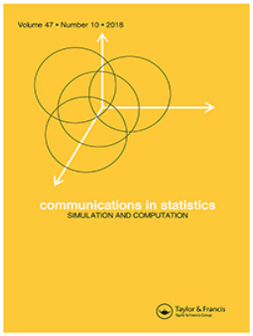

Communications in Statistics—Simulation and Computation®

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies

Henian Chen , Patricia Cohen & Sophie Chen

To cite this article: Henian Chen , Patricia Cohen & Sophie Chen (2010) How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies, Communications in Statistics—Simulation and Computation®, 39:4, 860-864, DOI: 10.1080/03610911003650383

To link to this article: https://doi.org/10.1080/03610911003650383

Published online: 06 Apr 2010.

Communications in Statistics—Simulation and Computation®, 39: 860–864, 2010

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1080/03610911003650383

Taylor & Francis
Taylor & Francis Group

How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies

HENIAN CHEN[1,2,3,4], PATRICIA COHEN[2,3,5], AND SOPHIE CHEN[6]

[1]Division of Biostatistics, New York State Psychiatric Institute, New York, NY, USA
[2]Division of Epidemiology, New York State Psychiatric Institute, New York, NY, USA
[3]Department of Psychiatry, College of Physicians and Surgeons, Columbia University
[4]Biostatistics, Office of Health Outcomes Research, Winthrop University Hospital, Stony Brook University School of Medicine, Mineola, NY, USA
[5]Department of Epidemiology, School of Public Health, Columbia University, New York, NY, USA
[6]Strategic Policy Branch, Health Canada, Ottawa, Canada

The odds ratio (OR) is probably the most widely used index of effect size in epidemiological studies. The difficulty of interpreting the OR has troubled many clinical researchers and epidemiologists for a long time. We propose a new method for interpreting the size of the OR by relating it to differences in a normal standard deviate. Our calculations indicate that $OR = 1.68$*, 3.47, and 6.71 are equivalent to Cohen's* $d = 0.2$ *(small), 0.5 (medium), and 0.8 (large), respectively, when disease rate is 1% in the nonexposed group; Cohen's* $d < 0.2$ *when* $OR < 1.5$*, and Cohen's* $d > 0.8$ *when* $OR > 5$*.*

Keywords Effect size; Epidemiological study; Logistic regression; Odds ratio; Risk factor.

Mathematics Subject Classification 62P10; 62H20.

1. Introduction

In recent years both scientists and the general public have expressed a growing concern about the difficulty of evaluating research findings on disease risk, especially

Received August 13, 2009; Accepted January 25, 2010
Address correspondence to Henian Chen, 222 Station Plaza North, Suite 510, Mineola, NY 11501, USA; E-mail: hchen@winthrop.org

regarding their implications for formulating health policy or altering health-related behavior or treatment (Koplan et al., 1999; Taubes, 1995; von Elm and Egger, 2004). The general press tends to overstate study findings about environmental or lifestyle factors associated with elevated risk (odds) for a specific disease; nonetheless, most such risks have far more subtle effects (Taubes, 1995).

Odds ratio (OR) is probably the most widely used statistic employed in risk factor research and is the predominant index of effect size used to demonstrate increased risk for disease in epidemiological studies (Bland and Altman, 2000). In fact, as indicated by Medline, publications noting "odds ratio" as a keyword more than tripled between 1995 and 2005, rising from 2105 to 7471, with a total of 49,270 such publications cited during that period (Cohen and Chen, 2009).

We reviewed *American Journal of Epidemiology* (AJE) to determine rate of use of OR in studies of risks for disease during 2005. AJE enjoys a wide circulation in epidemiology and in related fields, is frequently cited (often by the media), and has achieved recognition for the rigorous standards for publication, especially those pertaining to statistical and methodological issues. Over forty percent (44.1%) of original contributions (90 articles) used OR, with 171 ORs reported as significant effect size measures in the abstracts. We converted OR to 1/OR when OR < 1 in order to place all effects in a common frame for this discussion. Among these 171 reported ORs, ORs ranged from 1.17 to 290.00 and had a median value of 2.16. Clearly, there is widespread acceptance of OR as an indicator of risk for disease.

2. Odds Ratio (OR)

Odds ratio (OR) originally was proposed to determine whether the probability of an event (or disease) is the same or differs across two groups, generally a high-risk group and a low-risk group (Bland and Altman, 2000). The range of OR is from 0 to infinity: A value of 1 = no association with the specified risk (that is, the event or disease is equally likely in the high- and low-risk groups); as the value of OR increases or decreases away from 1, the association grows increasingly stronger. It is well known that, under certain circumstances (low population rates of "cases" $<10\%$) and with specific study designs (case-control studies), OR provides a good approximation to a risk ratio (RR; Hosmer and Lemeshow, 2000).

However, it also is recognized that OR does not give a good approximation of the RR when disease rates do not fall below 10% (Altman et al., 1998; Davies et al., 1998; Sinclair and Bracken, 1994). When based on the same data, an OR always will differ from zero more than the RR (Deeks, 1998; Sackett et al., 1996). In a case-control study when the sample prevalence $p = 0.5$, the OR equals exactly the square of RR (Kraemer, 2004). OR has little meaning in biomedical research unless it can approximate RR, and it is hard to conceive of a situation where the OR in a population would be the value of interest (Newcombe, 2006).

3. OR and Cohen's *d*

Cohen's d is the standardized mean difference between two group means, the effect size underlying power calculations for the two-sample t-test (Cohen, 1988). Cohen's $d = 0.2$, 0.5, and 0.8, often is cited as indicative of a small, medium, and large effect size, respectively.

Table 1
Cohen's d and the equivalent odds ratio (OR)

		Cohen's d								
		0.2			0.5			0.8		
P_0	Z_0	Z	P	OR	Z	P	OR	Z	P	OR
0.0100	−2.3263	−2.1263	0.0167	1.6814	−1.8263	0.0339	3.4739	−1.5263	0.0635	6.7128
0.0200	−2.0537	−1.8537	0.0319	1.6146	−1.5537	0.0601	3.1332	−1.2537	0.1050	5.7486
0.0300	−1.8808	−1.6808	0.0464	1.5733	−1.3808	0.0837	2.9535	−1.0808	0.1399	5.2592
0.0400	−1.7507	−1.5507	0.0605	1.5455	−1.2507	0.1055	2.8306	−0.9507	0.1709	4.9471
0.0500	−1.6449	−1.4449	0.0742	1.5228	−1.1449	0.1261	2.7416	−0.8449	0.1991	4.7233
0.0600	−1.5548	−1.3548	0.0877	1.5060	−1.0548	0.1458	2.6741	−0.7548	0.2252	4.5536
0.0700	−1.4758	−1.2758	0.1010	1.4926	−0.9758	0.1646	2.6177	−0.6758	0.2496	4.4191
0.0800	−1.4051	−1.2051	0.1141	1.4811	−0.9051	0.1827	2.5707	−0.6051	0.2726	4.3097
0.0900	−1.3408	−1.1408	0.1270	1.4709	−0.8408	0.2002	2.5309	−0.5408	0.2943	4.2167
0.1000	−1.2816	−1.0816	0.1397	1.4615	−0.7816	0.2172	2.4972	−0.4816	0.3150	4.1387

P_0: rate of outcome of interest in the nonexposed group.
Z_0: standard normal deviation for P_0.
Cohen's $d = Z - Z_0$ (standardized mean difference).
Z: standard normal deviation for P, $Z = Z_0 +$ Cohen's d.
P: rate of outcome of interest in the exposed group.
$\text{OR} = P(1 - P_0)/P_0(1 - P)$.

Table 1 shows the calculated ORs equivalent to Cohen's $d = 0.2$ (small), 0.5 (medium), and 0.8 (large) according to different disease rates in the nonexposed group. At a 1% disease rate in the nonexposed group, reference points reflecting a "weak association" OR, a "moderate association" OR, and a "strong association" OR are 1.68, 3.47, and 6.71, respectively. At a 5% disease rate in the nonexposed group, corresponding reference points are 1.52, 2.74, and 4.72.

We also calculated Cohen's d for a given OR according to different rates of disease in the nonexposed group. As shown in Table 2, Cohen's $d < 0.2$ when OR $<$ 1.5, indicating a small effect or weak association; but Cohen's $d > 0.8$ when OR $>$ 5, indicating a large effect or strong association (detailed calculations in Tables 1 and 2 are available from the first author upon request).

4. Discussion

As we know, a statistically significant outcome indicates only that there is some relationship between the risk factor and a disease. An indication of statistical significance, however, does not provide information about the strength of the association (effect size), although some misinterpret statistical significance to indicate effect magnitude. Meehl (1990, p. 123) pointed out that everything is more or less correlated with everything in the social sciences, a view that may appropriately generalize to some or even many areas of medical and public health research. It is quite possible with a large sample to have a statistically significant finding from a weak but true association (e.g., a small effect size) between a risk factor and a disease. Findings with lower probability (p) values (e.g., $p < 0.001$) are misinterpreted as having a stronger effect than those with higher p values (e.g., $p < 0.05$).

Table 2
Odds ratio (OR) and the equivalent Cohen's *d*

	P_0									
OR	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
1.1	***0.04***	***0.04***	***0.04***	***0.04***	***0.04***	***0.05***	***0..05***	***0.05***	***0.05***	***0.05***
1.2	***0.07***	***0.07***	***0.08***	***0.08***	***0.09***	***0.09***	***0.09***	***0.09***	***0.09***	***0.09***
1.3	***0.10***	***0.11***	***0.11***	***0.12***	***0.12***	***0.13***	***0.13***	***0.13***	***0.13***	***0.14***
1.4	***0.13***	***0.14***	***0.15***	***0.15***	***0.16***	***0.16***	***0.17***	***0.17***	***0.17***	***0.18***
1.5	***0.15***	***0.17***	***0.18***	***0.19***	***0.19***	0.20	0.20	0.21	0.21	0.21
1.6	***0.18***	0.20	0.21	0.22	0.22	0.23	0.24	0.24	0.25	0.25
1.7	0.20	0.22	0.24	0.25	0.25	0.26	0.27	0.27	0.28	0.28
1.8	0.23	0.25	0.26	0.27	0.28	0.29	0.30	0.30	0.31	0.31
1.9	0.25	0.27	0.29	0.30	0.31	0.32	0.33	0.33	0.34	0.34
2.0	0.27	0.29	0.31	0.32	0.34	0.35	0.35	0.36	0.37	0.37
3.0	0.44	0.48	0.51	0.53	0.55	0.56	0.58	0.59	0.60	0.61
4.0	0.56	0.62	0.65	0.68	0.71	0.73	0.74	0.76	0.77	0.78
5.0	0.66	0.73	0.77	***0.81***	***0.83***	***0.85***	***0.87***	***0.89***	***0.90***	***0.92***
6.0	0.75	***0.82***	***0.87***	***0.91***	***0.94***	***0.96***	***0.98***	***1.00***	***1.02***	***1.03***
7.0	***0.82***	***0.90***	***0.96***	***1.00***	***1.03***	***1.06***	***1.08***	***1.10***	***1.11***	***1.12***
8.0	***0.89***	***0.98***	***1.03***	***1.08***	***1.11***	***1.14***	***1.16***	***1.18***	***1.19***	***1.21***
9.0	***0.94***	***1.04***	***1.10***	***1.15***	***1.18***	***1.21***	***1.23***	***1.25***	***1.27***	***1.28***
10.0	***1.00***	***1.10***	***1.16***	***1.21***	***1.25***	***1.27***	***1.30***	***1.32***	***1.33***	***1.35***
15.0	***1.21***	***1.33***	***1.40***	***1.46***	***1.50***	***1.53***	***1.55***	***1.57***	***1.59***	***1.60***
20.0	***1.36***	***1.50***	***1.58***	***1.64***	***1.68***	***1.71***	***1.73***	***1.75***	***1.76***	***1.78***
25.0	***1.49***	***1.64***	***1.72***	***1.78***	***1.82***	***1.85***	***1.87***	***1.89***	***1.90***	***1.91***
30.0	***1.60***	***1.75***	***1.83***	***1.89***	***1.93***	***1.96***	***1.98***	***2.00***	***2.01***	***2.02***

P_0: rate of outcome of interest in the nonexposed group.
Z_0: standard normal deviation for P_0.
P_1: rate of outcome of interest in the exposed group, $P_1 = (\mathrm{OR} * P_0)/(1 - P_0 + \mathrm{OR} * P_0)$.
Z_1: standard normal deviation for P_1.
Cohen's $d = Z_0 - Z$.
Bold, italic values indicate Cohen's $d < 0.20$ or >0.80.

As logistic regression becomes more popular, OR is increasingly utilized in epidemiological studies (Chen et al., 2007). OR is used in case-control studies, for which the RR cannot be estimated. As noted above, OR is a good approximation to RR but only under certain circumstances (low population rates of "cases" <10%) and specific study designs (case-control studies; Davies et al., 1998; Hosmer and Lemeshow, 2000). In other words, OR does not give a good approximation of the RR when the sample rate of the outcome of interest is not very low (Altman et al., 1998; Sinclair and Bracken, 1994). Despite that caveat, however, OR is bound to be interpreted as RR, and it is unrealistic to expect news reporters or the public to understand this distinction (Schwartz et al., 1999). King and Zeng (2002) argued that "We have found no author who claims to be comfortable communicating with the general public using an odds ratio" (p. 1411).

The difficulty of interpreting the OR has troubled many clinical researchers and epidemiologists for a long time. We propose a new method for interpreting the size of the OR by relating it to differences in a normal standard deviate calculated from the respective probabilities being compared. Our calculations indicate that OR = 1.68, 3.47, and 6.71 are equivalent to Cohen's $d = 0.2$ (small), 0.5 (medium), and 0.8 (large), respectively, when disease rate is 1% in the nonexposed group; Cohen's $d < 0.2$ when OR < 1.5, and Cohen's $d > 0.8$ when OR > 5.

It would be useful to values with corresponding qualitative descriptors that estimate the strength of such associations; however, to date there is no consensus as to what those values of OR may be. Cohen (1988) suggested that $d = 0.2$, 0.5, and 0.8 are small, medium, and large on the basis of his experience as a statistician, but he also warned that these were only "rules of thumb." Better guidelines are needed to draw conclusions about strength of associations in studies of risks for disease when we use OR as the index of effect size in epidemiological studies.

References

Altman, D. G., Deeks, J. J., Sackett, D. L. (1998). Odds ratio should be avoided when events are common. *BMJ* 317:1318.

Bland, J. M., Altman, D. G. (2000). The odds ratio. *BMJ* 320:1468.

Chen, H., Cohen, P, Chen, S. (2007). Biased odds ratios from dichotomization of age. *Stat. Med.* 26:3487–3497.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

Cohen, P., Chen, H. (2009). How the reflection of linear correlation in odds ratios depends on the cut-off points. *Comm. Stat. Simul. Comput.* 38:612–620.

Davies, H. T. O., Crombie, I. K., Tavakoli, M. (1998). When can odds ratios mislead? *BMJ* 316:989–991.

Deeks, J. J. (1998). When can odds ratios mislead? *BMJ* 317:1155–1156.

Hosmer, D. W., Lemeshow, S. (2000). *Applied Logistic Regression*. 2nd ed. New York: Wiley.

King, G., Zeng, L. (2002). Estimating risk and rate levels, ratios and differences in case-control studies. *Stat. Med.* 21:1409–1427.

Koplan, J. P., Thacker, S. B., Lezin, N. A. (1999). Epidemiology in the 21st century: Calculation, communication and intervention. *Am. J. Publ. Health* 89:1153–1155.

Kraemer, H. C. (2004). Reconsidering the odds ratio as a measure of 2×2 association in a population. *Stat. Med.* 23:257–270.

Meehl, P. E. (1990). Appraising and amending theories: The strategy of lakatosian defense and principle that warrant it. *Psychol. Inq.* 1:108–141.

Newcombe, R. G. (2006). A deficiency of the odds ratio as a measure of effect size. *Stat. Med.* 25:4235–4240.

Sackett, D. L., Deeks, J. J., Altman, D. G. (1996). Down with odds ratios! *Evid. Base. Med.* 1:164–166.

Schwartz, L. M., Woloshin, S., Welch, H. G. (1999). Misunderstandings about the effects of race and sex on physicians' referrals for cardiac catheterization. *New Engl. J. Med.* 341:279–283.

Sinclair, J. C., Bracken, M. B. (1994). Clinically useful measures of effect in binary analyses of randomized trials. *J. Clin. Epidemiol.* 47:881–889.

Taubes, G. (1995). Epidemiology faces its limits. *Science* 269:164–169.

von Elm, E., Egger, M. (2004). The scandal of poor epidemiological research. *BMJ* 329:868–869.