

# On testing proportional odds assumptions for proportional odds models

Anqi Liu,<sup>1</sup> Hua He,<sup>2</sup> Xin M Tu,<sup>3</sup> Wan Tang <sup>1</sup>

**To cite:** Liu A, He H, Tu XM, *et al.* On testing proportional odds assumptions for proportional odds models. *General Psychiatry* 2023;**36**:e101048. doi:10.1136/gpsych-2023-101048

Received 03 March 2023  
Accepted 06 June 2023

## ABSTRACT

Proportional odds models are commonly used to model ordinal responses, but the proportional odds assumption may not hold in practice, leading to biased inference. Tests such as score, Wald and likelihood ratio (LR) have been proposed to evaluate the proportional odds assumption based on models without the assumption. Brant has proposed an independent binary model-based Wald-type test, and Wolfe and Gould have extended the idea to propose an LR-type test.

This paper provides a brief review of the Brant and Wolfe-Gould tests for evaluating the proportional odds assumption and evaluates their performance through simulation studies and a real data example. Sample programs are provided in SAS, SPSS and Stata to facilitate the implementation of these tests using standard statistical software packages.

This study highlights the importance of evaluating the proportional odds assumption when using proportional odds models for ordinal responses. The sample programs provided in this paper make it easy for researchers to apply these tests in their own analyses using standard statistical software packages.

## INTRODUCTION

Categorical variables are common in biomedical and psychosocial studies. For regression analysis of a binary response, logistic regression models may be the most popular. In a logistic model, the coefficients can be easily interpreted in terms of odds ratios (ORs). For an ordinal response where the response levels are ordered, it is common to model it through cumulative probabilities. In other words, the response is dichotomised based on the order using all possible cutpoints, and then regression models are applied to the resulting binary responses. More precisely, suppose the ordinal response levels of an ordinal response  $Y$  are labelled as  $1, 2, \dots, J$  according to their order, then for each  $j = 1, \dots, J - 1$ , we may dichotomise the outcome into two groups:  $y^{(j)} = 1$  if  $1 \leq y \leq j$  and  $y^{(j)} = 0$  if  $j + 1 \leq y \leq J$ . These  $J - 1$  dichotomised binaries all together convey the original level. Using the binaries allows us to model the ordinal outcome using models for binaries such as logistic models.

Let  $\gamma_j = \Pr(y^{(j)} = 1) = \Pr(Y \leq j)$  be the cumulative probability for the response to take a level up to  $j$ , then a cumulative logistic regression model can be specified as

$$\log\left(\frac{\gamma_j(\mathbf{x})}{1-\gamma_j(\mathbf{x})}\right) = \alpha_j + \beta_j^\top \mathbf{x}, \quad j = 1, \dots, J - 1, \quad (1)$$

where  $\mathbf{x}$  is the vector of independent variables. It is commonly assumed that all the  $\beta_j$  in model (1) are the same, resulting in the following proportional odds model:

$$\log\left(\frac{\gamma_j(\mathbf{x})}{1-\gamma_j(\mathbf{x})}\right) = \alpha_j + \beta^\top \mathbf{x}, \quad j = 1, \dots, J - 1. \quad (2)$$

The aforementioned equation is also called a model with parallel or equal slopes.<sup>1</sup>

Under model (2), for any two subjects with independent variables  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , the OR

$$\frac{\gamma_j(\mathbf{x}_1)/(1-\gamma_j(\mathbf{x}_1))}{\gamma_j(\mathbf{x}_2)/(1-\gamma_j(\mathbf{x}_2))} = \exp(\beta^\top (\mathbf{x}_1 - \mathbf{x}_2)), \quad j = 1, \dots, J - 1, \quad (3)$$

is independent of the cut point  $j$ . This property is called the *proportional odds property*, and model (2) is called a *proportional odds model*. This proportional odds property comes from the assumption that all  $\beta_j$  are the same. The proportional odds model may be the most popular model for ordinal response; however, the proportional odds assumption may be too strong. Thus, it is generally desired to test the proportional odds assumptions.

Based on model (1), the null hypothesis for testing the proportional odds assumption is given by

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{J-1}. \quad (4)$$

Score, Wald and likelihood ratio (LR) tests may all be applied to the hypothesis test. However, model (1) may not be estimable; when the coefficients of the covariates are different across different levels, the fitted probabilities for some levels may be negative. To overcome the issue, Brant proposed an approach which first estimates  $\beta_j$  separately based on the dichotomised binary responses and then compares the estimates through a



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Department of Biostatistics and Data Science, Tulane University, New Orleans, Louisiana, USA

<sup>2</sup>Department of Epidemiology, Tulane University, New Orleans, Louisiana, USA

<sup>3</sup>Division of Biostatistics and Bioinformatics, Herbert Wertheim School of Public Health and Human Longevity Science, UC San Diego, La Jolla, California, USA

## Correspondence to

Dr Wan Tang;  
wtang1@tulane.edu

Wald-like statistic based on their joint asymptotic distribution.<sup>2</sup> Wolfe and Gould generalised the idea of obtaining an LR test.<sup>3</sup>

In the section ‘Testing the proportional odds assumption’, we provide a brief description of these tests, as well as their availability in R, SAS, SPSS and Stata. In the section ‘Simulation studies’, simulation studies are carried out to assess the performances of these tests. Finally, a real data example is given in the section ‘Examples’, and the paper concludes with a discussion.

**TESTING THE PROPORTIONAL ODDS ASSUMPTION**

Let  $(\mathbf{x}_i, y_i), i = 1, 2, \dots, n$ , be an i.i.d. sample with  $\mathbf{x}_i$  being the vector of independent variables and  $y_i$  the ordinal response with outcome levels  $1, \dots, J$ . Let  $y_i^{(j)}$  be the dichotomised binary response for  $y_i$  at level  $j$  and  $\gamma_{ij} = \Pr(y_i^{(j)} = 1)$  be the cumulative probability for the response to take a level up to  $j$ , then  $\Pr(y_i = j) = \gamma_{ij} - \gamma_{ij-1}$ . Based on the cumulative model (1), the likelihood function is given by

$$L(\boldsymbol{\theta}; \mathbf{x}, \mathbf{y}) = \prod_{i=1}^n \prod_{j=1}^K (\gamma_{ij} - \gamma_{ij-1})^{I(y_i=j)},$$

where  $\boldsymbol{\theta}^\top = (\alpha_1, \beta_1^\top, \dots, \alpha_{J-1}, \beta_{J-1}^\top)$  includes all the parameters.

The maximum likelihood estimate (MLE) of the parameters of the cumulative model can be obtained by solving the following score equation:

$$\frac{\partial \log L}{\partial \beta_j^*} = \sum_{i=1}^n \left( I(y_i = j) \frac{1}{\pi_{ij}} - I(y_i = j+1) \frac{1}{\pi_{i,j+1}} \right) \gamma_{ij} (1 - \gamma_{ij}) \mathbf{x}_i^* = 0,$$

where  $\beta_j^* = (\alpha_j, \beta_j^\top)$ ,  $\mathbf{x}_i^* = (1, \mathbf{x}_i)$  and  $\pi_{ij} = \gamma_{ij} - \gamma_{ij-1}$ ,  $j = 1, \dots, J-1$ . The asymptotic variance of the MLE  $\hat{\boldsymbol{\theta}} = (\hat{\beta}_1^*, \dots, \hat{\beta}_{j-1}^*)$  may be estimated by  $I^{-1}$ , where  $I = (I_{jk})_{1 \leq j \leq J-1, 1 \leq k \leq J-1}$ , with the  $(j, k)$  th block  $I_{jk} = -\frac{\partial^2 \log L}{\partial \beta_j^* \partial \beta_k^{*T}}$ , is the Fisher information matrix. Note that the size of the block  $I_{jk}$  is  $(s+1) \times (s+1)$ , where  $s = \dim(\mathbf{x}_i)$ .

**Wald test**

Based on the asymptotic distribution of  $\hat{\boldsymbol{\theta}}$ , we can derive the Wald statistic for hypothesis (4). The null hypothesis (4) can be written in a matrix form

$$L * (\beta_1^\top, \beta_2^\top, \dots, \beta_{J-1}^\top)^\top = 0, \text{ where}$$

$$L = \begin{pmatrix} I_{s \times s} & 0_{s \times s} & \cdots & 0_{s \times s} & -I_{s \times s} \\ 0_{s \times s} & I_{s \times s} & \cdots & 0_{s \times s} & -I_{s \times s} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0_{s \times s} & 0_{s \times s} & \cdots & I_{s \times s} & -I_{s \times s} \end{pmatrix}$$

is an  $(J-2) \times (J-1)$  block matrix,  $I_{s \times s}$  is the  $s \times s$  identity matrix, and  $0_{s \times s}$  the  $s \times s$  matrix with all

entries 0. Let  $V$  be the estimated asymptotic variance of  $(\hat{\beta}_1^\top, \hat{\beta}_2^\top, \dots, \hat{\beta}_{j-1}^\top)$ , then the asymptotic variance of  $L * (\hat{\beta}_1^\top, \hat{\beta}_2^\top, \dots, \hat{\beta}_{j-1}^\top)$  is  $L * V * L^\top$ , and the Wald statistic can be defined as

$$s_{wald} = (\beta_1^\top, \beta_2^\top, \dots, \beta_{j-1}^\top) L^\top (L * V * L^\top)^{-1} L * (\beta_1^\top, \beta_2^\top, \dots, \beta_{j-1}^\top)^\top.$$

Under the null hypothesis of proportional odds assumption, the Wald statistic follows a  $\chi^2$  distribution with  $(J-2) s$  degrees of freedom (df).

**Likelihood test**

The likelihood test is defined based on the LR of the data set under models (1) and (2). More precisely, let  $l(\boldsymbol{\theta}; \mathbf{x}, \mathbf{y}) = \log(L(\boldsymbol{\theta}; \mathbf{x}, \mathbf{y}))$  be the log-likelihood function of the data, then the LR statistic is defined as

$$s_{LR} = 2[l(\hat{\boldsymbol{\theta}}; \mathbf{x}, \mathbf{y}) - l(\tilde{\boldsymbol{\theta}}; \mathbf{x}, \mathbf{y})], \tag{5}$$

where  $\hat{\boldsymbol{\theta}}$  and  $\tilde{\boldsymbol{\theta}}$  are the MLEs of the parameters for model (1) and model (2), respectively. Under the null hypothesis,  $s_{LR}$  asymptotically follows a  $\chi^2$  distribution with  $(J-2) s$  df.

**Score test**

The score test is defined based on the assumption that the null hypothesis is true. The score statistic is defined as

$$s_{score} = \frac{1}{n} \left( \frac{\partial \log L}{\partial \boldsymbol{\theta}}(\tilde{\boldsymbol{\theta}}; \mathbf{x}, \mathbf{y}) \right) \Gamma^{-1}(\tilde{\boldsymbol{\theta}}; \mathbf{x}, \mathbf{y}) \left( \frac{\partial \log L}{\partial \boldsymbol{\theta}}(\tilde{\boldsymbol{\theta}}; \mathbf{x}, \mathbf{y}) \right)^\top,$$

where both the score function and Fisher information matrix are evaluated at the MLE under the null hypothesis,  $\tilde{\boldsymbol{\theta}}$ . Under the null hypothesis, the score statistic also follows the  $\chi^2$  distribution with  $(J-2) s$  df.

Note that only the MLE under the proportional odds model is required for the score test. For the Wald and LR tests, we need to calculate the MLE for model (1). However, this MLE may not exist. In such cases, Wald and LR tests are not feasible.

**Brant test**

Brant proposed to fit the  $J-1$  logistic models included in model (1) for the dichotomised binary response  $y^{(j)}$  separately and then compare the estimated  $\beta_j^*$ s.<sup>2</sup> For each  $j = 1, \dots, J-1$ , let  $\hat{\beta}_j$  be the MLE of  $\beta_j$  based on the logistic model

$$\log \left( \frac{\gamma_j(\mathbf{x})}{1 - \gamma_j(\mathbf{x})} \right) = \alpha_j + \beta_j^\top \mathbf{x}. \tag{6}$$

The MLE  $\hat{\beta}_j$  almost always exists, unless there are data separation issues.<sup>1</sup> Based on the asymptotic linearity of  $\hat{\beta}_j$ , Brant derived their asymptotic joint distribution, and a Wald-type test for the proportionality can be obtained.

This approach is equivalent to modelling all the dichotomised binary  $y^{(j)}$ s together, but treating them as independent. The MLE  $\hat{\beta}_j$  based on the separate logistic regression models can be obtained from fitting a single

model by treating the dichotomised binary response from the subjects as independent:

$$\text{logit} \left( \Pr \left( y^{(j)} = 1 \right) \right) = \alpha_j + \beta_j^\top \mathbf{x}, j = 1, \dots, J - 1, \quad (7)$$

This model is called an independent binary model.

Note that some of the dichotomised binary outcomes may suffer from the issue of data separation, and the MLEs of the corresponding logistic models do not exist. In such cases, the Brant test cannot be computed.

### Wolfe-Gould test

Wolfe and Gould proposed an approximation to the LR statistics also using the MLE  $\hat{\beta}_j$  based on the separate logistic regression models.<sup>3</sup> The likelihood  $l(\hat{\theta}; \mathbf{x}, \mathbf{y})$  in (5) is replaced with  $l(\hat{\theta}; \mathbf{x}, \mathbf{y})$ , where  $\hat{\theta}$  is the MLE for independent binary model (7). Since  $\hat{\theta}$  is not as efficient as the MLE  $\hat{\theta}$ , Wolfe and Gould also proposed using an inefficient estimate for proportional model (2) by treating all the dichotomised binaries as independent. In other words, estimate the following independent binary model:

$$\text{logit} \left( \Pr \left( y^{(j)} = 1 \right) \right) = \alpha_j + \beta^\top \mathbf{x}, j = 1, \dots, J - 1. \quad (8)$$

Thus, the Wolfe-Gould statistic is defined as

$$l = 2[l(\hat{\theta}; \mathbf{x}, \mathbf{y}) - l(\tilde{\theta}; \mathbf{x}, \mathbf{y})],$$

where  $\tilde{\theta}$  and  $\hat{\theta}$  are the estimates for the independent binary models, with or without the proportional odds assumption, respectively. Note the parameters are estimated under the independent binaries, but the likelihoods are computed under the multinomial distribution; there is no guarantee that  $l(\hat{\theta}; \mathbf{x}, \mathbf{y})$  will be larger than  $l(\tilde{\theta}; \mathbf{x}, \mathbf{y})$ , and it may happen that the statistic may be negative. When the dichotomised binary outcomes have a data separation issue, the MLE of the corresponding logistic model does not exist, but the Wolfe-Gould statistic may still be calculated.

### Statistical software

Most statistical software packages can be used to fit proportional odds models. However, the availability of the tests for the proportional odds assumption varies.

In Stata, all the tests discussed in the last section are available in a user-developed module called ‘oparallel’.<sup>4</sup> After estimating a proportional odds model in Stata, one may simply call the command oparallel to request all the tests. One may also specify specific desired tests.

In SAS, the most popular procedure for cumulative logistic models is PROC logistic.<sup>5</sup> When a proportional odds model is fitted with the procedure, the score test for proportionality of the OR is automatically reported. It appears that no other statistics are directly available in the procedure. However, if the cumulative model without the proportional odds assumption can be estimated, then it is straightforward to compute the LR statistic based on the

definition. Using linear contrast statements, one can also obtain the Wald statistics. However, it appears the Brant and Wolfe-Gould tests are not yet available.

In SPSS, the LR test can be requested when a proportional odds model is fitted.<sup>6</sup> However, it seems that all other tests are not yet available.

Several packages in R can fit proportional odds models.<sup>7</sup> For example, packages ‘vglm’, ‘clm’ and ‘polr’ can be used to fit proportional odds models. However, it appears that tests for proportional odds assumption are not directly available, although the Wald and LR tests can be easily performed if the corresponding unequal coefficient cumulative model can be fitted. In addition, there is a package named ‘Brant’, which can provide the Brant statistic.

### SIMULATION STUDIES

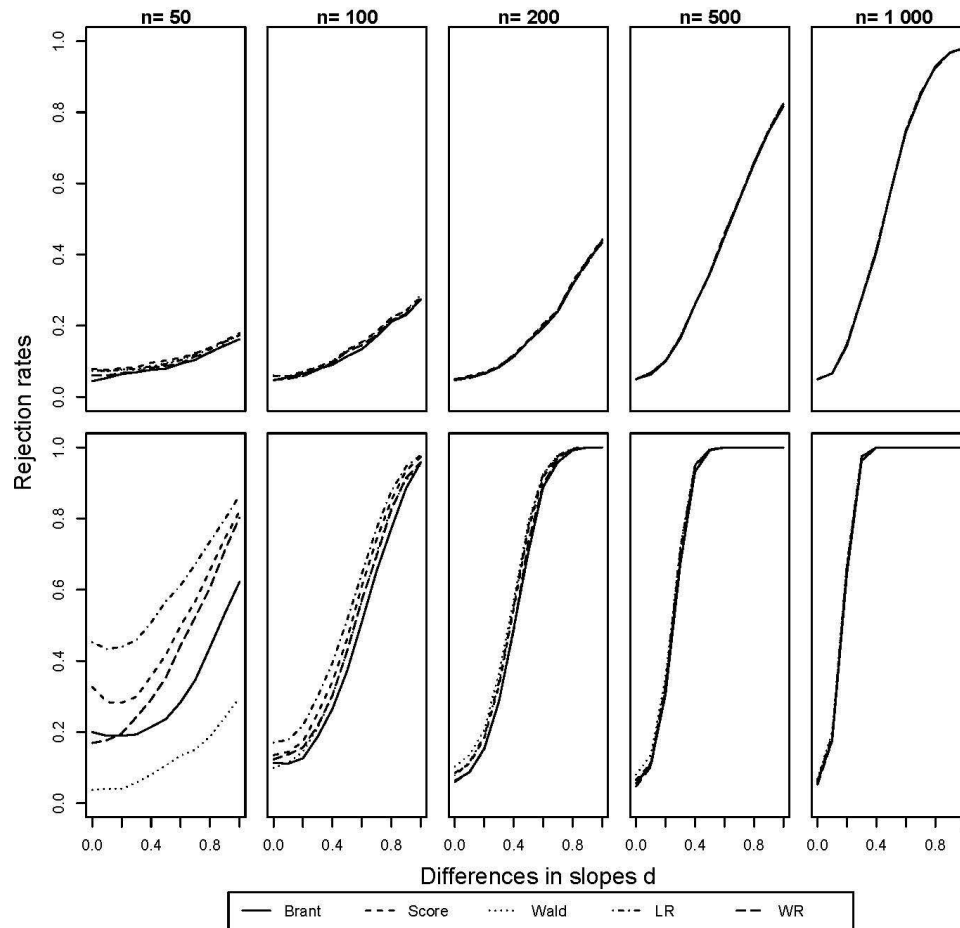
The performance of the different tests under various scenarios are assessed using Stata. We simulate data from cumulative logistic models to assess the performance in terms of type I errors and power. Different numbers of covariates (1–4), sample sizes (50, 100, 200, 500 and 1 000) and response levels (3–5) are considered. All simulations are performed with 1 000 Monte Carlo replicates.

The simulation is carried out using Stata.<sup>8</sup> Note that the original oparallel function will first check if model (1) is estimable. If the MLE for model (1) cannot be estimated, then the score, Wald and LR tests will all not be computed. Since the score test only requires the MLE under the null hypothesis, it may still be computable in such cases. Therefore, we removed the model check in the simulation study so that all the tests can be calculated whenever possible. For Brant and Wolfe-Gould tests, the original oparallel function also checks if any of the resulting binary responses may suffer from data separation. In such cases, the Brant test is not computable. However, since the likelihood may still be computed in perfectly predictable cases, the Wolfe-Gould statistic may still be computable. We also removed the check of data separation for the independent binary models from the original oparallel function.

We first simulate the covariates and then simulate the categorical responses from cumulative logistic models. To save space, we only report two scenarios: (1) there is only one covariate and the response has three levels; (2) there are four covariates and the response has four levels. For scenario 1, the following models are used:

$$x \sim \text{Unif}(0, 1), \log \left( \frac{\gamma_j(\mathbf{x})}{1 - \gamma_j(\mathbf{x})} \right) = \alpha_j + \beta_j x, \quad j = 1, 2, \quad (10)$$

where  $\alpha_1 = -0.5$ ,  $\alpha_2 = 0.5$ ,  $\beta_2 = 1$ , and  $\beta_1 = 1 - d$ , where  $d$  takes values from 0 to 1, in increments of 0.1. Thus, when  $d = 0$ , the proportional odds assumption is satisfied. As  $d$  increases, the model deviates from the proportional odds model. The probabilities for the three levels change from (0.500, 0.227, 0.272) when  $d = 0$  to (0.378, 0.350, 0.272) when  $d = 1$ .



**Figure 1** Rejection rates under nominal type I error 5% (0.05). LR, likelihood ratio; WR, Wolfe-Gould ratio.

For scenario 2, we simulate four independent variables, two continuous variables and two discrete variables:  $x_1 \sim N(0, 1)$ ,  $x_2 \sim Unif(0, 1)$ ,  $x_3 \sim Bernoulli(0.5)$ , and  $x_4 \sim Bernoulli(0.5)$ . The following cumulative models are used to simulate the response:

$$\log\left(\frac{\gamma_j(\mathbf{x})}{1-\gamma_j(\mathbf{x})}\right) = \alpha_j + \beta_{1j}x_1 + \beta_{2j}x_2 + \beta_{3j}x_3 + \beta_{4j}x_4, \quad j = 1, 2, 3, \quad (11)$$

where  $\alpha_1 = -1$ ,  $\alpha_2 = 0$ ,  $\alpha_3 = 1$ ,  $\beta_2 = (\beta_{12}, \beta_{22}, \beta_{32}, \beta_{42}) = (1, 0, -0.5, 0)$ ,  $\beta_1 = \beta_2 - (0, d, d, -d)$ , and  $\beta_3 = \beta_2 + (0, d, d, -d)$ . Again, when  $d = 0$ , the proportional odds assumption is satisfied. As  $d$  increases, the model deviates from the proportional odds model. Note that the first covariate,  $x_1$ , still satisfies the parallel coefficient assumption, but the other three do not. The probabilities for the four levels change from (0.261, 0.188, 0.202, 0.350) when  $d = 0$  to (0.235, 0.214, 0.244, 0.307) when  $d = 1$ .

Shown in [figure 1](#) are proportions of samples where the proportional odds assumption is rejected at the nominal type I error 5%; in other words, they are proportions of samples with p values less than 5% for each of the tests and sample sizes. When  $d = 0$ , the proportional odds assumption is satisfied, so these rejection rates should be less than 5%. When  $d > 0$ , the proportional odds assumption is violated, and these rejection rates are the power of the tests; the higher the rejection rates, the better.

The results in scenario 1 are shown in the plots in the first row. When the proportional odds assumption was satisfied ( $d = 0$ ), all the tests controlled the type I error very well, except when the sample size was 50, where the score and LR tests had rejection rates that were much higher than the nominal type I error 5%. When the models deviated from the proportional odds models, the power in general increased with the degree of deviations for all the tests. Also, the power in general increased with sample sizes for all the tests. The performance of the tests in terms of power was comparable except for in the case of sample sizes.

The results in scenario 2 are shown in the plots in the second row. In this scenario, we can see similar patterns; generally, the Brant and Wolfe-Gould tests controlled the type I better when sample sizes were small, and the power for all the tests were comparable when the sample size was large. Note that there were more covariates and more levels in the response; thus, larger sample sizes were required for good control of type I error. All the tests have elevated rejection rates when sample sizes were 50 and 100. Even for sample sizes of 200, the rejection rates were still much larger than the nominal level for the score, Wald and LR tests. Since the Brant and Wolfe-Gould tests control the type I better, they are recommended when the sample size is small.

**EXAMPLES**

In a study on depression among seniors, the depression outcome was diagnosed with three levels (non, minor and major depression).<sup>9</sup> As an illustrative example, we use the baseline information to study how the age, gender, marital status (ms, three levels) and medical burden (cirs, continuous) will predict the depression outcome ( $dep=0, 1$  and  $2$  for non, minor and major depressions, respectively) using the following proportional odds model:

$$dep \sim age + gender + ms + cirs.$$

Thus, the covariate vector has five dimensions (note that two dummy variables are needed for ms).

The Wolfe-Gould, Brant, score, LR and Wald statistics for the proportional odds assumption are 16.96, 15.6, 17.04, 16.70 and 17.71, respectively. Comparing with  $\chi^2$  distribution with 5 df, we obtain p values 0.005, 0.008, 0.004, 0.005 and 0.003 for the respective tests. Thus, all the tests suggest that the proportional odds assumption is unlikely to hold for the data. If the proportional odds model is applied, we may obtain biased inference on the relationship between the depression outcome and the predictors. In this case, not only are the p values but also the regression coefficients difficult to interpret because the proportional odds model does not fit the data. When the proportional odds assumption is violated, we need to assess the cause of the violation and develop an appropriate model for the data.

The Stata program, for the example:

```
ologit dep gender age cirs i.ms
oparallel
```

Using the available SAS procedure, we calculated the Wald, score, LR and Wolfe-Gould statistics. The calculation of the Brant is more complicated, and it may be preferable to develop a SAS macro for it.

The following code fit the proportional odds model, which provides the score statistic and the likelihood for the proportional model:

```
proc logist data=dos;
class ms;
model dep = age gender ms cirs;
run;
```

The following code fit the corresponding model with unequal slopes (all  $\beta_j$  can be different). The maximum likelihood under the model is reported. Thus, the LR statistic can be computed by combining with the output from the proportional odds model. The test statement computes the Wald statistic.

```
proc logist data=dos;
class ms;
model dep=age gender ms cirs/UNEQUALSLOPES;
```

```
prop: test age_0=age_1, gender_0=gender_1,cirs_0=
cirs_1, ms1_0=ms1_1,ms2_0=ms2_1; run;
```

To fit the proportional odds model in SPSS, one may choose ‘Analyse’, then ‘Regression’ and ‘Ordinal’ from the menu system. After the dependent and independent variables are selected, click on the ‘output’ button and select the item ‘Test of parallel lines’. Alternatively, one may use the following SPSS program. The ‘TPARALLEL’ option offers the LR test for the proportional odds assumption.

```
PLUM dep BY Gender ms WITH age cirsttl
/LINK=LOGIT
/PRINT=FIT PARAMETER SUMMARYTPARALLEL.
```

**DISCUSSION**

Proportional odds models should not be applied blindly. Hypothesis testing may be performed to assess the validity of the proportional odds assumption. The score test is widely used in practice as it is the only test directly available in SAS; however, its controlling of type I error is not very good. For small samples, generally, it may reject much more often than the nominal type I error level indicates. The same is true for the Wald and LR tests. On the other hand, the Brant and Wolfe-Gould tests generally control the type I error quite well. Thus, the Brant and Wolfe Gould tests are usually recommended when the sample size is small.

When the proportional odds assumption is rejected, we may assess the cause of the violation. For a really large sample, we may obtain significant results even when the proportional odds assumption is slightly violated. In such a case, it may still be practical to apply the proportional odds models. Otherwise, we may need to revise the model. For example, model (1), those models with only some components of  $\beta_j$  assumed to be the same across different  $j$ s (so-called *partial proportional odds* model<sup>10</sup>) or other models for ordinal responses may be considered. See the book by Tang *et al*<sup>1</sup> for a detailed discussion on modelling of categorical responses.

**Contributors** WT conceived the initial idea and drafted and finalised the manuscript. AL and WT searched the literature on related topics, performed the analyses and assisted in manuscript preparation. HH and XMT researched the statistical issues, directed simulation studies, and helped draft parts of the manuscript and finalise the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

**Funding** The project described was partially supported by the National Institutes of Health (NIH) (grant UL1TR001442). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Wan Tang <http://orcid.org/0000-0003-0404-1494>

#### REFERENCES

- 1 Tang W, He H, Tu XM. *Applied categorical and count data analysis*. 2nd edn. Chapman & Hall/CRC, 2023.
- 2 Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* 1990;46:1171–8.
- 3 Wolfe R, Gould W. An approximate likelihood-ratio test for ordinal response models. *Stata Technical Bulletin* 1998;7:42.
- 4 Buis M. OPARALLEL: STATA Module providing post-estimation command for testing the parallel regression assumption. Available: <https://ideas.repec.org/c/boc/bocode/s457720.html> [Accessed 12 Feb 2023].
- 5 SAS Institute. *SAS/STAT 15.1 User's Guide*. Cary, NC, USA, 2018.
- 6 SPSS Inc. *IBM SPSS Statistics Base V27*. Armonk, NY, 2020.
- 7 R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria, 2022.
- 8 StataCorp. *Stata statistical software: release 18*. College Station, TX: StataCorp LP, 2023.
- 9 Lyness JM, Kim J, Tang W, et al. The clinical significance of subsyndromal depression in older primary care patients. *Am J Geriatr Psychiatry* 2007;15:214–23.
- 10 Peterson B, Harrell FE. Partial proportional odds models for ordinal response variables. *Applied Statistics* 1990;39:205.



*Anqi Liu is a master student at the Department of Biostatistics and Data Science, School of Public Health and Tropical Medicine, Tulane University in USA. She is currently working as a Graduate Student Researcher in Tulane. She got her B.Sc. degree in Statistics from the University of California, USA in 2021. Her main research interests include clinical trials and statistical methods for population science research.*