Environmental Health Insights

METHODOLOGY

Informational Odds Ratio: A Useful Measure of Epidemiologic Association in Environment Exposure Studies

Jimmy T. Efird[1,2], Suzanne Lea[1], Amanda Toland[3] and Christopher J. Phillips[4]

[1]Department of Public Health, Brody School of Medicine, Greenville, NC. [2]Center for Health Disparities Research, East Carolina University, Greenville, NC. [3]Department of Molecular Virology, Immunology and Medical Genetics, The Ohio State University, Columbus, OH. [4]Department of Defense Center for Deployment Health Research, Naval Health Research Center, San Diego, CA. Corresponding author email: jimmy.efird@stanfordalumni.org

Abstract: The informational odds ratio (IOR) measures the post-exposure odds divided by the pre-exposure odds (ie, information gained after knowing exposure status). A desirable property of an adjusted ratio estimate is collapsibility (ie, the combined crude ratio will not change after adjusting for a variable that is not a confounder). Adjusted traditional odds ratios (TORs) are not collapsible. In contrast, Mantel-Haenszel adjusted IORs generally are collapsible. IORs are a useful measure of disease association in environmental case-referent studies, especially when the disease is common in the exposed and/or unexposed groups.

Keywords: informational odds ratio, collapsibility, pre- and post-exposure odds

Environmental Health Insights 2012:6 17–25

doi: 10.4137/EHI.S9236

This article is available from http://www.la-press.com.

Introduction

A central theme of environmental epidemiology is to quantify the occurrence (eg, incidence, prevalence) and/or outcome (eg, morbidity, mortality) of disease among a population exposed to a putative environmental hazard. The exposed population is then compared with a non-exposed population to determine if exposure is associated with disease. The environmental hazard may be behavioral in nature (eg, cigarette smoking, methamphetamine use, fat in diet), the consequence of modern lifestyle (eg, job stress, inadequate sleep), a by-product of industry (eg, air population, groundwater contamination, mercury in fish), or attributable to other sources in one's surroundings (eg, automobile exhaust, pesticide spraying, off-gassing of indoor building materials). Furthermore, the timing of the exposure may be short-lived, long-term, retrospective, prospective, current (ecologic), and/or ongoing. A short-term exposure to a very hazardous agent may convey the same impact on health as the continuous exposure to a relatively minor hazard. Gene-environment interaction also may play an important role in the underlying disease process.[1]

Different epidemiologic measures are available to gauge the association between environmental exposure and disease. The application of a particular measure depends on the underlying properties of the measure and the respective context of the study.[2] A frequently used measure of disease association in environmental exposure studies is the traditional odds ratio (TOR). This measure is defined as the odds for disease given exposure divided by the odds for disease given no exposure (Fig. 1). TORs have the distinct advantage of being invariant to rotation. That is, the disease TOR [ie, (a/b)/(c/d)] is equal to the exposure TOR [ie, (a/c)/(b/d)]. Furthermore, when disease is rare among both the exposed and non-exposed groups, TORs often are used in retrospective analyses as an approximate measure of relative risk (RR) [ie, TOR ≈ RR = (a/e)/(c/f)].[3]

An alternative measure of disease association closely related to the TOR is the informational odds ratio (IOR). The IOR measures the probability for exposure given disease divided by the probability for exposure given no disease (Fig. 1). Using Bayes theorem, it is easy to see that the IOR is equivalent to the post-exposure odds divided by the pre-exposure odds (Fig. 2).[4] The IOR resembles the traditional odds ratio (TOR) except that the probability terms in the denominator (ie, $P(D)/P(\bar{D})$) are not conditioned on the absence of exposure (ie, $P(D|\bar{E})/P(\bar{D}|\bar{E})$). When defined in the context of a receiver operator curve (ROC), the IOR also may be computed by multiplying the TOR by the likelihood ratio for a negative exposure (LR^-) (ie, $P(\bar{E}|D)/P(\bar{E}|\bar{D})$) (Fig. 3). Referring to Figure 1, TOR = (a/b)/(c/d) = 2.58 and LR^- = (c/d)/(g/h) = 0.56. Accordingly, IOR = 2.58*0.56. = 1.44. The IOR is interpreted as an outcome measure of information gained after knowing exposure status and may be used in case-referent studies independent of whether the disease is rare or common. When exposure is rare in both disease and non-disease groups, TOR ≈ IOR.

A desirable property of an adjusted ratio estimate is collapsibility (ie, the combined crude ratio will not change after adjusting for a variable that is not a confounder). TORs are not collapsible.[5,6] Applying standard techniques, we illustrate two approaches for computing a common IOR and 100(1−α)% confidence intervals (CIs) and compare the measures with respect to collapsibility.

Disease → ↓Exposure	D	$\bar{D}$	Total
E	a = 2352	b = 1600	e = 3952
$\bar{E}$	c = 912	d = 1600	f = 2512
Total	g = 3264	h = 3200	i = 6464

$$TOR = \frac{P(D|E)/P(\bar{D}|E)}{P(D|\bar{E})/P(\bar{D}|\bar{E})} = \left(\frac{a/b}{c/d}\right) = 2.58$$

$$IOR = \left(\frac{P(E|D)}{P(E|\bar{D})}\right) = \left(\frac{a/g}{b/h}\right) = 1.44$$

Figure 1. Computing TOR and IOR from a 2 × 2 contingency table.

$$IOR = \left(\frac{P(E|D)}{P(E|\bar{D})}\right) = \left(\frac{\frac{P(D|E)P(E)}{P(D)}}{\frac{P(\bar{D}|E)P(E)}{P(\bar{D})}}\right) = \left(\frac{\frac{P(D|E)}{P(\bar{D}|E)}}{\frac{P(D)}{P(\bar{D})}}\right) = \frac{\text{Post-exposure odds}}{\text{Pre exposure odds}} = \left(\frac{a/b}{g/h}\right).$$

Figure 2. Equivalence between IOR and the post-exposure odds divided by the pre-exposure odds.

Methods

95% robust (Normal theory) CI estimate for IOR

Given a single stratum (j), a large-sample (asymptotically consistent) estimate for var{log(IOR_j)} may be derived using the delta-method (based on a first order Taylor series) and is seen to equal $(1/a - 1/g + 1/b - 1/h)$ (Fig. 4).[7,8] The latter is equivalent to the robust "sandwich" estimate for var{log(IOR_j)}.[9,10] IORs are ratios of probabilities and confidence intervals are computed in an analogous manner as risk ratios.[11] Applying the central limit theorem (CLT), the computational formula for a 100(1−α)% robust (normal theory) CI estimate for IOR_j is given in Figure 5.[12] The 95% CI_{Robust} estimate for the crude IOR shown in Figure 1 is given as (1.38–1.50).

Covariate adjusted (pooled) estimate and 100(1−α)% confidence interval (CI) for stratified IOR

A summary estimate or common IOR for a series of 2 × 2 tables may be easily computed by taking the weighted average of stratum-specific IORs, given a fixed-effects model (ie, barring chance, the treatment effect is similar in all strata). Two main weighting techniques for pooling data across stratum are traditionally used in practice to compute combined relative-effect estimates.[13] Below, the methods are presented in the context of estimating a covariate-adjusted IOR and corresponding 100(1−α)%.

Woolf method

Assuming IORs are not significantly heterogeneous for k (j = 1 to k) strata and applying Woolf's weighted least squares method, the logarithm of the covariate adjusted (pooled) estimate for a stratified IOR [ie, log(IOR_{Woolf})] may be obtained by weighting the logarithm of each stratum-specific IOR_j estimate inversely proportional to its estimated variance (Fig. 6).[14,15] A 100(1−α)% normal theory CI estimate for IOR_{Woolf} is given in Figure 7.

Mantel-Haenszel method

The IOR also may be expressed as the cross-frequency for the a^{th} cell (ie, a*h/i) of a 2 × 2 table divided by cross-frequency for the c^{th} cell (ie, b*g/i). Given a series of 2 × 2 tables (stratum) indexed by (j = 1 to k), the weighted Mantel-Haenszel estimate for the common IOR is then computed by separately summing the cross-frequency terms in the numerator and denominator of the IOR estimate over each of the (k) stratum (Fig. 8).[16] Here again, we have assumed that the IORs are not significantly heterogeneous for k (j = 1 to k) strata. The term $(\hat{w})$ defined in Figure 4, which denotes the inverse $\widehat{\text{var}\{\log(IOR)\}}$ estimate, also may be written as a function of the cross-frequencies for the a^{th} and c^{th} cell (Fig. 9). A pooled estimate for $(\hat{w})$ is then computed by separately summing the terms in the numerator and denominator over each of the (k) stratum (j = 1 to k) (Fig. 10).[17] Applying the central limit theorem, a robust 100(1−α)% normal theory CI estimate for IOR_{MH} is given in Figure 11. Note,

$$IOR = \left(\frac{\frac{P(D|E)}{P(\bar{D}|E)}}{\frac{P(D)}{P(\bar{D})}}\right) = \left(\frac{\frac{P(D|E)}{P(\bar{D}|E)}}{\frac{P(D|\bar{E})}{P(D|\bar{E})}}\right) * \left(\frac{\frac{P(D|\bar{E})}{P(D)}}{\frac{P(\bar{D}|\bar{E})}{P(\bar{D})}}\right) = TOR * \frac{P(\bar{E}|D)}{P(\bar{E}|\bar{D})} = TOR * \frac{1 - \text{sensitivity}}{\text{Specificity}} = TOR * LR^{-}.$$

Figure 3. Relationship between IOR, TOR and LR^{-}.

Delta-method: $\mathrm{Var}(f(X) \approx [f'(\mu_X)]^2 \mathrm{Var}(X)$. If $f(X) = \log(X)$, then $f'(X) = 1/x$.

Assuming A~Bin (g,a/g) and B~Bin(h,b/h) and applying the delta-method, it follows that

$$\mathrm{Var}\{\log(\widehat{\mathrm{IOR}})\} = \mathrm{Var}\left\{\log\left(\frac{a/b}{g/h}\right)\right\} = \mathrm{Var}\left\{\log\left(\frac{a}{g}\right)\right\} + \mathrm{Var}\left\{\log\left(\frac{b}{h}\right)\right\}$$
$$= \left(\frac{g}{a}\right)^2 \left(\frac{\frac{a}{g}\left(1-\frac{a}{g}\right)}{g}\right) + \left(\frac{h}{b}\right)^2 \left(\frac{\frac{b}{h}\left(1-\frac{b}{h}\right)}{h}\right)$$
$$= \frac{1}{a} - \frac{1}{g} + \frac{1}{b} - \frac{1}{h} = \left(\frac{1}{\hat{w}}\right).$$

Figure 4. Derivation of $\mathrm{Var}\{\log(\widehat{\mathrm{IOR}})\}$ using the delta-method.

the $\mathrm{IOR}_{\mathrm{MH}}$ estimate will always be bounded by the minimum and maximum of the stratum specific IORs estimates, since it represents a weighted average of the individual stratum. If the disease ratios g_j/h_j are constant across strata, the Mantel-Haenszel estimate for IOR will equal the combine crude IOR.[17] When b_j/h_j are not constant across strata the variance estimate of the combined crude IOR will not be consistent and the Mantel-Haenszel estimate is generally recommended as the measure of association in this case.[17]

Results

Comparison of the Woolf and Mantel-Haenszel methods with respect to collapsibility

A confounding variable is an extraneous variable that masks the true influence of a putative causal variable on the effect (outcome) being studied. By definition, it must be related to both the cause and effect variables.[3] Consider the association between "crystal meth" (methamphetamine) use and cardiomyopathy in young patients.[18] Crystal meth users tend to be cigarette smokers and cigarette smoking potentially is associated with cardiomyopathy.[19] Failing to adjust for cigarette smoking may confound the association between crystal meth use and cardiomyopathy. An estimate is collapsible if the combined crude estimate does not change after adjusting for a variable that is not a confounder. It is well known that adjusted TORs are not collapsible.[5,6]

Consider the stratified data shown in Figures 12 and 13 corresponding to the collapsed data presented in Figure 1. If Exposure (E) represents the causal factor and Death (D) the effect, then Sex (S) is not a confounding variable since it is not related to Death on either the TOR or IOR scale (ie, $\widehat{\mathrm{TOR}}_{\mathrm{crude}} = 1.0$, $\widehat{\mathrm{IOR}}_{\mathrm{crude}} = 1.0$). However, if Sex (S) represents the causal factor and Death (D) the effect, then Exposure (E) is a confounder

By the central limit theorem (CLT), $\dfrac{\log(\widehat{\mathrm{IOR}}) - \log(\mathrm{IOR})}{\sqrt{\mathrm{Var}\left\{\log(\widehat{\mathrm{IOR}})\right\}}} = \dfrac{\log(\widehat{\mathrm{IOR}}) - \log(\mathrm{IOR})}{\sqrt{1/\hat{w}}} \rightarrow N(0,1).$

Accordingly, a 100(1−α)% confidence interval (CI) for $\widehat{\mathrm{IOR}}$ is given as $[e^L, e^U]$, where

$U = \log(\widehat{\mathrm{IOR}}) - \left(\dfrac{z_{1-\frac{\alpha}{2}}}{\sqrt{\hat{w}}}\right)$, $L = \log(\widehat{\mathrm{IOR}}) + \left(\dfrac{z_{1-\frac{\alpha}{2}}}{\sqrt{\hat{w}}}\right)$, and $Z_{1-\frac{\alpha}{2}}$ is the appropriate value from the standard Normal distribution for the 100(1−α/2) percentile.

Figure 5. Computing a robust 100(1−α)% confidence interval estimate for IOR.

$$\log_e\left(\widehat{IOR}_{Woolf}\right) = \frac{\sum \hat{w}_j \log_e \widehat{IOR}_j}{\sum \hat{w}_j}, \quad \text{where } \hat{w}_j = 1/\text{var}\left\{\log\left(\widehat{IOR}_j\right)\right\}.$$

Figure 6. Woolf's weighted least squares estimate for the logarithm of IOR.

An 100(1−α)% normal theory confidence interval (CI) for $\widehat{IOR}_{Woolf}$ is given as $[e^L, e^U]$, where

$$U = \log\left(\widehat{IOR}_{Woolf}\right) - \left(\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{\sum \hat{w}_j}}\right), \; L = \log\left(\widehat{IOR}_{Woolf}\right) + \left(\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{\sum \hat{w}_j}}\right), \text{ and } Z_{1-\frac{\alpha}{2}} \text{ is the appropriate value}$$

from the standard Normal distribution for the 100(1−α/2) percentile.

Figure 7. Computing an 100(1−α)% confidence interval estimate for IOR_{Woolf}.

$$\widehat{IOR}_{MH} = \frac{\sum \frac{a_j h_j}{i_j}}{\sum \frac{b_j g_j}{i_j}}.$$

Figure 8. Mantel-Haenszel estimate for a common IOR.

$$\hat{w} = \text{var}\left\{\log\left(\widehat{IOR}\right)\right\}^{-1} = \left(\frac{1}{a} - \frac{1}{g} + \frac{1}{b} - \frac{1}{h}\right)^{-1} = \frac{\left(\frac{ah}{i}\right)\left(\frac{bg}{i}\right)}{\left(\frac{(ghe - abi)}{i^2}\right)}.$$

Figure 9. Expressing $\hat{w}$ in terms of the cross-frequencies for the a^{th} and b^{th} cell.

$$\hat{w}_{pooled} = \frac{\sum\left(\frac{a_j h_j}{i_j}\right)\sum\left(\frac{b_j g_j}{i_j}\right)}{\sum\left\{\frac{\left(g_j h_j e_j - a_j b_j i_j\right)}{i_j^2}\right\}}.$$

Figure 10. Computing a pooled version for $\hat{w}$ over (k) stratum (j = 1 to k).

An 100(1−α)% normal theory confidence interval (CI) estimate for $\widehat{IOR}_{MH}$ is given as $[e^L, e^U]$, where

$$U = \log\left(\widehat{IOR}_{MH}\right) - \left(\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{\sum \hat{w}_{pooled}}}\right), \; L = \log\left(\widehat{IOR}_{MH}\right) + \left(\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{\sum \hat{w}_{pooled}}}\right), \text{ and } Z_{1-\frac{\alpha}{2}} \text{ is the appropriate value}$$

from the standard Normal distribution for the 100(1−α/2) percentile.

Figure 11. Computing a robust 100(1−α)% confidence interval estimate for IOR_{MH}.

Males

Disease → ↓ Exposure	D	$\bar{D}$	Total
E	a = 1356	b = 1040	e = 2396
$\bar{E}$	c = 276	d = 560	f = 836
Total	g = 1632	h = 1600	i = 3232

TOR = 2.65 (95% CI_{Exact} = 2.24–3.13)

IOR = 1.28 (95% CI_{Robust} = 1.23–1.33)

Females

Disease → ↓ Exposure	D	$\bar{D}$	Total
E	a = 996	b = 560	e = 1556
$\bar{E}$	c = 636	d = 1040	f = 1676
Total	g = 1632	h = 1600	i = 3232

TOR = 2.91 (95% CI_{Exact} = 2.51–3.36)

IOR = 1.74 (95% CI_{Robust} = 1.61–1.88)

All patients

Disease → ↓ Sex	D	$\bar{D}$	Total
Male	a = 1632	b = 1600	e = 3232
Female	c = 1632	d = 1600	f = 3232
Total	g = 3264	h = 3200	i = 6464

TOR = 1.00 (95% CI_{Exact} = 0.91–1.10)

IOR = 1.00 (95% CI_{Robust} = 0.95–1.05)

All patients

Exposure → ↓ Sex	E	$\bar{E}$	Total
Male	a = 2396	b = 836	e = 3232
Female	c = 1556	d = 1674	f = 3232
Total	g = 3952	h = 2512	i = 6464

TOR = 3.09 (95% CI_{Exact} = 2.78–3.43)

IOR = 1.82 (95% CI_{Robust} = 1.71–1.94)

Figure 12. Contingency tables corresponding to data in Figure 1 stratified by sex.

because it is related to both Death ($\widehat{TOR}_{crude} = 2.58$, $\widehat{IOR}_{crude} = 1.44$) and Sex ($\widehat{TOR}_{crude} = 3.09$, $\widehat{IOR}_{crude} = 1.54$). Referring to Figure 14, we see that neither $\widehat{TOR}_{Woolf} = 2.79$ nor $\widehat{TOR}_{MH} = 2.79$ are collapsible with respect to sex because both adjusted estimates differ from the combined $\widehat{TOR}_{crude} = 2.58$. However, referring to Figure 15 we see that the adjusted Mantel-Haenszel estimate for this example is collapsible with respect to sex (ie, $\widehat{IOR}_{MH} = 1.44 = \widehat{IOR}_{crude}$). On the other hand the adjusted Woof estimate is not is collapsible with respect to sex (ie, $\widehat{IOR}_{Woolf} = 1.37 \# \widehat{IOR}_{crude}$). The IOR_{Woolf} estimate is based on a non-linear (logarithmic) weighted estimate of stratum-specific IORs and accordingly the combined crude IOR does not necessarily remain constant after adjusting for a variable that is not a confounder. In our simple example, we see that the results obtained by the Mantel-Haenszel method are identical to those obtained from a Poisson regression model using robust variance estimation.

Exact confidence intervals for IOR

When sample sizes are small, an exact unconditional CI estimate may be computed for the IOR. However, due to the discrete nature of the problem, the resulting CI estimates tend to be very wide. Consider the case when exposure is rare in both disease and non-disease groups (ie, TOR ≈ IOR). In the example shown in Figure 16, we see that the standard exact CI estimate for the IOR[20] is considerably wider than the standard exact CI estimate for the TOR[21] even though one would expect the coverage to be nearly equal. Furthermore, as illustrated in Figure 17, the standard exact CI for the IOR estimate is neither asymptotically efficient nor consistent. A pseudo "continuity-adjusted" exact confidence interval based on the Farrington-Manning score statistic provides better coverage in some cases, however the resulting CIs may be too narrow when one or more cell sizes are very small, as illustrated in Figure 16 (IOR = 1.0, CI_{MF} = 0.0594–11.1435).[22] By parallel analogy, the above small-sample concerns identically apply to RR estimates. Methods for improving the nominal coverage (ie, at least 1–α) of unconditional exact marginal effect estimates have been suggested in the literature.[23]

Discussion

A desirable property of an adjusted ratio estimate is that the combined crude ratio will not change after adjusting for a variable that is not a confounder

Exposed

Disease → ↓ Sex	D	$\bar{D}$	Total
Male	a = 1356	b = 1040	e = 2396
Female	c = 996	d = 560	f = 1556
Total	g = 2352	h = 1600	i = 3952

TOR = 0.73 (95% CI_{Exact} = 0.64–0.84)
IOR = 0.89 (95% CI_{Robust} = 0.84–0.93)

Non-exposed

Disease → ↓ Sex	D	$\bar{D}$	Total
Male	a = 276	b = 560	e = 836
Female	c = 636	d = 1040	f = 1676
Total	g = 912	h = 1600	i = 2512

TOR = 0.81 (95% CI_{Exact} = 0.67–0.93)
IOR = 0.86 (95% CI_{Robust} = 0.77–0.97)

All patients

Disease → ↓ Exposure	D	$\bar{D}$	Total
E	a = 2352	b = 1600	e = 3952
$\bar{E}$	c = 912	d = 1600	f = 2512
Total	g = 3264	h = 3200	i = 6464

TOR = 2.58 (95% CI_{Exact} = 2.32–2.86)
IOR = 1.44 (95% CI_{Robust} = 1.38–1.50)

All patients

Sex→ ↓ Exposure	Male	Female	Total
E	a = 2396	b = 1556	e = 3952
$\bar{E}$	c = 836	d = 1676	f = 2512
Total	g = 3232	h = 3232	i = 6464

TOR = 3.09 (95% CI_{Exact} = 2.78–3.43)
IOR = 1.54 (95% CI_{Robust} = 1.48–1.60)

Figure 13. Contingency tables corresponding to data in Figure 1 stratified by exposure.

Characteristic ↓	TOR_{Crude} (95% CI)	TOR_{Woolf} (95% CI)	TOR_{MH}† (95% CI)	TOR_{LR}‡ (95% CI)
Exposure				
$\bar{E}$	1.00 referent	1.00 referent	1.00 referent	1.00 referent
E	2.58 (2.32–2.86)	2.79 (2.51–3.11)	2.79 (2.50–3.11)	2.79 (2.51–3.11)
Sex				
Female	1.00 referent	1.00 referent	1.00 referent	1.00 referent
Male	1.00 (0.91–1.10)	0.76 (0.68–0.84)	0.76 (0.68–0.84)	0.76 (0.68–0.84)

Notes: †Adjusted Mantel-Haenszel estimate; ‡Adjusted logistic regression estimate.

Figure 14. Crude and adjusted TOR estimates corresponding to data in Figures 1, 12 and 13.

Characteristic ↓	IOR_{Crude} (95% CI)	IOR_{Woolf} (95% CI)	IOR_{MH}† (95% CI)	IOR_{PR}‡ (95% CI)
Exposure				
$\bar{E}$	1.00 referent	1.00 referent	1.00 referent	1.00 referent
E	1.44 (1.38–1.50)	1.37 (1.32–1.42)	1.44 (1.39–1.50)	1.44 (1.39–1.50)
Sex				
Female	1.00 referent	1.00 referent	1.00 referent	1.00 referent
Male	1.00 (0.95–1.05)	0.88 (0.84–0.93)	0.88 (0.84–0.92)	0.88 (0.84–0.92)

Notes: †Adjusted Mantel-Haenszel estimate; ‡Adjusted Poisson regression estimate.

Figure 15. Crude and adjusted IOR estimates corresponding to data in Figures 1, 12 and 13.

Disease → ↓ Exposure	D	$\bar{D}$	Total
E	a = 1	b = 2	e = 3
$\bar{E}$	c = 250	d = 500	f = 750
Total	g = 251	h = 502	i = 753

TOR = 1.00 (95% CI_{Exact} = 0.0169–19.2944)
IOR = 1.00 (95% CI_{Exact} = 0.0001–29.3570)
IOR = 1.00 (95% CI_{FM} = 0.0594–11.1435)

Figure 16. Comparison of exact confidence interval procedures for TOR and IOR.

Disease → ↓ Exposure	D	$\bar{D}$	Total
E	a = 1440	b = 480	e = 1920
$\bar{E}$	c = 1760	d = 2720	f = 4480
Total	g = 3200	h = 3200	i = 6400

IOR = 3.00 (95% $CI_{Asymptotic}$ = 2.7392–3.2856)
IOR = 3.00 (95% CI_{Exact} = 0.2696–1171.4405)

Figure 17. Comparison of asymptotic and exact confidence interval procedures for IOR.

(ie, collapsibility). It is well known in the literature that adjusted TORs are not collapsible. This is illustrated in Figure 14, where both the TOR_{Woolf} and TOR_{MH} sex adjusted estimates differed from the combined crude TOR, even though sex is not a confounding variable. In prospective (cohort) studies, the association between a putative exposure and disease adjusting for other important model variables may be computed using the generally collapsible Mantel-Haenszel RR estimate. When disease is rare among both the exposed and non-exposure groups in a case-referent study, the TOR and RR estimates will be approximately equal. However, the outcome of interest in some retrospective environment exposure studies may be fairly common and the TOR estimate will not equal the combined crude estimate after adjusting for a variable that is not a confounder.

The IOR is a useful measure of association in environmental case-referent studies, especially when the outcome under consideration is known to occur frequently. Similar to RRs, Mantel-Haenszel adjusted IORs are generally collapsible (criteria for simple and strict collapsibility are discussed in the literature[6,24,25]). The IOR measures how much more (or less) likely patients with the disease have a particular exposure than those without disease (ie, the post-exposure odds divided by the pre-exposure odds).[11] Similar to other relative effect estimates IORs are logarithmic, meaning that a value of 1.0 corresponds to no association between exposure and disease, while an IOR greater/less than unity indicates a positive/negative association with disease.

Acknowledgments

Katherine T. Jones (ECU) offered valuable editorial assistance during the writing of this manuscript.

Author Contributions

Conceived and designed the experiments: JTE. Analysed the data: JTE. Wrote the first draft of the manuscript: JTE. Contributed to the writing of the manuscript: JTE, SL, AT, CJP. Agree with manuscript results and conclusions: JTE, SL, AT, CJP. Jointly developed the structure and arguments for the paper: JTE, SL, AT, CJP. Made critical revisions and approved final version: JTE, SL, AT, CJP. All authors reviewed and approved of the final manuscript.

Disclosures and Ethics

As a requirement of publication author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship and contributorship, conflicts of interest, privacy and confidentiality and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

References

1. Efird J. An efficient gatekeeper algorithm for detecting GxE. *Cancer Inform*. 2010;12:115–20.
2. Behrens T, Pigeot I, Ahrens W. Epidemiologische und statistische methoden der risikoabschätzung. *Bundesgesundheitsbl*. 2009;52:1151–60.
3. Wassertheil-Smoller S. *Biostatistics and Epidemiology*. New York, NY: Springer-Verlag, 1990.
4. Katz M. A probability graph describing the predictive value of a highly sensitive diagnostic test. *N Engl J Med*. 1944;291:1115–6.
5. Cummings P. The relative merits of risk ratios and odds ratios. *Arch Pediatr Adoles Med*. 2009;163:438–45.
6. Wermuth N. Parametric collapsibility and the lack of moderating effects in contingency tables with a dichotomous response variable. *JR Statist Soc B*. 1987;49:353–64.
7. Oehlert G. A note on the delta method. *Amer Stat*. 1992;46:27–9.
8. Katz D, Baptista J, Azen S, Pike M. Obtaining confidence intervals for risk ratio in cohort studies. *Biometrics*. 1978;34:469–74.

9. Kauermann G, Carroll R. A note on the efficiency of sandwich covariance matrix estimation. *JASA*. 2001;96:1387–96.
10. Zou G. A modified Poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159:702–6.
11. Deeks J, Altman D. Diagnostic tests 4: likelihood ratios. *Br Med J*. 2004;329: 168–9.
12. Le Cam L. The central limit theorem around 1935. *Statist Sci*. 1986;1: 78–91.
13. Morris J, Gardner M. Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. *Br Med J*. 1988;296:1313–6.
14. Woolf B. On estimating the relation between blood group and disease. *Ann Hum Genet*. 1955;19:251–3.
15. Grizzle J, Starmer F, Koch G. Analysis of categorical data by linear models. *Biometrics*. 1969;25:489–504.
16. Woodward M. *Epidemiology: Study Design and Data Analysis*. 2nd ed. Boca Raton: Chapman & Hall/CRC, 2005.
17. Greenland S, Robins J. Estimation of a common effects parameter from sparse follow-up data. *Biometrics*. 1985;41:55–68.
18. Yeo K, Wijetunga M, Ito H, et al. The association of methamphetamine use and cardiomyopathy in young patients. *Am J Med*. 2007;120:165–71.
19. Hartz A, Anderson A, Brooks H, Manley J, Parent G, Barboriak J. The association of smoking with cardiomyopathy. *N Engl J Med*. 1984;311: 1201–6.
20. Santner T, Snell M. Small-sample confidence intervals for p_1-p_2 and p_1/p_2 in 2×2 contingency tables. *JASA*. 1980;75:386–94.
21. Thomas D. Algorithm AS-36. Exact confidence limits for the odds ratio in a 2×2 table. *Appl Stat*. 1971;20:105–10.
22. Chan I, Zhang Z. Test-based exact confidence intervals for the difference of two binomial proportions. *Biometrics*. 1999;55:1202–9.
23. Mukhopadhyay P. Exact tests and exact confidence intervals for the ratio of two binomial proportions. Ph.D. Dissertation, NC State University, 2003.
24. Whittemore A. Collapsibility of multidimensional contingency tables. *JR Statist Soc B*. 1978;40:328–40.
25. Geng Z. Collapsibility of relative risk in contingency tables with a response variable. *JR Statist Soc B*. 1992;54:585–93.