Against all Odds

Michael P. LaValley
Boston University
http://people.bu.edu/mlava/

ACR/ARHP Annual Scientific Meeting
San Antonio Texas 10/20/2004
Outline

- Measures of Risk
  - Relative Risk (RR)
  - Odds Ratio (OR)
- Obtaining Adjusted RR
  - Logistic Regression with Transformation
  - Binomial Regression
  - Modified Poisson Regression
- Summary
Measures of Risk

- Measure relationship between two binary variables
  - Binary variables: 0/1 or No/Yes
- Usually expressed as
  - At how much greater risk of X is one group than another?
- Example
  - At how much greater risk of osteoarthritis (OA) are women than men?
Measures of Risk

- Often want to adjust for differences between groups in other factors
  - Remove the effects of the other factors from the group difference
- Example
  - At how much greater risk of OA are women than men after adjusting for age and body mass index?
Relative Risk

**Relative Risk (RR)**

- Ratio of the probabilities of the occurrence of the outcome of interest in group 1 to group 2

\[
RR = \frac{Pr_1}{Pr_2}
\]

- \(Pr_1\) is the probability of the outcome in group 1
- \(Pr_2\) is the probability of the outcome in group 2
Odds

- Odds (used in odds ratio)
  - Odds are the probability of occurrence divided by the probability of non-occurrence

\[ Odds_1 = \frac{Pr_1}{1 - Pr_1} \]

- Odds_2 defined using Pr_2
Odds

- Odds are used in gambling
  - ‘The odds are two to one for Seabiscuit to win’
  - $2:1 \rightarrow \text{odds} = 2 \rightarrow \text{Pr} = 0.67$

- Translating odds to probabilities
  - Odds = 3.0 $\leftrightarrow$ Pr = 0.75
  - Odds = 2.0 $\leftrightarrow$ Pr = 0.67
  - Odds = 1.0 $\leftrightarrow$ Pr = 0.50
  - Odds = 0.5 $\leftrightarrow$ Pr = 0.33
Odds Ratio

- **Odds Ratio (OR)**
  - Ratio of the odds of the occurrence of the event of interest in group 1 to group 2

\[
OR = \frac{Odds_1}{Odds_2} = \frac{\frac{Pr_1}{(1-Pr_1)}}{\frac{Pr_2}{(1-Pr_2)}}
\]
RR and OR Comparison

- RR and OR are *ratio* measures
  - 1.0 is the point of no difference between groups (the null value)
  - Are greater than 1 if group 1 is at increased risk relative to group 2
  - Are less than 1 if group 1 is at decreased risk relative to group 2
  - Reciprocals are the same distance from the null value
    - E.g. 2 and $\frac{1}{2}$ are equivalent group differences
The RR is more understandable
  - When the RR=2 then the probability of the outcome in group 1 is twice that of group 2
  - This is not true for the odds ratio
Most people are more comfortable with probabilities or percentages that with odds
RR and OR Comparison

- However, the OR has some advantages
  - In case-control studies the OR can be estimated but not the RR
  - The OR is symmetric to which outcome level is chosen as being of interest, the RR is not
RR and OR Comparison

- When are the RR and OR Similar?
  - If the probability of the event is small, the odds and the probability are close
    \[ \text{Odds}_1 = \frac{\Pr_1}{1 - \Pr_1} \approx \Pr_1 \]
  - When the probability of the event is small in both the OR is a good approximation to the RR
  - Rule of thumb for small: \( \Pr < 0.1 \)
RR and OR Comparison

- The OR is always more extreme (farther from 1) than the RR.
- When the events of interest are common, the OR can be much larger than the RR.
Which is Better?

- For case-control studies need to present the OR
- For cohort studies and clinical trials the RR is better to report
  - Reduces the chance of incorrect interpretation
  - Becoming preferred to report RR in medical journals
Osteoarthritis in Framingham

- In the Framingham Osteoarthritis study, prevalence of osteoarthritis (OA) was measured in 1992-93
- Female sex is an established risk factor for OA
- At how much greater risk of osteoarthritis are women than men in this study?
Osteoarthritis in Framingham

- Subset of 840 subjects to evaluate the prevalence of OA in women versus men
  - 538 women
  - 302 men
  - 513 (61%) no OA
  - 327 (39%) with OA

<table>
<thead>
<tr>
<th></th>
<th>No OA</th>
<th>OA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>316</td>
<td>222</td>
<td>538</td>
</tr>
<tr>
<td>Men</td>
<td>197</td>
<td>105</td>
<td>302</td>
</tr>
<tr>
<td>Total</td>
<td>513</td>
<td>327</td>
<td>840</td>
</tr>
</tbody>
</table>
### Osteoarthritis in Framingham

- In men
  - 197 (65%) no OA
  - 105 (35%) with OA
- In women
  - 316 (59%) no OA
  - 222 (41%) with OA
- RR = 0.41/0.35 = 1.19
- OR = 1.32

<table>
<thead>
<tr>
<th></th>
<th>No OA</th>
<th>OA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>316</td>
<td>222</td>
<td>538</td>
</tr>
<tr>
<td>Men</td>
<td>197</td>
<td>105</td>
<td>302</td>
</tr>
<tr>
<td>Total</td>
<td>513</td>
<td>327</td>
<td>840</td>
</tr>
</tbody>
</table>
Osteoarthritis in Framingham

- Women have 1.19 times the risk of OA compared to men.
- Women have 1.30 times the odds of OA compared to men.
- If we interpret OR as an RR, we would mistakenly conclude women are at 1.3 times the risk of OA.
Suppose we look at No OA as the outcome

- RR for No OA is $0.59/0.65 = 0.91$
- But RR for OA is 1.19 and $1/1.19 = 0.84$

The RR implies that sex plays a larger role for OA than for No OA!
Osteoarthritis in Framingham

- RR is not symmetric around the null value for both outcome levels
  - RR for No OA ≠ 1/RR for OA
- OR is symmetric
  - OR for No OA = 1/OR for OA
- Usually the outcome to choose is clear and this isn’t a problem. But some situations aren’t clear
  - E.g. use ‘lived’ or ‘died’?
Adjusted RR

- Logistic regression provides adjusted OR
- But, until recently it has been difficult to obtain adjusted RR
- Three methods for getting adjusted RR
  - Logistic regression with transformation
  - Binomial regression
  - Modified Poisson regression
Logistic Regression

- Logistic regression is widely used regression method for binary outcomes
- Logistic regression coefficients are log(OR)
- Provides adjusted OR if adjustors are used as additional predictors
Logistic Regression

- If outcome probabilities are < 0.1 for all values of the predictors then the OR are good approximations to RR
- Otherwise Zhang and Yu proposed a formula to convert OR to RR

\[ RR = \frac{OR}{(1 - Pr_2) + (Pr_2 \times OR)} \]
Logistic Regression

- However the conversion formula has been criticized*
  - Leads to confidence intervals for RR that are too small
  - Gives biased estimate if some regression predictors are confounders
  - Doesn’t work if there are interactions in the regression model

*See McNutt et al.
Binomial Regression

- Binomial regression is a rarely used regression method for binary outcomes
- Binomial regression coefficients are $\log(\text{RR})$
- Provides adjusted RR if adjustors are used as additional predictors
Binomial Regression

- This model often fails due to numerical problems
- Especially failure prone if
  - Correlated predictors
  - One or more continuous predictors
Modified Poisson Regression

- Poisson regression is a method for count outcomes
  - Count outcomes: 0, 1, 2, 3, ….
- Poisson regression coefficients are log(RR)
- Provides adjusted RR if adjustors are used as additional predictors
- Poisson regression is conservative for binary outcomes
  - Less likely to be significant
  - Confidence intervals too wide
Modified Poisson Regression

- Modification due to Zou
  - Adjust variability with generalized estimating equations (GEE)
  - Uses variability in the data to adjust model
- This has been shown to work very well
- Software implementation
  - SAS in Lundquist
  - STATA in Barros and Hirakata
Osteoarthritis in Framingham

- We found greater risk of OA in women than men
  - Could this be due to age differences between women and men?
  - Could this be due to differences in body mass index between women and men?
- Use regression models with sex, age, and body mass index
Osteoarthritis in Framingham

- After adjustment for age and body mass index
  - Logistic OR = 1.45
  - Transformed Logistic RR = 1.25
  - Binomial RR = 1.20*
  - Modified Poisson RR = 1.23

*Failed to arrive at final estimate
Adjusted Effect of Sex on OA
Framingham Study

Relative Risk (95%CI)
Odds Ratio (95% CI)
Summary

- Medical literature is moving toward reporting RR instead of OR whenever possible
  - Need to keep in mind that the RR changes in non-intuitive ways when outcomes are switched
- When reporting OR make it clear that it is not the RR
- Modified Poisson regression will become standard method for obtaining adjusted RR
References

- Simon SD. Understanding the odds ratio and the relative risk. *Journal of Andrology* **22**:533-6, 2001
References

- Zou G. A modified poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology* **159**:702-6, 2004