Relative Effect Sizes for Measures of Risk

Jake Olivier, Melanie Bell, Warren May

November 2015
Motivating Examples

Effect Size Phi and Relative Risk

Extensions to Other Measures

Discussion
What is a small, medium or large odds ratio?

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Dead</th>
<th>Alive</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ileostomy</td>
<td>4</td>
<td>19</td>
<td>17%</td>
</tr>
<tr>
<td>Ileoproctostomy</td>
<td>2</td>
<td>35</td>
<td>5%</td>
</tr>
</tbody>
</table>

- 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 $OR = 3.68^{1}$

---

What is a small, medium or large odds ratio?

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Dead</th>
<th>Alive</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ileostomy</td>
<td>4</td>
<td>19</td>
<td>17%</td>
</tr>
<tr>
<td>Ileoproctostomy</td>
<td>2</td>
<td>35</td>
<td>5%</td>
</tr>
</tbody>
</table>

- 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 $OR = 3.68^1$
- $p = 0.132$

---

What is a small, medium or large odds ratio?

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Dead</th>
<th>Alive</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ileostomy</td>
<td>4</td>
<td>19</td>
<td>17%</td>
</tr>
<tr>
<td>Ileoproctostomy</td>
<td>2</td>
<td>35</td>
<td>5%</td>
</tr>
</tbody>
</table>

- 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 $OR = 3.68$
- $p = 0.132$
- Is this result unimportant? Would you choose one method over the other if you were the patient?

---

What is a small, medium or large odds ratio?

<table>
<thead>
<tr>
<th>Helmet</th>
<th>Alcohol</th>
<th>% Alcohol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>No</td>
<td>18</td>
<td>125</td>
</tr>
</tbody>
</table>

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’

What is a small, medium or large odds ratio?

<table>
<thead>
<tr>
<th>Helmet</th>
<th>Alcohol</th>
<th>% Alcohol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>No</td>
<td>18</td>
<td>125</td>
</tr>
</tbody>
</table>

- 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

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
- Vehciles overtook, on average, closer when helmeted
- Reanalyze data with passing distance categorized by one metre rule\(^4\)
  - Close overtaking increases lateral forces (\(< 1\, m \equiv \text{‘unsafe’}\))
  - \( 2 \times 2 \) table for helmet wearing vs. safe/unsafe passing distance


Motorists are more aggressive to helmeted cyclists

- Ian Walker, University of Bath
- 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
  - Close overtaking increases lateral forces ($< 1m \equiv \text{‘unsafe’}$)
  - $2 \times 2$ table for helmet wearing vs. safe/unsafe passing distance
  - $OR_{adj} = 1.13$ (adj for kerb distance, vehicle size, city)

---


Relative Effect Sizes

- Broad effect size recommendations exist for other measures\(^5\)
- Simplest case: Cohen’s \(d\) for difference in means

\[
\text{Population} \hspace{10em} \text{Sample}
\]

\[
\delta = \frac{\mu_1 - \mu_2}{\sigma} \hspace{10em} d = \frac{\bar{x}_1 - \bar{x}_2}{s}
\]

- \(d = 0.2\) (small)
- \(d = 0.5\) (medium)
- \(d = 0.8\) (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)
Cohen’s $d$ for Correlation

- Using $d$ as an \textit{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

---


Motivating Examples

Effect Size Phi and Relative Risk

Extensions to Other Measures

Discussion
# 2×2 Contingency Tables

<table>
<thead>
<tr>
<th></th>
<th>$X = 0$</th>
<th>$X = 1$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 0$</td>
<td>$\pi_{00}$</td>
<td>$\pi_{01}$</td>
<td>$\pi_{0+}$</td>
</tr>
<tr>
<td>$Y = 1$</td>
<td>$\pi_{10}$</td>
<td>$\pi_{11}$</td>
<td>$\pi_{1+}$</td>
</tr>
<tr>
<td>Total</td>
<td>$\pi_{+0}$</td>
<td>$\pi_{+1}$</td>
<td>1.0</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>$X = 0$</th>
<th>$X = 1$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 0$</td>
<td>$n_{00}$</td>
<td>$n_{01}$</td>
<td>$n_{0+}$</td>
</tr>
<tr>
<td>$Y = 1$</td>
<td>$n_{10}$</td>
<td>$n_{11}$</td>
<td>$n_{1+}$</td>
</tr>
<tr>
<td>Total</td>
<td>$n_{+0}$</td>
<td>$n_{+1}$</td>
<td>$n$</td>
</tr>
</tbody>
</table>
Effect Size for $2 \times 2$ Tables

- Pearson’s correlation coefficient

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

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

- Cohen gives recommendations of

- $\phi = 0.1$ (small)
- $\phi = 0.3$ (medium)
- $\phi = 0.5$ (large)
Relative Risk

- For $2 \times 2$ 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 $\pi_{11}$ in terms of the marginal probabilities and $\phi$

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

- So, clearly the relative risk can be written as a function of $\phi$ and the marginal probabilities only
Consequence of Fixed Margins

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

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

- This also constrains the relative risk when transforming from $\phi$

- For example, when $\pi_{+1} = 0.5$ and $\pi_{1+} = 0.1$,

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

- 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., $\alpha = \phi / \phi_{\text{max}}$

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

- 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 $\alpha$ and $\pi + 1$ can be used.
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 $\alpha$ and $\pi + 1$ can be used
Motivating Examples

Effect Size Phi and Relative Risk

Extensions to Other Measures

Discussion
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 \(\pi_{1+}\), 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\(^8\).

“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.

---

“Practical” Odds Ratio Effect Sizes

▶ For planning purposes or when the event is rare \(\pi_{1+} < 0.10\) 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

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

▶ Average odds ratio

\[
\bar{OR} = \frac{1}{\pi_{1+}} \int_0^{\pi_{1+}} OR_\alpha d\pi_{1+} = 1 - \frac{\log((1 - \pi_{1+})(1 - \alpha))}{\pi_{1+}(1 - \alpha) + \alpha}(RR_\alpha - 1)
\]

“Practical” Odds Ratio Effect Sizes

<table>
<thead>
<tr>
<th></th>
<th>$\alpha$</th>
<th>$OR_{min}$</th>
<th>$OR_{mid}$</th>
<th>$OR$</th>
<th>$OR_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.1</td>
<td>1.22</td>
<td>1.31</td>
<td>1.32</td>
<td>1.49</td>
</tr>
<tr>
<td>Medium</td>
<td>0.3</td>
<td>1.86</td>
<td>2.27</td>
<td>2.38</td>
<td>3.45</td>
</tr>
<tr>
<td>Large</td>
<td>0.5</td>
<td>3.00</td>
<td>4.20</td>
<td>4.70</td>
<td>9.00</td>
</tr>
</tbody>
</table>
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 \( \Rightarrow \) 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 \( \implies \) 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

<table>
<thead>
<tr>
<th>Study</th>
<th>OR</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payne (2007)</td>
<td>3.68</td>
<td>Large</td>
</tr>
<tr>
<td>Heng (2006)</td>
<td>$1/0.19=5.26$</td>
<td>Large</td>
</tr>
<tr>
<td>Walker (2007)</td>
<td>1.13</td>
<td>Trivial?</td>
</tr>
</tbody>
</table>
## Revisit Examples

<table>
<thead>
<tr>
<th>Study</th>
<th>OR</th>
<th>Effect Size</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payne (2007)</td>
<td>3.68</td>
<td>Large</td>
<td>60</td>
</tr>
<tr>
<td>Heng (2006)</td>
<td>$1/0.19=5.26$</td>
<td>Large</td>
<td>160</td>
</tr>
<tr>
<td>Walker (2007)</td>
<td>1.13</td>
<td><em>Trivial?</em></td>
<td>2355</td>
</tr>
</tbody>
</table>
Motivating Examples

Effect Size Phi and Relative Risk

Extensions to Other Measures

Discussion
Discussion

- Binary correlation coefficient (and equivalences) are unusable as effect size measures
- Relative risk and odds ratio are not constrained by $\phi_{\text{max}}$ 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

---


¹⁰Russell 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