Assessing an ordinal outcome:
Proportional odds analysis

OKLAHOMA
Clinical and Translational Science Institute
Located at The University of Oklahoma Health Sciences Center

Statistical concepts for clinical investigators
David M. Thompson
Oklahoma Shared Translational and Clinical Resource
September 6, 2016

Dichotomizing an ordinal outcome scale like the mRS or the GCS (Roozenbeek, et al., 2011; Bath et al., 2012) ignores some of the information that the scale provides. A “proportional odds analysis” is one of several approaches under investigation to produce more powerful assessment of outcomes that are measured on an ordinal scale. These approaches try to use all the information contained among the ordered categories, and to avoid the loss of information inherent when one dichotomizes the scale at a threshold.

Proportional odds analysis is performed using a “proportional odds logistic regression,” also called ordinal logistic regression (Valenta, Pitha, Poledne, 2006; Rozenbeek et al., 2011). With reference to the percentages reported in Table 1a of:

Goyal M, Demchuk AM, Menon BK, Eesa M, Rempel JL, et al. Randomized Assessment of Rapid Endovascular Treatment of Ischemic Stroke. N Engl J Med. 2015 Feb 11. [Epub ahead of print]

I used SAS PROC LOGISTIC to a produce a proportional odds analysis, and obtained results that were close to those the article reported:

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	0	1	-2.6795	0.2261	140.4258	<.0001
Intercept	1	1	-1.5711	0.1808	75.4876	<.0001
Intercept	2	1	-0.8381	0.1644	25.9777	<.0001
Intercept	3	1	-0.1661	0.1572	1.1156	0.2909
Intercept	4	1	0.7211	0.1641	19.3004	<.0001
intervention		1	0.9976	0.2073	23.1681	<.0001

Estimate of Common Odds Ratio

Effect	Point Estimate	95% Wald Confidence Limits	
intervention	2.712	1.807	4.071

Interpreting the common odds ratio

The analysis produces a “common odds ratio,” which the authors describe in the footnote to Figure 1A as “the odds of improvement of 1 point on the modified Rankin scale.” Similarly, they state that the “common odds ratio (indicating the odds of improvement of 1 point on the modified Rankin scale) of 2.6 (95% confidence interval [CI], 1.7 to 3.8) [favors] the intervention (P<0.001).”

Given how the proportional odds logistic regression model defines the common odds ratio, *it is not strictly accurate to describe it in terms of an improvement of one level in an ordinal outcome like the mRS.*

The authors' interpretation is on firmer ground in the abstract, where they describe "the common odds ratio as a measure of the likelihood that the intervention would lead to lower scores on the modified Rankin scale than would control care..."

Here's how it works. When the odds are truly proportional, all of the following odds ratios are the same:

> The odds that an mRS of 0 (as opposed to a score of 1 or higher) is observed in someone in the intervention group is 2.6 times higher than those same odds for someone in the control group.
>
> The odds that an mRS of 0 or 1 (as opposed to a score of 2 or higher) is observed in someone in the intervention group is 2.6 times higher than those same odds for someone in the control group.
>
> The odds that an mRS of 0,1 or 2 (as opposed to a score of 3 or higher) is observed in someone in the intervention group is 2.6 times higher than those same odds for someone in the control group.
>
> The odds that an MRS of 0,1, 2 or 3 (as opposed to a score of 4 or higher) is observed in someone in the intervention group is 2.6 times higher than those same odds for someone in the control group.

You get the idea; the common odds ratio applies, no matter what threshold we attach to the mRS.

Of course, this interpretation's validity rests on the assumption that the odds at different threshold values or cutpoints are truly proportional. A footnote to Table 2 reports that "the proportional odds assumption was tested and found to be valid." I verified this in SAS PROC LOGISTIC.

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
1.5334	4	0.8207

The non-significant p value means that evidence from the sample is insufficient to bring into question the null hypothesis, which is that the odds of the outcome at each cutpoint are proportional between the intervention and control groups. Therefore, a single or common odds ratio can describe the group's relationship regardless of the cutpoint that is used.

References

Bath PM, Lees KR, Schellinger PD, Altman H, Bland M, Hogg C, Howard G, Saver JL; European Stroke Organisation Outcomes Working Group. Statistical analysis of the primary outcome in acute stroke trials. Stroke. 2012 Apr;43(4):1171-8.

Roozenbeek B, Lingsma HF, Perel P, Edwards P, Roberts I, Murray GD, Maas AI, Steyerberg EW; IMPACT (International Mission on Prognosis and Clinical Trial Design in Traumatic Brain Injury) Study Group; CRASH (Corticosteroid Randomisation After Significant Head Injury) Trial Collaborators. The added value of ordinal analysis in clinical trials: an example in traumatic brain injury. Critical Care. 2011;15(3):R127. doi: 10.1186/cc10240.

Valenta Z, Pitha J, Poledne R. Proportional odds logistic regression--effective means of dealing with limited uncertainty in dichotomizing clinical outcomes. Stat Med. 2006 Dec 30;25(24):4227-34. Online: http://onlinelibrary.wiley.com/doi/10.1002/sim.2678/pdf