Bias

Zuber D. Mulla

Explanations when you Observe or Don't Observe an Association

- *Truth*

- Chance
- Bias
- Confounding

From *Epidemiology in Medicine* (Hennekens & Buring)

Bias

- **When you detect an association or fail to detect an association, it is possible that some aspect of the design or conduct of a study has introduced a bias into the results**

- ***Systematic error***

Bias

- Key word: “Different”

Bias

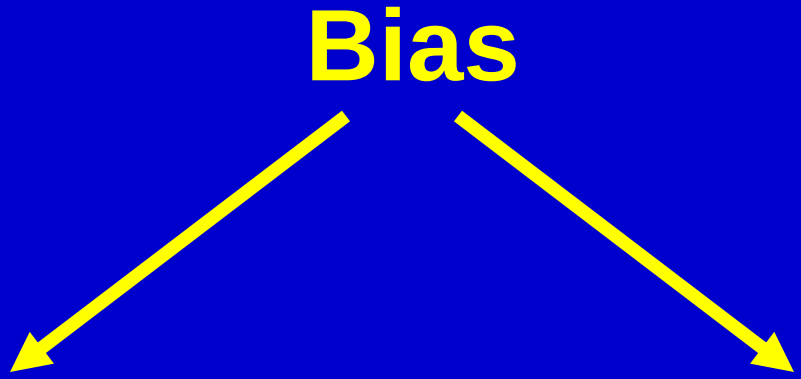
- Unlike “chance” and “confounding,” which can be evaluated quantitatively, the effects of bias are more difficult to evaluate and may be impossible to take into account in the analysis.

Are there quantitative methods that take account of biases?

- "Most biases can be fully analyzed only if additional "validation" data are available, but such data are usually absent or very limited." *Modern Epidemiology* 2nd edition by Rothman and Greenland

- Sensitivity analysis formulae

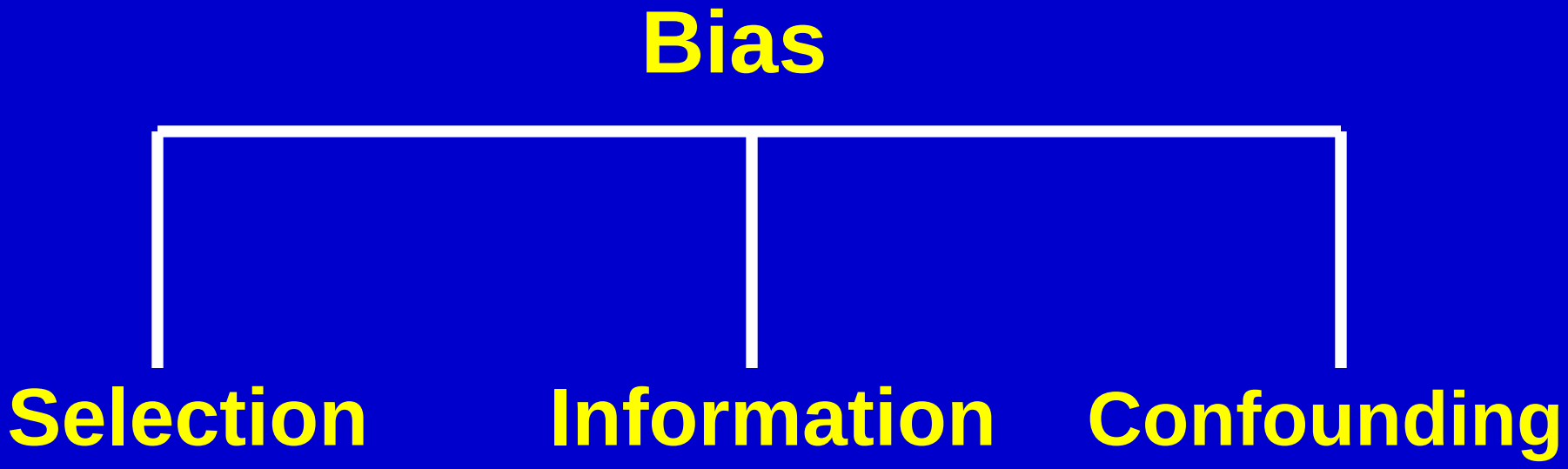
General Classes



Selection bias

Observation (Information bias)

General Classes (some books)



Selection Bias

Occurs when the association between exposure and disease (E:D) differs for those who participate and those who do not participate in the study.

Selection Bias

- Can occur whenever the identification of individual subjects for inclusion into the study on the basis of either exposure (cohort) or disease (case-control study) status depends in some way on the other axis of interest

Classic Example

- **Hospital-based case-control studies**
- **Exposure: Oral contraceptives (OC)**
- **Outcome: Thromboembolism**

Classic Example
(see text by Hennekens & Buring)

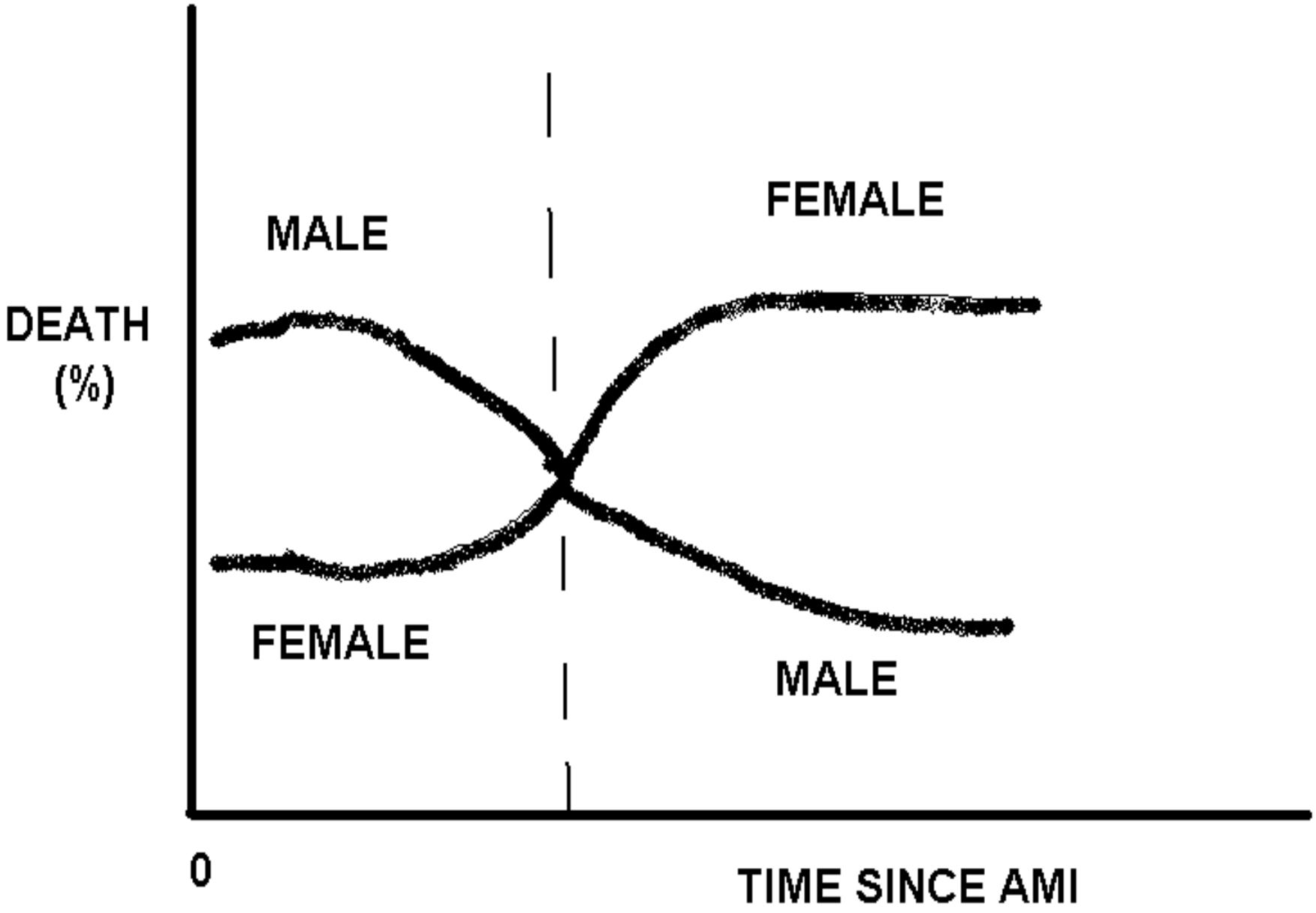
- When the first hospital-based case-control studies were reported, there was some concern that since physicians were aware of this hypothesis, a proportion of the women in the hospital case series had been hospitalized for evaluation of this disease because they were currently using OCs

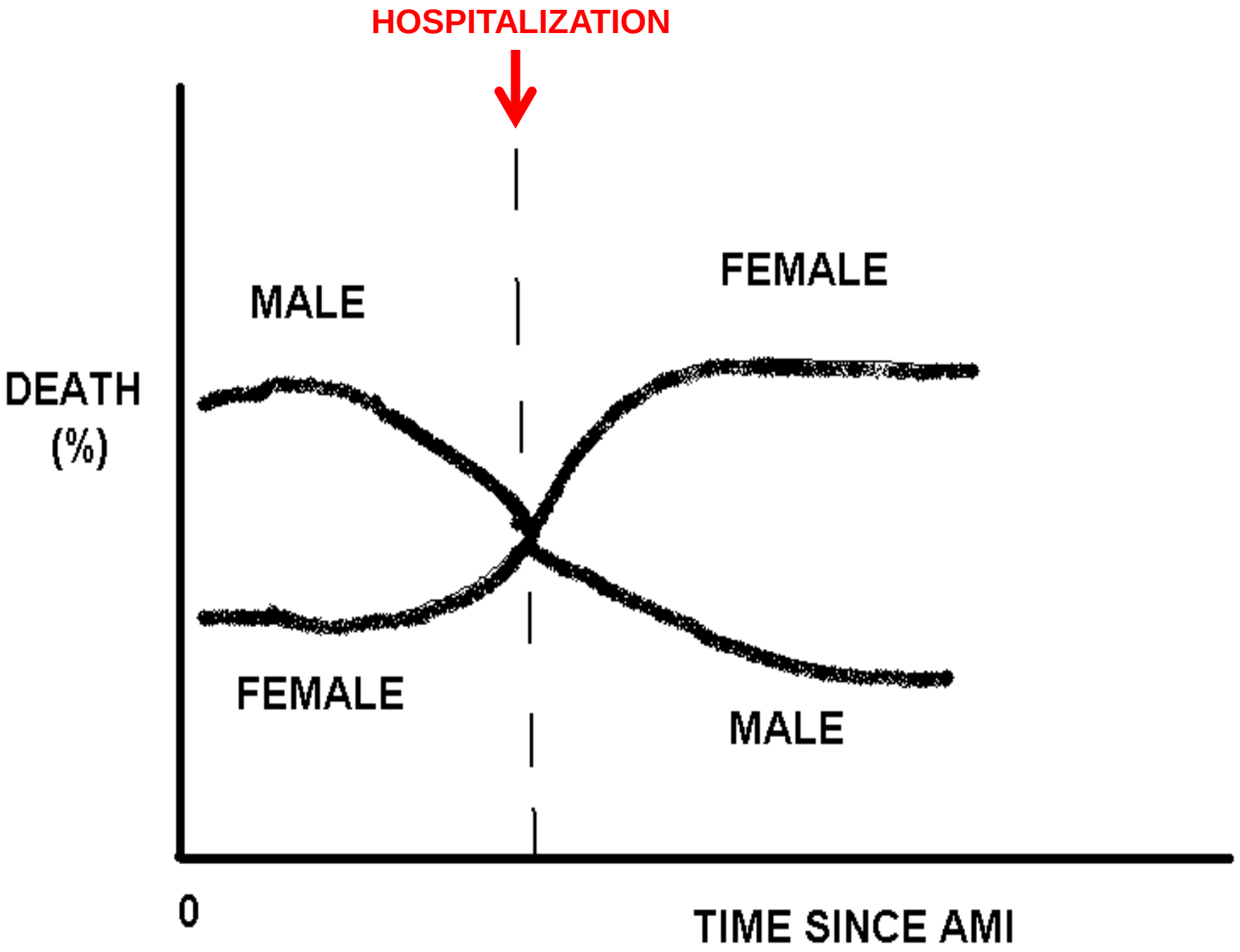
Classic Example (continued)

- If so, any increased frequency of current OC use among women hospitalized for thrombo-embolism might actually be due in part to the fact that hospitalization and the determination of the diagnosis were both influenced by a history of OC use

Another Example of Selection Bias

- Effect of gender on short-term survival after a heart attack [acute myocardial infarction (AMI)]


MALE
FEMALE
DEATH
(%)
FEMALE
MALE
0
TIME SINCE AMI


HOSPITALIZATION
MALE
FEMALE
DEATH
(%)
FEMALE
MALE
0
TIME SINCE AMI

Detection Bias: A type of selection bias

- Occurs when the exposure produces a greater or lesser likelihood of detecting the disease of interest

Detection Bias

- Exposure: Estrogen
- Outcome: Uterine cancer
- Exposed women had bleeding and therefore more likely to seek medical care and have a diagnostic evaluation than nonexposed

Detection Bias

- **Exposure: Vasectomy**
- **Outcome: Prostate cancer**

Response Bias: A type of selection bias

- Occurs when the responders differ, in a systematic way, from the source population

- Case-control study
- Controls: Daytime at home

Observation/Information Bias

- Results from systematic differences in the way data on exposure or outcome are obtained from the various study groups

Recall Bias
(A type of information bias)

- Arises when persons with a particular adverse health outcome remember and report their previous exposure experience differently from those who are not similarly affected

Recall Bias

- Sick individuals may be more likely to remember and report exposures than healthy individuals

- Problematic in case-control studies

Recall Bias

- **"Systematic error due to differences in accuracy or completeness of recall to memory of past events or experiences."**

  A Dictionary of Epidemiology **5th edition, edited by Porta**

Recall Bias in a Case-Control Study

- **Differential misclassification of exposure**

- **Now equally poor recall by cases and controls is not recall bias.**

Example of Recall Bias (see Hennekens & Buring)

- Case-control study of Down's syndrome
- Exposure: Spermicide use by mothers
- Mothers of cases continually thinking of possible causes/exposures

Example of Recall Bias

- One solution: Instead of healthy controls, use controls who have congenital heart disease
- Congenital heart disease not associated with spermicide use
- Now, exposure information provided by mothers of cases and controls would be subject to the same recall bias

Interviewer Bias

- Differential probing
- In case-control studies, the investigator should be blind to case status when eliciting exposure information
- In cohort studies, should be blind to exposure status

Losses to Follow-Up

- Possible problem in prospective cohort studies and clinical trials
- Is loss related to BOTH exposure and disease?

Misclassification

- Occurs whenever subjects are erroneously categorized with respect to either exposure or disease status

Misclassification

- The effect of misclassification depends on whether the misclassification with respect to exposure (or disease) is dependent on the person's disease (or exposure) status

Misclassification

- **Differential: Nonrandom**
- **Nondifferential: Random**

Differential Misclassification

- This arises if the proportions of subjects misclassified differ between the study groups

Differential Misclassification

- **Measurement error that depends on the values of other variables is referred to as *differential misclassification***

- **Source: *Modern Epi*, 2nd edition, Rothman and Greenland**

Differential Misclassification

- **The bias caused by differential misclassification can either exaggerate or underestimate an effect.**

- **Source: *Modern Epi*, 2nd edition, Rothman and Greenland**

Nondifferential Misclassification

- **Nondifferential exposure or disease misclassification occurs when the proportion of subjects misclassified on exposure does not depend on disease status or when the proportion of subjects misclassified on disease does not depend on exposure.**

Source: ***Modern Epi*****, 2nd edition, Rothman and Greenland**

Nondifferential Misclassification

- **The bias caused by non-differential misclassification of a binary exposure or outcome variable is *usually* toward the null value.**

- **Source: *Modern Epi*, 2nd edition, Rothman and Greenland**

Nondifferential Misclassification of Exposure

- Source: *Modern Epi*, 2nd edition, Rothman and Greenland

Hypothetical Case-Control Study, N=1280

	Case	Control
Exposed	240	240
Unexposed	200	600

Correct data: Odds ratio = 3.0

From *Modern Epi*, 2nd edition

Nondifferential Misclassification

- **Now suppose that the exposure is measured by a questionnaire that results in an exposure measure that has 100% specificity but only 80% sensitivity.**

- **Source: *Modern Epi*, 2nd edition, Rothman and Greenland**

Fake Data, N=1280

Sensitivity	Specificity	Odds Ratio
100%	*100%*	*3.0*
80%	100%	2.6
40%	60%	1.0
0.0%	0.0%	0.33

Adapted from *Modern Epi*, 2nd edition, Rothman and Greenland, p. 129

Nondifferential Misclassification

- In the example above, the misclassification is nondifferential, because the sensitivity and specificity of the exposure measurement method is the same for both cases and controls

Biased Odds Ratio

Nondifferential Misclassification

- This example shows that if the misclassification is severe enough, the bias can obliterate the association and even reverse the direction of association (in this example, from a harmful exposure to a protective exposure)

Example of Possible Differential Misclassification of Exposure

- Case-control study of demented individuals
- Proxies used only for cases (not the controls)
- What may happen?

Control of Bias in the Design Phase

- 1. Choice of Study Population

- For example, for a hospital-based case-control study, use hospital controls

Control of Bias

- This will increase comparability with the cases in terms of willingness to participate, the presence of selective factors that influenced the subjects' choice of a particular hospital, and awareness of exposures

Control of Bias in the Design Phase

- 2. Data Collection Methods

- Use standardized questionnaires
- Train data collectors/interviewers
- Method of data collection should be similar for all study groups. Don't, e.g., cases=interview, controls=chart review

Control of Bias in the Design Phase

- Okay to use multiple sources of data, but be fair to both groups.

- In fact, use multiple sources of data whenever possible as a way to provide an independent verification of exposure or disease status

Control of Bias in the Design Phase

- 2. Data Collection Methods
- Favor objective outcomes
- Blinding

Control of Bias in the Design Phase

- Use “foils” as a check for recall bias
- Example: Case-control study of acute myocardial infarction (AMI)
- Have a question on something that is not a risk factor for AMI

Control of Bias in the Design Phase

- Clinical trials of treatments: How to monitor compliance / exposure?

- Count pills
- Weigh subjects
- Labs

References Cited in This Lecture

- Hennekens CH, Buring JE. *Epidemiology in Medicine*. Little, Brown and Company, Boston; 1987.

- Porta M. *A Dictionary of Epidemiology Fifth Edition*. Oxford University Press, New York; 2008.

- Rothman KJ, Greenland S. *Modern Epidemiology Second Edition*. Lippincott Williams & Wilkins, Philadelphia; 1998.