EPID600 (Spring 2007) module on Case-control Studies

Objectives:

- List the characteristics, advantages, and disadvantages of case control studies.
- Define the term odds ratio (OR); calculate and interpret an OR from tabular data in an article.
- Identify when the OR is a good estimate of the RR.
- Distinguish and explain advantages and disadvantages of (a) hospital-based and (b) population-based controls.
- Differentiate crude from adjusted OR's.
- Distinguish between prevalent and incidence cases.
- Explain the relation of controls in a case-control study to the study base (source population) for the cases.
- Explain why using other diseased persons as controls helps diminish recall bias in studies.
- Interpret 95% confidence intervals.

Instructions:

1. **Read**: Aschengrau and Seage, ch. 9 - Case-control studies . Answer the practice questions at the end of the chapter or at http://publichealth.jbpub.com/aschengrau/student_resources.cfm and check your answers (recommended, but optional) (animated flashcards, weblinks, and Powerpoint slides from the authors] can also be found at that URL)

2. Look over the case study questions and then read the case study reading: Fontham ETH, Correa P, Wu-Williams A, *et al.* (1991). Lung cancer in nonsmoking women: A multicenter case control study. *Cancer Epid Biomarkers & Prev.* 1:35-43. (Notes: HCFA = Health Care Financing Administration, now the Centers for Medicare & Medicaid Services; abstract, full text)

3. (Optional, but earns credit) Before lab, submit the answers to the starred case study questions (numbers 1, 3, 5, 8, and 10).

4. Read the lecture slides and attend the lecture (or read the speaker notes).

5. Work on the rest of the case study questions in **lab** and afterwards.

6. Agree on the answers, so the facilitator can submit the group's consensus answers by the following Sunday evening (EST).

Case Study Questions (NOTE: For some of these questions there may not be one "right answer".)

**1. Succinctly state the research question addressed in the study by Fontham *et al.*?

2. The authors refer to the existing body of data as suggesting that there is a "small but significant elevation in risk of lung cancer" (35, col 2). What do you think they mean by "small" and "significant" in this context?

**3. For each of the "unresolved issues" listed in the introduction, indicate why that issue is important to resolve in order to answer the research question. Could any of these account for the finding of an association between ETS and lung cancer?

4. What are advantages of the study's being population-based?

**5. What is the key information that the control group is intended to provide?

6. What are relative advantages and disadvantages of population controls versus hospital controls?

7. What population group will be unavailable for inclusion in the population-based control group in this study. Is this unavailability likely to be an important source of bias?

**8. Interviews were conducted with 84% (431 of 514) of eligible cases and 72-73% of controls. What concern does this raise about the validity of the overall results? In other words, if interviews had been conducted on 100 % of eligible cases, how different might the observed odds ratio have been from the one given in the journal article's abstract? Is it likely that the 16% non-interviewed cases and the 28% non-interviewed controls would produce serious distortion? What further data would you want to evaluate this distortion, rather than just speculate on it?

9. Explain the Fontham *et al.* statement (page 36, column 2, lines 15-16) that colon cancer controls "provided an opportunity to examine the issue of recall bias associated with a recent diagnosis of cancer."

10. Use the data on "Education" in Table 4 of the journal article to calculate the odds ratio for the association between less than high school education vs. (some college + college + graduate education) and lung cancer in nonsmoking women. Consider "exposed" as women with less than high school education; consider "nonexposed" as the combination of women with some college, college, and graduate education. **Use the population controls as the control group.

a. What are the values of cells (A, B, C, and D) of the 2 by 2 table?

b. What is the value of the odds ratio?

c. Can you spot the incorrect number in Table 4?

11. Interpret this odds ratio in a sentence. Can this odds ratio be used as an estimate of the risk ratio or incidence density ratio?

12. The 95% confidence interval around the odds ratio calculated in 10b above is (1.67-3.06). What information does this confidence interval provide?

13. In Table 5 of the journal article, the adjusted odds ratio for lung carcinomas (outcome) and ETS (exposure) obtained with the population controls was 1.20 (0.93- 1.55). Changing only the outcome to adenocarcinoma of the lung yields an adjusted odds ratio of 1.36 (1.02 - 1.84).

a. Explain what is meant by the "adjusted" odds ratio.

b. How do you interpret the findings both in Table 5 and in Figure 1 of the journal article, in which adjusted odds ratios for adenocarcinomas of the lung are most often larger than those for all lung cancer?

12/10/2001 vs/lka, 2/18/2003 vs, 12/8-11/2003vs

Lung Cancer in Nonsmoking Women: A Multicenter Case-Control Study[1]

Elizabeth T. H. Fontham,[2] Pelayo Correa, Anna Wu-Williams, Peggy Reynolds, Raymond S. Greenberg, Patricia A. Buffler, Vivien W. Chen, Peggy Boyd, Toni Alterman, Donald F. Austin, Jonathan Liff, and S. Donald Greenberg

Department of Pathology, Louisiana State University Medical Center, New Orleans, Louisiana 70112 [E. T. H. F., P. C., V. W. C.]; Department of Family and Community Medicine, University of Southern California, Los Angeles, California 90033 [A. W-W.]; California Department of Health Services, Emeryville, California [P. R., D. F. A.]; Division of Epidemiology, Emory University School of Public Health, Atlanta, Georgia 30322 [R. S. G., J. L.]; School of Public Health, University of Texas Health Science Center, Houston, Texas 77030 [P. A. B., T. A.]; California Public Health Foundation, Berkeley, California 94720 [P. B.]; and Department of Pathology, Baylor College of Medicine, Houston, Texas 77030 [S. D. G.]

Abstract

The association between exposure to environmental tobacco smoke and lung cancer in female lifetime nonsmokers was evaluated using data collected during the first 3 years of an ongoing case-control study. This large, multicenter, population-based study was designed to minimize some of the methodological problems which have been of concern in previous studies of environmental tobacco smoke and lung cancer. Both a cancer control group and a population control group were selected in order to evaluate recall bias. A uniform histopathological review of diagnostic material was conducted for case confirmation and detailed classification. Biochemical determination of current exposure to tobacco and screening of multiple sources of information to determine lifetime nonuse were utilized to minimize misclassification of smokers as nonsmokers.

A 30% increased risk of lung cancer was associated with exposure to environmental tobacco smoke from a spouse, and a 50% increase was observed for adenocarcinoma of the lung. A statistically significant positive trend in risk was observed as pack-years of exposure from a spouse increased, reaching a relative risk of 1.7 for pulmonary adenocarcinoma with exposures of 80 or more pack-years. The predominant cell type of the reviewed, eligible lung cancer cases was adenocarcinoma (78%). Results were very similar when cases were compared to each control group and when separate analyses were conducted for surrogate and personal respondents. Other adult-life exposures in household, occupational, and social settings were each associated with a 40-60% increased risk of adenocarcinoma of the lung. No association was found between risk of any type of lung cancer and childhood exposures from a father, mother, or other household members.

Introduction

Approximately one decade has passed since the initial reports of increased risk of lung cancer in nonsmoking women married to smokers (1, 2). The ensuing studies have provided a body of data which suggests a small but significant elevation in risk of lung cancer associated with exposure to ETS[3] (3-22). In reported prospective studies exposure has been assessed by the spouse's smoking history, primarily that of husbands. In case-control studies, the primary ETS exposure assessed has also been that from a spouse, although exposures from parents, other household exposures, and the workplace have been examined in some studies.

In general, these studies have included fewer than 100 nonsmoking lung cancer cases whose self-reported smoking status has not been validated by biochemical determination or other means. Reviews of available studies of ETS and lung cancer in nonsmokers by the National Research Council (23), the International Agency for Cancer Research (24), and others (25, 26) have concluded that although misclassification is unlikely to account for all of the observed increased risk, some misclassification of current or former smokers as nonsmokers is likely (0.5-5.0%). Because smokers tend to marry smokers, misreporting may introduce some bias in the estimation of the magnitude of the observed effect.

This study was undertaken in 1985 in an effort to address a number of unresolved issues related to ETS:

(a) Misclassification of Smoking Status. Multiple sources of information are utilized to ascertain nonsmoking status (medical record, physician, and then the study subject or surrogate). Study respondents are questioned twice (at contact to set up the interview and at the beginning of the interview). Self-reported current nonsmoking status is corroborated by measurement of urinary cotinine.

(b) Histopathological Specificity. Microscopic diagnostic slides are reviewed by one pulmonary pathologist both to confirm eligibility of cases as primary lung carcinomas and to provide a detailed review (subtype, differ-

Received 5/2/91.
[1] This research was supported by Grant CA40095 from the National Cancer Institute.
[2] To whom requests for reprints should be addressed, at Department of Pathology, LSU Medical Center, 1901 Perdido St., New Orleans, LA 70112.

[3] The abbreviations used are: ETS, environmental tobacco smoke; SEER, Surveillance, Epidemiology, and End Result; OR, odds ratio; CI, confidence interval.

Principles of Epidemiology for Public Health (EPID600)

Study designs: Case-control studies

Victor J. Schoenbach, PhD home page

Department of Epidemiology
School of Public Health
University of North Carolina at Chapel Hill

www.unc.edu/epid600/

Plan for this lecture

- Confidence intervals and significance tests (read only)
- Incidence density and cumulative incidence (brief)
- Attributable risk (brief)
- Theoretical overview of case-control studies as a complement to the traditional perspective

Confidence intervals & significance tests

- Everything you've been told so far about confidence intervals and statistical significance is misleading, including this statement.
- I am not licensed to teach statistics, so what I say on this topic mustn't leave this room!

Confidence intervals

- "a plausible range of values for the unknown population parameter"

  Michael Oakes, *Statistical inference*, p.52
- Exact interpretation is problematic
- We are more confident that a 95% interval covers the parameter than a 90% interval, but the 95% interval is wider (provides a less precise estimate)

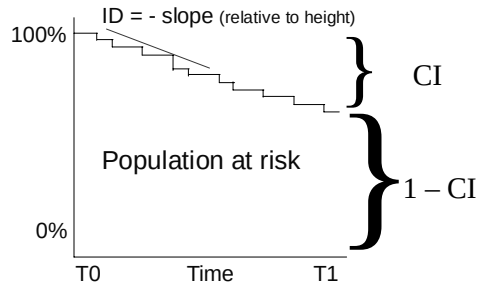
Significance tests

"It might be argued that the significance test, if properly understood, does no harm. This is, perhaps, fair comment, but anyone who appreciates the force of the case presented in this chapter will realize that equally, it does very little good."

Michael Oakes, *Statistical inference*, p.72

Incidence density and cumulative incidence

ID and CI in a closed cohort



Attributable risk

Attributable risk

Assume that we know a causal factor for a disease.

Conceptually, the "attributable risk" for that factor is:

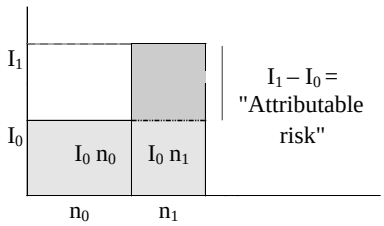
1. difference in risk or incidence between exposed and unexposed people or
2. difference in risk or incidence between total population and unexposed people

Attributable risk

Attributable risk can be presented as:

1. an "absolute" number, e.g., "80,000, or 20 per 100 cases/year of stroke are attributable to smoking"
2. a "relative" number, e.g., "20% of stroke cases are attributable to smoking".

For relative measures, think of % of cases

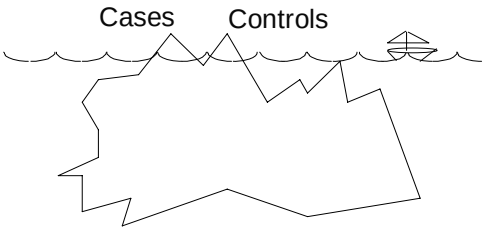


Case-control studies

Case-control studies

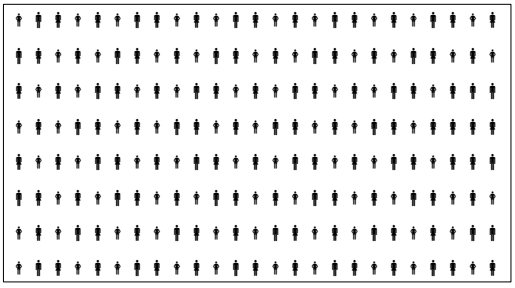
- Traditional view: compare
  - people who get the disease
  - people who do not get the disease
- "Controls" a misnomer, derived from faulty analogy to controls in experiment
- Modern conceptualization: controls are a "window" into the "study base"

Case-control studies

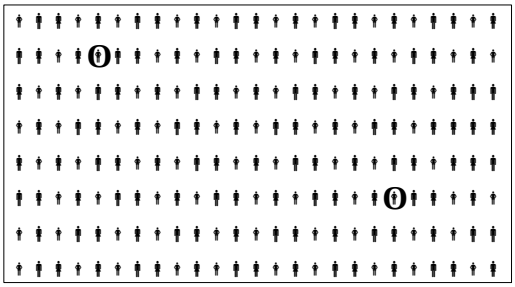


Population at risk

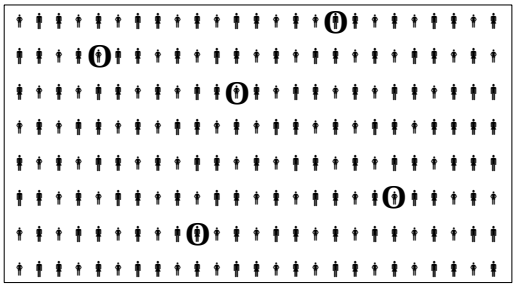
Population at risk (N=200)



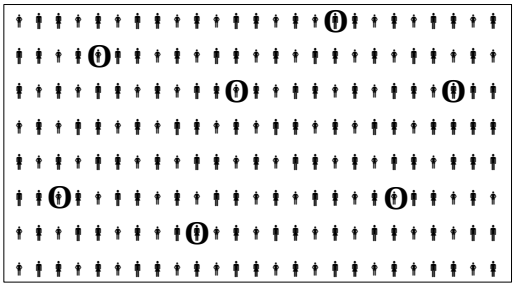
Week 1



Week 2



Week 3



Incidence rate ("incidence density")

$$ID = \frac{\text{Number of new cases}}{\text{Population time}}$$

Incidence rate ("incidence density")

$$ID = \frac{\text{Number of new cases}}{\text{Population time}} = \frac{7}{?}$$

Incidence rate ("incidence density")

Population time at risk:

200 people for 3 weeks = 600 person-wks

But 2 people became cases in 1st week

3 people became cases in 2nd week

2 people became cases in 3rd week

Only 193 people at risk for 3 weeks

Incidence rate ("incidence density")

Assume that:

2 people who became cases in 1st week were at risk for 0.5 weeks each = 2 @ 0.5 = 1.0

3 people who became cases in 2nd week were at risk for 1.5 weeks each = 3 @ 1.5 = 4.5

2 people who became cases in 3rd week were at risk for 2.5 weeks each = 2 @ 2.5 = 5.0

Incidence rate ("incidence density")

Total population-time =

Cases occuring during week 1: 1.0 p-w

Cases occuring during week 2: 4.5 p-w

Cases occuring during week 3: 5.0 p-w

Non-cases: 193 x 3 = 579.0 p-w

589.5 p-w

Incidence rate ("incidence density")

$$ID = \frac{\text{Number of new cases}}{\text{Population time}}$$

$$ID = \frac{7}{589.5} = 0.0119 \text{ cases / person-wk}$$

average over 3 weeks

Incidence proportion ("cumulative incidence")

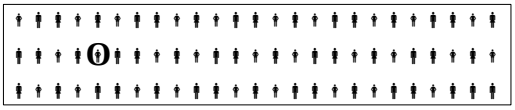
$$CI = \frac{\text{Number of new cases}}{\text{Population at risk}}$$

Incidence proportion ("cumulative incidence")

$$CI = \frac{\text{Number of new cases}}{\text{Population at risk}}$$

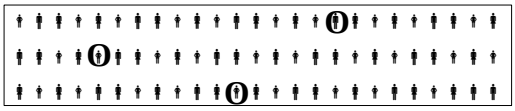
$$\text{3-week CI} = \frac{7}{200} = 0.035$$

Can estimate incidence in people who are "exposed"



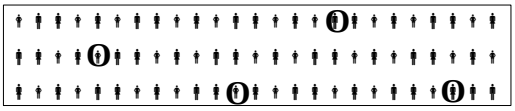
Week 1

Can estimate incidence in people who are "exposed"



Week 2

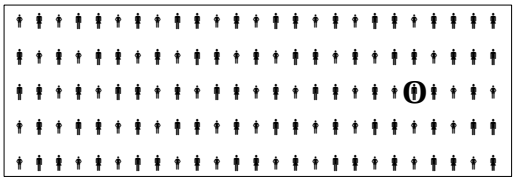
Can estimate incidence in people who are "exposed"



Week 3

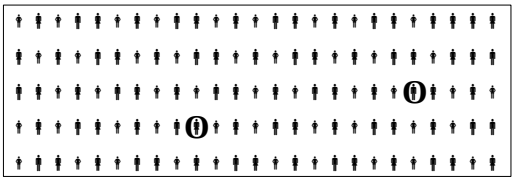
Can estimate incidence in people who are "unexposed"

Week 1



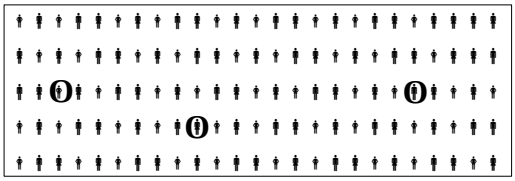
Can estimate incidence in people who are "unexposed"

Week 2

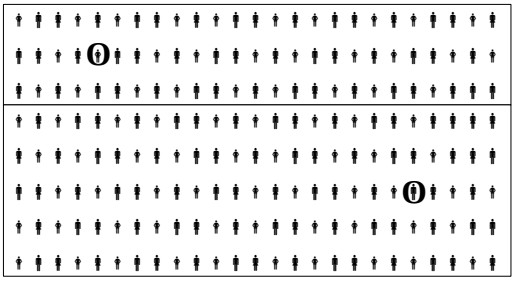


Can estimate incidence in people who are "unexposed"

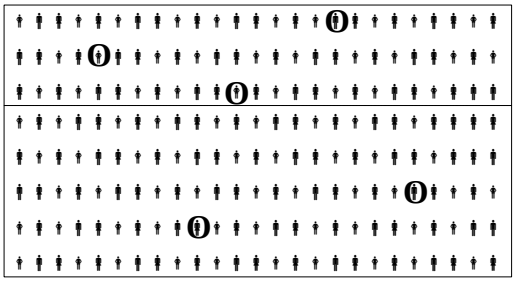
Week 3



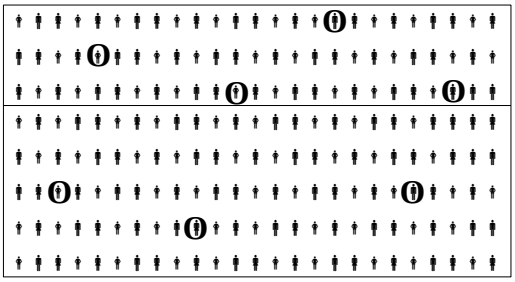
Entire population, week 1



Entire population, week 2



Entire population, week 3



Incidence rate ("incidence density")

$$ID = \frac{\text{Number of new cases}}{\text{Population time}}$$

Incidence rate in exposed during three weeks

$$ID = \frac{\text{Number of new cases}}{\text{Population time}}$$

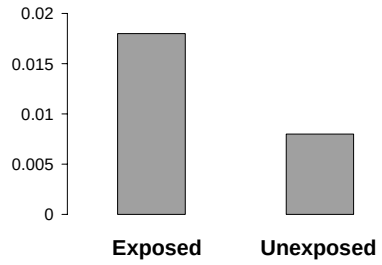
$$ID = \frac{4}{74.5 + 73 + 71.5} = \frac{4}{219} = 0.018 / wk$$

Incidence rate in unexposed during three weeks

$$ID = \frac{\text{Number of new cases}}{\text{Population time}}$$

$$ID = \frac{3}{124.5 + 123.5 + 122.5} = \frac{3}{370.5} = 0.008 / wk$$

Compare incidence rates in exposed and unexposed



Difference between incidence rates in exposed and unexposed

Incidence rate difference (IRD, IDD)

= (0.018 – 0.008) / wk

= 0.010 / week

How to interpret?

Difference in incidence rates

Incidence rate difference (IDD)

= (0.018 – 0.008) / wk

= 0.010 / week

"The rate in the exposed was 0.010 / week greater than the rate in the unexposed."

Relative difference in incidence rates in exposed and unexposed

Relative incidence rate difference

$$= \frac{0.018 - 0.008}{0.008} = 2.25 - 1 = 1.25$$

How to interpret?

Relative difference in incidence rates

Relative incidence rate difference

$$= \frac{0.018 - 0.008}{0.008} = 2.25 - 1 = 1.25$$

"The rate in the exposed was 125% greater than the rate in the unexposed."

Ratio of incidence rates in exposed and unexposed

$$\text{Incidence rate ratio (IRR, IDR)} = \frac{0.018}{0.008} = 2.25$$

How to interpret?

Ratio of incidence rates

$$\text{Incidence rate ratio (IRR, IDR)} = \frac{0.018}{0.008} = 2.25$$

"The rate in the exposed was 2.25 times the rate in the unexposed." [not "times greater"]

Estimating IDR and CIR with the Odds Ratio

Odds

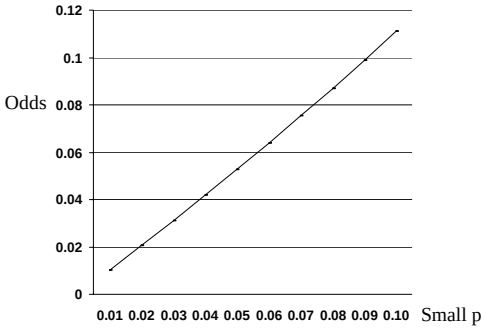
odds = probability / (1 – probability)

odds = risk / (1 – risk)

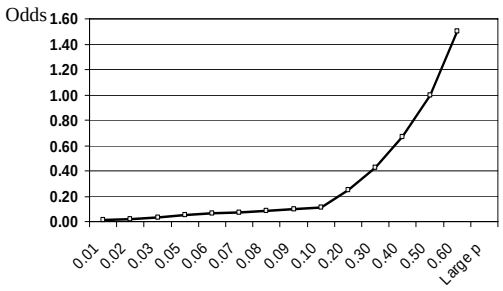
(most commonly)

Risk	0.010	0.050	0.100	0.20	0.80
Odds	0.010	0.053	0.111	0.25	4.00

For small probability, odds ≈ probability



Odds = probability / (1 – probability)



Odds

Any ratio of two natural numbers can be regarded as an odds:

Exposed: 20

Unexposed: 30

Total: 50

Odds of exposure: 20/50 divided by 30/50

Incidence rate ratio (a.k.a. "incidence density ratio") can be expressed as a ratio of odds

Incidence rate ratio

$$IDR = \frac{\text{ID in exposed}}{\text{ID in unexposed}} = \frac{ID_1}{ID_0}$$

Incidence rate ratio

$$IDR = \frac{ID_1}{ID_0} = \frac{\text{Exposed cases / Exp PT}}{\text{Unexposed cases / Unexp PT}}$$

(PT = "population time")

Incidence density ratio is a ratio of exposed to unexposed cases . . .

$$IDR = \frac{ID_1}{ID_0} = \frac{\text{Exposed cases / Exp PT}}{\text{Unexposed cases / Unexp PT}}$$

$$= \frac{\text{Exposed cases / Unexposed cases}}{\text{Exposed PT / Unexposed PT}}$$

. . . divided by a ratio of exposed to unexposed population-time

$$IDR = \frac{ID_1}{ID_0} = \frac{\text{Exposed cases / Exp PT}}{\text{Unexposed cases / Unexp PT}}$$

$$= \frac{\text{Exposed cases / Unexposed cases}}{\text{Exposed PT / Unexposed PT}}$$

Ratio of exposed to unexposed cases = "exposure odds" in cases

$$IDR = \frac{\text{Exposed cases / Unexposed cases}}{\text{Exp Person-time / Unexp Person-time}}$$

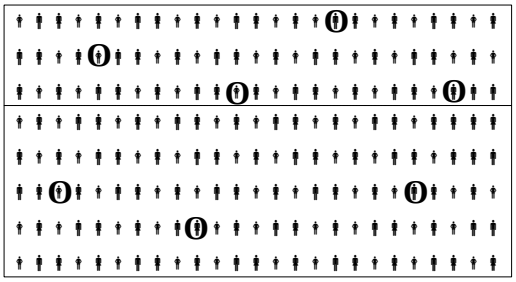
$$= \frac{\text{"Exposure odds" in cases}}{\text{"Exposure odds" in population}}$$

Ratio of exposed to unexposed population-time = "exposure odds"

$$IDR = \frac{\text{Exposed cases / Unexposed cases}}{\text{Exp Person-time / Unexp Person-time}}$$

$$= \frac{\text{"Exposure odds" in cases}}{\text{"Exposure odds" in population}}$$

Count exposed and unexposed cases



From before:

Exposed:

$$ID = \frac{4}{74.5 + 73 + 71.5} = \frac{4}{219} = 0.018 \text{ / wk}$$

Unexposed:

$$ID = \frac{3}{124.5 + 123.5 + 122.5} = \frac{3}{370.5} = 0.008 \text{ / wk}$$

Exposure odds in cases

= Exposed cases / Unexposed cases

= (4/7) / (3/7) = 4 / 3

= 1.33

The exposure odds in cases are 1.33

From before:

Exposed:

$$ID = \frac{4}{74.5 + 73 + 71.5} = \frac{4}{219} = 0.018 / wk$$

Unexposed:

$$ID = \frac{3}{124.5 + 123.5 + 122.5} = \frac{3}{370.5} = 0.008 / wk$$

Exposure odds in population-time

= Exp. person-time / Unexp. person-time

= (74.5 + 73 + 71.5) / (124.5 + 123.5 + 122.5)

= 219 / 370.5

= 0.59

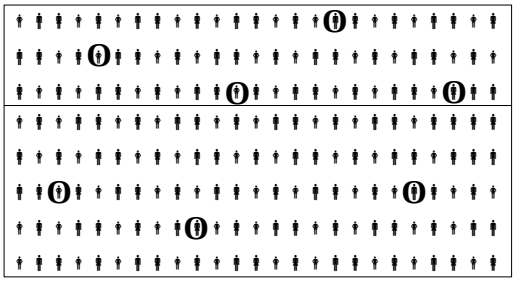
The exposure odds in population-time are 0.59

So incidence density ratio can be expressed as a ratio of odds

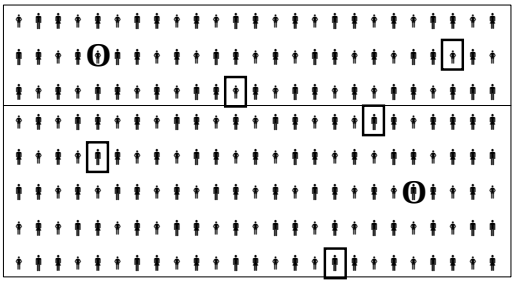
$$IDR = \frac{\text{“Exposure odds” in cases}}{\text{“Exposure odds” in population}}$$

$$= \frac{1.33}{0.59} = 2.25 \text{ (same as earlier)}$$

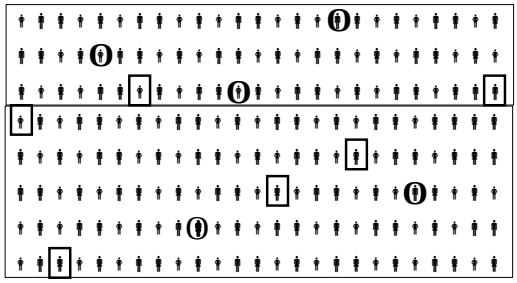
Can we estimate exposure odds in the population by taking a sample?



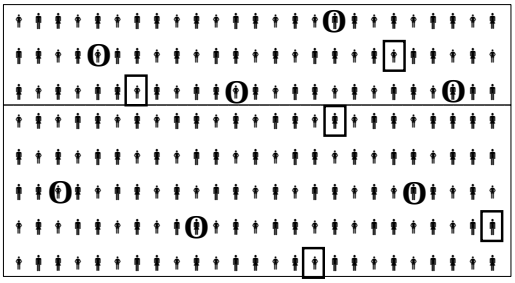
Take sample of “risk set” – “density controls” – week 1



Take sample of "risk set" – "density controls" – week 2



Take sample of "risk set" – "density controls" – week 3



Now estimate incidence density ratio using the "density controls"

$$IDR = \frac{ID_1}{ID_0} = \frac{\text{Exposed cases / Unexposed cases}}{\text{Exposed PT / Unexposed PT}}$$

Estimating the population odds with the odds in the control group

$$IDR = \frac{\text{Exposed cases / Unexposed cases}}{\text{Exposed controls / Unexposed controls}}$$

We call this the "exposure odds ratio" (OR).

Exposure odds ratio = estimates incidence density (rate) ratio

$$OR = \frac{1.33}{6 / 10} = \frac{1.33}{0.60} = 2.22$$

Odds ratio from a 2 x 2 table

	Cases	Controls
Exposed	a	b
Unexposed	c	d

$$\text{Odds ratio} = \frac{a / c}{b / d} = \frac{ad}{bc}$$

Odds ratio from a 2 x 2 table

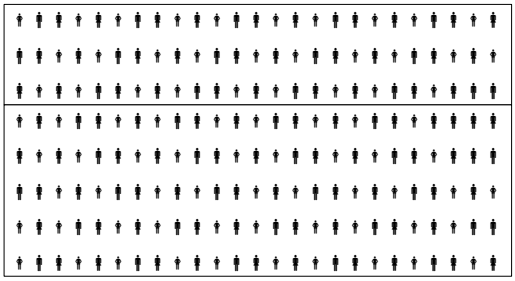
	Cases	Controls
Exposed	4	6
Unexposed	3	10

$$\text{Odds ratio} = \frac{4 / 3}{6 / 10} = \frac{(4)(10)}{(6)(3)}$$

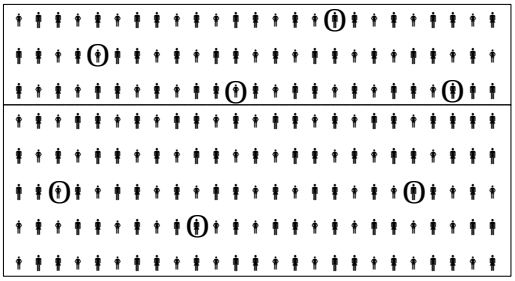
What about incidence proportion in a cohort? ("cumulative incidence")

$$\text{3-week CI} = \frac{\text{Number of new cases}}{\text{Population at risk}}$$

Population at risk – baseline



Entire population, week 3



Cumulative incidence in <u>exposed</u>

$$\text{3-week CI} = \frac{\text{Number of new } \underline{\text{exposed}} \text{ cases}}{\underline{\text{Exposed}} \text{ population at risk}}$$

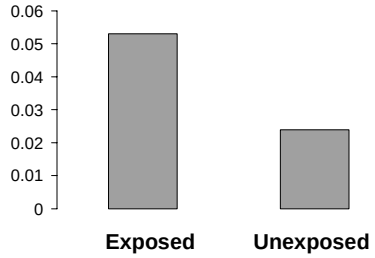
$$\text{3-week CI} = \frac{4}{75} = 0.053$$

Cumulative incidence in <u>unexposed</u>

$$\text{3-week CI} = \frac{\text{Number of new } \underline{\text{unexposed}} \text{ cases}}{\underline{\text{Unexposed}} \text{ population at risk}}$$

$$\text{3-week CI} = \frac{3}{125} = 0.024$$

Compare incidence proportions in exposed and unexposed



Difference in incidence proportions between exposed and unexposed

3-wk cumulative incidence difference (CID)

= (0.053 – 0.024)

= 0.029

How to interpret?

Difference in incidence proportions

3-wk cumulative incidence difference (CID)

= (0.053 – 0.024)

= 0.029

"The 3-week cumulative incidence in the exposed was 0.029 greater than in the unexposed."

Relative difference in incidence proportions for exposed and unexposed

3-wk cumulative incidence relative difference

$$= \frac{(0.053 - 0.024)}{0.024} = \frac{0.029}{0.024} = 1.22$$

How to interpret?

Relative difference in incidence proportions

3-wk cumulative incidence relative difference

$$= \frac{(0.053 - 0.024)}{0.024} = \frac{0.029}{0.024} = 1.22$$

"The 3-week CI in the exposed was 122% greater than in the unexposed."

Ratio of incidence proportions for exposed and unexposed

3-week cumulative incidence ratio (CIR)

= (0.053 / 0.024) = 2.22

How to interpret?

Ratio of incidence proportions ("relative risk", "risk ratio")

3-week cumulative incidence ratio (CIR)

= (0.053 / 0.024) = 2.22

"The 3-week CI in the exposed was 2.2 times that in the unexposed." [not "times greater than"]

Ratio of incidence proportions, a.k.a. cumulative incidence ratio

$$CIR = \frac{CI_1}{CI_0} = \frac{\text{Exposed cases / Exp PAR}}{\text{Unexposed cases / Unexp PAR}}$$

Cumulative incidence ratio can also be expressed as a ratio of odds of exposure in cases divided by . . .

$$CIR = \frac{CI_1}{CI_0} = \frac{\text{Exposed cases / Exp PAR}}{\text{Unexposed cases / Unexp PAR}}$$

$$= \frac{\text{Exposed cases / Unexposed cases}}{\text{Exp PAR / Unexp PAR}}$$

. . . an odds of exposure in the population at risk (PAR)

$$CIR = \frac{CI_1}{CI_0} = \frac{\text{Exposed cases / Exp PAR}}{\text{Unexposed cases / Unexp PAR}}$$

$$= \frac{\text{Exposed cases / Unexposed cases}}{\text{Exp PAR / Unexp PAR}}$$

So cumulative incidence ratio is also an odds ratio

$$CIR = \frac{CI_1}{CI_0} = \frac{\text{Exposed cases / Unexposed cases}}{\text{Exp PAR / Unexp PAR}}$$

$$= \frac{\text{Exposure odds in cases}}{\text{Exposure odds in population at risk}}$$

Exposure odds in cases

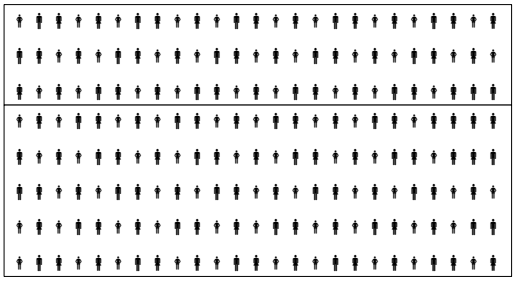
= Exposed cases / Unexposed cases

= 4 / 3

= 1.33

(same as for incidence density ratio)

Population at risk (before cases occur)



Exposure odds in population at risk (before cases occur)

= Exposed PAR / Unexposed PAR

= 75 / 125

= 0.60

(slightly different from odds of exposure for person-time)

So cumulative incidence ratio is also an odds ratio

$$CIR = \frac{CI_1}{CI_0} = \frac{\text{Exposure odds in cases}}{\text{Exposure odds in pop. at risk}}$$

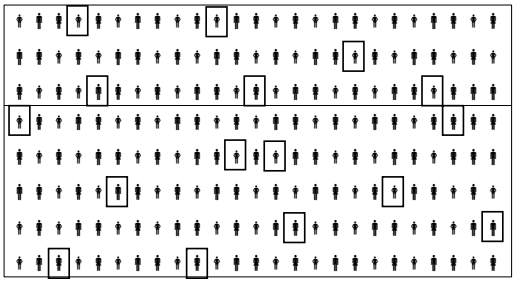
$$\text{3-week CIR} = \frac{1.33}{0.60} = 2.22$$

So if can estimate exposure odds in PAR, may not need to analyze entire cohort

"Case-cohort" or "case-base" design

Can be very advantageous if, for example, one wants to analyze specimens stored at baseline.

Sample from population at risk (before cases occur)



Exposure odds in controls

= Exp. controls / Unexp. controls

= 6 / 10

= 0.60

(expected value for the estimated odds)

Cumulative incidence ratio

$$CIR = \frac{CI_1}{CI_0} = \frac{\text{Exposed odds in cases}}{\text{Exposure odds in population}}$$

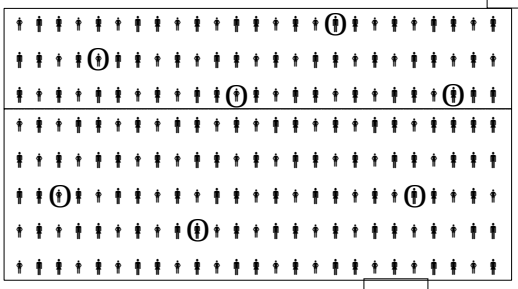
$$\text{3-wk CIR} = \text{exposure OR} = \frac{1.33}{0.6} = 2.22$$

What if cannot sample population at risk?

- Draw controls from noncases at <u>end of follow-up period</u>

Non-cases at end of follow-up

71



122

Exposure odds in population at risk (<u>after</u> cases occur)

= Exposed <u>noncases</u> / Unexp. <u>noncases</u>

= 71 / 122

= 0.58

(slightly different from ratio of person-time and ratio of population at risk)

Odds ratio

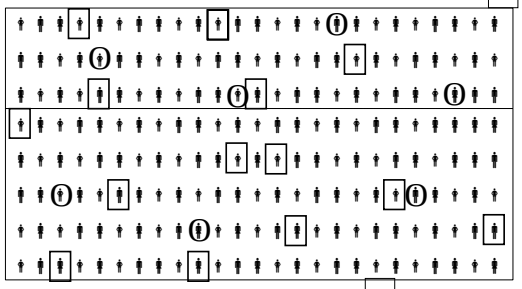
$$OR = \frac{\text{Exposed odds in cases}}{\text{Exposure odds in population}}$$

$$= \frac{1.33}{0.58} = 2.29$$

(slightly larger than CIR)

Draw controls from noncases

5



9

Exposure odds in "cumulative" controls

= Exposed noncases / Unexp. noncases

= 5 / 9

(about) = 0.58

(Note: 5/9=0.555, but a larger "sample" would produce 0.58)

Odds ratio

$$OR = \frac{\text{Exposed odds in cases}}{\text{Exposure odds in population}}$$

$$= \frac{1.33}{0.58} = 2.29$$

(slightly larger than CIR)

Case-control design is an efficient sampling technique

- Much more efficient, especially for rare outcomes
- Validity depends upon whether controls provide a clear view of population from which cases arise
- Susceptible to various sources of bias

New Software for Psychics

"Notice to user: By breaking the seal of this envelope, you accept the terms of the enclosed license agreement."

– Adobe Font Pack for Windows

Source: Willmott, Don, Abort, Retry, Fail? *PC Magazine*, June 14, 1994, 482.