Epidemiologic Methods 2

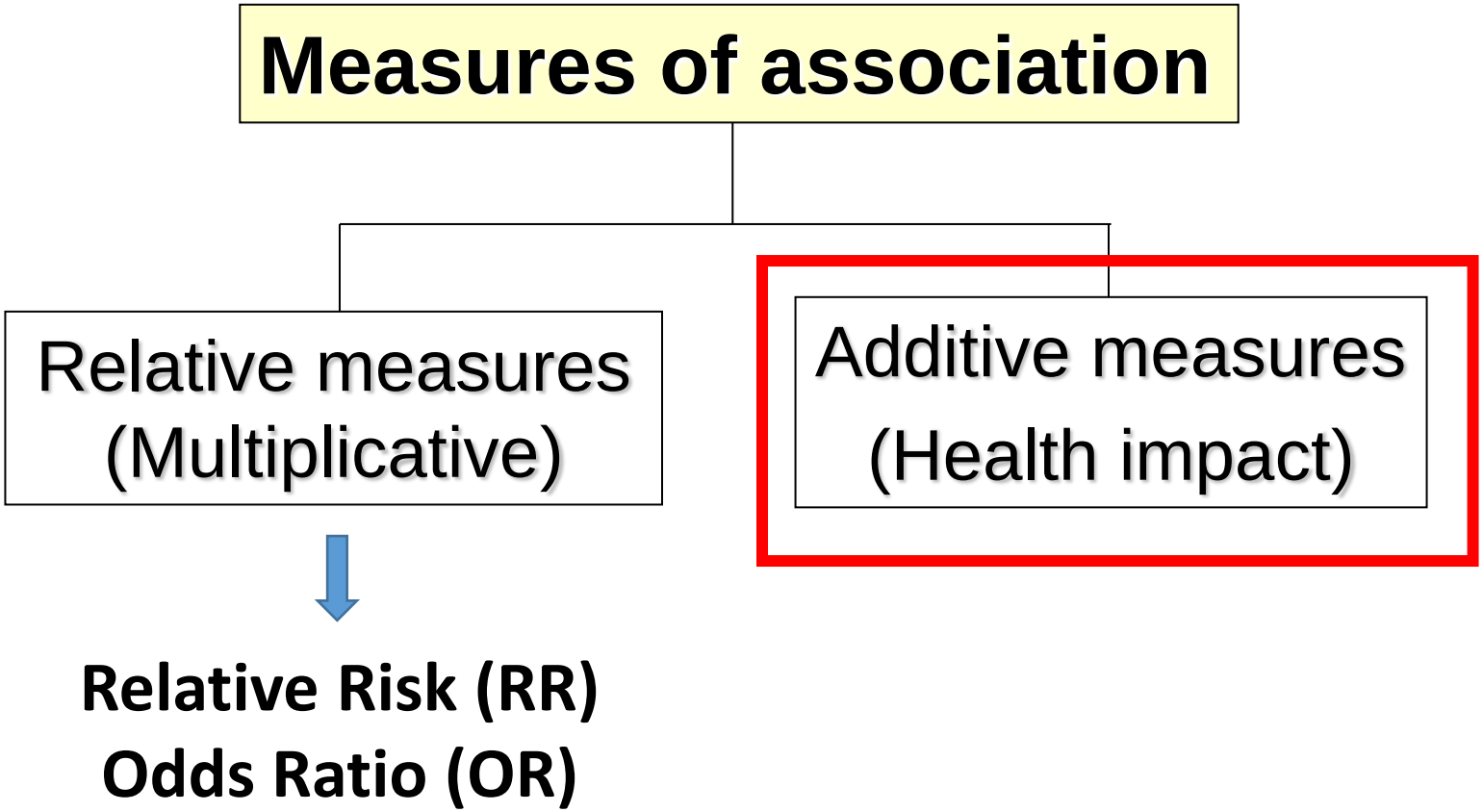
Research Skill 2

Kriengkrai Srithanaviboonchai MD, MPH
Department of Community Medicine
29 October 2020

Topics

- Association
- Internal and External validity
- Causation

Association


Measures of association
Relative measures (Multiplicative)
Additive measures (Health impact)
Relative Risk (RR)
Odds Ratio (OR)

Attributable risk (AR_{exp})

- Other names: Excess risk, Risk difference
- the difference between the incidence rates in exposed and non-exposed groups

AR_{exp} = Incidence in exposed – Incidence in un-exposed

- Excess risk of the outcome in the exposed group compared with the non-exposed group.

Association between contraceptive pills use and urinary tract infection (UTI)

Pills use	UTI		Total
	Yes	No	
Yes	27	455	482
No	31	1,831	1,862
Total	**58**	**2,286**	**2,344**

AR = Incidence in exposed - Incidence in un-exposed

= a / (a + b) – c / (c + d) = (27/482) - (31/1,862)

= 0.056 – 0.017 = 0.039

Attributable risk percent ($\%AR_{exp}$)

- Other names: Attributable rate percent, Attributable proportion, Etiologic fraction
- Percentage of disease in the exposed group that can be attributed to the exposure.
- Proportion of disease which can be reduced among exposed the risk factor is removed.

$$\%AR_{exp} = \frac{AR_{exp}}{\text{Incidence in exposed}} \text{ X } 100$$

Association between contraceptive pills use and urinary tract infection (UTI)

Pills use	UTI		Total
	Yes	No	
Yes	27	455	482
No	31	1,831	1,862
Total	**58**	**2,286**	**2,344**

$$\%AR = \frac{AR}{\text{Incidence in exposed}} \times 100$$

$$= \frac{0.039}{0.056} \times 100 \qquad = 69.64\%$$

Population attributable risk (PAR)

- Proportion of disease in the study population that is attributable to the exposure

PAR = Incidence in population – Incidence in unexposed

or

= (AR) * (proportion of exposed in the population)

Association between contraceptive pills use and urinary tract infection (UTI)

Pills use	UTI		Total
	Yes	No	
Yes	27	455	482
No	31	1,831	1,862
Total	**58**	**2,286**	**2,344**

PAR =Incidence in population - Incidence in un-exposed

= a + c / (a+b+c+d) – c / (c + d) = (58/2,344) - (31/1,862)

= 0.025 – 0.017 = 0.008

Population attributable risk (%PAR)

- Percentage of disease in the population that can be attributed to the exposure.
- Proportion of disease which can be reduced in population if the risk factor is removed.

$$\%PAR = \frac{PAR}{\text{Incidence of disease in population}} \times 100$$

$$= \frac{\{Pe(RR-1)\}}{\{Pe(RR-1)\} + 1} \times 100$$

Pe=proportion of exposure in population

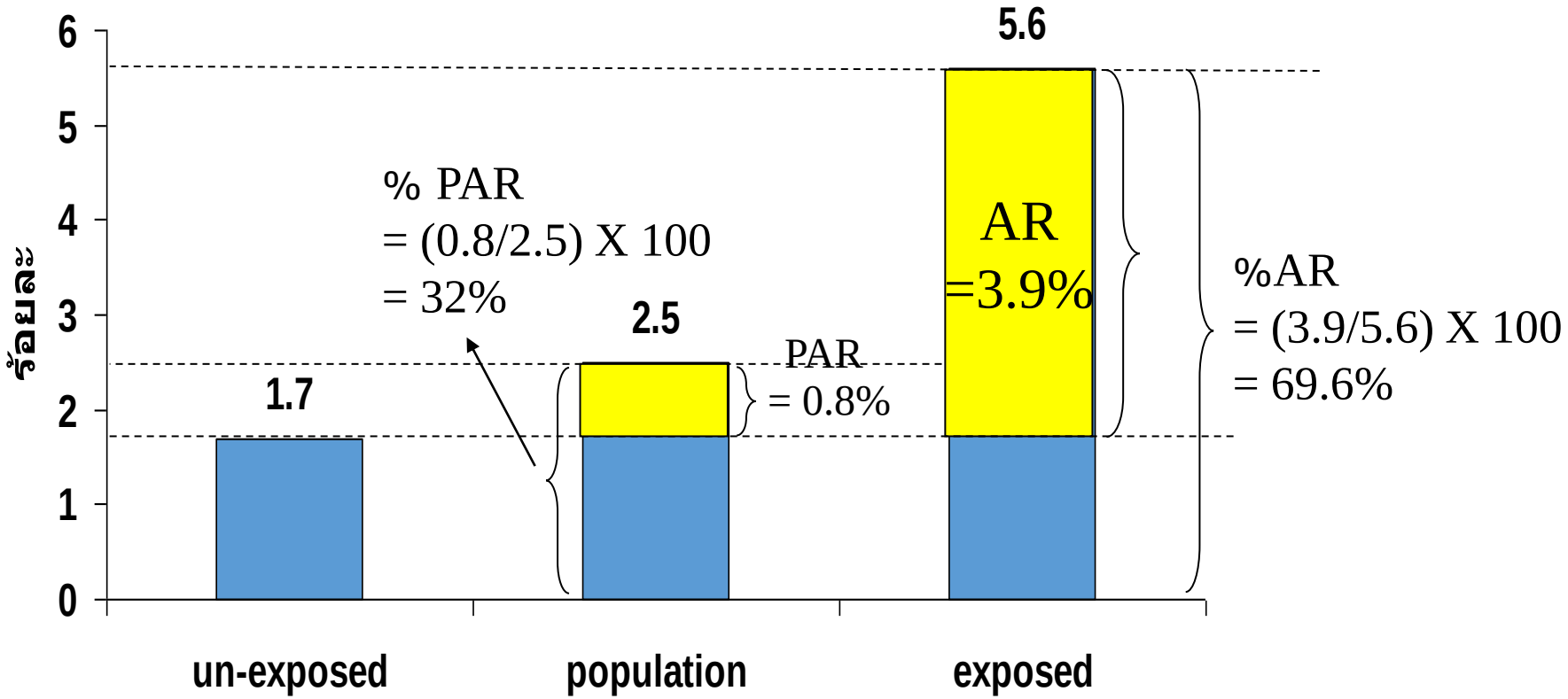
Association between contraceptive pills use and urinary tract infection (UTI)

Pills use	UTI		Total
	Yes	No	
Yes	27	455	482
No	31	1,831	1,862
Total	**58**	**2,286**	**2,344**

$$\%PAR = \frac{PAR}{\text{Incidence in population}} \times 100$$

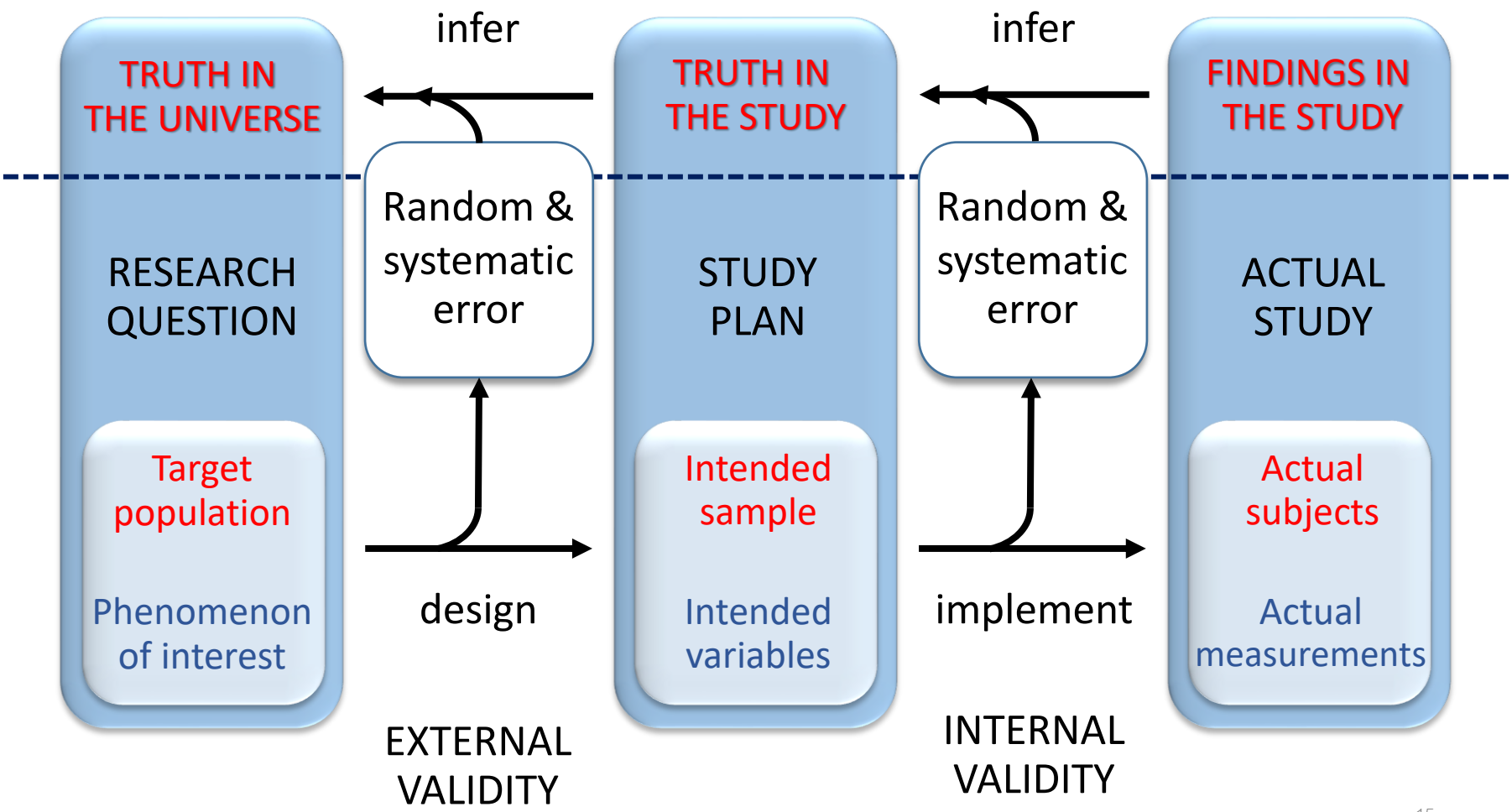
$$= \frac{0.008}{0.025} \times 100 \qquad = 32\%$$

Diagram shows relationships between AR, %AR, PAR, %PAR



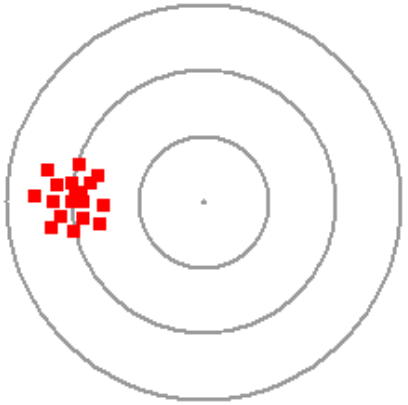
Internal and external validity

How research works



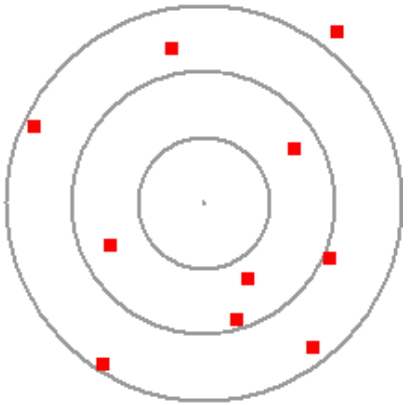
Internal validity

- Systemic error
  - Selection bias
  - Information bias
  - Confounding



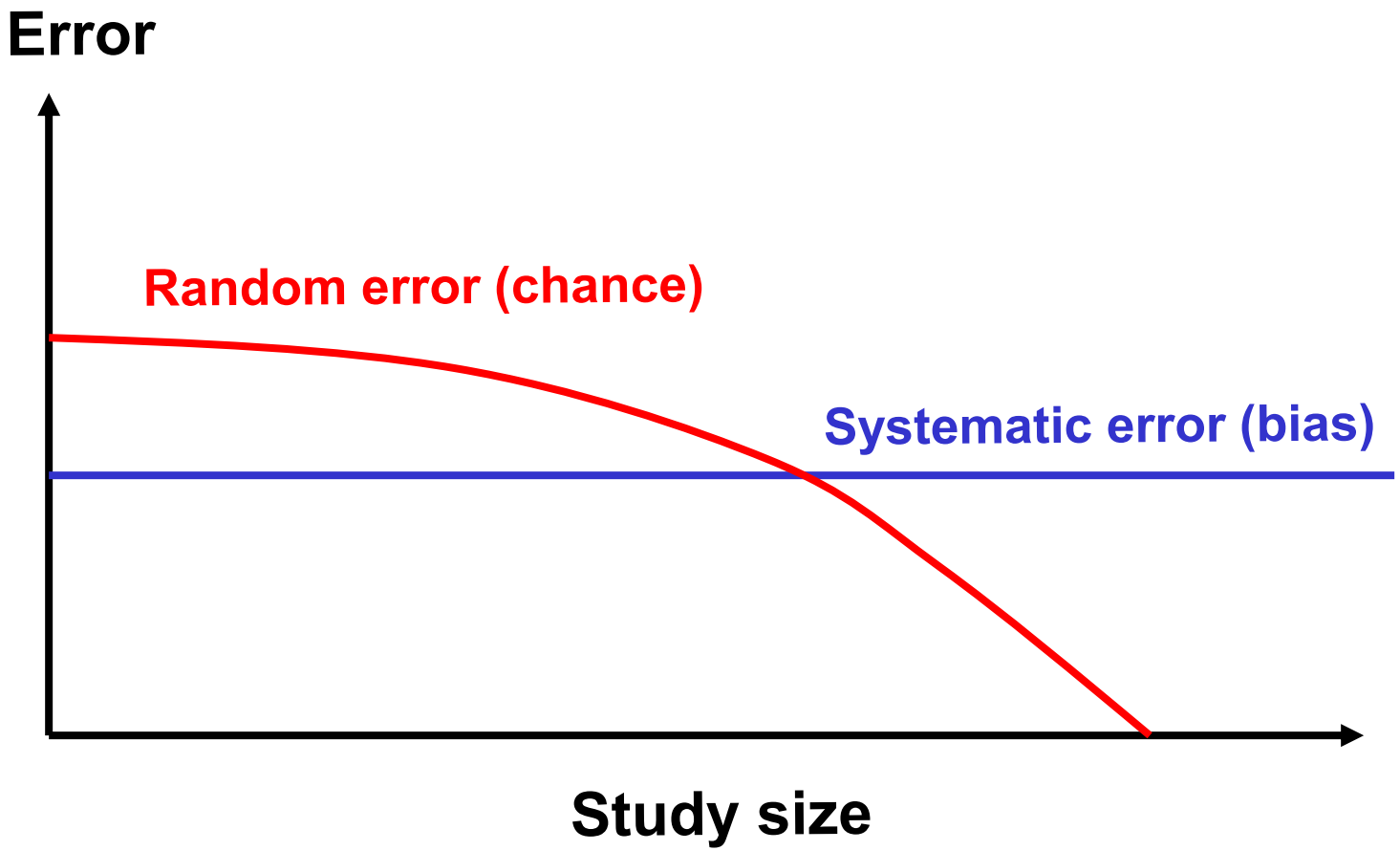
Systematic Error

- Random error (chance)
  - Inadequate sample size



Random Error

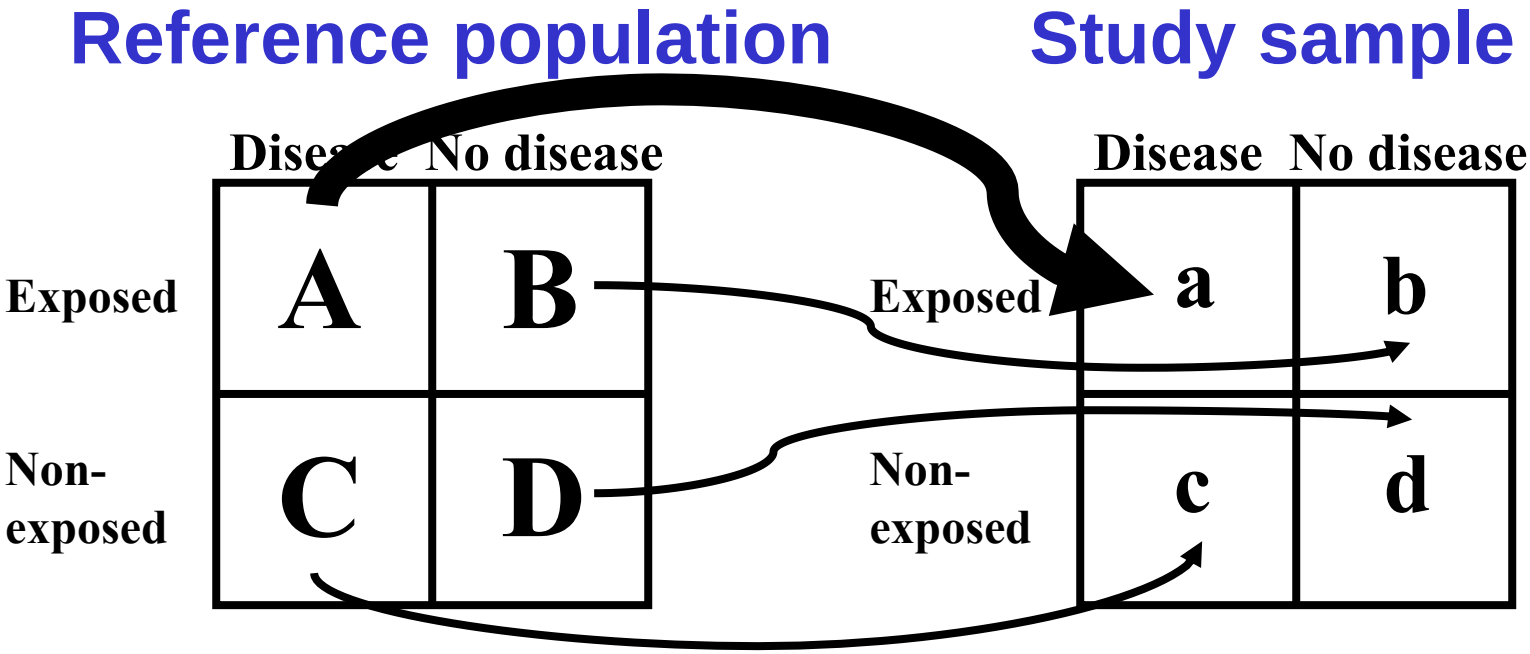
Errors in Epidemiological studies



Source: Rothman, 2002

Selection bias

- Bias of study results due to bias in selecting the study subjects
- Some types of population were sampled into study more than others



Healthy worker effect (HWE)

- “Workers usually exhibit lower overall death rates than the general population because the severely ill and chronically disabled are ordinarily excluded from employment” – Last, 1995.
- The reduction of mortality or morbidity of workers when compared with the general population leads to underestimation of the exposed risk among workers.

Mortality (SMR) of gas workers compared with national experience (Doll et al. 1965)

Heavy exposure	105
Intermediate exposure	90
No exposure	84

SMR in general population = 100

Healthy worker effect (HWE)

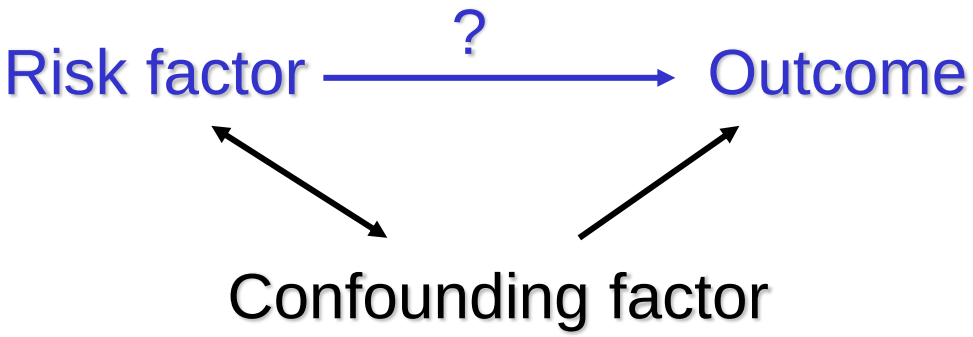
- Healthy hire: Employees in certain professions need to be able-bodied. Further, hiring of workers may be affected by personal habits and physical conditioning such as weight, alcohol intake, smoking, or gender depending on the situation
- Survival effect: Over time, the health status of workers drops, and they leave the workforce. The ones who remain over time are the healthiest of those who started.
- Advantages of working: Workers tend to remain employed as a result of their improved access to healthcare, regular screening for disease, early treatment of conditions, and physical exercise.

Information bias

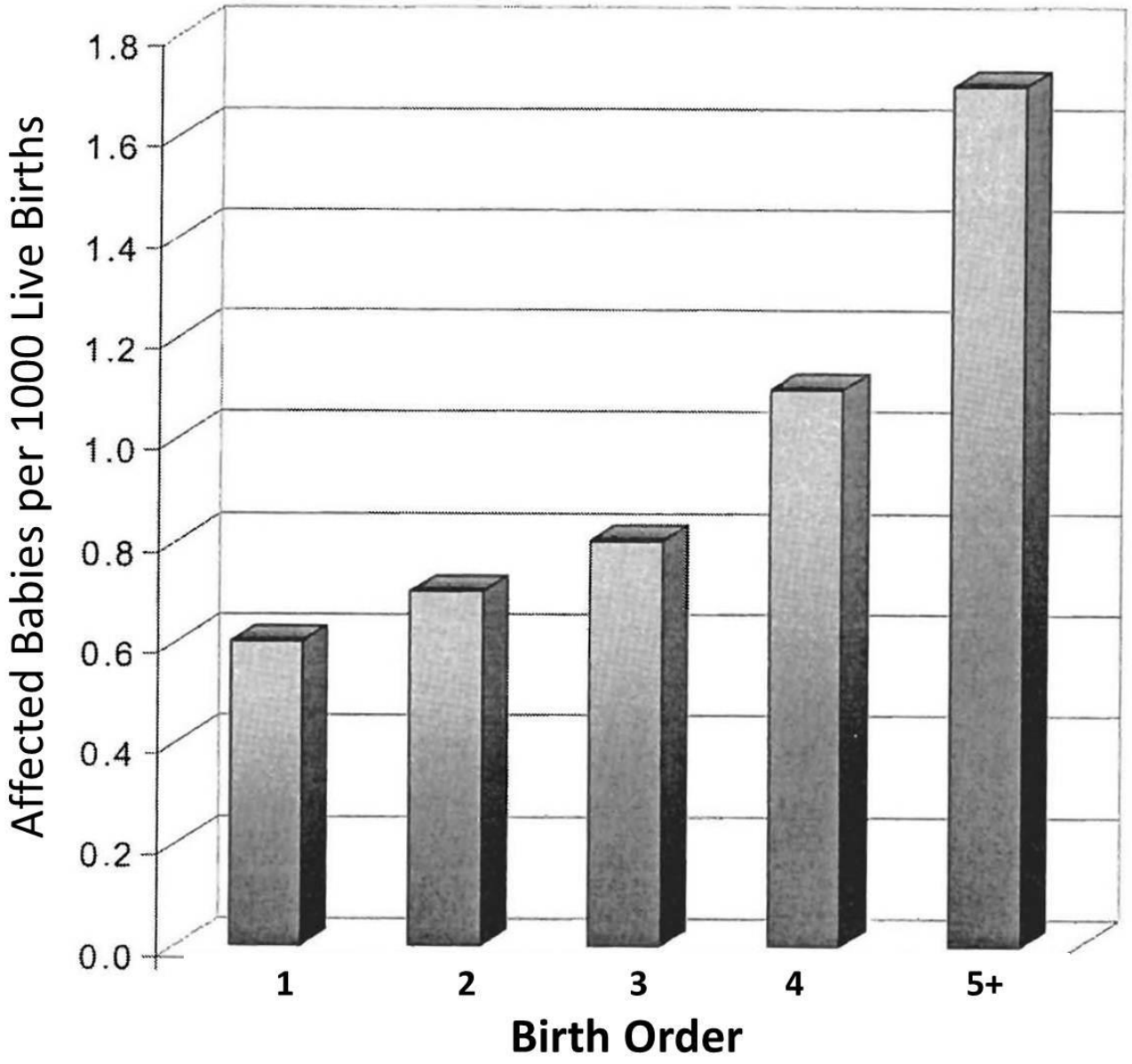
- Bias during data collection process
- Systematic different in collecting data regarding risk factor (intervention) or outcome
- Example
  - Instrument bias
  - Interviewer bias
  - Recall bias
  - End digit preference bias

Confounding

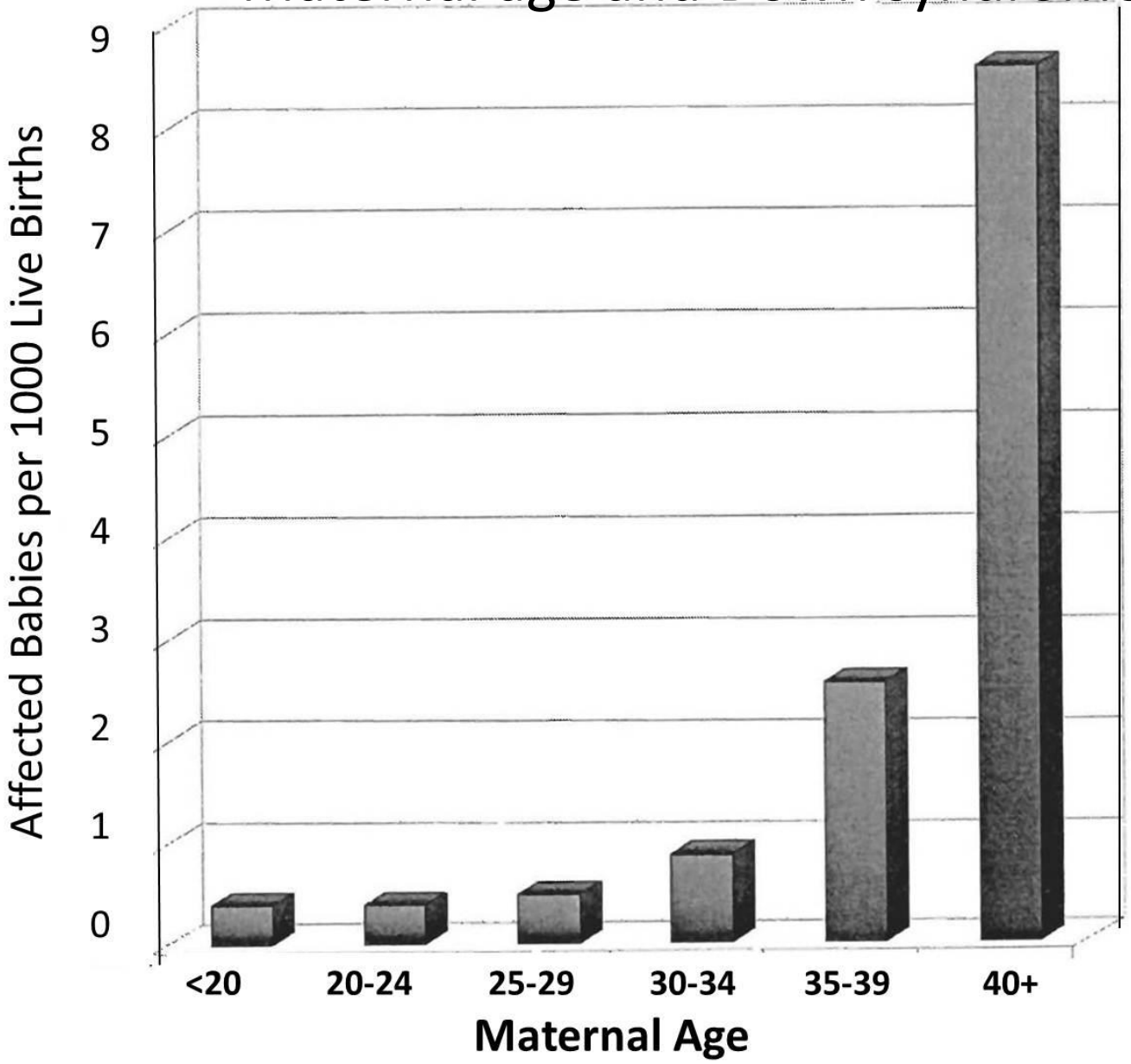
- Wrongly conclude the association between the variables under study due to the effect of other variable (Confounding factor)
- The confounding factor must:
  - Be risk factor of outcome under study
  - Associate with the risk factor



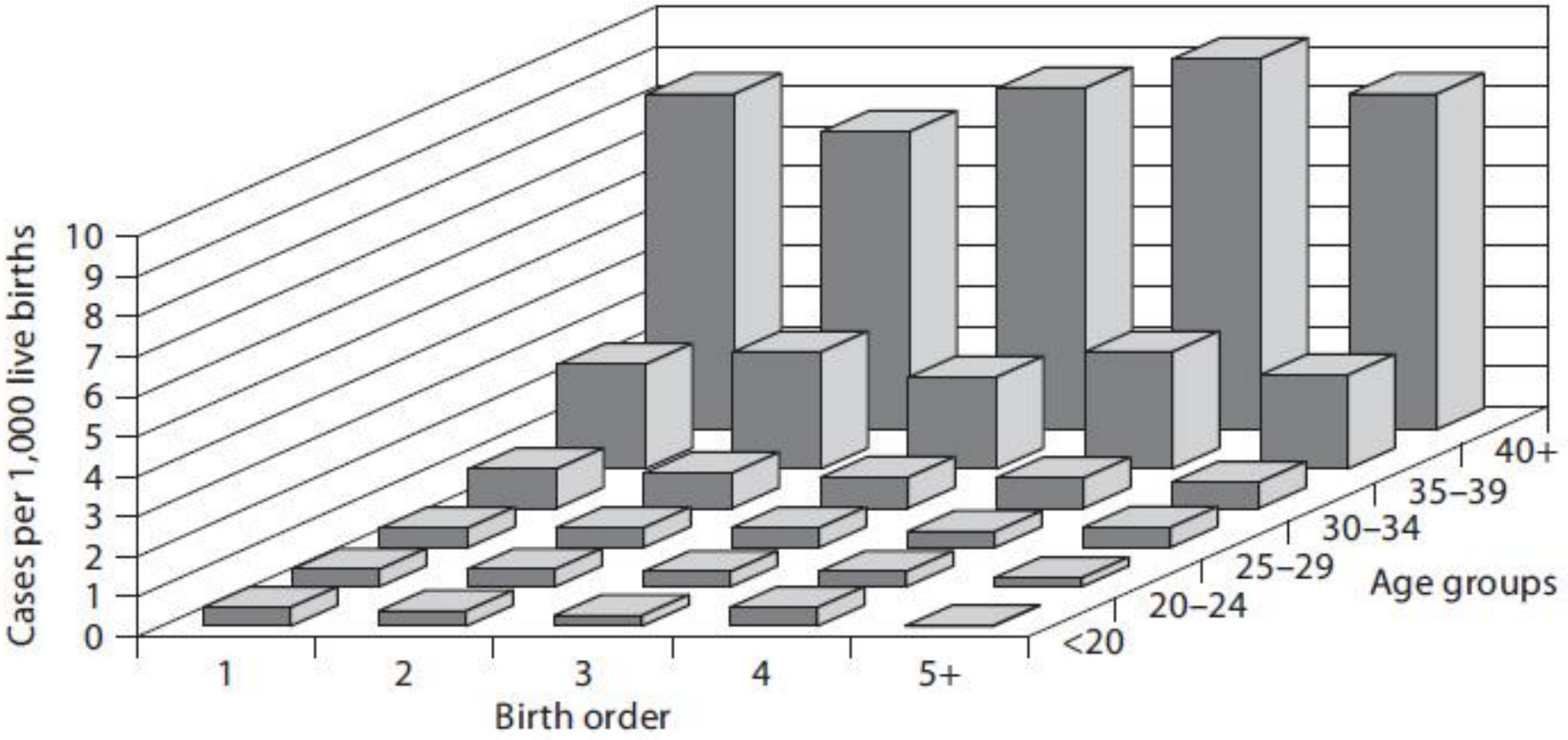
Birth order and Down syndrome



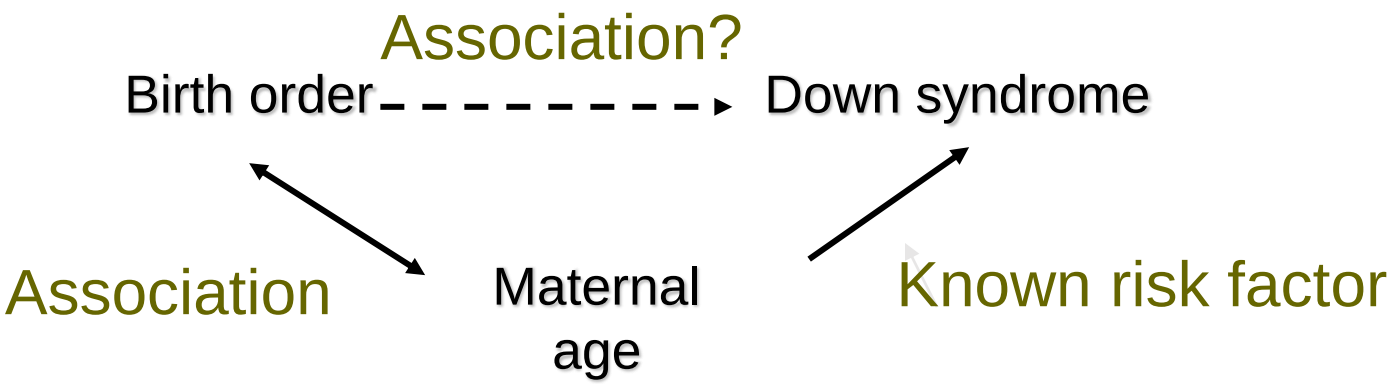
Maternal age and Down syndrome



Birth order and Down syndrome stratified by age groups



Confounding



Example

Question: "Does treatment A improve survival?"

E = Treatment

D = Survival

	Survived	Died		% survived
Treatment A	**18**	**12**	**30**	**60%**
Placebo	**12**	**18**	**30**	**40%**

RR = 1.5

"Patients getting treatment A were 1.5 times more likely to survive than patients receiving placebo"

Example

- **This analysis only examined the exposure and disease**
- **We call this "crude analysis", or a "crude association"**
- ***Is there some other variable that distorts the real association?***

Example

Early stage disease

	Survived	Died		% survived
Treatment A	**17**	**3**	**20**	**85%**
Placebo	**9**	**1**	**10**	**90%**

Late stage disease

	Survived	Died		% survived
Treatment A	**1**	**9**	**10**	**10%**
Placebo	**3**	**17**	**20**	**15%**

Example

Early stage disease

	Survived	Died	
Treatment A	**17**	**3**	**20**
Placebo	**9**	**1**	**10**

RR = 0.94

Late stage disease

	Survived	Died	
Treatment A	**1**	**9**	**10**
Placebo	**3**	**17**	**20**

RR = 0.67

Example

- **The potential confounder is: stage of disease at the time of therapy**
- **Early stage is associated with better survival (C ⟶ D)**
- **More persons in early stage received treatment A than those in late stage (C ⟶ E)**
- **In both cases, treatment A is not associated with better survival**

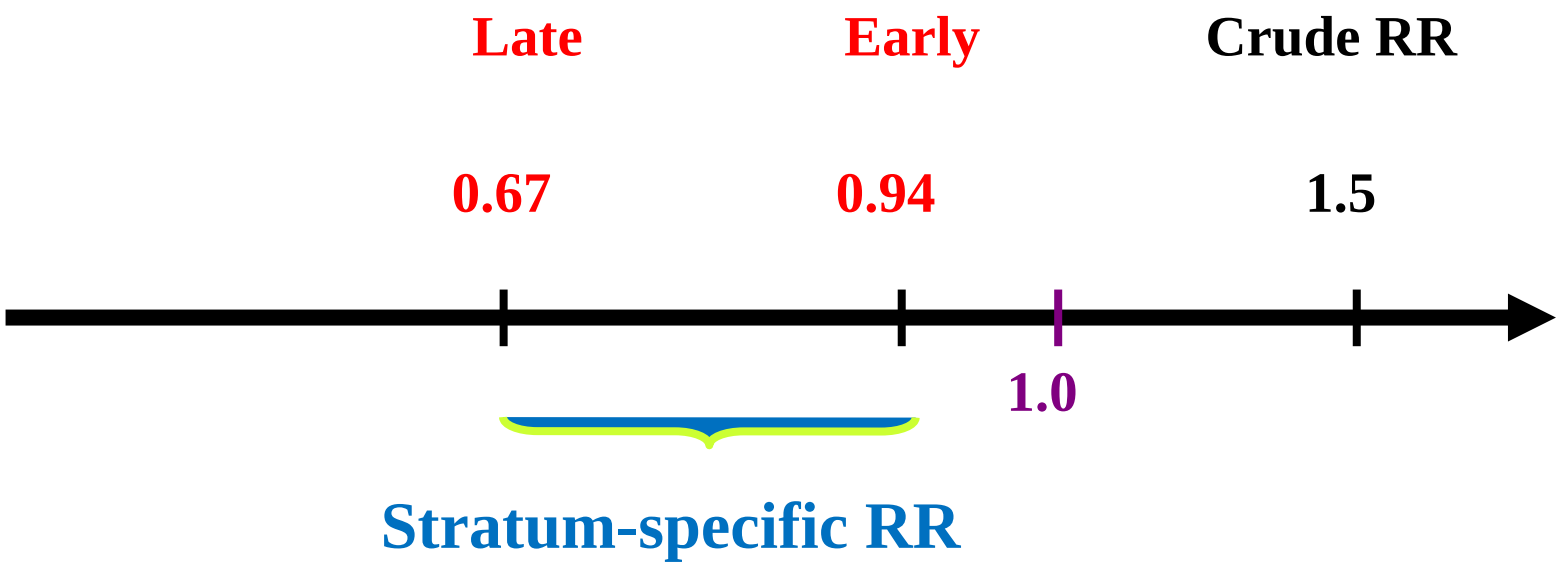
How to Identify Confounding

- **Examine 2 x 2 table and E-D association (RR or OR) for each value of confounding variable**
- **This is called "stratum-specific analysis", or "stratified analysis"**

  e.g. Strata 1= Early stage
  Strata 2= Late stage

- **If crude RR is not inside the range of stratum-specific RR, AND, stratum specific RR are "similar" to each other, then confounding is occurring**

Example of confounding



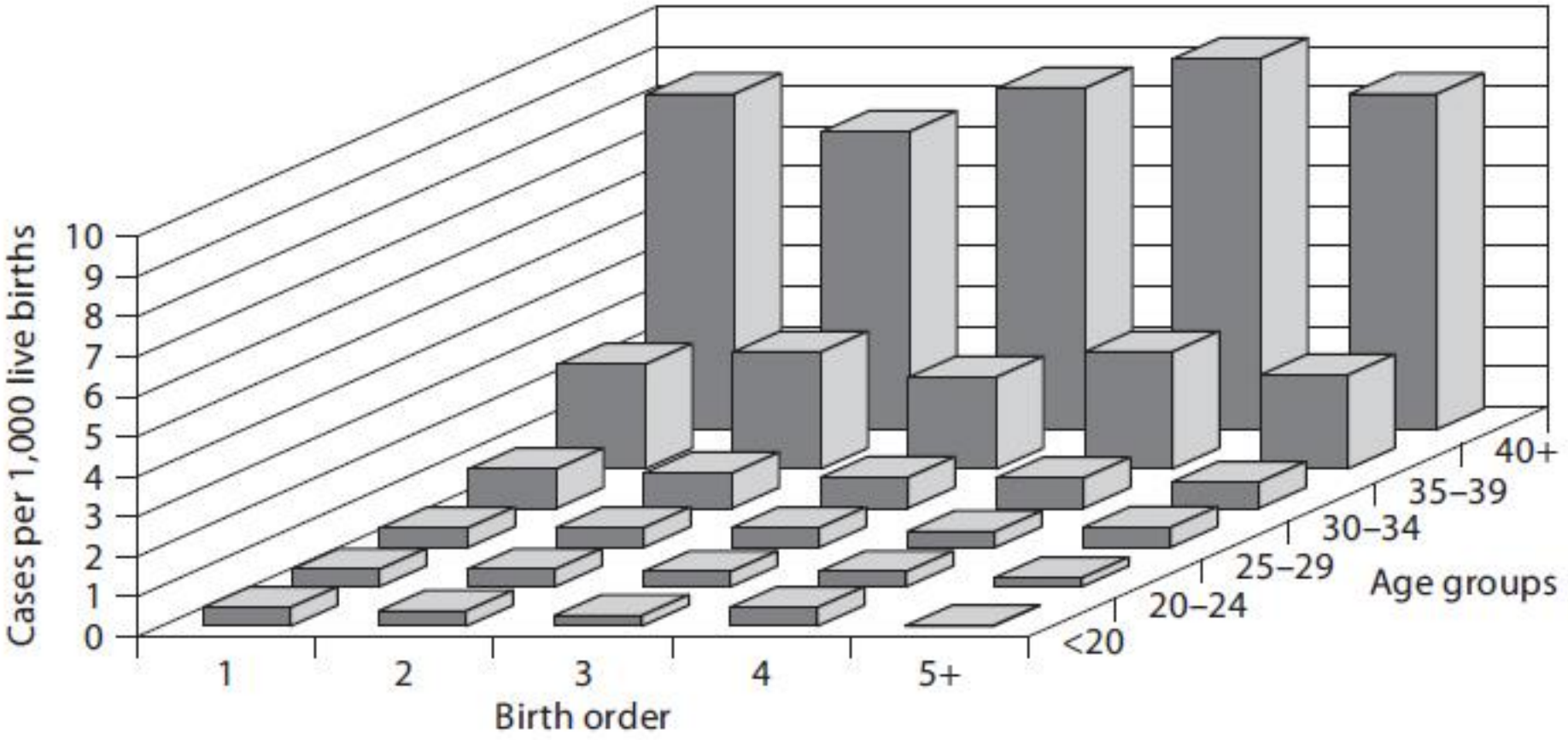
Effects of Confounding

- **Confounding can result in either overestimation or underestimation of the "true" association between E and D**
- **In this example, confounding led to an overestimation of the association**

Control of Confounding

- Design
  - Randomization
  - Restriction
  - Matching
- Analysis
  - Stratification
  - Adjustment

Stratification



Stratified by age group

Adjustment

- **We want to show the “true” relationship between treatment A and survival, controlling for stage of disease**
- **We must calculate an “adjusted RR” that removes the effect of the confounder**
- **Most computer programs will do this for you**

The Mantel- Haenszel Adjusted RR or OR

- **Calculates a weighted average of each of the 2 x 2 tables (the different strata)**
- **Takes out influence of third variable (e.g stage of disease) by making the result "as if" there were equal numbers of subjects in each group**
- **Computers can do this**

Adjusted RR

Early stage

	Survived	Died	
Treatment A	17	3	20
Placebo	9	1	10

RR = 0.94

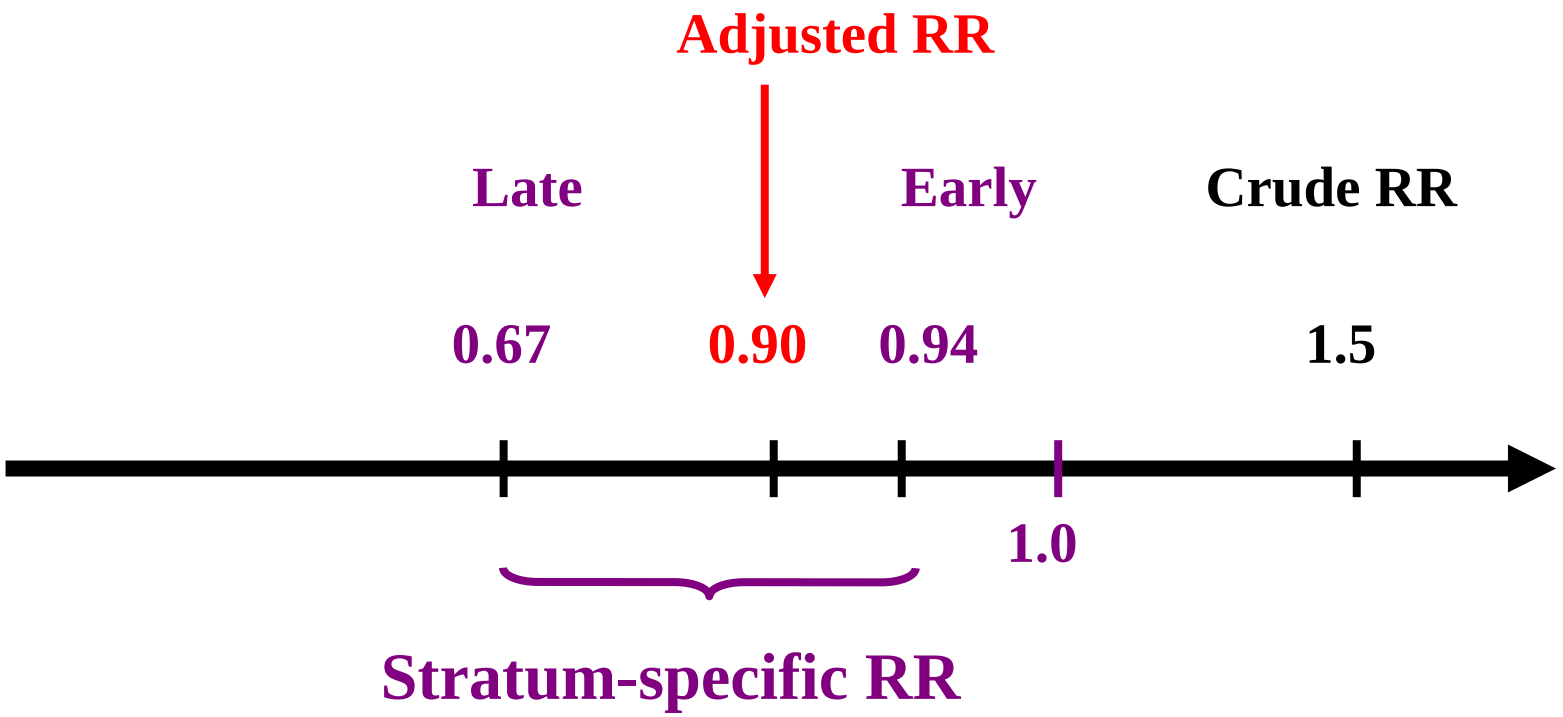
Late stage

	Survived	Died	
Treatment A	1	9	10
Placebo	3	17	20

RR = 0.67

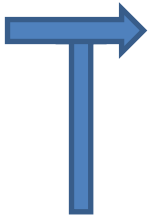
RR_{MH} = 0.90

Example of Adjusting for Confounding



Results of the National Health Examination Survey V (2013-2014)

Diagnosed with HT	n/N	(%)	OR (95% CI)	AOR (95% CI)
Marital status				
- Single	280/2,056	(13.6)	1	1
- Married	3,007/12,576	(23.9)	2.00 (1.75-2.28)	1.04 (0.89-1.20)
- Widow/divorce	1,188/3,203	(37.1)	3.74 (3.24-4.32)	1.17 (0.98-1.38)



Adjusted for age

Interaction (Effect Modification)

Effect Modification

The relation between an exposure and an outcome/disease is modified by the presence or absence of a third factor.

Example

<u>Question</u>: "Are women at greater risk of hip fracture than men?"

<u>Study design</u>: Case Control study (n=387)

E	**= exposure**	**=**	**female**
D	**= disease**	**=**	**hip fracture (HF)**
EM	**= effect modifier**	**=**	**age**

Example (N=387)

	Hip fracture	No hip fracture	
Women	**128**	**111**	**239**
Men	**66**	**82**	**148**
	194	**193**	

OR = (128x82) / (66x111) = 1.4

"The odds of suffering from a hip fracture are 1.4 times greater for women than men"

Example

- **This analysis only examined the exposure and disease**
- **We call this "crude analysis" (simple analysis)**
- **Is there some other variable that modifies exposure-disease association?**
- **Age?**

N=387 (stratified by age)

Young (n=188)

	HF+	HF-
Women	54	62
Men	38	34
	92	96

OR = 0.78

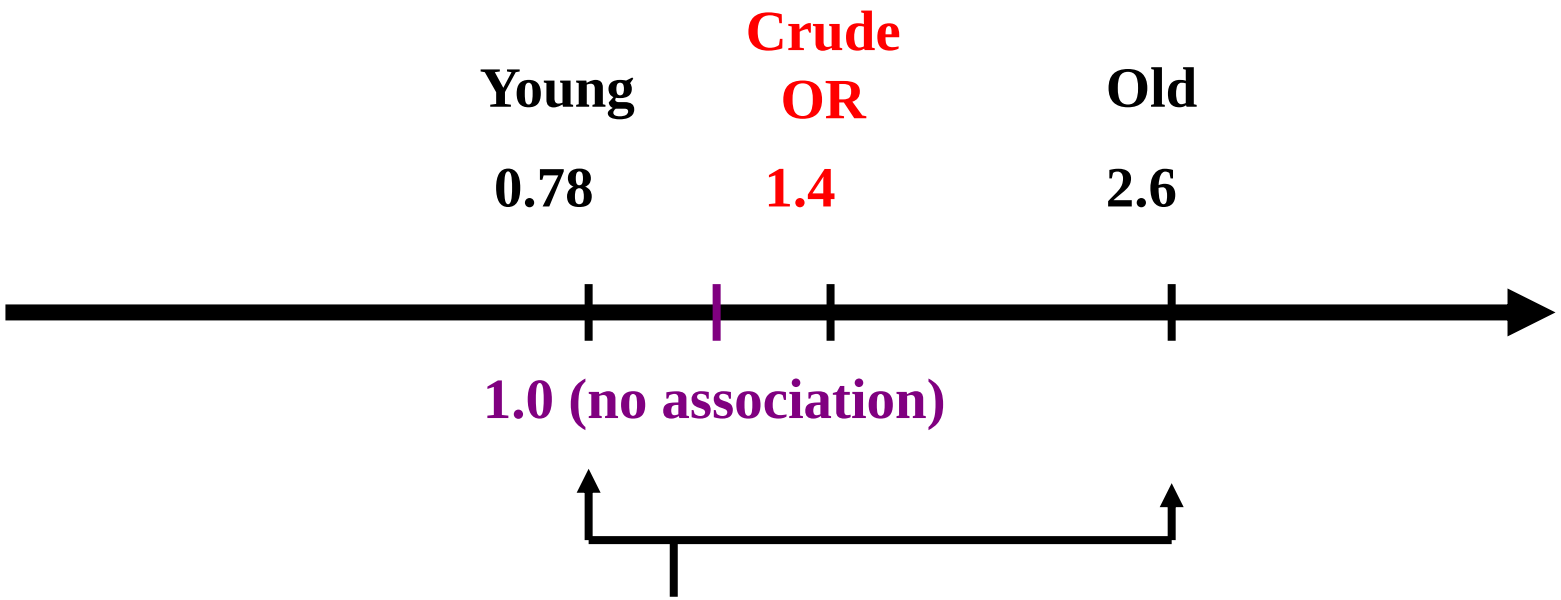
Old (n=199)

	HF+	HF-
	74	49
	28	48
	102	97

OR = 2.6

"The association between sex and hip fracture is modified by age."

Effect Modification



Effect Modification

- **The effect should be shown (present stratum specific RR/ORs rather than crude RR/OR-do not combine RR/OR!!!!).**
  - **Crude OR (1.4) OR Combined OR does not accurately describe the relation between sex and HF**
  - **Stratum specific ORs (0.78 and 2.6) better describe the true relation between sex and HF**

Confounding ≠ Effect Modification

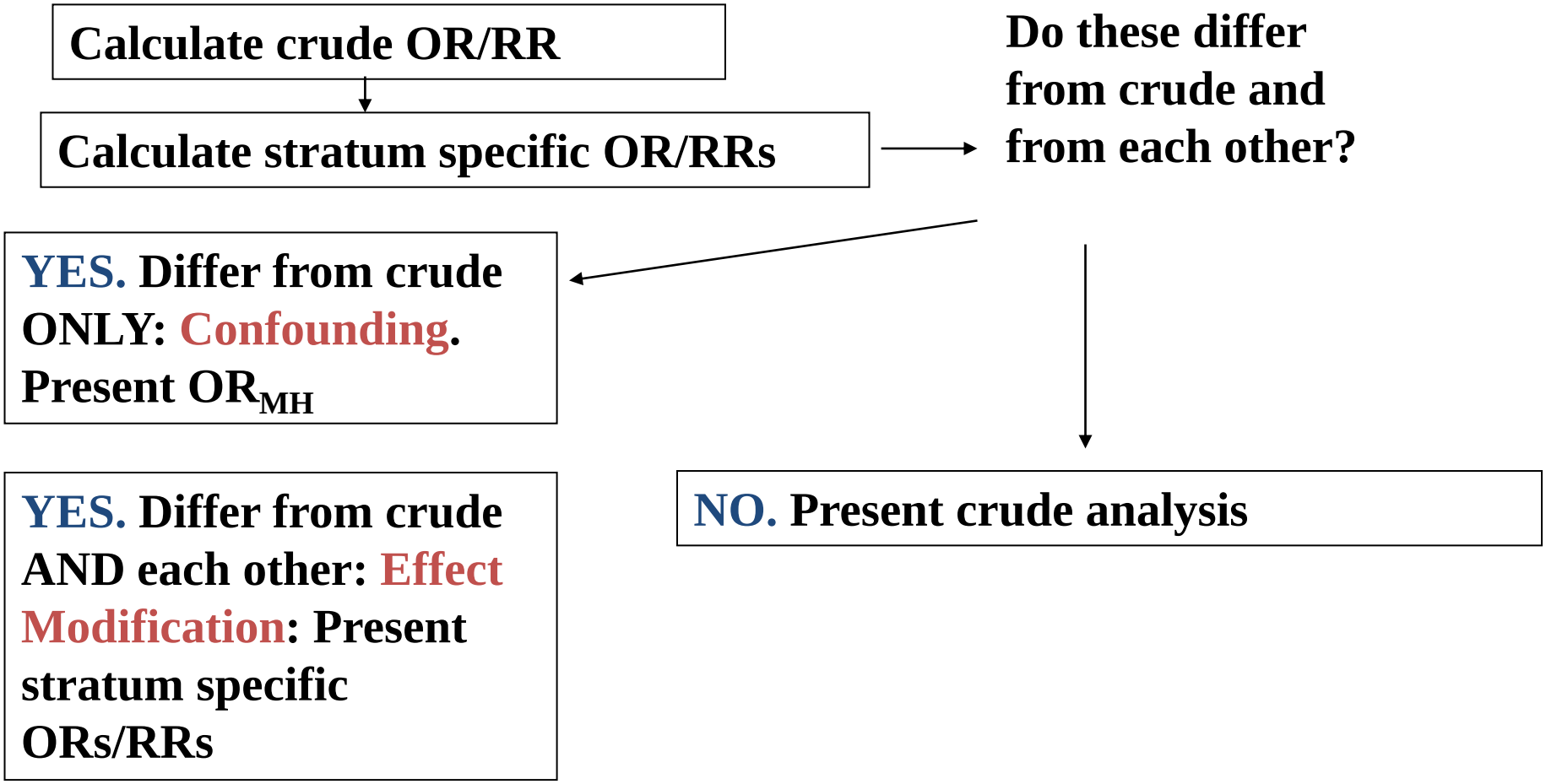
Confounding

- **Bias**
- **Control confounding**
- **Something to eliminate**

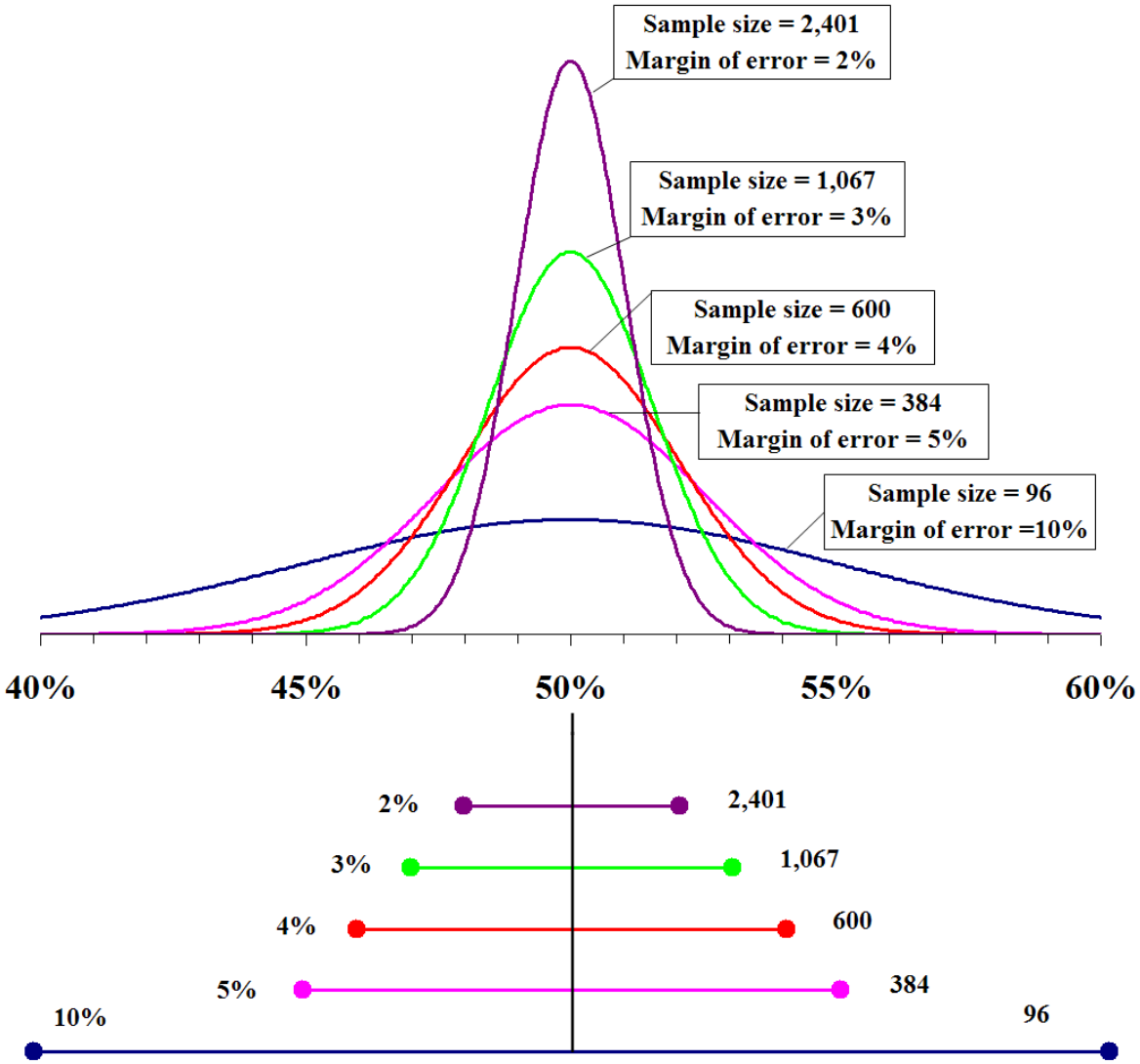
Effect Modification

- **Description of effect**
- **Report effect modification**
- **Something to show**

Assessment and Control of Confounding and Effect Modification with Stratified Analysis



Random error (chance)



External validity, Generalizability

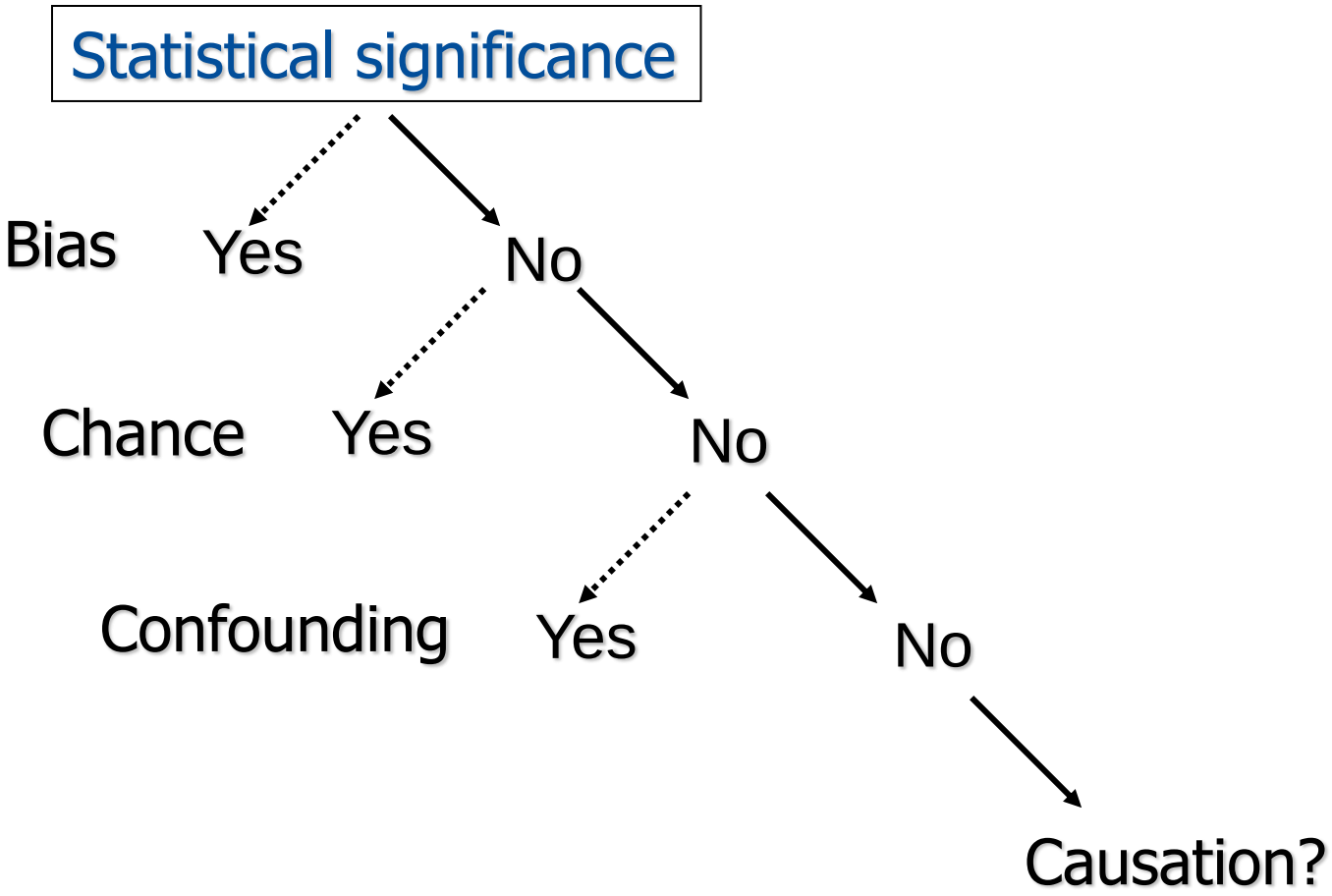
- How much would the study results still be true when apply to other populations or situations
  - Infer the study result from subjects to accessible population (geographical and time)
  - Infer the study result from accessible population to target population (demographic and clinical characteristics)
- Biomedical association has higher external validity when compared to behavior or social association

Example

Study result	External validity
A new COVID-19 vaccine showed efficacy in prevention of new infections.	high
A new surgical technique to remove gall bladder is better than the standard technique.	medium
A sex education curriculum helped reduce sexual risk behavior in a high school youth in Chiang Mai	low

Causation

Association & causation



Cause and effect association

- Koch's postulates
- Hill's criteria of causality
- Concept of sufficient cause and component causes

Koch's postulates

Robert Koch

1843-1910

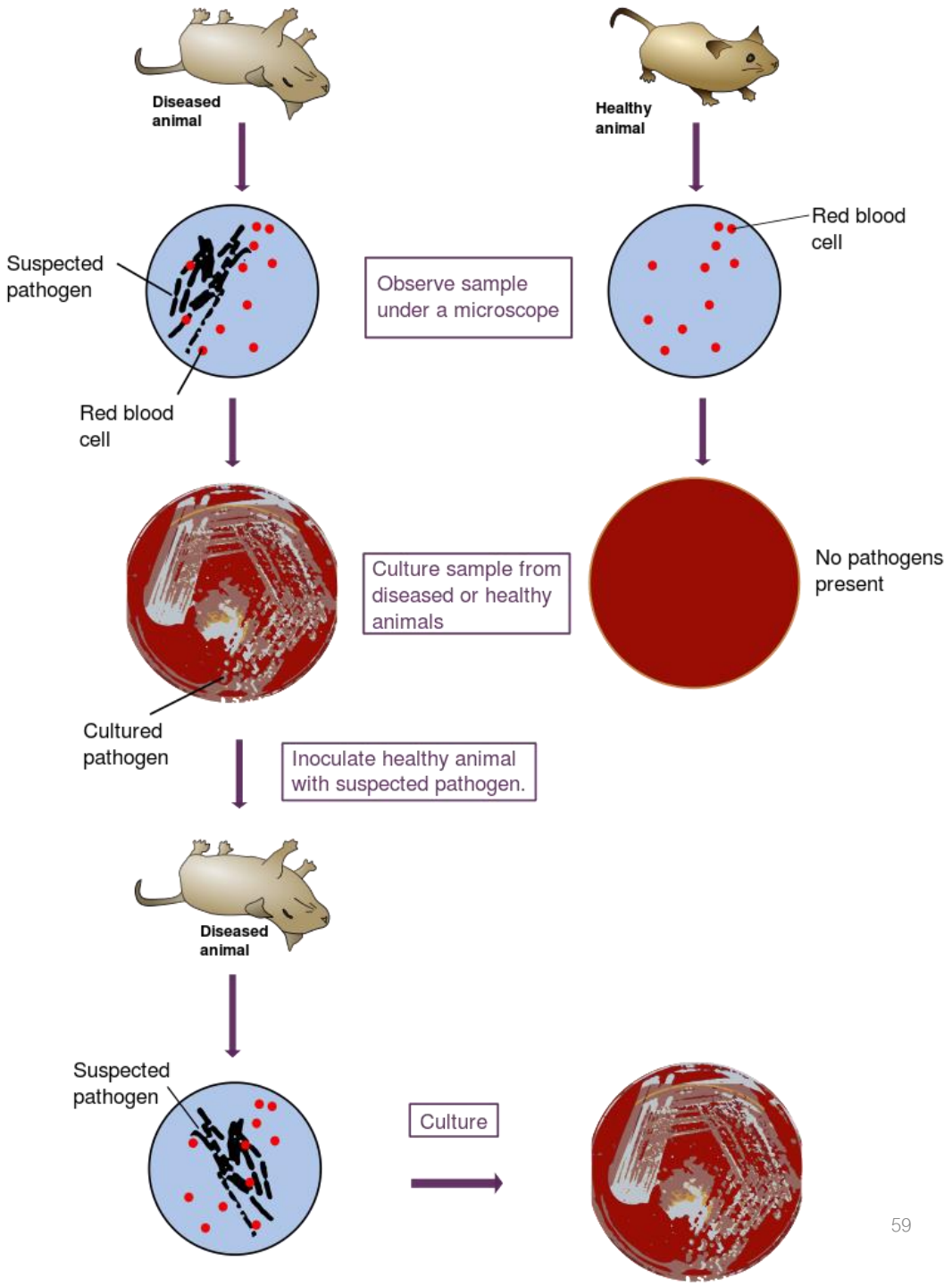
Koch's Postulates:

① The microorganism must be found in abundance in all organisms suffering from the disease, but should not be found in healthy organisms.

② The microorganism must be isolated from a diseased organism and grown in pure culture.

③ The cultured microorganism should cause disease when introduced into a healthy organism.

④ The microorganism must be reisolated from the inoculated, diseased experimental host and identified as being identical to the original specific causative agent.



Hill's criteria of causality

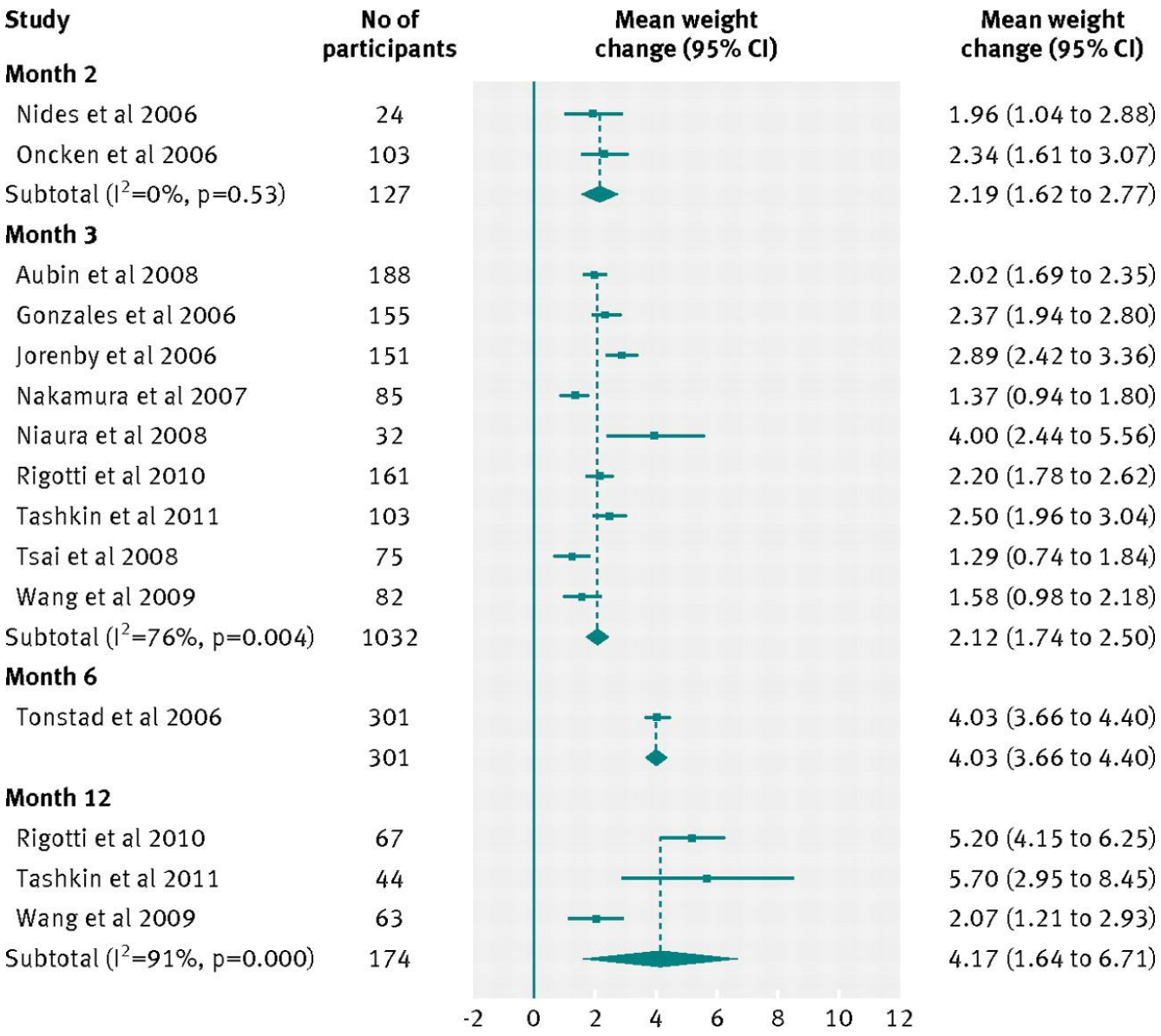
Sir Austin Bradford Hill

1897-1991

Hill's criteria of causality

- **Strength**: The stronger the association between the exposure to a treatment and the clinical outcome, the less likely it is influenced by an external variable.
- **Consistency**: Consistent findings observed by different persons in different places with different samples strengthens the likelihood of an effect.

Weight gain in smoker after quitting cigarette (Aubin HJ, et al. 2012)

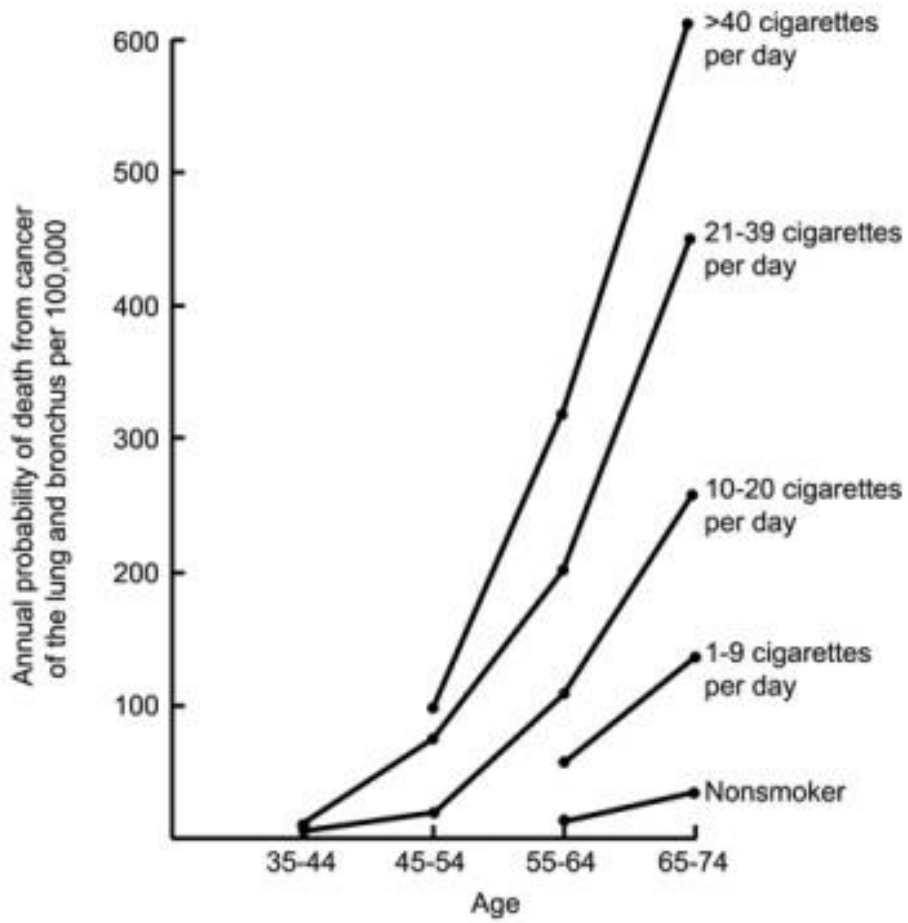


Hill's criteria of causality

- **Strength**: A small association does not mean that there is not a causal effect, though the larger the association, the more likely that it is causal.
- **Consistency**: Consistent findings observed by different persons in different places with different samples strengthens the likelihood of an effect.
- **Specificity**: Causation is likely if a very specific population at a specific site and disease with no other likely explanation. The more specific an association between a factor and an effect is, the bigger the probability of a causal relationship.

Hill's criteria of causality

- **Temporality**: The effect has to occur after the cause (and if there is an expected delay between the cause and expected effect, then the effect must occur after that delay).
- **Biological gradient**: Greater exposure should generally lead to greater incidence of the effect. However, in some cases, the mere presence of the factor can trigger the effect. In other cases, an inverse proportion is observed: greater exposure leads to lower incidence.



Death rates from cancer of the lung and bronchus in nonsmokers and smokers of various numbers of cigarettes per day.

Source: Kahn (1966).

Hill's criteria of causality

- **Temporality**: The effect has to occur after the cause (and if there is an expected delay between the cause and expected effect, then the effect must occur after that delay).
- **Biological gradient**: Greater exposure should generally lead to greater incidence of the effect. However, in some cases, the mere presence of the factor can trigger the effect. In other cases, an inverse proportion is observed: greater exposure leads to lower incidence.
- **Plausibility:** There is stronger support for causality if there is a likely biological or pharmacological mechanism that can explain the association between exposure to treatment and the outcome.

Hill's criteria of causality

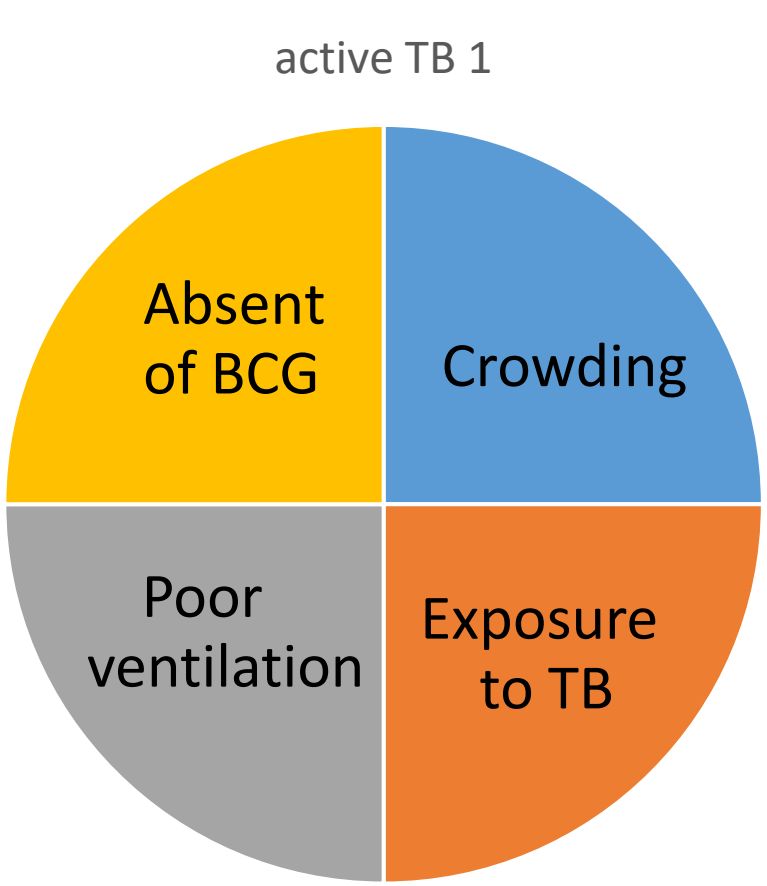
- **Coherence**: Causality between exposures to an intervention and the observed findings is supported when the association is coherent with current knowledge of the disease. Vice versa, conflicting or lack of supporting evidence would count against coherence.
- **Experiment**: If experimental manipulation of the exposure-outcome association impacts the outcome, experimental evidence is given. This delivers the strong support for causation.
- **Analogy**: The effect of similar factors may be considered as association.

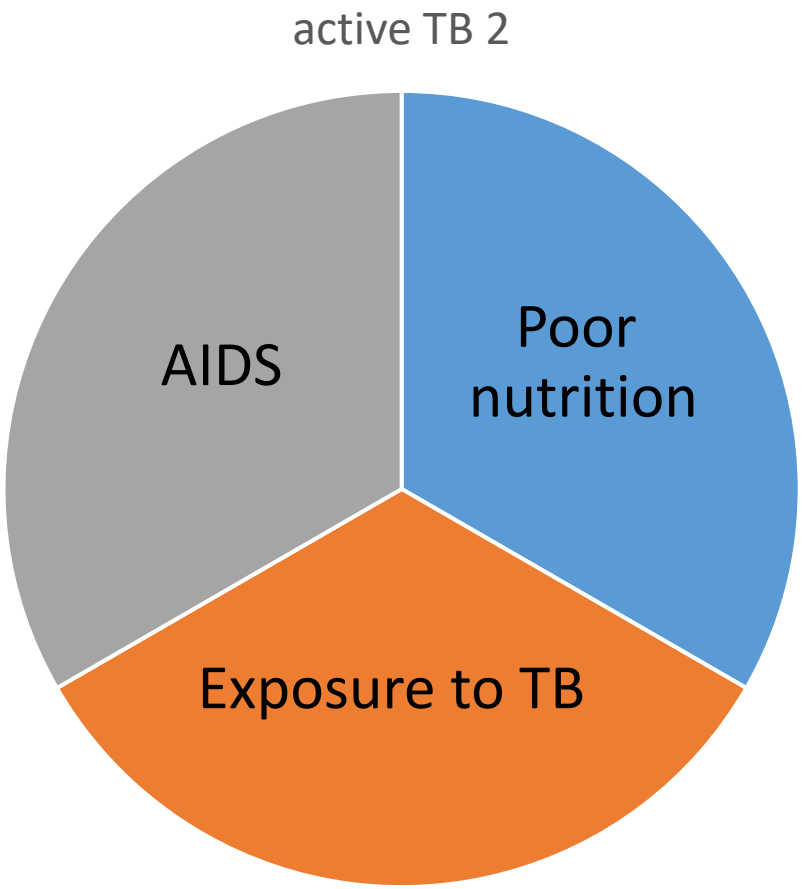
Concept of sufficient cause and component causes

- **Cause:** an agent, event, condition or characteristic which plays an essential role in producing an occurrence of the disease
- **Component causes:** disease outcomes have multiple contributing determinants that act together to produce a given instance of disease
- **Sufficient cause:** a minimum set of factors and circumstances that, if present in a given individual, will produce the disease

Modern Epidemiology 2nd edition (Rothman, 1998)

Example: TB


active TB 1
Absent of BCG
Crowding
Poor ventilation
Exposure to TB


active TB 2
AIDS
Poor nutrition
Exposure to TB

Sufficient cause model

A = necessary cause

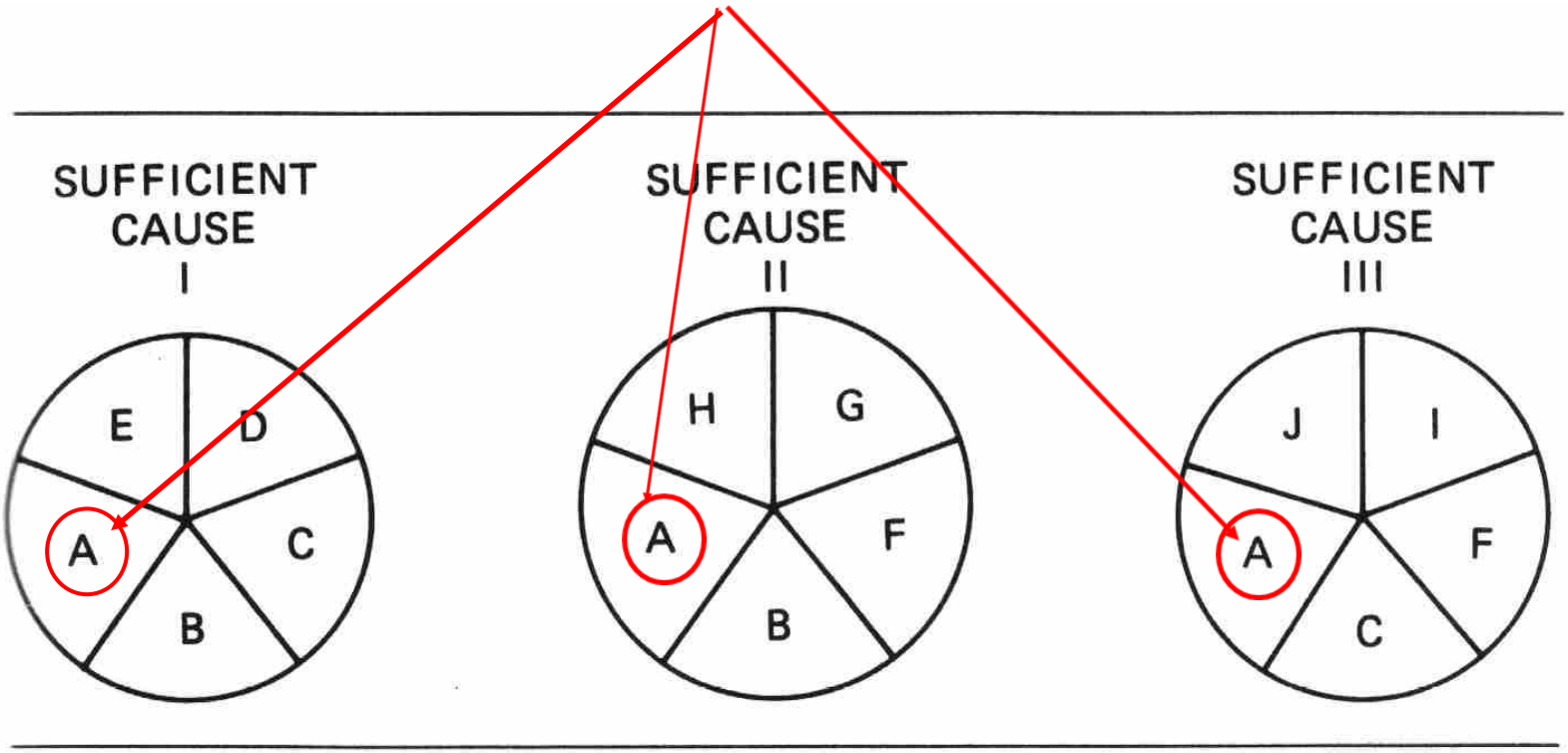
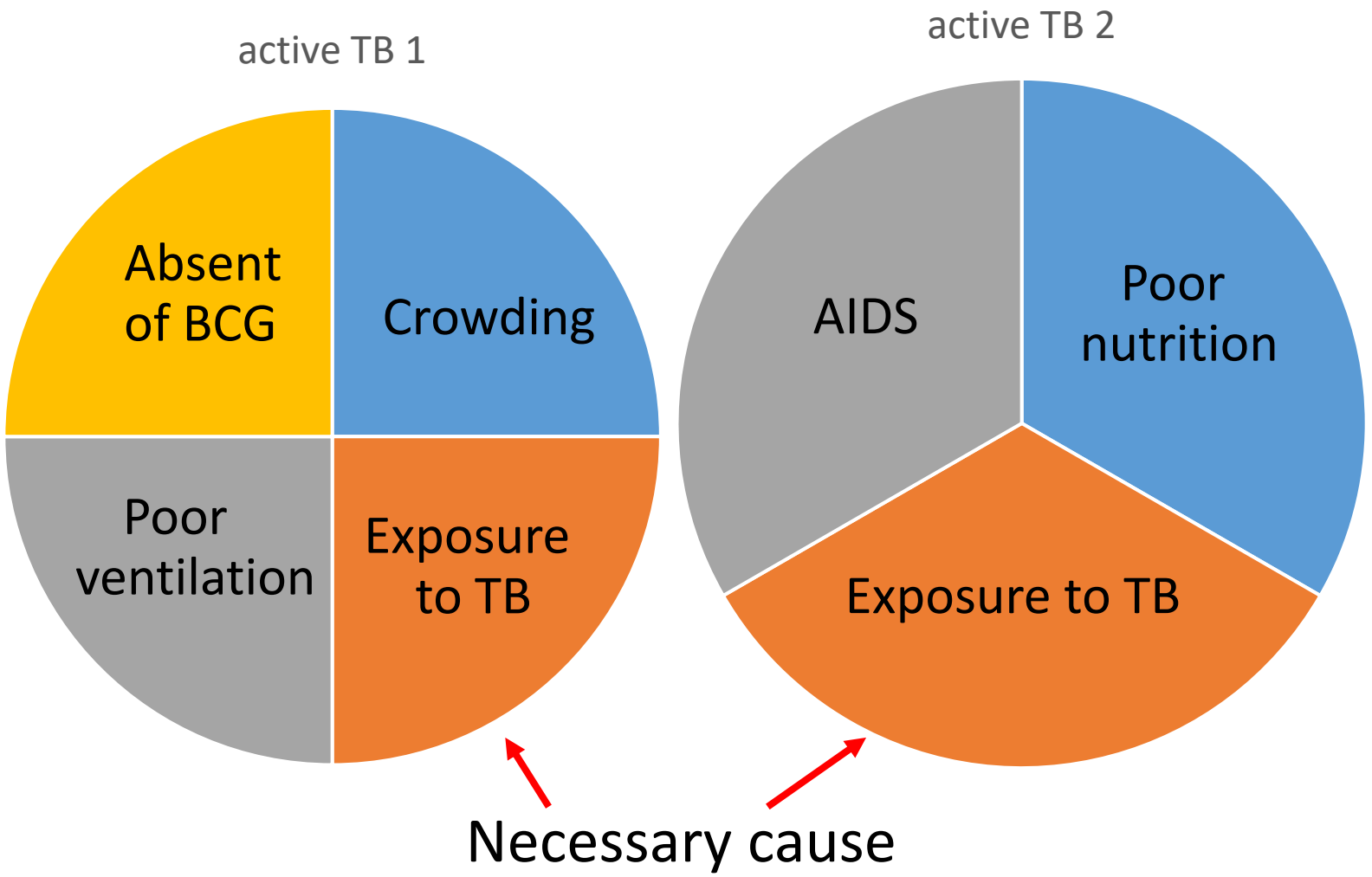


Fig. 2-1. Conceptual schematization of three sufficient causes for a disease [Rothman, 1976].

Example: TB


active TB 1
Absent of BCG
Crowding
Poor ventilation
Exposure to TB
active TB 2
AIDS
Poor nutrition
Exposure to TB
Necessary cause