

COMMENTARY

Requiem for odds ratios

Edward C. Norton PhD^{1,2}  | Bryan E. Dowd PhD³  | Melissa M. Garrido PhD^{4,5}  |
Matthew L. Maciejewski PhD^{6,7,8} 

¹Department of Health Management and Policy, University of Michigan School of Public Health, Ann Arbor, Michigan, USA

²Department of Economics, University of Michigan, Ann Arbor, Michigan, USA

³Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA

⁴Department of Health Law, Policy and Management, Boston University School of Public Health, Boston, Massachusetts, USA

⁵Partnered Evidence-Based Policy Resource Center (PEPREc), Boston VA Healthcare System, Boston, Massachusetts, USA

⁶Center of Innovation to Accelerate Discovery & Practice Transformation, Durham VA Medical Center, Durham, North Carolina, USA

⁷Department of Population Health Sciences, Duke University, Durham, North Carolina, USA

⁸Division of General Internal Medicine, Department of Medicine, Duke University, Durham, North Carolina, USA

Correspondence

Edward C. Norton, Department of Health Management and Policy, University of Michigan, 1415 Washington Heights, SPH II, Ann Arbor, MI 48109, USA.
Email: ecnorton@umich.edu

Funding information

Department of Veterans Affairs, Grant/Award Number: RCS 10-391

1 | INTRODUCTION

Health Services Research encourages authors to report marginal effects instead of odds ratios for logistic regression with a binary outcome. Specifically, in the instructions for authors, Manuscript Formatting and Submission Requirements, section 2.4.2.2 Structured abstract and keywords, it reads “Reporting of odds ratios is discouraged (marginal effects preferred) except in case-control studies” (see the *HSR* website <https://www.hsr.org/authors/manuscript-formatting-submission-requirements>).

We applaud this decision. We also encourage other journals to make the same decision. It is time to end the reporting of odds ratios in the scientific literature for most research studies, except for case-control studies with matched samples.

HSR's decision is due to increasing recognition that odds ratios are not only confusing to non-researchers,^{1,2} but that researchers themselves often misinterpret them.^{3,4} Odds ratios are also of limited utility in meta-analyses. Marginal effects, which represent the difference in the probability of a binary outcome between comparison groups, are more straightforward to interpret and compare. Below, we illustrate the difficulties in interpreting odds ratios, outline the conditions that must be met for odds ratios to be compared directly, and explain how marginal effects overcome these difficulties.

2 | ODDS RATIOS ARE DIFFICULT TO INTERPRET

Consider a hypothetical prospective cohort study of whether a new hospital-based discharge program affects the 30-day readmission rate, a binary outcome, observed for each patient who is discharged alive. The program's goal is to help eligible patients avoid unnecessary readmissions, and patients are randomized into participating in the program or not. Suppose that a carefully designed study estimates the logistic regression coefficient (the log odds) on the discharge program to be -0.2 , indicating that readmission rates are lower for patients who participate in the discharge program than patients who do not. When writing about the results, the researcher must decide how to report the magnitude of the change and has several choices for how to do so.

One option is to report the odds ratio, which in this case is $0.82 = \exp(-0.2)$, and then compare it with other published odds ratios in the literature. However, this estimated odds ratio of 0.82 depends on an unobservable scaling factor that makes its interpretation conditional on the data and on the model specification.^{3,5} As odds ratios are scaled by different unobservable factors and are conditional on different model specifications, the estimated odds ratio cannot be

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Health Services Research* published by Wiley Periodicals LLC on behalf of Health Research and Educational Trust.

compared with any other odds ratio.^{6,7} Even within a single study, odds ratios based on models including different sets of covariates cannot be compared. It would be more accurate to report that, “The estimated odds ratio is 0.82, conditional on the covariates included in the regression, but a different odds ratio would be found if the model included a different set of explanatory variables.” Due to an unobserved scaling factor that is included in every estimated odds ratio, odds ratios are not generalizable.

3 | WHEN CAN ODDS RATIOS BE COMPARED DIRECTLY?

Odds ratios from different covariate specifications within the same study or between different studies can almost never be compared directly. The explanation for this requires an understanding of how logistic regression differs from linear regression.³ In least squares regression, adding covariates that predict the outcome—but are independent of other covariates (and are therefore not mediators or confounders)—does not change either the estimated parameters or the marginal effects. Adding more independent covariates to a linear regression just reduces the amount of unexplained variation, which reduces the error variance (σ^2), and results in smaller standard errors for each parameter or marginal effect because of improved precision. For example, in a perfectly executed randomized controlled trial (RCT), the assignment to treatment is independent of all covariates, and the covariates are balanced in the treatment and comparison groups. In a perfectly executed RCT, the estimated treatment effect from a least squares regression should be the same whether covariates are included or not. The only difference in the estimated treatment effect with or without covariate adjustment is the standard errors. Including covariates corrects for any imbalance in the covariates resulting from sampling variation. Adding covariates thus improves statistical significance while leaving the expected value of the estimated treatment effects unchanged.

This result does not carry over to logistic regression (or to probit regression). In contrast to linear regression applied to the RCT, adding covariates will change the estimated coefficients in a logistic regression of a binary outcome from the same RCT, even when those added covariates are not confounders.³⁻⁷ Therefore, the estimated odds ratios also change unlike the linear regression where the estimated coefficients do not change. The reason that the odds ratios change is because the estimated coefficients in a logistic regression are scaled by an arbitrary factor equal to the square root of the variance of the unexplained part of binary outcome, or σ . That is, logistic regressions estimate β/σ , not β (for the full mathematical derivation, see Norton and Dowd³). Furthermore and more problematic, σ is unknown to the researcher.

Because the estimated coefficients in a logistic regression are scaled by an arbitrary factor σ , the odds ratios are also scaled by an arbitrary factor (odds ratio = $\exp(\beta/\sigma)$). Ideally, this arbitrary scaling factor σ would be invariant to changes in covariate specification, but it is not. In fact, this scaling factor changes when more explanatory variables are added to the logistic regression model, because the

added variables explain more of the total variation and reduce the unexplained variance and reduce σ . Therefore, adding more independent explanatory variables to the model will increase the odds ratio of the variable of interest (e.g., treatment) due to dividing by a smaller scaling factor (σ), which does not occur when representing the strength of association via relative risks or absolute risks.

In the same perfectly executed RCT, including additional covariates to a logistic regression on a binary outcome would change the magnitude of the estimated treatment effect (log odds, β/σ) and the corresponding odds ratio ($\exp(\beta/\sigma)$). As a result, the interpretation of the odds ratio depends on the covariates included in the model. A comparison of ORs from prior literature is not meaningful if either the covariate specification is different or if the sample is different because the unknown σ is different for each study.

4 | MARGINAL EFFECTS CAN BE COMPARED ACROSS DIFFERENT MODEL SPECIFICATIONS AND STUDIES

In the readmission example above, a clearer option would be to report marginal effects in terms of a percentage point change in the probability of readmission, along with the base readmission rate for context.⁸

In health services research, the most common way of reporting marginal effects is through average marginal effects—the average of the marginal effects computed for each person. These are interpreted as the mean percentage point difference—not the percent difference—in outcome probabilities that accompany a change in the treatment variable’s value. For binary treatments, an alternative is to present the predicted probabilities of the outcome when the treatment variable equals 0 and 1.

Marginal effects are much less sensitive to the unknown scaling factor and exhibit little change when independent covariates are added to the logistic regression model. When averaged, many of these small changes cancel out.³ The magnitude of average marginal effects can be compared across different studies, whereas the magnitude of odds ratios cannot. For this reason, marginal effects are preferable to report from logistic regression from RCTs and nonrandomized studies.

5 | CONSEQUENCES OF USING ODDS RATIOS VERSUS MARGINAL EFFECTS

By extension from odds ratios not being comparable across studies due to unknown scaling factors being different, they have limited utility in systematic reviews and meta-analyses. Marginal effects overcome these difficulties.

Similarly, marginal effects are preferable to odds ratios or coefficients when using logistic regression to generate predictive models that will be applied to other populations. The magnitude of the unknown scaling factor in odds ratios or log odds will differ across populations, limiting the generalizability of a predictive model to a population other than the one in which it is tested and trained.

6 | CONCLUSION

The choice of how to report results from a logistic regression is important because logistic regression is one of the most common statistical tools in the health services research toolkit. It is also important that researchers—especially researchers who study public policies and quality of care—communicate their results and conclusions clearly to other researchers, policymakers, and the public. Therefore, HSR's stand on odds ratios will help improve interpretation and communication.

We commend *Health Services Research* for deciding to discourage the reporting of odds ratios in most studies. We agree wholeheartedly with this decision, which keeps *Health Services Research* at the forefront of best practices.

FUNDING INFORMATION

Dr. Maciejewski was also supported by a Research Career Scientist award from the Department of Veterans Affairs (RCS 10-391).

ORCID

Edward C. Norton  <https://orcid.org/0000-0003-4555-0631>

Bryan E. Dowd  <https://orcid.org/0000-0002-3384-5144>

Melissa M. Garrido  <https://orcid.org/0000-0002-8986-3536>

Matthew L. Maciejewski  <https://orcid.org/0000-0003-1765-5938>

REFERENCES

1. Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! *Evidence Based Med.* 1996;1(6):164-166.
2. Tajeu GS, Sen B, Allison DB, Menachemi N. Misuse of odds ratios in obesity literature: an empirical analysis of published studies. *Obesity.* 2012;20(8):1726-1731.
3. Norton EC, Dowd BE. Log odds and the interpretation of logit models. *Health Serv Res.* 2018;53(2):859-878.
4. Norton EC, Dowd BE, Maciejewski ML. Odds ratios—current best practice and use. *JAMA.* 2018;320(1):84-85.
5. Mroz TA, Zayats YV. Arbitrarily normalized coefficients, Information sets, and false reports of “biases” in binary outcome models. *Rev Econ Stat.* 2008;90(3):406-413.
6. Allison PD. Comparing logit and probit coefficients across groups. *Soc Methods Res.* 1999;28(2):186-208.
7. Mood C. Logistic regression: why we cannot do what we think we can do, and what we can do about it. *Europ Soc Rev.* 2010;26(1):67-82.
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.

How to cite this article: Norton EC, Dowd BE, Garrido MM, Maciejewski ML. Requiem for odds ratios. *Health Serv Res.* 2024;59(4):e14337. doi:[10.1111/1475-6773.14337](https://doi.org/10.1111/1475-6773.14337)