At Odds: Concerns Raised by Using Odds Ratios for Continuous or Common Dichotomous Outcomes in Research on Physical Activity and Obesity

Gina S. Lovasi*, Lindsay J. Underhill, Darby Jack, Catherine Richards, Christopher Weiss and Andrew Rundle

Columbia University, Mailman School of Public Health, Department of Epidemiology, 722 W 168th St, Room 804, New York, NY 10032, USA

Abstract: Purpose: Research on obesity and the built environment has often featured logistic regression and the corresponding parameter, the odds ratio. Use of odds ratios for common outcomes such obesity may unnecessarily hinder the validity, interpretation, and communication of research findings. Methods: We identified three key issues raised by the use of odds ratios, illustrating them with data on walkability and body mass index from a study of 13,102 New York City residents. Results: First, dichotomization of continuous measures such as body mass index discards theoretically relevant information, reduces statistical power, and amplifies measurement error. Second, odds ratios are systematically higher (further from the null) than prevalence ratios; this inflation is trivial for rare outcomes, but substantial for common outcomes like obesity. Third, odds ratios can lead to incorrect conclusions during tests of interactions. The odds ratio in a particular subgroup might higher simply because the outcome is more common (and the odds ratio inflated) compared with other subgroups. Conclusion: Our recommendations are to take full advantage of continuous outcome data when feasible and to use prevalence ratios in place of odds ratios for common dichotomous outcomes. When odds ratios must be used, authors should document outcome prevalence across exposure groups.

Keywords: Physical activity, obesity, logistic regression, odds ratio, prevalence ratio, dichotomization.

INTRODUCTION

Physical activity, obesity, logistic regression, odds ratio, prevalence ratio, dichotomization.

DICHOTOMIZATION OF CONTINUOUS MEASURES DISCARDS INFORMATION

The use of logistic regression frequently involves dichotomizing continuous measures such as physical activity or body mass index (BMI) (Fig. 1). Although dichotomization may be the best choice for some questions and audiences [10], we have identified three key problems with dichotomizing continuous outcomes.

*Address correspondence to this author at the Columbia University, Mailman School of Public Health, Department of Epidemiology, 722 W 168th St, Room 804, New York, NY 10032, USA; Tel: 212-342-0446; Fax: 212-842-5169; E-mail: glovasi@columbia.edu

Fig. (1). Odds ratios commonly reported in active living research. Notes: Box plots indicate the interquartile range (the rectangle is bounded by the 75th and 25th percentiles) and the interdecile range (lines extend to the 90th and 10th percentiles); variation in the distribution is shown for the entire Active Living Research Literature Database [7], the subset for which authors had reviewed and confirmed the database entry, and for four commonly reported outcomes: meeting physical activity recommendations, inactivity, overweight, and obesity.

First, theories about how environments influence physical activity and body weight generally suggest that the relationships are continuous. Conceptual models or frameworks guide the selection and organization of measures
found that the sample size would have to be almost doubled models until the sample size was at least 1,248. Thus, we was not consistently reached for our logistic regression randomly selected participants, but statistical significance outcome of obesity (BMI density [23]. We re-analyzed subsets of the data using linear units for each 10,000 people/km² increase in residential sample size (N = 13,102) reported a decrease of 0.4 BMI and body mass index [21, 22]. This study, which had a large increase in the sample size may be required to depends on other study characteristics as well [18, 19], but a power loss information is discarded [15-17]. A questionnaire, motion improvements for an entire population.

These exposures may nonetheless be important for their proximate, or along some other categorization scheme. Since environmental association rather than referencing a threshold. Information is discarded, reducing statistical power. Unbiased measurement error will increase extreme values, altering the proportion exceeding a threshold

---

**Table 1. Problems Arising from Over–Use of Odds Ratios, Along with Recommended Strategies and their Anticipated Benefits**

<table>
<thead>
<tr>
<th>Category</th>
<th>Potential Problems</th>
<th>Recommended Strategies</th>
<th>Anticipated Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dichotomization of continuous measures</td>
<td>Theory implies continuous, subtle influence of environment</td>
<td>Use continuous measures to take full advantage of collected data</td>
<td>Flexibility to examine the form of the association is preserved</td>
</tr>
<tr>
<td></td>
<td>impact than referencing a threshold. Information is discarded, reducing statistical power</td>
<td>Consider generalized linear models and elasticities</td>
<td>Smaller sample sizes are needed to detect a statistically significant effect</td>
</tr>
<tr>
<td></td>
<td>Unbiased measurement error will increase extreme values, altering the</td>
<td>Minimize measurement error and be cautious when discussing results derived from</td>
<td></td>
</tr>
<tr>
<td></td>
<td>proportion exceeding a threshold</td>
<td>an imprecise, dichotomized measure</td>
<td></td>
</tr>
<tr>
<td>Odds ratio interpretation</td>
<td>Odds ratios are difficult to interpret and communicate</td>
<td>Use prevalence ratios in place of, or to complement, odds ratios</td>
<td>Research results can be stated clearly and accurately in terms of how much more likely the outcome was in the exposed group</td>
</tr>
<tr>
<td></td>
<td>when they approximate prevalence ratios</td>
<td>Provide outcome prevalence information to give context when odds ratios must be presented</td>
<td>Informed readers can use prevalence information to estimate prevalence ratios from published odds ratios</td>
</tr>
<tr>
<td>Interactions on an odds ratio scale</td>
<td>Interactions of scientific interest may be missed on an odds ratio scale</td>
<td>Consider using continuous outcome data to maximize statistical power to detect an</td>
<td>Interaction analyses are more likely to yield scientifically interesting results</td>
</tr>
<tr>
<td></td>
<td>Interactions on the odds ratio scale may result from a difference in outcome</td>
<td>interaction</td>
<td>Hypotheses regarding effect modification will receive support or be dismissed on the basis of valid tests</td>
</tr>
<tr>
<td></td>
<td>prevalence across subgroups, even if the prevalence ratios are the same</td>
<td>Conduct the interaction analysis that corresponds to the scientific question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>within each subgroup</td>
<td>Provide outcome prevalence data for each subgroup if odds ratios are being compared</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>across subgroups</td>
<td></td>
</tr>
</tbody>
</table>

[11-14] which range from personal to societal, from distal to proximate, or along some other categorization scheme. Since numerous influences are proposed, each presumably has only a small effect on behavior or health. An exposure that is causally related to BMI will only occasionally cause an individual to cross a threshold value such as a BMI of 30. These exposures may nonetheless be important for their potential to generate incremental behavioral and health improvements for an entire population.

In addition, dichotomization is problematic because information is discarded [15-17]. A questionnaire, motion sensor, or anthropometry protocol captures a wide range of variation. Dichotomization ignores much of this variation and leads to a decrease in statistical power. Power loss depends on other study characteristics as well [18, 19], but a large increase in the sample size may be required to compensate for the dichotomization of an outcome variable [20]. As an example, consider a study of residential density and body mass index [21, 22]. This study, which had a large sample size (N = 13,102) reported a decrease of 0.4 BMI units for each 10,000 people/km² increase in residential density [23]. We re-analyzed subsets of the data using linear models of BMI and logistic regression for the dichotomous outcome of obesity (BMI≥30); methods and adjustments were otherwise identical. We were able to detect statistical significance for all continuous models with at least 783 randomly selected participants, but statistical significance was not consistently reached for our logistic regression models until the sample size was at least 1,248. Thus, we found that the sample size would have to be almost doubled before a continuous association of interest was detected in logistic regression. Dichotomization of continuous measures may thus contribute to Type II error [17, 24].

A final concern about dichotomization is that it exaggerates misclassification. Physical activity and adiposity are difficult to measure, and common approaches have limited reliability and validity [25-27]. Even unbiased measurement error (mean error=0) in a continuous measure will affect the proportion exceeding a threshold [28]. One can visualize this by considering the distribution of BMI from the population of 13,102 adults discussed earlier [21, 22]. These data, based on heights and weights measured by trained staff, indicate that 28.8% of the study participants were obese. If we add random, unbiased error of up to 10 BMI units in either direction, we would find that 37.8% of participants met our criteria for obesity (Fig. 2). The existence of nonrandom error such as social desirability bias may further complicate the picture by having differential effects across the BMI distribution or across other groups of interest.

Our recommendation is to use statistical approaches that take full advantage of continuous outcome data; useful strategies may include linear models, generalized linear models, zero inflated Poisson models, or proportional hazards models. In addition, studies should be designed to minimize measurement error, and should interpret with caution the proportion above a threshold in the presence of measurement error in the underlying continuous variable.
ODDS RATIOS MAY MISLEAD WHEN THE OUTCOME IS COMMON

For rare outcomes affecting <10% of the population [8, 29], the odds ratio approximates the prevalence ratio (also referred to as the probability ratio, risk ratio, or relative risk). However, for common outcomes odds ratios are systematically more extreme (further from the null) than the corresponding prevalence ratios [29-31]. For the common magnitudes of association, odds ratios are markedly different from the underlying prevalence ratios, being 50% to 400% further from the null value of 1 (Fig. 3). In our study of obesity in New York City [21], participants living in low-density neighborhoods (defined as the lowest quartile of population density) had a 30% higher odds of obesity (OR=1.3). But the corresponding prevalence ratio of 1.2 indicates that the probability of obesity was only 20% higher in the lowest density quartile compared with the other three quartiles. For stronger associations or more common outcomes, the difference would be larger.

Odds ratios and prevalence ratios contain essentially similar information, but are numerically different. If described and interpreted correctly, the difference between these approaches and the parameters of interest need not present a problem. A problem often arises, however, when investigators try to explain the magnitude of odds ratios [8, 9, 32]. The magnitude of association may become particularly important when research is used to assess attributable risk, drive cost benefit analyses, or shape policy goals. It is very tempting to interpret an odds ratio of 3 in an obesity study as meaning that obesity is three-times as likely in the exposed group. However, an odds ratio of 3 may correspond to a prevalence ratio of only 2 (Fig. 3).

As above, we caution against dichotomizing continuous measures, preferring methods that use all of the theoretically relevant data. However, we recognize that there are circumstances that may encourage or compel an investigator to use a dichotomous version of a continuous measure [10]. For example, clinical and policy audiences may prefer a message framed in terms of reducing obesity risk, rather than decreasing body mass index. Relative risk regression [29, 33, 34] can be used in place of logistic regression (Box 1). A simple formula is available for estimating prevalence ratios from published odds ratios [30], but requires that the outcome prevalence in the reference group be known. This formula offers only an approximation for adjusted models, in comparison to regression methods that directly estimate prevalence ratios with adjustment on the same scale.

THE ODD MEANING OF INTERACTIONS ON AN ODDS RATIO SCALE

The loss of statistical power due to dichotomization of a continuous outcome may undermine one’s ability to detect effect modification. More importantly, apparent interactions may appear in analyses using odds ratios that would not be evident in analyses based on prevalence ratios; conversely an interaction on the prevalence ratio scale may be obscured by using odds ratios [35-37].

A higher odds ratio in a particular subgroup might be observed simply because the outcome is more common in that group. As an example, consider fast food restaurant
proximity and obesity prevalence among each of 4 age groups. Suppose obesity prevalence varies from 10 to 40 percent among age groups. If the prevalence ratio in each age group was 2.0, odds ratios would be 2.25 in the low-prevalence group and 6.0 in the high-prevalence group (see right side of Fig. 3). This odds ratio “interaction” is difficult to explain, potentially misleading, and not well aligned with a scientific interest the pattern of association between fast food restaurants and the probability of obesity. This interaction fallacy [35] not only affects interaction analyses within a single study, but also has potential to bias meta-analyses that integrate effect estimates from multiple studies [19, 24, 31, 37].

The proposed alternative of using prevalence ratios rather than odds ratios (Table 1, Box 1) should be given strong consideration when assessing interactions. When odds ratios are used to define an interaction, outcome prevalence should be shown by subgroup.

CONCLUSIONS

For research on physical activity, obesity, or other common outcomes, odds ratios should be viewed critically because of the information lost through dichotomization of continuous measures and the mismatch between the odds ratio scale and the scientific questions of interest. Continuous outcomes should be used to take full advantage of the collected data, particularly in the context of small sample sizes or substantial measurement error. When a dichotomous outcome must be used, prevalence ratios are easier to understand and communicate. When odds ratios must be used, presentation of outcome prevalence can facilitate interpretation.

ACKNOWLEDGEMENTS

The authors acknowledge the Robert Wood Johnson Active Living Research program (grant #68507), the National Institute of Environmental Health Science (grant #5R01ES014229) and the Robert Wood Johnson Foundation’s Health & Society Scholars Program for their financial support.

CONFLICT OF INTEREST

The authors have no conflicts of interest to report.
REFERENCES


Received: November 30, 2011 Revised: March 29, 2012 Accepted: March 30, 2012

© Lovasi et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.