

Relative Risk and Odds Ratios

- In lecture 4, we discussed hypothesis testing and confidence intervals for comparing two binomial proportions.
- In medical studies, results for such data are frequently expressed in terms of relative risk or odds ratios, depending on whether the study is prospective or retrospective.
- One example we discussed is the following:

Example: To test the effectiveness of a new pain-relieving drug, 80 patients at a clinic were given a pill containing a drug and 80 others were given a placebo. In the first group, 56 of the patients showed improvement. In the second group, 38 of the patients showed improvement. How effective is the drug compared to the placebo?

- This is an example of a **prospective randomized trial**.
- The data can be presented in a 2 by 2 table

	Group	
Improved	Drug	Placebo
No	24	42
Yes	56	38
Total	80	80

- Note that the drug/placebo (exposure) totals are fixed in this study, while the improved yes/no (disease status) totals are not.

- A second example follows: Rosenberg et al (1980) studied the relationship between coffee drinking and myocardial infarction (MI) in women aged 30-49 years. 487 hospitalized cases of MI and 980 controls hospitalized for an acute condition (trauma, acute cholecystitis, acute respiratory diseases and appendicitis) were asked about their coffee consumption, with the following results

	Cups per Day		Total
	≥ 5	< 5	
MI	152	335	487
Control	183	797	980

- This is a **retrospective case-control study**.
- Note that the MI/Control (disease status) totals are fixed by the study design, and the Cups per Day (exposure) totals are not.

Relative Risk

- In a **prospective study** we gain information about the prevalence of the disease in the the two (or more) groups.
- The ratio of the probability of a disease in two groups, often denoted 'exposed' and 'unexposed' is the *relative risk*.
- If D^+ and D^- represent presence and absence of the disease, and E^+ and E^- represent presence and absence of the exposure, then

$$RR = \frac{P(D^+|E^+)}{P(D^+|E^-)}$$

- If exposure has no effect, $RR = 1$.
- If prevalence higher in the exposed group, $RR > 1$.
- If prevalence lower in the exposed group, $RR < 1$.
- For a 2 by 2 table of counts, with row totals not fixed by the study design,

		Exposure		Total
		+	-	
Disease	+	a	b	a+b
	-	c	d	c+d
Total		a+c	b+d	

RR is estimated by

$$RR = \frac{a/(a+c)}{b/(b+d)}$$

Example: In the pain example, we can equate pain with D^+ , so

$$RR = \frac{24/80}{42/80} = \frac{24}{42} = .5714$$

Another example: The probability that a man over 35 dies of lung cancer in the US is .002679 for current smokers, and .000154 for nonsmokers. The relative risk of death for smokers versus nonsmokers is

$$RR =$$

Confidence interval for Relative Risk

- We can calculate a confidence interval for the risk ratio.
- Because the distribution of risk ratios is skewed to the right, we calculate the interval on the (natural) log scale and then transform using exponents to the original scale.
- The standard error of $\ln(\widehat{RR})$ is estimated by

$$\hat{se}(\ln(\widehat{RR})) = \sqrt{\frac{1}{a} - \frac{1}{a+c} + \frac{1}{b} - \frac{1}{b+d}}$$

- So the $(1 - \alpha)100\%$ confidence interval on the log scale is

$$\ln(\widehat{RR}) \pm z_{\alpha/2} \hat{se}(\ln(\widehat{RR})).$$

- We raise e to each of these powers to get the confidence interval for the relative risk.

Example: In the pain example,

- The standard error for $\ln(\widehat{RR})$ is

$$\hat{se}(\ln(\widehat{RR})) = \sqrt{\frac{1}{24} - \frac{1}{80} + \frac{1}{42} - \frac{1}{80}} = .2012$$

- The 95% confidence interval for $\ln(\widehat{RR})$ is

$$\begin{aligned} \ln(.5714) &\pm 1.96(.2012) \\ -.5596 &\pm .3943 \\ (-.9539 &, -.1653) \end{aligned}$$

- The 95% confidence interval for \widehat{RR} is

$$\begin{aligned} (\exp(-.9539) &, \exp(-.1653)) \\ (.3852 &, .8476) \end{aligned}$$

- Note that this interval does not include 1, which indicates significant evidence (at the .05 level) against the null hypothesis of no difference in pain reduction between the drug and placebo group.

Odds Ratios

- In **retrospective studies**, samples are taken for those affected and unaffected by a disease, and their exposure to a factor recorded.
- There is therefore no information about disease prevalence in the data.
- Instead, the comparison is between the rates of exposure in the affected and unaffected groups.
- Once again, the data can be arranged in a 2 by 2 table

		Exposure		Total
		+	-	
Disease	+	a	b	a+b
	-	c	d	c+d
Total		a+c	b+d	

- In this case, the Disease group totals (row totals) are fixed by the design.
- The *odds* for an event is the ratio of the probability it occurs to the probability that it doesn't.
- For example, the odds for exposure are

$$P(E^+) / (1 - P(E^+)) = P(E^+) / P(E^-)$$

- For the 2 by 2 table of counts, the odds for exposure in the Disease + group is

$$\frac{a/(a+b)}{b/(a+b)} = \frac{a}{b}$$

- The *odds ratio* is the ratio of the odds in two groups.

$$OR = \frac{P(E^+|D^+)/P(E^-|D^+)}{P(E^+|D^-)/P(E^-|D^-)} = \frac{P(D^+|E^+)/P(D^-|E^+)}{P(D^+|E^-)/P(D^-|E^-)}$$

- If the odds are the same in the two groups $OR = 1$.
- If the odds are higher in the Disease+ group, $OR > 1$.
- If the odds are lower in the Disease- group. $OR < 1$.
- For the 2 by 2 table, the odds ratio is estimated by

$$OR = \frac{a/b}{c/d} = \frac{ad}{bc}$$

Example: The following table resulted from a study on breast cancer and oral contraceptive use

		OC use		Total
		+	-	
Cancer	+	273	716	989
	-	2641	7260	9901
Total		2914	7976	

- The disease margin was fixed in this study (prevalence not 989/9901).
- The odds ratio is estimated as

$$OR = \frac{273(7260)}{716(2641)} = 1.0481$$

- Women who used oral contraceptives have an odds of developing breast cancer that is only 1.05 times of the nonusers.

Confidence interval for the Odds Ratio

- Once again the distribution of the odds ratio is skewed to the right, so we calculate the confidence interval on the log scale and then transform back.
- The standard error for $\ln(\hat{OR})$ is

$$\hat{se}(\ln(\widehat{OR})) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

- So the $(1 - \alpha)100\%$ confidence interval on the log scale is

$$\ln(\widehat{OR}) \pm z_{\alpha/2} \hat{se}(\ln(\widehat{OR}))$$

- We raise e to each of these powers to get the confidence interval for the odds ratio.

Example: For the breast cancer/OC use data:

- $\ln(\widehat{OR}) = \ln(1.0481) = .0470$
- The standard error for $\ln(\widehat{OR})$ is

$$\begin{aligned} se(\widehat{OR}) &= \sqrt{\frac{1}{273} + \frac{1}{716} + \frac{1}{2641} + \frac{1}{7260}} \\ &= .0747 \end{aligned}$$

- The 95% confidence interval for $\ln(\widehat{OR})$ is

$$\begin{aligned} &.0470 \pm 1.96(.0747) \\ &.0470 \pm .1464 \\ &(-.0993 \quad , \quad .1934) \end{aligned}$$

- The 95% confidence interval for \widehat{OR} is

$$\begin{aligned} &(exp(-.0993) \quad , \quad exp(-.1934)) \\ &(.9054 \quad , \quad 1.2133) \end{aligned}$$

- This confidence interval includes one, which indicates there is no significant difference between the groups.

Cross-sectional studies

- In a cross-sectional study, data is collected on disease status and exposure on a group of subjects.
- The data can be summarized as before

		Exposure		Total
		+	-	
Disease	+	a	b	a+b
	-	c	d	c+d
Total		a+c	b+d	

- This time, neither margin is fixed by the design.
- We can calculate both the RR and the OR.
- The relative risk for the disease is

$$RR = \frac{a/(a+c)}{b/(b+d)}$$

- The odds ratio for disease is

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc}$$

- The odds ratio for exposure is

$$OR = \frac{a/b}{c/d} = \frac{ad}{bc}$$

- These are the same!!
- Note that if the disease is rare, then a is much smaller than c and b is much smaller than d .
- In this case we can remove the a and b from the two denominators in RR, giving

$$RR \approx \frac{a/c}{b/d} = \frac{ad}{bc}$$

which is the odds ratio!

- So for rare diseases, RR and OR are close.
- Also the OR for disease is the same as the OR for exposure.
- So the OR for exposure calculated from a retrospective case-control study is an estimate of the RR for disease.