

# Lecture 3: Measures of effect: Risk Difference Attributable Fraction Risk Ratio and Odds Ratio

Dankmar Böhning

Southampton Statistical Sciences Research Institute  
University of Southampton, UK

March 2 - 4, 2015

## Measures of differences in disease occurrence

### Risk difference

### Choosing an Effect Measure and Public Health Impact

### Attributable Fraction

### Risk Ratio

### Odds Ratio

### Odds ratio and study type

We have seen earlier how to measure diseases and their distributions using prevalence and incidence.

Now we are concerned with differences in disease occurrence in different populations.

Common measures are

1. risk difference (RD)
2. relative risk difference or attributable fraction (AF)
3. risk ratio (RR)
4. odds ratio (OR)

In this lecture we will look at the first two.

The risk ratio and odds ratio will be covered in the next lecture.

The **Risk Difference** (RD) is the difference between disease risk in an **exposed** population and risk in an **non-exposed** population.

Let  $p_1$  = disease risk in an **exposed** population

$p_0$  = disease risk in an **non-exposed** population.

$$RD = p_1 - p_0$$

$RD$  is a number between -1 and 1.

### Example 1

In a study of two toothpastes, 10 out of 100 caries-free children using a new toothpaste (exposure) develop caries after 1 year. In another group of 100 caries-free children using a standard toothpaste, 25 develop caries.

$$\widehat{RD} = \frac{10}{100} - \frac{25}{100} = -0.15$$

## Example 2

In a group of 1000 persons with heavy sun-exposure, there are 40 cases of skin cancer. In a comparative, equally sized, non-exposed group there are 10 cases of skin cancer.

$$\widehat{RD} = \frac{40}{1000} - \frac{10}{1000} = 0.03$$

## Exercise 1

In a cohort study evaluating radiation exposures, 52 tumours developed among 2872 exposed individuals and 6 tumours developed among 5049 unexposed individuals within the observation period.

What is the risk difference?

$$\widehat{RD} = \hat{p}_1 - \hat{p}_0 =$$

## Distribution of number of diseased

Suppose that in a cohort study,  
 $Y_1$  out of  $n_1$  exposed individuals and  
 $Y_0$  out of  $n_0$  non-exposed individuals  
developed the disease.

Assume that the probability  $p_1$  of developing the disease is the **same** for everyone in the exposed group

Similarly, assume that the probability  $p_0$  of developing the disease is the **same** for everyone in the non-exposed group

Then  $Y_1 \sim B(n_1, p_1)$  distribution

And  $Y_0 \sim B(n_0, p_0)$  distribution

## Variance of RD

A reasonable estimate for the RD is

$$\widehat{RD} = \hat{p}_1 - \hat{p}_0 = \frac{Y_1}{n_1} - \frac{Y_0}{n_0}$$

From which we get,

$$\begin{aligned} \text{Var}(\widehat{RD}) &= \text{Var}\left(\frac{Y_1}{n_1} - \frac{Y_0}{n_0}\right) \\ &= \text{Var}\left(\frac{Y_1}{n_1}\right) + \text{Var}\left(\frac{Y_0}{n_0}\right) \end{aligned}$$

and since both  $Y_1$  and  $Y_2$  follow binomial distributions,

$$\text{Var}(\widehat{RD}) = \frac{p_1(1-p_1)}{n_1} + \frac{p_0(1-p_0)}{n_0}$$

## A confidence interval for RD

$$SD(\widehat{RD}) = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_0(1-p_0)}{n_0}}$$

Estimating  $p_1$  and  $p_0$  by  $\hat{p}_1 = Y_1/n_1$  and  $\hat{p}_0 = Y_0/n_0$

A 95% confidence interval for RD is

$$\begin{aligned} & \widehat{RD} \pm 2SD(\widehat{RD}) \\ &= \widehat{RD} \pm 2\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_0}} \end{aligned}$$



## Example 1 (revisited)

Here we had that 10 children out of 100 using a new toothpaste developed caries while 25 out of 100 using the standard toothpaste developed caries.

The estimated RD was shown to be  $\widehat{RD} = \frac{10}{100} - \frac{25}{100} = -0.15$

A 95% CI for RD is  $\widehat{RD} \pm 2SD(\widehat{RD})$

$$\begin{aligned}
 &= \widehat{RD} \pm 2\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_0(1 - \hat{p}_0)}{n_0}} \\
 &= -0.15 \pm 2\sqrt{\frac{0.1(1 - 0.1)}{100} + \frac{.25(1 - 0.25)}{100}} \\
 &= -0.15 \pm 2\sqrt{0.002775} \\
 &= -0.15 \pm 2 \times 0.0526783 = (-0.255, -0.045)
 \end{aligned}$$

## Exercise 1 (revisited)

Here we had a cohort study on radiation exposure where 52 tumours developed among 2872 exposed and 6 tumours developed among 5049 unexposed individuals.

The risk difference was  $\widehat{RD} = \hat{p}_1 - \hat{p}_0 =$

A 95% CI for the risk difference is:

$$\widehat{RD} \pm 2 \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_0(1 - \hat{p}_0)}{n_0}}$$

=

**Interpretation:**

# Choosing an Effect Measure and Public Health Impact

## An illustration

case A):  $p_1 = 1/10$ ,  $p_0 = 1/100$

$$RR = 10, RD = 0.1 - 0.01 = 0.09 \sim 0.1$$

case B):  $p_1 = 1/100$ ,  $p_0 = 1/1000$

$$RR = 10, RD = 0.01 - 0.001 = 0.009 \sim 0.01$$

the risk ratio is independent of the absolute disease occurrence,  
whereas the risk difference is containing this!

## NNT or NNE

the **number needed to be exposed** *NNE* is defined as the number of people needed to be exposed to get **one case more** than under non-exposure:

$$NNE = \frac{1}{p_1 - p_0}$$

## motivation:

$np_1$  expected number of cases under exposure

$np_0$  expected number of cases under non-exposure

$$np_1 - np_0 = 1?$$

$$n(= NNE) = \frac{1}{p_1 - p_0}$$

## illustration

case A):  $p_1 = 1/10$ ,  $p_0 = 1/100$

$$RR = 10, RD = 0.1 - 0.01 = 0.09 \sim 0.1$$

$$NNE = 10$$

case B):  $p_1 = 1/100$ ,  $p_0 = 1/1000$

$$RR = 10, RD = 0.01 - 0.001 = 0.009 \sim 0.01$$

$$NNE = 100$$

## interpretation:

the smaller the  $NNE$  the higher the public health impact of the exposure.

## Attributable Fraction (AF):

The attributable fraction (AF) or **relative risk difference** is a measure that **combines** RD and prevalence

**AF due to exposure:** Assume that exposure **increases** risk.

That is assume  $p_1 > p_0$ .

$$AF = \frac{RD}{p_1} = \frac{p_1 - p_0}{p_1}$$

**interpretation:** Let  $n$  be the total number of cases and controls

$$AF = \frac{np_1 - np_0}{np_1}$$

$$= \frac{(\# \text{ cases if everyone exposed}) - (\# \text{ cases if everyone non-exposed})}{\# \text{ cases if everyone exposed}}$$

$AF$  = proportion of cases due to exposure  
= proportion of avoidable cases due to exposure

### $AF$ is a relative measure:

Effects with similar risks will have similar attributable fractions.

Scenario A):  $p_1 = 1/10$ ,  $p_0 = 1/100$

$$RD = 0.1 - 0.01 = 0.09 \sim 0.1$$

$$AF = 0.09/0.1 = 0.90$$

Scenario B):  $p_1 = 1/100$ ,  $p_0 = 1/1000$

$$RD = 0.01 - 0.001 = 0.009 \sim 0.01$$

$$AF = 0.009/0.01 = 0.90$$



## Preventive fraction

If exposure **decreases** risk the preventive fraction is instead calculated:

$$\frac{p_0 - p_1}{p_0}$$

## Population attributable fraction (PAF)

This is the proportion of cases occurring in the total population which can be explained by the exposure

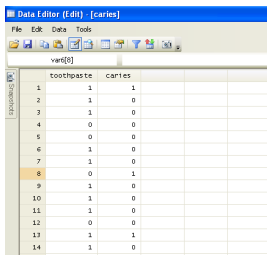
Let the proportion exposed be  $p$

$$PAF = \frac{p(p_1 - p_0)}{pp_1 + (1 - p)p_0}$$

# In STATA

## Example 1: Caries Study

Data in rectangular format:



var[0]

	toothpaste	caries
1	1	1
2	1	0
3	1	0
4	0	0
5	0	0
6	1	0
7	1	0
8	0	1
9	1	0
10	1	0
11	1	0
12	0	0
13	1	1
14	1	0

cs caries toothpaste

	toothpaste Exposed	unexposed	Total
Cases	10	25	35
NonCases	90	75	165
Total	100	100	200
Risk	.1	.25	.175
Point estimate	[95% Conf. Interval]		
Risk difference	-.15	-.2532475	-.0467525
Risk ratio	.4	.2028594	.7887236
Prev. frac. ex.	.6	.2112764	.7971406
Prev. frac. pop	.3		

ch12(1) = 7.79 Pr>chi2 = 0.0052

csi 10 25 90 75

## Risk ratio (RR):

The risk ratio or **relative risk** is the ratio of disease risk in an **exposed** to disease risk in an **non-exposed** population.

$$RR = \frac{p_1}{p_0}$$

where  $p_1$  is disease risk in **exposed** and  $p_0$  is disease risk in **non-exposed** population.

- $RR$  is a number between 0 and  $\infty$ .

## Interpretation:

For example,  $RR=2$  means that disease occurrence is 2 times more likely in exposure group than in non-exposure group.

$RR=1$  means **no effect** of exposure.

## Example 1

In a study of two toothpastes, 10 out of 100 caries-free children using a new toothpaste (exposure) develop caries after 1 year. In another group of 100 caries-free children using a standard toothpaste, 25 develop caries.

$$\widehat{RR} = \frac{10}{100} / \frac{25}{100} = 0.40$$

## Example 2

In a group of 1000 persons with heavy sun-exposure, there are 40 cases of skin cancer. In a comparative, equally sized, non-exposed group there are 10 cases of skin cancer.

$$\widehat{RR} = \frac{40}{1000} / \frac{10}{1000} = 40$$

## Exercise 1

In a cohort study evaluating radiation exposures, 52 tumours developed among 2872 exposed individuals and 6 tumours developed among 5049 unexposed individuals within the observation period.

What is the risk ratio?

$$\widehat{RR} = \frac{\hat{p}_1}{\hat{p}_0} =$$

## Estimator of RR

Suppose that in a cohort study,  
 $Y_1$  out of  $n_1$  exposed individuals and  
 $Y_0$  out of  $n_0$  non-exposed individuals  
developed the disease.

Assume that the probability  $p_1$  of developing the disease is the **same** for everyone in the exposed group

Similarly, assume that the probability  $p_0$  of developing the disease is the **same** for everyone in the non-exposed group

Then a plausible **estimator of the risk ratio** is

$$\widehat{RR} = \frac{\frac{Y_1}{n_1}}{\frac{Y_0}{n_0}} = \frac{Y_1 n_0}{Y_0 n_1}$$

## Variance of RR

Technically it is easier to work with the logarithm of the risk ratio.

$$\log(RR) = \log(p_1) - \log(p_0)$$

Applying the  **$\delta$  method**, an approximate variance is

$$\begin{aligned} \text{Var}(\widehat{\log RR}) &= \begin{pmatrix} \frac{1}{p_1} & \frac{1}{p_0} \end{pmatrix} \begin{pmatrix} \text{Var}(\hat{p}_1) & 0 \\ 0 & \text{Var}(\hat{p}_0) \end{pmatrix} \begin{pmatrix} \frac{1}{p_1} \\ \frac{1}{p_0} \end{pmatrix} \\ &= \frac{1}{p_1^2} \frac{p_1(1-p_1)}{n_1} + \frac{1}{p_0^2} \frac{p_0(1-p_0)}{n_0} \end{aligned}$$

Estimating  $p_1$  by  $Y_1/n_1$  and  $p_0$  by  $Y_0/n_0$  and simplifying, we get

$$\text{Var}(\widehat{\log RR}) = \frac{1}{Y_1} - \frac{1}{n_1} + \frac{1}{Y_0} - \frac{1}{n_0}$$

## A confidence interval for RR

$$SD(\widehat{\log RR}) = \sqrt{\frac{1}{Y_1} - \frac{1}{n_1} + \frac{1}{Y_0} - \frac{1}{n_0}}$$

**Consequently**, a 95% confidence interval for the **log relative risk** is

$$\begin{aligned} & \widehat{\log RR} \pm 2SD(\widehat{\log RR}) \\ &= \widehat{\log RR} \pm 2\sqrt{\frac{1}{Y_1} - \frac{1}{n_1} + \frac{1}{Y_0} - \frac{1}{n_0}} \end{aligned}$$

and back on the **relative risk scale**, a 95% CI for  $RR$  is

$$\exp\left(\widehat{\log RR} \pm 2\sqrt{\frac{1}{Y_1} - \frac{1}{n_1} + \frac{1}{Y_0} - \frac{1}{n_0}}\right)$$



## Example 1 (revisited)

Here we had that 10 children out of 100 using a new toothpaste developed caries while 25 out of 100 using the standard toothpaste developed caries.

The estimated RR was shown to be

$$\widehat{RR} = \frac{10}{100} / \frac{25}{100} = 0.4$$

A 95%CI for  $\log(RR)$  is

$$\begin{aligned} & \widehat{\log RR} \pm 2\sqrt{\frac{1}{Y_1} - \frac{1}{n_1} + \frac{1}{Y_0} - \frac{1}{n_0}} \\ &= \log 0.4 \pm 2\sqrt{\frac{1}{10} - \frac{1}{100} + \frac{1}{25} - \frac{1}{100}} \end{aligned}$$

$$\begin{aligned} &= -0.92 \pm 2\sqrt{0.12} \\ &= -0.92 \pm 2 \times 0.3464 = (-1.6128, -0.2272) \end{aligned}$$

Hence a 95%CI for the **risk ratio** is

$$(\exp(-1.6128), \exp(-0.2272)) = (0.1993, 0.7968)$$

This shows that the new toothpaste **significantly** reduces the risk of developing caries.

## Exercise 1 (revisited)

Here we had a cohort study on radiation exposure where 52 tumours developed among 2872 exposed and 6 tumours developed among 5049 unexposed individuals.

The risk ratio was  $\widehat{RR} = \frac{\hat{p}_1}{\hat{p}_0}$

A 95% CI for RR is:

**Interpretation:**

## AF and RR:

Assume that  $p_1 > p_0$ :

$$\begin{aligned} AF &= RD/p_1 = \frac{p_1 - p_0}{p_1} \\ &= 1 - \frac{p_0}{p_1} \\ &= 1 - \frac{1}{RR} \end{aligned}$$

Hence an **estimate of AF is available if an estimate of RR** is available.

## Estimator of $RR$ for person-time data

Suppose that in a cohort study people are under risk with different person-times

$Y_1$  events in  $T_1$  person time units in the exposed group

and

$Y_0$  events in  $T_0$  person time units in the non-exposed group

Assume that the probability  $p_1$  ( $p_0$ ) of developing the disease is the **same** for everyone in the exposed (non-exposed) group

Then a plausible **estimator of the risk ratio** is ratio of the incidence densities

$$\widehat{RR} = \frac{\frac{Y_1}{T_1}}{\frac{Y_0}{T_0}} = \frac{Y_1 T_0}{Y_0 T_1}$$

## Example

A cohort study is conducted to evaluate the relationship between dietary fat intake and the development in prostate cancer in men. In the study, 100 men with high fat diet are compared with 100 men who are on low fat diet. Both groups start at age 65 and are followed for 10 years. During the follow-up period, 10 men in the high fat intake group are diagnosed with prostate cancer and 5 men in the low fat intake group develop prostate cancer. The incidence density is  $\widehat{ID} = 10/(1,000) = 0.01$  in the high fat intake group and  $\widehat{ID} = 5/(1,000) = 0.005$  in the low fat intake group.

Hence

$$\widehat{RR} = \frac{Y_1/T_1}{Y_0/T_0} = 0.01/0.005 = 2$$

# Lecture 3: Measures of effect: Risk Difference Attributable Fraction Risk Ratio and Odds Ratio

## Risk Ratio

Review  
# Command \_irc  
1 cci 66 27 14 15  
2 iri 10 5 1000 1000

Variables  
Variable Label  
There are no items to show.

Properties  
Variables  
Name  
Label

File Edit Data Graphics Statistics User Window Help

Review  
# Command \_irc  
1 cci 66 27 14 15  
2 iri 10 5 1000 1000

Exposed Unexposed

Cases 10 5  
Person-time 1000 1000

☐ Calculate test-based confidence intervals  
25 Confidence level

OK Cancel Submit

Total 80  
Point estimate  
Odds ratio 2.6190  
Attr. frac. ex. .61818  
Attr. frac. pop. .43870

6557  
7228 (exact)  
1291 (exact)  
0247

. iri 10 5 1000 1000

	Exposed	Unexposed	Total
Cases	10	5	15
Person-time	1000	1000	2000
Incidence rate	.01	.005	.0075
Point estimate	[95% Conf. Interval]		
Inc. rate diff.	.005	-.0025909	.0125909
Inc. rate ratio	2	.6228596	7.457296 (exact)
Attr. frac. ex.	.5	-.6054984	.8659031 (exact)
Attr. frac. pop	.3333333		
(midp) Pr(k>=10) =			0.1051 (exact)
(midp) 2*Pr(k>=10) =			0.2101 (exact)

Command

\\soton.ac.uk\udel\personalfiles\users\debi1110\mydocuments

CAP NUM OVR

31 / 47

## Odds

The odds of an outcome is the number of times the outcome occurs to the number of times it does not.

Suppose that  $p$  is the probability of the outcome, then

$$odds = \frac{p}{1 - p}$$

It follows that  $p = \frac{odds}{odds+1}$

## Examples

- ▶  $p = 1/2 \Rightarrow odds = 1$
- ▶  $p = 1/4 \Rightarrow odds = 1/3$
- ▶  $p = 3/4 \Rightarrow odds = 3/1 = 3$



## Odds Ratio

$$\begin{aligned} OR &= \frac{\text{odds( in exposure )}}{\text{odds( in non-exposure )}} \\ &= \frac{p_1/(1 - p_1)}{p_0/(1 - p_0)} \end{aligned}$$

## Properties of Odds Ratio

- ▶  $0 < OR < \infty$
- ▶  $OR = 1$  if and only if  $p_1 = p_0$

## Examples

$$\text{risk} = \begin{cases} p_1 = 1/4 \\ p_0 = 1/8 \end{cases} \quad \text{effect measure} = \begin{cases} OR = \frac{p_1/(1-p_1)}{p_0/(1-p_0)} = \frac{1/3}{1/7} = 2.33 \\ RR = \frac{p_1}{p_0} = 2 \end{cases}$$

$$\text{risk} = \begin{cases} p_1 = 1/100 \\ p_0 = 1/1000 \end{cases} \quad \text{eff. meas.} = \begin{cases} OR = \frac{1/99}{1/999} = 10.09 \\ RR = \frac{p_1}{p_0} = 10 \end{cases}$$

## Fundamental Theorem of Epidemiology

$$p_0 \text{ small} \Rightarrow OR \approx RR$$

**benefit:**  $OR$  is interpretable as  $RR$  which is easier to deal with

## Example: Radiation Exposure and Tumor Development

	cases	non-cases	
E	52	2820	2872
NE	6	5043	5049

### odds and *OR*

odds for disease given exposure:

$$\frac{52/2872}{2820/2872} = 52/2820$$

odds for disease given non-exposure:

$$\frac{6/5049}{5043/5049} = 6/5043$$

## Example, cont'd

	cases	non-cases	
E	52	2820	2872
NE	6	5043	5049

odds ratio for disease :

$$OR = \frac{52/2820}{6/5043} = \frac{52 \times 5043}{6 \times 2820} = 15.49$$

or,  $\log OR = \log 15.49 = 2.74$

for comparison

$$RR = \frac{52/2872}{6/5049} = 15.24$$

	cases	non-cases
E	a	b
NE	c	d

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

CI for OR: Using

$$Var(\log OR) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

A 95% CI for  $\log OR$  is  $\log OR \pm 2\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$

As for  $RR$ , the exponent of these limits will provide the CI for  $OR$

## In STATA

## Example: Radiation Exposure and Tumor Development

Stata/IC 11.0 [Results]

File Edit Data Graphics Statistics User Window Help

Review

Command

1 logit S2 6 2820 5043, woolf

STATA (R)  
Statistics/Data Analysis 11.0

Copyright 1984-2009  
StataCorp  
4905 Lakeway Drive  
College Station, Texas 77843 USA  
800-STATA-PC <http://www.stata.com>  
979-696-4600 [stata@stata.com](mailto:stata@stata.com)  
979-696-4601 (fax)

single-user stata perpetual license:  
Serial number: 30110518370  
Licensed to: Fazil Baksh  
Reading University

NOTES:  
1. (/mh option or -set memory-) 10.00 MB allocated to data

. cci S2 6 2820 5043, woolf

	Exposed	Unexposed	Total	Proportion Exposed
Cases	52	6	58	0.8966
Controls	2820	5043	7863	0.3586
Total	2872	5049	7921	0.3626

Point estimate [95% Conf. Interval]

Odds ratio	15.49858	6.648811	36.12766 (woolf)
Attr. frac. ex.	.935478	.8495972	.9723204 (woolf)
Attr. frac. pop	.8387044		

chi2(1) = 72.08 Pr>chi2 = 0.0000

Command

NEL

CAP NUM OVR

## OR and study type cohort/CT

	case	non-case	
exposed	$p_1$	$1 - p_1$	$n_1$
non-exposed	$p_0$	$1 - p_0$	$n_0$

odds ratio for disease:

$$OR = \frac{p_1 / (1 - p_1)}{p_0 / (1 - p_0)}$$

for comparison, relative risk for disease:

$$RR = \frac{p_1}{p_0}$$

## OR and study type case-control

	case	non-case	
exposed	$q_1$	$q_0$	
non-exposed	$1 - q_1$	$1 - q_0$	
	$m_1$	$m_0$	

relative risk for disease is not estimable:

$$RR = \frac{p_1}{p_0}$$

relative risk for exposure is estimable but not of interest:

$$RR_e = \frac{q_1}{q_0}$$

since unfortunately

$$RR \neq RR_e$$



## Illustration

	case	non-case	
exposed	500	199,500	200,000
non-exposed	500	799,500	800,000
	1,000	999,000	1,000,000

relative risk for disease:

$$RR = \frac{p_1}{p_0} = 4$$

relative risk for exposure:

$$RR_e = \frac{q_1}{q_0} = \frac{5/10}{1,955/9990} = 2.5 \neq RR$$

## Odds Ratio for disease

$$OR = \frac{p_1/(1-p_1)}{p_0/(1-p_0)}$$

notations:

$$P(D|E) = p_1, P(D|NE) = p_0$$

$$P(E|D) = q_1, P(E|ND) = q_0, P(D) = p$$

then

$$\begin{aligned} p_1 = P(D|E) &= \frac{P(E|D)P(D)}{P(E)} = \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|ND)P(ND)} \\ &= \frac{q_1 p}{q_1 p + q_0(1-p)} \\ p_1/(1-p_1) &= \frac{q_1 p}{q_1 p + q_0(1-p)} / \left( 1 - \frac{q_1 p}{q_1 p + q_0(1-p)} \right) = \frac{q_1 p}{q_0(1-p)} \quad (1) \end{aligned}$$

notations:

$$P(D|E) = p_1, P(D|NE) = p_0$$

$$P(E|D) = q_1, P(E|ND) = q_0, P(D) = p$$

then

$$\begin{aligned} p_0 = P(D|NE) &= \frac{P(NE|D)P(D)}{P(NE)} = \frac{P(NE|D)P(D)}{P(NE|D)P(D) + P(NE|ND)P(ND)} \\ &= \frac{(1-q_1)p}{(1-q_1)p + (1-q_0)(1-p)} \\ p_0 / (1 - p_0) &= \frac{(1-q_1)p}{(1-q_1)p + (1-q_0)(1-p)} / \left( 1 - \frac{(1-q_1)p}{(1-q_1)p + (1-q_0)(1-p)} \right) \\ &= \frac{(1-q_1)p}{(1-q_0)(1-p)} \quad (2) \end{aligned}$$

## Odds Ratio

$$\begin{aligned} OR &= \frac{p_1/(1-p_1)}{p_0/(1-p_0)} \\ &= \frac{(1)}{(2)} \\ &= \left( \frac{q_1 p}{q_0 (1-p)} \right) / \left( \frac{(1-q_1) p}{(1-q_0) (1-p)} \right) \\ &= \frac{q_1/(1-q_1)}{q_0/(1-q_0)} = OR_e \end{aligned}$$

**disease odds ratio = exposure odds ratio**

## Illustration

	case	non-case	
exposed	500	199,500	200,000
non-exposed	500	799,500	800,000
	1,000	999,000	1,000,000

odds ratio for disease:

$$OR = \frac{5/1995}{5/7995} = 4.007$$

odds ratio for exposure:

$$OR_e = \frac{5/5}{1995/7995} = 4.007$$

also, if disease occurrence is low (low prevalence):  $OR \approx RR$

## Sun exposure and lip cancer occurrence

a case-control study was done in a population of 50-69 year old men

	case	non-case	
exposed	500	199,500	200,000
non-exposed	66	14	80
	27	15	42
	93	29	122

$$OR = \frac{66/27}{14/15} = \frac{66 \times 15}{27 \times 14}$$

## Lecture 3: Measures of effect: Risk Difference Attributable Fraction Risk Ratio and Odds Ratio

### Odds ratio and study type

Stata/IC 12.1 - [Results]

File Edit Data Graphics Statistics User Window Help

Statistics/Data Analysis

cci - Case-control studies

Exposed Unexposed

Cases 66 27

Controls 14 15

Exact confidence intervals  
 Cornfield approximation  
 Woolf approximation  
 Test-based confidence intervals

Fisher's exact p Confidence level

OK Cancel Submit

50-user Stata network per  
 Serial number: 301  
 Licensed to: iSo  
 Uni

Notes:

. cci 66 27 14 15

	Exposed	Unexposed	Total	Proportion Exposed
Cases	66	27	93	0.7097
Controls	14	15	29	0.4828
Total	80	42	122	0.6557

Point estimate [95% Conf. Interval]

Odds ratio	2.619048	1.016247	6.717228 (exact)
Attr. frac. ex.	.6181818	.0159877	.8511291 (exact)
Attr. frac. pop	.4387097		

chi2(1) = 5.04 Pr>chi2 = 0.0247

Variables

Variable Label

There are no items to show.

Properties

Variables

Name Label

Command

77845 USA  
[p://www.stata.com](http://www.stata.com)  
[ta@stata.com](mailto:ta@stata.com)

47 / 47