Misuse of Odds Ratios in Obesity Literature: An Empirical Analysis of Published Studies

Gabriel S. Tajeu1, Bisakha Sen1, David B. Allison2 and Nir Menachemi1

Odds ratios (ORs) are widely used in scientific research to demonstrate the associations between outcome variables and covariates (risk factors) of interest, and are often described in language suitable for risks or probabilities, but odds and probabilities are related, not equivalent. In situations where the outcome is not rare (e.g., obesity), ORs no longer approximate the relative risk ratio (RR) and may be misinterpreted. Our study examines the extent of misinterpretation of ORs in Obesity and International Journal of Obesity. We reviewed all 2010 issues of these journals to identify all articles that presented ORs. Included articles were then primarily reviewed for correct presentation and interpretation of ORs; and secondarily reviewed for article characteristics that may have been associated with how ORs are presented and interpreted. Of the 855 articles examined, 62 (7.3%) presented ORs. ORs were presented incorrectly in 23.2% of these articles. Clinical articles were more likely to present ORs correctly than social science or basic science articles. Studies with outcome variables that had higher relative prevalence were less likely to present ORs correctly. Overall, almost one-quarter of the studies presenting ORs in two leading journals on obesity misinterpreted them. Furthermore, even when researchers present ORs correctly, the lay media may misinterpret them as relative RRs. Therefore, we suggest that when the magnitude of associations is of interest, researchers should carefully and accurately present interpretable measures of association—including RRs and risk differences—to minimize confusion and misrepresentation of research results.

Obesity (2012) 20, 1726–1731. doi:10.1038/oby.2012.71

INTRODUCTION
Odds ratios (ORs) and risk ratios (RRs) are commonly reported measures of association in the health-care literature (1–3). Some authors have offered a preference for RRs because they may provide a more straightforward, intuitive way to interpret results than ORs (1–3). However, RR cannot be easily calculated in certain situations such as in case–control and meta-analysis studies. In these cases and/or other situations where RR is simply not calculated, the OR is often used (1–4). ORs asymptotically approximate RRs as the prevalence of an outcome of interest observed in a study approaches zero (3). However, in practical terms, when the prevalence of an outcome is >10%, ORs no longer approximate RRs and may be misinterpreted. Our study examines the extent of misinterpretation of ORs in Obesity and International Journal of Obesity. We reviewed all 2010 issues of these journals to identify all articles that presented ORs. Included articles were then primarily reviewed for correct presentation and interpretation of ORs; and secondarily reviewed for article characteristics that may have been associated with how ORs are presented and interpreted. Of the 855 articles examined, 62 (7.3%) presented ORs. ORs were presented incorrectly in 23.2% of these articles. Clinical articles were more likely to present ORs correctly than social science or basic science articles. Studies with outcome variables that had higher relative prevalence were less likely to present ORs correctly. Overall, almost one-quarter of the studies presenting ORs in two leading journals on obesity misinterpreted them. Furthermore, even when researchers present ORs correctly, the lay media may misinterpret them as relative RRs. Therefore, we suggest that when the magnitude of associations is of interest, researchers should carefully and accurately present interpretable measures of association—including RRs and risk differences—to minimize confusion and misrepresentation of research results.
Methods and Techniques

The difference between whites and blacks was in actuality only 7% (27; see Supplementary Table S1 online). A more recent instance occurred in Science Daily, where authors correctly reported the ORs, but incorrectly interpreted them as an increased percent in risk as in the following example: “results indicate that short sleep was associated with obesity, with the adjusted ORs for black Americans (1.78) and white Americans (1.43) showing that blacks had a 35 percent greater risk than whites of obesity associated with short sleep,” (28). This type of misinterpretation is problematic given that scientific obesity studies are intended to influence policy makers, clinicians, and other researchers.

Misuse of ORs has been examined in the field of obstetrics and gynecology (1) and in medical and epidemiology journals (2,29), but to our knowledge, no previous study has examined the extent of potential misuse of ORs in the obesity literature. The purpose of this article is to review the recent obesity literature for potential misuse and or misinterpretation of ORs. Specifically, we are interested in estimating the prevalence of misuse (using it in studies where the outcome of interest has an occurrence rate >10%); and the prevalence of misinterpretation of ORs (incorrectly quantifying the effect size) in obesity studies. To do so, we focus on two of the most prominent journals focusing on obesity research and publishing primarily original research (i.e., not reviews): Obesity and International Journal of Obesity.

Methods and Procedures

Every article in each of the 2010 issues of International Journal of Obesity and Obesity were searched for articles that presented ORs. Our literature review consisted of two phases. First, we identified articles that were likely to use ORs by electronically searching the titles and abstracts of articles in these journals for the terms “odds ratio,” “OR,” and “logistic,” in reference to logistic regression. Second, the full text of these articles was further assessed using a coding sheet designed to extract information that may relate to the use and misuse of ORs in these articles (see Supplementary Data online). Upon each article review, we extracted a number of variables including the journal the article was found in, whether the first author had a university appointment or not, whether the research was conducted in the United States or internationally, whether the research was basic, clinical, or social science, cellular or noncellular, animal or human, experimental (e.g., randomized-controlled trial) or non-experimental (e.g., observational).

Information on sample size, prevalence of the outcome of interest, the reported main effect of the primary OR, and the OR calculation method were also extracted from each included article. In addition, we looked at whether $P_0$, the prevalence of the outcome of interest in the unexposed group, was calculated and reported. $P_0$ is useful because along with the OR it can be used to estimate the RR of an outcome (3). In some cases where $P_0$ was not calculated, we were able to impute the value, if sufficient data were provided. We also looked at whether “risk differences” were reported. Widely used in the economics literature, risk differences estimate the absolute change observed in an outcome variable given a change in a covariate (30,31), and, like RRs, risk differences are more intuitively interpretable than are ORs (14). In order to ensure inter-rater reliability, the first 16% of articles examined were coded separately by three authors and discussed for discrepancies (<1% disagreement). Next, one author coded the remaining articles and at least one additional author reviewed a random sample of 15 articles for agreement; again little or no disagreement existed.

Ultimately, we were concerned with whether the articles we evaluated presented ORs correctly. We defined correct use of ORs as cases when authors only presented direction of relationship and statistical significance or interpreted ORs as RRs only when the outcome was rare, ≤10% (It is a general rule of thumb and mathematically validated that ORs approximate RRs in studies where outcome prevalence rates are below and up to 10%, (1–3,32,33). However, we are aware that some researchers find 10% too forgiving a cutoff and have suggested a cutoff point <10% given that in studies where ORs are below 0.2 and above 2, they begin to diverge from RRs at outcome prevalence rates below 10% (34).).

ORs were defined as incorrect when authors interpreted ORs as RRs where the outcome of interest was not rare (e.g., >10%). We looked for statements by the authors such as “increased risk,” “X% higher,” “X times as likely,” that denoted misinterpretation of their studies’ ORs as RRs (An “RR” of X should be interpreted as the risk being $(X − 1) × 100$ percent higher. ORs themselves lack any similar intuitive interpretation. We refer readers to the following citations for more information on how to correctly interpret odds ratios (3,31). In this context, we also note that there is another potential nuanced misinterpretation. An RR (or OR) of X (when X > 1) is often interpreted as “an X-fold increase in the risk (or odds);” when in fact, it is an $(X − 1)$-fold increase. We are aware of this interpretation problem but did not study it in our analysis.). All of these data were then entered into a database.

Statistical Analysis

After all articles were coded, results were entered into STATA, version 11 (StataCorp, College Station, TX) for analysis. We conducted χ² or Fisher’s exact tests as appropriate to determine whether there were differences in the association between categorical variables and correct OR usage in our sample. Fisher’s exact tests were performed because of the presence of small cell sizes. Next, logistic regression was used to determine the relationship between the dependent variable, correct use of ORs, and the following independent variables: journal, author appointment and location, article type, and prevalence of study outcome variable. The prevalence of the main outcome variable was categorized as; <20%, 20–49%, >50%, or missing. Article type was categorized as “basic science,” “clinical science,” or “social science.” “Basic science” articles were studies conducted at the cellular level (e.g., effect of tissue composition on enzymes, effects of antibodies on total cholesterol). “Clinical science” articles were intended to have prevention, treatment, or diagnosis implications (e.g., genetic makeup as a risk factor for higher BMI, predicting metabolic disease risk based on abdominal volume). “Social science” articles were concerned with behavioral factors related to obesity (e.g., association of leisure time activities with obesity, effect of exercise protocols on weight loss). Lastly, we calculated the risk differences using the “mfx” command in STATA, version 11 (StataCorp).

Results

A total of 855 articles from the International Journal of Obesity and from Obesity were examined, and 62 (7.3% of 855) articles all of whom utilized logistic regression were included in the current analysis. Relative to how many total articles are
published, both journals contributed similar proportions of included articles to our study. Table 1 displays the characteristics of these articles. Of these articles 23 (37.1%) came from International Journal of Obesity, and 39 (62.9%) came from Obesity. The majority of articles was published by university-based authors (75.8%), came from non-US researchers (67.7%), had observational study designs (95.2%), and focused on human subjects (98.4%). The most frequent article focused on a social science topic (53.2%); followed by clinical science (27.4%) and basic science (19.4%). The mean sample size of all articles was 59,992 subjects with a median of 3,310, and a range between 61 and 1.7 million. A high percentage of the articles (82.8%) reported a prevalence rate for their outcome of interest with a median prevalence of 27%.

ORs were presented correctly in 76.8% of the articles. In univariate analysis, several characteristics of articles were associated with correct presentation of ORs by authors (Table 2). Mean prevalence rates for articles reporting ORs correctly were lower than their counterparts (28.9% vs. 43.9%; \(P = 0.058\)). Moreover, articles published on clinical science topics were more likely than social science and basic science topics to report ORs correctly (94.1 vs. 72.4 vs. 60.0; \(P = 0.065\)).

In a logistic regression model that controlled for journal, article type, appointment of first author, prevalence of outcome, and location of research, several article characteristics were still significantly associated with correctly reporting ORs (Table 3). Article type was associated with authors reporting ORs correctly. Specifically, when compared with the basic science articles, clinical articles (OR = 140.7; risk difference +43%, \(P = 0.007\)) and social science articles (OR = 16.2; risk difference +38%, \(P = 0.045\)) had higher odds of reporting ORs correctly. Also, as the prevalence of the outcome variable increased, the odds of correctly reporting the ORs decreased (Table 3).

**DISCUSSION**

ORs are vulnerable to incorrect interpretation. When the prevalence of an outcome is >10%, such as in many obesity studies, correct interpretation of ORs becomes particularly difficult. Even when researchers present ORs correctly, they can be misinterpreted by the mainstream press. In this analysis, we examined recent literature from two prominent obesity journals in order to estimate the prevalence of misuse and misinterpretation of ORs.

We found that almost one in four studies that present ORs discuss their results incorrectly. Although the problems we outlined regarding the misuse and limitations of ORs have been discussed in many different fields (1–11), only one previous study is comparable to ours in that it estimated the prevalence of misuse and misinterpretation of ORs in leading journals in a particular field. Specifically, Holcomb et al. estimated an OR misuse rate of 26% (39 out of 151 articles), in Obstetrics & Gynecology and the American Journal of Obstetrics and Gynecology (1). This is similar to our own findings herein.

In our study, we also found that certain article characteristics were associated with incorrect presentation of ORs. For example, we found that an increase in prevalence rate of an outcome variable was associated with increased odds of reporting the ORs incorrectly. Notably, this was not due to one of our definitions for correct presentation of ORs (interpreted ORs as RRs only when the outcome was rare, ≤10%), as none of the articles we examined with prevalence rates ≤10% presented ORs as

| Table 1 Article characteristics included in the current study (Total N = 62) |
|-------------------------------------------------|
| Article characteristics                          | Frequency (%) |
| Journal                                         |               |
| Obesity                                        | 39 (62.9)     |
| International Journal of Obesity               | 23 (37.1)     |
| First author appointment                        |               |
| University                                     | 47 (75.8)     |
| Other                                          | 15 (24.2)     |
| Location                                       |               |
| US                                             | 20 (32.3)     |
| Foreign                                        | 42 (67.7)     |
| Article type                                    |               |
| Basic science                                   | 12 (19.4)     |
| Clinical science                                | 17 (27.4)     |
| Social science                                  | 33 (53.2)     |
| Study design                                    |               |
| RCT                                            | 3 (4.8)       |
| Observational                                  | 59 (95.2)     |
| OR reported                                    |               |
| No                                             | 7 (11.3)      |
| Yes                                            | 55 (88.7)     |
| OR significant                                  |               |
| No                                             | 1 (1.9)       |
| Yes                                            | 53 (98.1)     |
| Prevalence in unexposed (\(P_0\)) reported     |               |
| No                                             | 49 (80.3)     |
| Yes                                            | 12 (19.7)     |
| Prevalence in unexposed (\(P_0\)) imputable    |               |
| No                                             | 39 (79.6)     |
| Yes                                            | 10 (20.4)     |
| Marginal effects reported                       |               |
| No                                             | 62 (100.0)    |
| Yes                                            | 0 (0.00)      |
| OR reported correctly                           |               |
| No                                             | 13 (23.2)     |
| Yes                                            | 43 (76.8)     |
| Sample size                                    |               |
| Mean                                           | 59,992        |
| Median                                         | 3,310         |
| SD                                             | 231,642       |

OR, odds ratio; RCT, randomized–controlled trial.
RRs. This tendency to misuse ORs as study outcome prevalence rates increase is particularly worrisome given that high prevalence rates are more susceptible to over or under inflating the OR in relation to the RR. We believe that as researchers contributing to the obesity literature we should be cognizant about the correct use of these techniques.

Article type was also associated with correct use of ORs in our study. Authors of articles categorized as basic science had higher odds of incorrectly presenting ORs when compared with authors of clinical science and social science articles. This may reflect the type of training and exposure to OR methods received by authors in these various areas of study. In fact, our finding illuminates the often confusing nature of ORs. Specifically, we found that clinical studies had 140 greater odds than basic science articles to correctly present ORs. If incorrectly interpreted as a RR, this would suggest a 100+ fold increase in the correct use of ORs by authors of clinical studies.

Notably, the corresponding risk difference articulated as a risk difference was 43%: a measure that is several orders of magnitude lower. Although none of the article types were associated with 100% correct use of ORs, journal editors and authors of basic science articles could consider paying closer attention to the use and presentation of ORs in these types of articles.

Despite the advantages of mathematically converting ORs to RRs (3,14,15) there are some issues with RRs that are worth...
pointing out. First, when comparing more than two groups, the RR for each group changes if the reference group changes. Also, RRs cannot take into account the differences in prevalence rates across groups. For instance, if the prevalence of disease A in one group is 3% of the sample and in the other group it is 1.5%, and the prevalence of disease B in one group is 40% and in the other group is 20%, then the RR for the first group relative to the second for both diseases will be 2. In order to overcome these limitations, researchers in other fields, most notably economists (14), have used the risk difference or marginal effect. Risk differences are not subject to the limits of RRs and would yield a result of 0.015 in the first case and 0.20 in the second case, thus helping illustrate the fact that the second disease has a higher prevalence in the overall sample than the first. However, in spite of the advantages of risk differences, none of the articles reviewed in our study presented them. We believe that for studies likely to be of interest to many stakeholders, providing the RR (its limitations notwithstanding) or risk differences in addition to ORs is more useful to academic and lay readers than providing ORs alone.

Limitations

There are several limitations to our study. Although we examined the most recent articles of two of the preeminent obesity journals, our study only looked at one year of research from each journal. Because of this, there is no way to determine whether the misuse of ORs in these journals is getting better or worse over time. The use of only 1 year of each journal also resulted in another limitation: a small sample size. However, even with a small sample size, we were able to find statistically significant findings between some article characteristics and the correct presentation of ORs. Furthermore, our findings regarding the prevalence of incorrect use of ORs are similar to other literature on the topic. We recognize it was not possible to collect all possible article characteristic. For example, we did not collect information on “funding source,” which may have been interesting to some readers. Although we targeted Obesity and International Journal of Obesity because of their exclusive publication of articles related to obesity, our focus on obesity journals represents a limitation to the generalizability of our findings given that articles related to obesity are published in many general public health journals, generic clinical journals, and other outlets outside of specialized obesity journals.

Conclusion

The obesity epidemic has had major political, social, and economic impacts in communities. As such, it is incumbent upon those in the scientific community to ensure that results of original research are presented in a clear, accurate, and interpretable manner (35–40). We advocate the prudent use of ORs in general; and when doing so, including clear statements, which specify that associations with odds (and not risks) are being estimated, presenting risk differences, and/or converting ORs to RRs. In cases where the outcome prevalence rates are >10%, researchers should be especially cautious about the potential misinterpretation of ORs among the media and other diverse groups of stakeholders to their research.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/oby

ACKNOWLEDGMENTS

This work was supported in part by P30DK056336 National Institutes of Health (D.B.A.). The opinions expressed are those of the authors and not necessarily any organization with which we are affiliated.

DISCLOSURE

D.B.A. has, anticipates, or has had financial interests with the Frontiers Foundation; Vivus, Inc; Kraft Foods: University of Wisconsin; University of Arizona; Paul, Weiss, Wharton & Garrison LLP; and Sage Publications. All other authors declared no conflict of interest. See the online ICMJE Conflict of Interest Forms for this article.

© 2012 The Obesity Society

REFERENCES

1. Holcomb WL Jr, Chaiworapongsa T, Luke DA, Burgdorf KD. An odd measure of risk: use and misuse of the odds ratio. Obstet Gynecol 2001;98:668–688.
2. Katz KA. The (relative) risks of using odds ratios. Arch Dermatol 2006;142:761–764.
3. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690–1691.
4. Davies HT, Orombie IK, Tavakol M. When can odds ratios mislead? Br Med J 1998;316:989–991.
5. Axelson O, Fredriksson M, Ekborg K. Use of the prevalence ratio v the prevalence odds ratio in view of confounding in cross sectional studies. Occup Environ Med 1995;52:494.
6. Lee J, Chia KS. Estimation of prevalence rate ratios for cross sectional data: an example in occupational epidemiology. Br J Ind Med 1993;50:861–862.
7. McNutt LA, Hafner JP, Xue X. Correcting the odds ratio in cohort studies of common outcomes. JAMA 1989;262:529.
8. Nemes S, Jonasson JM, Genell A, Steinbeck G. Bias in odds ratios by logistic regression modelling and sample size. BMC Med Res Methodol 2009;9:56.
9. Strömberg U. Prevalence odds ratio v prevalence ratio–some further comments. Occup Environ Med 1995;52:143.
10. Hughes K. Odds ratios in cross-sectional studies. Int J Epidemiol 1995;24:463–4, 468.
11. Sinclair JC, Bracken MB. Clinically useful measures of effect in binary data. J Clin Epidemiol 1994;47:881–889.
12. Fleiss J, Levin B, Paik M. Statistical Methods for Rates and Proportions. Wiley-Interscience: Hoboken, NJ, 2003.
13. Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res 1993;2:121–145.
14. Kleinman LC, Norton EC. What's the Risk? A simple approach for estimating adjusted risk measures from nonlinear models including logistic regression. Health Serv Res 2009;44:288–302.
15. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702–706.
16. Holland PW. A note on the covariance of the Mantel-Haenszel log-odds-ratio estimator and the sample marginal rates. Biometrics 1989;45:1009–1016.
17. Zocchetti C, Consonni D, Bertazzi PA. Relationship between prevalence rate ratios and odds ratios in cross-sectional studies. Int J Epidemiol 1997;26:220–223.
18. Baier LA, Wilczynski NL, Haynes RB. Tackling the growth of the obesity literature: obesity evidence spreads across many journals. Int J Obes (Lond) 2010;34:1526–1530.
19. Crawford AG, Cote C, Couto J et al. Prevalence of obesity, type II diabetes mellitus, hyperlipidemia, and hypertension in the United States: findings from the GE Centricity Electronic Medical Record database. Popul Health Manag 2010;13:151–161.
20. Eaton DK, Kann L, Kinchen S et al.; Centers for Disease Control and Prevention (CDC). Youth risk behavior surveillance - United States, 2007. MMWR Surveill Summ 2010;59:1–142.
21. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA 2010;303:235–241.
22. Ogden CL, Carroll MD, Curtin LR, Lamb MM, Flegal KM. Prevalence of high body mass index in US children and adolescents, 2007-2008. JAMA 2010;303:242–249.
23. Ogden CL, Carroll MD, Curtin LR et al. Prevalence of overweight and obesity in the United States, 1999-2004. JAMA 2006;295:1549–1555.
24. Yanovski SZ, Yanovski JA. Obesity prevalence in the United States—up, down, or sideways? N Engl J Med 2011;364:987–989.
25. Verghese A. Showing doctors their biases. NY Times (Print) 1999:A21.
26. Schulman KA, Berlin JA, Harless W et al. The effect of race and sex on physicians’ recommendations for cardiac catheterization. N Engl J Med 1999;340:618–626.
27. Schwartz LM, Woloshin S, Welch HG. Misunderstandings about the effects of race and sex on physicians’ referrals for cardiac catheterization. N Engl J Med 1999;341:279–283; discussion 286–287.
28. Science News. Race and short sleep duration increase the risk for obesity. Science Daily. <http://www.sciencedaily.com/releases/2009/06/090608071812.htm> (2009).
29. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157:940–943.
30. Associates, F.a. Farin Client Care. (Accessed 9/6/11) <http://clientcare.farin.com/ Farin_Foresight/terms.asp> (2011).
31. Menachemi N, Yeager VA, Bilello L et al. Florida doctors seeing Medicaid patients show broad interest in federal incentives for adopting electronic health records. Health Aff (Millwood) 2011;30:1461–1470.
32. Knol MJ, Duinhoven RG, Grobbee DE, Moons KG, Groenwold RH. Potential misinterpretation of treatment effects due to use of odds ratios and logistic regression in randomized controlled trials. PLoS ONE 2011;6:e21248.
33. Viera AJ. Odds ratios and risk ratios: what’s the difference and why does it matter? South Med J 2008;101:730–734.
34. Shrier I, Steele R. Understanding the relationship between risks and odds ratios. Clin J Sport Med 2006;16:107–110.
35. Cofield SS, Corona RV, Allison DB. Use of causal language in observational studies of obesity and nutrition. Obes Facts 2010;3:353–356.
36. Cope MB, Allison DB. White hat bias: examples of its presence in obesity research and a call for renewed commitment to faithfulness in research reporting. Int J Obes (Lond) 2010;34:84–8; discussion 83.
37. Delgado-Nogueran M, Tort S, Bonfill X, Gich I, Alonso-Coello P. Quality assessment of clinical practice guidelines for the prevention and treatment of childhood overweight and obesity. Eur J Pediatr 2009;168:789–793.
38. Kiesges LM, Dzewaltowski DA, Glasgow RE. Review of external validity reporting in childhood obesity prevention research. Am J Prev Med 2008;34:216–223.
39. Sugerman HJ, Kral JG. Evidence-based medicine reports on obesity surgery: a critique. Int J Obes (Lond) 2005;29:735–745.
40. Thomas O, Thabane L, Douketis J et al. Industry funding and the reporting quality of large long-term weight loss trials. Int J Obes (Lond) 2008;32:1531–1536.