Relative Effect Sizes for Measures of Risk

Jake Olivier, Melanie Bell, Warren May

November 2015

What is a small, medium or large odds ratio?

Procedure	Dead	Alive	Mortality
Ileostomy	4	19	17%
Ileoproctostomy	2	35	5%

- Participants were all patients receiving a total abdominal colectomy from two hospitals over a five year period
- The odds ratio for 30 day mortality was $OR = 3.68$[1]

[1] Payne JA, Snyder DC, Olivier J & Salameh JR. (2007) Total Abdominal Colectomy: Patient Satisfaction and Outcomes. The American Surgeon, 73(7): 709-711.

What is a small, medium or large odds ratio?

Procedure	Dead	Alive	Mortality
Ileostomy	4	19	17%
Ileoproctostomy	2	35	5%

- Participants were all patients receiving a total abdominal colectomy from two hospitals over a five year period
- The odds ratio for 30 day mortality was $OR = 3.68$[1]
- $p = 0.132$

[1] Payne JA, Snyder DC, Olivier J & Salameh JR. (2007) Total Abdominal Colectomy: Patient Satisfaction and Outcomes. The American Surgeon, 73(7): 709-711.

What is a small, medium or large odds ratio?

Procedure	Dead	Alive	Mortality
Ileostomy	4	19	17%
Ileoproctostomy	2	35	5%

- Participants were all patients receiving a total abdominal colectomy from two hospitals over a five year period
- The odds ratio for 30 day mortality was $OR = 3.68$[1]
- $p = 0.132$
- Is this result unimportant? Would you choose one method over the other if you were the patient?

[1] Payne JA, Snyder DC, Olivier J & Salameh JR. (2007) Total Abdominal Colectomy: Patient Satisfaction and Outcomes. The American Surgeon, 73(7): 709-711.

What is a small, medium or large odds ratio?

	Alcohol		
Helmet	Yes	No	% Alcohol
Yes	0	17	0%
No	18	125	12.6%

- Case-control study of bicyclists presenting to trauma centre in Singapore[2]
- Authors conclude "[a]lcohol consumption did not correlate with...helmet wearing" and report a p-value of 'NS'

[2] Heng KWJ, Lee AHP, Zhu S, Tham KY, Seow E. (2006) Helmet use and bicycle-related trauma in patients presenting to an acute hospital in Singapore. Singapore Med J 47(5): 367-372.

What is a small, medium or large odds ratio?

	Alcohol		
Helmet	Yes	No	% Alcohol
Yes	0	17	0%
No	18	125	12.6%

- Case-control study of bicyclists presenting to trauma centre in Singapore[2]
- Authors conclude "[a]lcohol consumption did not correlate with...helmet wearing" and report a p-value of 'NS'
- $OR_{CC} = 0.19$ (continuity corrected)
- Helmet wearing was associated with an 81% reduction in the odds of alcohol consumption

[2] Heng KWJ, Lee AHP, Zhu S, Tham KY, Seow E. (2006) Helmet use and bicycle-related trauma in patients presenting to an acute hospital in Singapore. Singapore Med J 47(5): 367-372.

Motorists are more aggressive to helmeted cyclists

- Ian Walker, University of Bath[3]
- Two sensors on a bicycle: one for overtaking distance and the other the distance to the kerb
- Alternated between wearing and not wearing a helmet
- Vehicles overtook, on average, closer when helmeted
- Reanalyze data with passing distance categorized by one metre rule[4]
  - Close overtaking increases lateral forces ($< 1m \equiv$'unsafe')
  - 2×2 table for helmet wearing vs. safe/unsafe passing distance

---

[3] Walker, I., 2007. Drivers overtaking bicyclists: Objective data on the effects of riding position, helmet use, vehicle type and apparent gender. Acc. Anal. Prev. 39, 417-425.

[4] Olivier J, Walter SR (2013) Bicycle Helmet Wearing Is Not Associated with Close Motor Vehicle Passing: A Re-Analysis of Walker, 2007. PLoS One 8: e75424.

Motorists are more aggressive to helmeted cyclists

- Ian Walker, University of Bath[3]
- Two sensors on a bicycle: one for overtaking distance and the other the distance to the kerb
- Alternated between wearing and not wearing a helmet
- Vehicles overtook, on average, closer when helmeted
- Reanalyze data with passing distance categorized by one metre rule[4]
  - Close overtaking increases lateral forces ($< 1m \equiv$'unsafe')
  - 2×2 table for helmet wearing vs. safe/unsafe passing distance
  - $OR_{adj} = 1.13$ (adj for kerb distance, vehicle size, city)

---

[3] Walker, I., 2007. Drivers overtaking bicyclists: Objective data on the effects of riding position, helmet use, vehicle type and apparent gender. Acc. Anal. Prev. 39, 417-425.

[4] Olivier J, Walter SR (2013) Bicycle Helmet Wearing Is Not Associated with Close Motor Vehicle Passing: A Re-Analysis of Walker, 2007. PLoS One 8: e75424.

Relative Effect Sizes

- Broad effect size recommendations exist for other measures[5]
- Simplest case: Cohen's d for difference in means

Population

$$\delta = \frac{\mu_1 - \mu_2}{\sigma}$$

Sample

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

  - $d = 0.2$ (small)
  - $d = 0.5$ (medium)
  - $d = 0.8$ (large)

[5] Cohen J. Statistical Power Analysis for the Behavioral Sciences. Academic Press: San Diego, CA, 1988.

Cohen's *d* for Correlation

- Using d as an *anchor*, Cohen extended recommendations to related hypothesis tests
- For example,
  - 2 sample t-test (equal variances)
  - $\iff$ point biserial correlation
  - $\implies$ Pearson's correlation coefficient
- Convert d to r (equal sample sizes)

$$r = \frac{d}{\sqrt{d^2 + 4}}$$

- Modified recommendations
  - $r = 0.1$ (small)
  - $r = 0.3$ (medium)
  - $r = 0.5$ (large)

Cohen's d for Correlation

- Using d as an *anchor*, Cohen extended recommendations to related hypothesis tests
- For example,
  - 2 sample t-test (equal variances)
  - $\iff$ point biserial correlation
  - $\implies$ Pearson's correlation coefficient
- Convert d to r (equal sample sizes)

$$r = \frac{d}{\sqrt{d^2 + 4}}$$

- Modified recommendations
  - $r = 0.1$ (small)
  - $r = 0.3$ (medium)
  - $r = 0.5$ (large)
- Unmodified $r = 0.1, 0.24, 0.37$

Existing Odds Ratio Recommendations

- For uniform margins (i.e., $\pi_{1+} = \pi_{+1} = 0.5$), Cohen's recommendations are equivalent to
  - $OR = 1.49$ (small)
  - $OR = 3.45$ (medium)
  - $OR = 9.00$ (large)
- Ferguson[6] recommends 2.0, 3.0 and 4.0
  - Doesn't seem to be based on anything
- Haddock et al.[7] consider odds ratios greater than 3.0 to be *large*
  - Only a rule of thumb

[6] Ferguson, C.J., 2009. An Effect Size Primer: A Guide for Clinicians and Researchers. Professional Psychology: Research and Practice 40(5), 532-538.

[7] Haddock C, Rindskopf D, Shadish W (1998) Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. Psychological Methods 3: 339–353.

2×2 Contingency Tables

	$X=0$	$X=1$	*Total*
$Y=0$	π_{00}	π_{01}	π_{0+}
$Y=1$	π_{10}	π_{11}	π_{1+}
Total	π_{+0}	π_{+1}	1.0

- $\pi_{ij} = P\left(Y=i, X=j\right)$ for $i, j \in \{0, 1\}$
- If X and Y are independent, $\pi_{ij} = \pi_{i+}\pi_{+j}$
- marginal probabilities π_{1+} and π_{+1}
- Alternatively, can be expressed by $n_{ij} = n \times \pi_{ij}$

	$X=0$	$X=1$	*Total*
$Y=0$	n_{00}	n_{01}	n_{0+}
$Y=1$	n_{10}	n_{11}	n_{1+}
Total	n_{+0}	n_{+1}	n

Effect Size for 2×2 Tables

- Pearson's correlation coefficient

$$r = \phi = \frac{\sum_{\ell=1}^{n} \left(X_\ell - \bar{X}\right)\left(Y_\ell - \bar{Y}\right)}{\sqrt{\sum_{\ell=1}^{n} \left(X_\ell - \bar{X}\right)^2 \sum_{\ell=1}^{n} \left(Y_\ell - \bar{Y}\right)^2}}$$
$$= \frac{\pi_{11} - \pi_{1+}\pi_{+1}}{\sqrt{\pi_{1+}\pi_{+1}\left(1 - \pi_{1+}\right)\left(1 - \pi_{+1}\right)}}$$

- Cohen gives recommendations of
  - $\phi = 0.1$ (small)
  - $\phi = 0.3$ (medium)
  - $\phi = 0.5$ (large)

Relative Risk

- For 2×2 tables, the relative risk is

$$RR = \frac{\pi_{11}(1-\pi_{+1})}{\pi_{10}\pi_{+1}} = \frac{\pi_{11} - \pi_{11}\pi_{+1}}{\pi_{1+}\pi_{+1} - \pi_{11}\pi_{+1}}$$

- We can represent π_{11} in terms of the marginal probabilities and ϕ

$$\pi_{11} = \pi_{1+}\pi_{+1} + \phi\sqrt{\pi_{1+}\pi_{+1}\left(1-\pi_{1+}\right)\left(1-\pi_{+1}\right)}$$

- So, clearly the relative risk can be written as a function of ϕ and the marginal probabilities only

Consequence of Fixed Margins

- For $\pi_{1+} < \pi_{+1}$, ϕ is bounded above by

$$\phi_{\max} = \max_{\pi_{11}} \phi = \sqrt{\frac{\pi_{1+}\left(1 - \pi_{+1}\right)}{\pi_{+1}\left(1 - \pi_{1+}\right)}}$$

- This also constrains the relative risk when transforming from ϕ
- For example, when $\pi_{+1} = 0.5$ and $\pi_{1+} = 0.1$,

$$\phi_{\max} = \frac{1}{3}$$

- This problem is exacerbated when $\pi_{1+} \downarrow 0$

Effect Sizes Relative to PhiMax

- Cohen's recommendations are, in fact, increments of perfect correlation
- Choose increments of maximum possible correlation instead, i.e., $\alpha = \phi/\phi_{\text{max}}$

Effect Sizes Relative to PhiMax

- Cohen's recommendations are, in fact, increments of perfect correlation
- Choose increments of maximum possible correlation instead, i.e., $\alpha = \phi/\phi_{\mathsf{max}}$

$$RR_\alpha = 1 + \frac{\alpha}{(1-\alpha)\pi_{+1}}$$

Effect Sizes Relative to PhiMax

- Cohen's recommendations are, in fact, increments of perfect correlation
- Choose increments of maximum possible correlation instead, i.e., $\alpha = \phi/\phi_{\text{max}}$

$$RR_\alpha = 1 + \frac{\alpha}{(1-\alpha)\pi_{+1}}$$

- For 1:1 allocation, i.e., $\pi_{+1} = 0.5$
  - $RR_\alpha = 1.22$ (small)
  - $RR_\alpha = 1.86$ (medium)
  - $RR_\alpha = 3.00$ (large)
- Obviously, other inputs for α and π_{+1} can be used

Odds Ratio

- In a similar manner, the odds ratio is

$$OR_\alpha = 1 + \frac{\pi_{+1}}{\pi_{+1} - \pi_{1+}(\pi_{+1}(1-\alpha) + \alpha)}(RR_\alpha - 1)$$

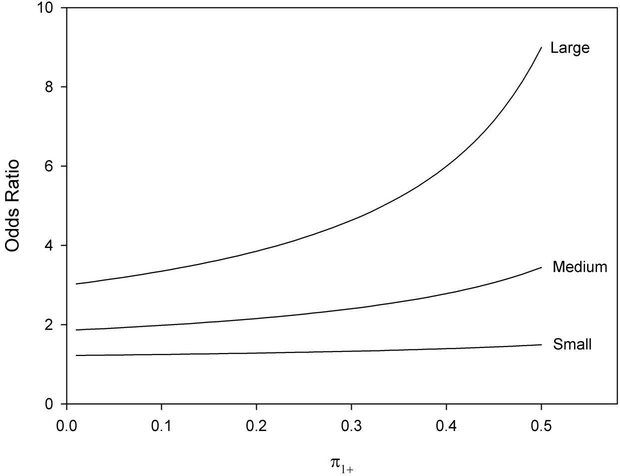
- In terms of π_{1+}, the extreme values are

$$OR_{min} = \lim_{\pi_{1+} \to 0} OR_\alpha = RR_\alpha$$

and when $\pi_{1+} = \pi_{+1} = \pi$

$$OR_{max} = 1 + \frac{RR_{\alpha - 1}}{(1-\pi)(1-\alpha)}$$

1:1 Participant Allocation ($\pi_{+1} = 0.5$)



"Practical" Odds Ratio Effect Sizes

- For planning purposes or when the event is *rare* (sometimes $\pi_{1+} < 0.10$ is used), the minimum odds ratio (or relative risk) would be an acceptable conservative approach[8]

[8] Olivier J, Bell ML (2013) Effect sizes for 2×2 contingency tables. PLoS One 8: e58777.

"Practical" Odds Ratio Effect Sizes

- For planning purposes or when the event is *rare* (sometimes $\pi_{1+} < 0.10$ is used), the minimum odds ratio (or relative risk) would be an acceptable conservative approach[8]
- What about more common outcomes? Odds ratios are known to overaccentuate relative risk.

[8] Olivier J, Bell ML (2013) Effect sizes for 2×2 contingency tables. PLoS One 8: e58777.

"Practical" Odds Ratio Effect Sizes

- For planning purposes or when the event is *rare* (sometimes $\pi_{1+} < 0.10$ is used), the minimum odds ratio (or relative risk) would be an acceptable conservative approach[8]
- What about more common outcomes? Odds ratios are known to overaccentuate relative risk.
- Midpoint odds ratio

$$OR_{mid} = 1 + \frac{2}{2 - (\pi_{+1}(1-\alpha) + \alpha)}(RR_\alpha - 1)$$

- Average odds ratio

$$\begin{aligned}\overline{OR} &= \frac{1}{\pi_{+1}} \int_0^{\pi_{+1}} OR_\alpha d\pi_{1+} \\ &= 1 - \frac{\log((1-\pi_{+1})(1-\alpha))}{\pi_{+1}(1-\alpha) + \alpha}(RR_\alpha - 1)\end{aligned}$$

[8] Olivier J, Bell ML (2013) Effect sizes for 2×2 contingency tables. PLoS One 8: e58777.

"Practical" Odds Ratio Effect Sizes

	α	OR_{min}	OR_{mid}	$\overline{OR}$	OR_{max}
Small	0.1	1.22	1.31	1.32	1.49
Medium	0.3	1.86	2.27	2.38	3.45
Large	0.5	3.00	4.20	4.70	9.00

Hazard Ratio

- The hazard function

$$h(t) = \lim_{\Delta t \downarrow 0} \frac{P(t < T \leq t + \Delta t)}{P(T > t)}$$

  is the instantaneous probability of death at time t given that subjects have survived up to time t

- HR is the ratio of instantaneous risks $\Longrightarrow$ *instantaneous relative risk*
- Under certain conditions, $HR = RR$ and can be interpreted in a similar fashion
- For example, accelerated failure time model where everyone observed for 1 unit of time

Hazard Ratio

- The hazard function

$$h(t) = \lim_{\Delta t \downarrow 0} \frac{P(t < T \leq t + \Delta t)}{P(T > t)}$$

  is the instantaneous probability of death at time t given that subjects have survived up to time t
- HR is the ratio of instantaneous risks $\Longrightarrow$ *instantaneous relative risk*
- Under certain conditions, $HR = RR$ and can be interpreted in a similar fashion
- For example, accelerated failure time model where everyone observed for 1 unit of time

$$HR_{mle} = RR$$

McNemar's Test

- Exact binomial test for $H_0 : \pi = 1/2$ where

$$P = \frac{b}{b+c} \equiv \frac{\pi_b}{\pi_b + \pi_c}$$

- Mantel-Haenszel Odds Ratio is the effect size

$$OR_{MH} = \frac{b}{c} \equiv \frac{\pi_b}{\pi_c}$$

- For single proportion, Cohen uses effect size g

$$g = P - 0.5$$

- Can write OR_{MH} in terms of g

$$OR_{MH} = 1 + \frac{4g}{1-2g}$$

McNemar's Test

- Exact binomial test for $H_0 : \pi = 1/2$ where
$$P = \frac{b}{b+c} \equiv \frac{\pi_b}{\pi_b + \pi_c}$$
- Mantel-Haenszel Odds Ratio is the effect size
$$OR_{MH} = \frac{b}{c} \equiv \frac{\pi_b}{\pi_c}$$
- For single proportion, Cohen uses effect size g
$$g = P - 0.5$$
- Can write OR_{MH} in terms of g
$$OR_{MH} = 1 + \frac{4g}{1-2g}$$
- Using Cohen's recommendations for $g \in \{0.05, 0.15, 0.25\}$, we get
  - $OR_{MH} = 1.22$ (small)
  - $OR_{MH} = 1.86$ (medium)
  - $OR_{MH} = 3.00$ (large)

Revisit Examples

Study	OR	Effect Size
Payne (2007)	3.68	Large
Heng (2006)	$1/0.19{=}5.26$	Large
Walker (2007)	1.13	*Trivial?*

Revisit Examples

Study	OR	Effect Size	Sample Size
Payne (2007)	3.68	Large	60
Heng (2006)	1/0.19=5.26	Large	160
Walker (2007)	1.13	*Trivial?*	2355

Discussion

- Binary correlation coefficient (and equivalences) are unusable as effect size measures
- Relative risk and odds ratio are not constrained by ϕ_{max} and more generalizable (logistic regression)
- *RR*, *OR* for *rare* events, *HR*, McNemar's test effect sizes anchored to Cohen's recommendations are 1.22, 1.86 and 3.00
- Can be modified for other participant allocation ratios
- Larger *OR* for common events

Is this a *good* idea?

- Is this just a "guess masquerading as mathematics"?[9]
- "This is an elaborate way to arrive at the same sample size that has been used in past social science studies of large, medium, and small size (respectively). The method uses a standardized effect size as the goal. Think about it: for a "medium" effect size, you'll choose the same n regardless of the accuracy or reliability of your instrument, or the narrowness or diversity of your subjects. Clearly, important considerations are being ignored here. "Medium" is definitely not the message!"[10]

---

[9] Julious SA, Campbell MJ (2012) Tutorial in biostatistics: sample sizes for parallel group clinical trials with binary data. Statistics in Medicine 31, 2904–2936.

[10] Russell Lenth, University of Iowa, http://homepage.stat.uiowa.edu/ rlenth/Power/

Acknowledgements

I would like to thank my co-authors

- Melanie Bell, Mel and Enid Zuckerman College of Public Health, University of Arizona
- Warren May, Department of Medicine, University of Mississippi Medical Center

Thank You!

Questions?

email: j.olivier@unsw.edu.au
blog: injurystats.wordpress.com