JAMA Guide to Statistics and Methods

Regression Models for Ordinal Outcomes

Benjamin French, PhD; Matthew S. Shotwell, PhD

In the December 1, 2020, issue of *JAMA*, Self et al[1] reported a randomized clinical trial that evaluated whether treatment with hydroxychloroquine improved clinical outcomes of adults hospitalized with COVID-19 compared with placebo. The primary outcome was the patient's clinical status 14 days after randomization, assessed with an ordinal 7-category scale ranging from worst ("death") to the best ("discharged from the hospital and able to perform normal activities"). The term "ordinal" is applied to an outcome measure for which its mutually exclusive categories can be ordered by their clinical preference. The primary outcome was analyzed with a multivariable ordinal logistic regression model, which is a regression model for an ordinal dependent variable. The authors found that there was not a statistically significant difference between the hydroxychloroquine and placebo groups in clinical status 14 days after randomization.

Use of Regression Models for Ordinal Outcomes

Why Are Ordinal Outcomes Used?

Ordinal outcomes are those for which their categories (in this case, specific clinical states at 14 days) can be naturally rank ordered, but the degree of difference between categories may not be quantifiable.[2] For example, the event "discharged from the hospital and able to perform normal activities" is better than "death," but there is no numerical quantity to measure how much better. Analyses of ordinal outcomes seek to maximize use of available information by exploiting the inherent rank ordering of their categories.

Important information can be lost if the rank ordering of the outcome categories is ignored. An arbitrary binary dichotomization of an ordinal outcome (ie, collapsing the ordered categories into 2 levels) sacrifices granularity and reduces statistical power.[3] In addition, a composite binary outcome—an indicator of whether any event occurs—treats each component event equally and ignores their relative severity. Although binary outcomes are simple to analyze using standard logistic regression models, this loss of information is inconsistent with the principles of ethical conduct of research, which require that investigators maximize the possible benefits of research, including the most efficient use of study data. At the same time, an ordinal outcome should not be analyzed as a continuous variable (eg, using a linear regression model) because an ordinal outcome does not take on numeric values and likely does not satisfy the assumptions required for such models. However, in some situations, quantitative outcomes can be treated as ordinal outcomes for statistical analysis purposes.[4]

Description of Regression Models for Ordinal Outcomes

Ordinal logistic regression models are tailored for the analysis of ordinal outcomes.[5] A nonparametric Mann-Whitney test (also known as the *Wilcoxon rank sum test*) could evaluate the null hypothesis that the distribution of the ordinal outcome is similar across groups, but such a test does not allow adjustment for baseline characteristics. Similar to the Mann-Whitney test, the ordinal logistic regression model makes few assumptions about the distribution of the ordinal outcome categories so the model can accommodate outcomes that exhibit a skewed distribution. A commonly used ordinal regression model is the *cumulative logit model*, which consists of a set of logistic regression models for each possible binary dichotomization of the ordinal outcome. For example, for the 7-category scale used by Self et al,[1] there were 6 possible dichotomizations, and therefore 6 logistic regression models. For each of these models, the association of an independent variable (eg, treatment with hydroxychloroquine vs placebo) is quantified by an odds ratio. Across these logistic regression models, the odds ratios are often assumed to be equal. The same assumption is made for the associations of any other independent variables (baseline clinical characteristics such as age and sex). Thus, the association of each independent variable with the ordinal outcome is represented by a single odds ratio. Odds ratios greater than 1 indicate more frequent outcomes in the higher-ranked categories, and vice versa.

The assumption that the association of each independent variable with the outcome of interest is represented by a single odds ratio is known as the *proportional odds assumption*. Although statistical hypothesis tests are available to evaluate the proportional odds assumption, these tests can falsely reject the null hypothesis that the assumption is satisfied, leading to an incorrect conclusion that the analysis is invalid. Instead, graphical summaries can be used to qualitatively evaluate the proportional odds assumption.[6]

Other types of ordinal regression models might be preferred in some settings. A partial proportional odds model assumes proportional odds for a subset of the independent variables, which might be preferred in studies focused on a particular intervention (eg, treatment with hydroxychloroquine vs placebo) for which a single odds ratio is of scientific interest, while not requiring proportional odds for other independent variables (eg, baseline clinical characteristics) or for situations in which the proportional odds assumption is violated for 1 or more independent variables.[7] A continuation ratio model would be preferred when the outcome categories represent successive stages of disease progression.

As with any multivariable analysis, the results of an ordinal logistic regression model should be presented along with appropriate summary statistics and graphical summaries. For an ordinal outcome, a stacked bar plot with a sequential color scale, along with percentages for each category, can illustrate the distribution of the outcome for each group (eg, Figure 2 in the study by Self et al[1]).

Assuming proportional odds, the ordinal logistic regression model provides an adjusted odds ratio and CI comparing the ordinal outcome between groups. With outcome categories ordered from worst to best, an odds ratio greater than 1 indicates better outcomes. The odds ratio represents the proportional association of the independent variable with the odds of better outcomes; this interpretation applies for each dichotomization of the ordinal outcome. A hypothesis test in which the odds ratio is equal to 1 can evaluate whether the association is statistically significant.

What Are the Limitations of Ordinal Logistic Regression Models?

Limitations of ordinal logistic regression models include those inherent to ordinal outcomes (eg, the model does not quantify differences between outcome categories, only their ordering) and the assumptions required to formulate the model (eg, the proportional odds assumption), although statistical hypothesis tests regarding the association of independent variables with the ordinal outcome can be valid even if the proportional odds assumption is violated. In addition, ordinal logistic regression models provide a relative, rather than absolute, contrast in the outcome between groups. Therefore, similar to binary logistic regression models, it is more challenging to interpret the results with respect to absolute risk[8] and to translate the results into a meaningful quantity, such as the number needed to treat.[9] However, the statistical analysis of clinical data should be used to first establish whether a treatment effect exists; subsequent work may then be needed to establish effective strategies for communicating risk differences to clinicians and patients.

How Were Ordinal Logistic Regression Models Used in the Study by Self et al?

In the trial by Self et al,[1] the primary outcome was clinical status 14 days after randomization, assessed with an ordinal scale consisting of 7 mutually exclusive categories: death; hospitalized, receiving extracorporeal membrane oxygenation (ECMO) or invasive mechanical ventilation; hospitalized, receiving noninvasive mechanical ventilation or nasal high-flow oxygen therapy; hospitalized, receiving supplemental oxygen without positive pressure or high flow; hospitalized, not receiving supplemental oxygen; not hospitalized and unable to perform normal activities; and not hospitalized and able to perform normal activities. This ordinal scale captures a full spectrum of clinical outcomes that can be experienced by patients diagnosed with COVID-19. The authors used a multivariable ordinal logistic regression model, assuming proportional odds, to compare the ordinal clinical status outcome between the hydroxychloroquine and placebo groups, adjusted for prespecified baseline characteristics.

How Should the Ordinal Regression Models Be Interpreted in the Study by Self et al?

Self et al[1] found no significant difference in the ordinal clinical status outcome between the hydroxychloroquine and placebo groups (adjusted odds ratio, 1.02 [95% CI, 0.73-1.42]).

The authors conducted a variety of sensitivity analyses (eg, restricting the analysis to confirmed cases of SARS-CoV-2 infection or to patients who received at least 1 dose of the study drug) and a post hoc analysis in which the model additionally included site-specific random intercepts to account for the correlation due to clustering of patients within enrolling sites, with consistent results.[10] As a whole, these results strongly support the authors' conclusion that hydroxychloroquine did not improve patients' clinical status at day 14.

ARTICLE INFORMATION

Author Affiliations: Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee (French, Shotwell); Statistical Editor, *JAMA Pediatrics* (French).

Corresponding Author: Benjamin French, PhD, Department of Biostatistics, Vanderbilt University Medical Center, 2525 West End Ave, Ste 1100, Nashville, TN 37203 (b.french@vumc.org).

Section Editors: Robert Golub, MD, Executive Deputy Editor; and Roger J. Lewis, MD, PhD, Department of Emergency Medicine, Harbor-UCLA Medical Center and David Geffen School of Medicine at UCLA.

Published Online: August 4, 2022. doi:10.1001/jama.2022.12104

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Self WH, Semler MW, Leither LM, et al; National Heart, Lung, and Blood Institute PETAL Clinical Trials Network. Effect of hydroxychloroquine on clinical status at 14 days in hospitalized patients with COVID-19: a randomized clinical trial. *JAMA*. 2020;324(21):2165-2176. doi:10.1001/jama.2020.22240

2. Stevens SS. On the theory of scales of measurement. *Science*. 1946;103(2684):677-680. doi:10.1126/science.103.2684.677

3. Roozenbeek B, Lingsma HF, Perel P, et al; 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. *Crit Care*. 2011;15(3):R127. doi:10.1186/cc10240

4. Liu Q, Shepherd BE, Li C, Harrell FE Jr. Modeling continuous response variables using ordinal regression. *Stat Med*. 2017;36(27):4316-4335. doi:10.1002/sim.7433

5. McCullagh P. Regression models for ordinal data. *J R Stat Soc Series B Stat Methodol*. 1980;42(2):109-142.

6. Harrel F. Assessing the proportional odds assumption and its impact. *Statistical Thinking* blog. March 9, 2022. Accessed April 29, 2022. https://www.fharrell.com/post/impactpo

7. Schildcrout JS, Harrell FE Jr, Heagerty PJ, et al. Model-assisted analyses of longitudinal, ordinal outcomes with absorbing states. *Stat Med*. 2022;41(14):2497-2512. doi:10.1002/sim.9366

8. Norton EC, Dowd BE, Maciejewski ML. Marginal effects—quantifying the effect of changes in risk factors in logistic regression models. *JAMA*. 2019;321(13):1304-1305. doi:10.1001/jama.2019.1954

9. Saver JL, Lewis RJ. Number needed to treat: conveying the likelihood of a therapeutic effect. *JAMA*. 2019;321(8):798-799. doi:10.1001/jama.2018.21971

10. Detry MA, Ma Y. Analyzing repeated measurements using mixed models. *JAMA*. 2016;315(4):407-408. doi:10.1001/jama.2015.19394