups (Proportional Odds Assumption Violated)

Conditional Probability Distribution	Sample Size	Proportional Odds	Binary LR	OLS	Robust
Equal Group Sizes ($N_1 = N_2$)					
I	80	17.2	15.4	20.7	22.4
	160	30.2	29.1	36.4	39.2
	240	45.2	38.9	51.4	53.7
	400	63.3	56.2	74.0	75.2
II	80	37.8	25.6	36.9	38.7
	160	67.1	49.0	66.6	68.0
	240	85.1	66.5	84.3	85.5
	400	97.0	84.9	96.7	97.1
III	80	20.3	22.0	19.3	21.6
	160	31.8	40.0	30.4	32.7
	240	44.7	58.1	44.2	45.2
	400	64.9	78.8	64.7	65.3
IV	80	16.7	25.2	16.4	17.2
	160	25.2	41.1	25.4	25.5
	240	34.4	57.3	35.2	35.7
	400	51.8	80.6	54.1	54.2
Unequal Group Sizes ($N_1 = 3N_2$)					
I	80	21.6	8.6	16.1	18.5
	160	38.4	21.4	30.6	32.2
	240	52.5	30.8	44.9	46.6
	400	74.7	45.6	62.8	64.9
II	80	30.6	13.4	29.6	31.4
	160	55.0	33.5	54.2	57.3
	240	74.2	51.2	74.1	75.7
	400	90.9	71.7	90.6	90.7
III	80	16.3	11.5	15.5	16.7
	160	23.2	27.2	23.0	24.0
	240	33.9	44.1	32.6	33.4
	400	50.0	62.9	49.4	49.9
IV	80	15.5	22.8	14.4	16.5
	160	23.6	32.4	23.0	24.1
	240	29.1	44.5	29.3	29.5

Conditional Probability Distribution	Sample Size	Proportional Odds	Binary LR	OLS	Robust
	400	42.8	67.2	44.3	43.7

Note: Binary LR = Binary logistic regression; OLS = Ordinary least squares linear regression; Robust = Robust Regression

Table 5

Summary of Ideal Analytic Technique to Maximize Power

	Expected distribution of outcomes on the mRS (intervention v. control)	Randomization ratio	Proportional odds assumption
Proportional Odds logistic regression	More good outcomes and less bad outcomes across the scale	n:1	Valid
Binary logistic regression	More good outcomes but no difference in bad outcomes OR more bad outcomes but no difference in good outcomes	1:1	n/a
Ordinary least squares regression	More good outcomes and less bad outcomes across the scale	1:1	Valid
Robust linear regression	Not recommended because of truncation	---	---