MRC Epidemiology Unit

UNIVERSITY OF CAMBRIDGE

Aetiological Epidemiology

Ken Ong

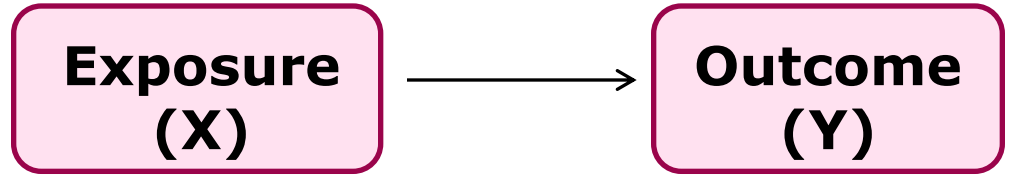
MRC Epidemiology Unit

CDS April 2019

Wellcome Trust – MRC IMS Institute of Metabolic Science

Cambridge Institute of Public Health

Aetiological Epidemiology: Does **X** cause **Y**?



Terminology:
"Risk difference"

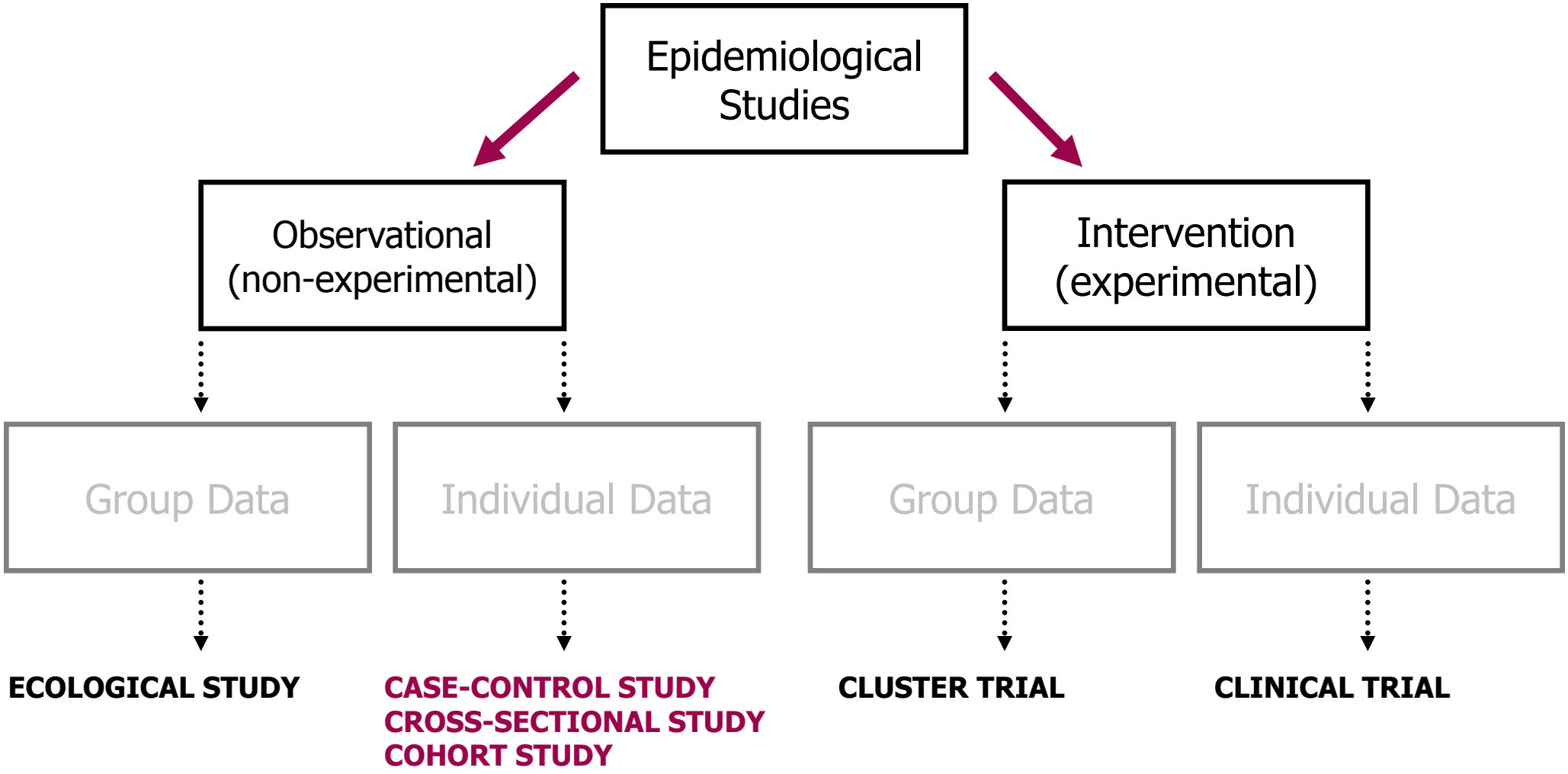
a) Absolute difference…

b) Relative difference:
- *Relative risk*
- *Odds ratio*
- *Hazard ratio*
- *Incidence rate ratio*

Terminology:
"Risk"

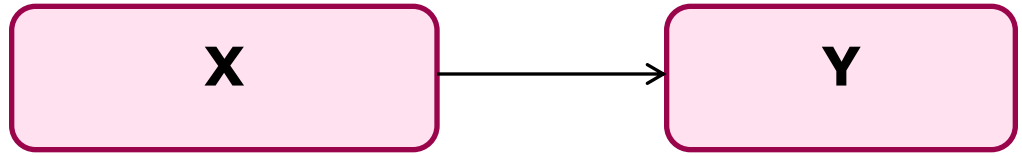
- *Prevalence*
- *Point prevalence*
- *Incidence*
- *Cumulative incidence*

Study Design Overview

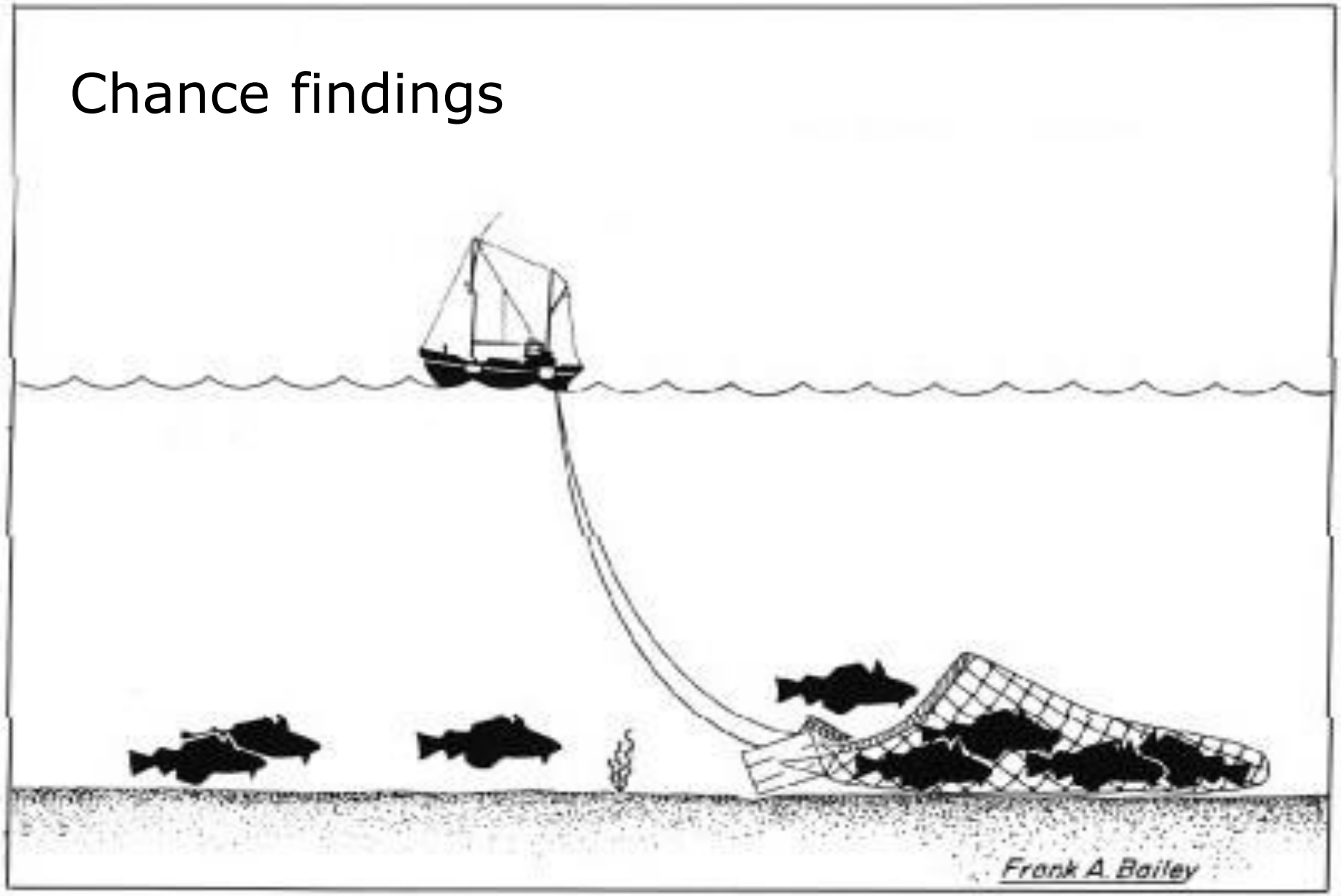


Observed Associations – why aren't they all causal?

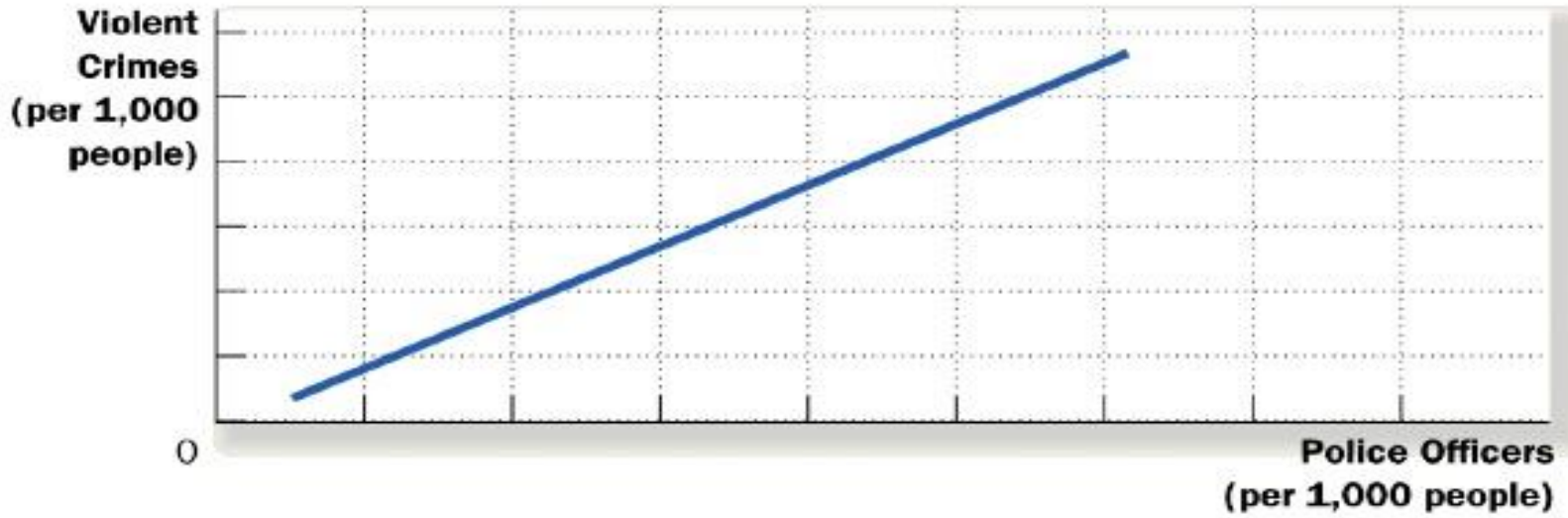
What are the possible reasons for an observed association between X & Y?



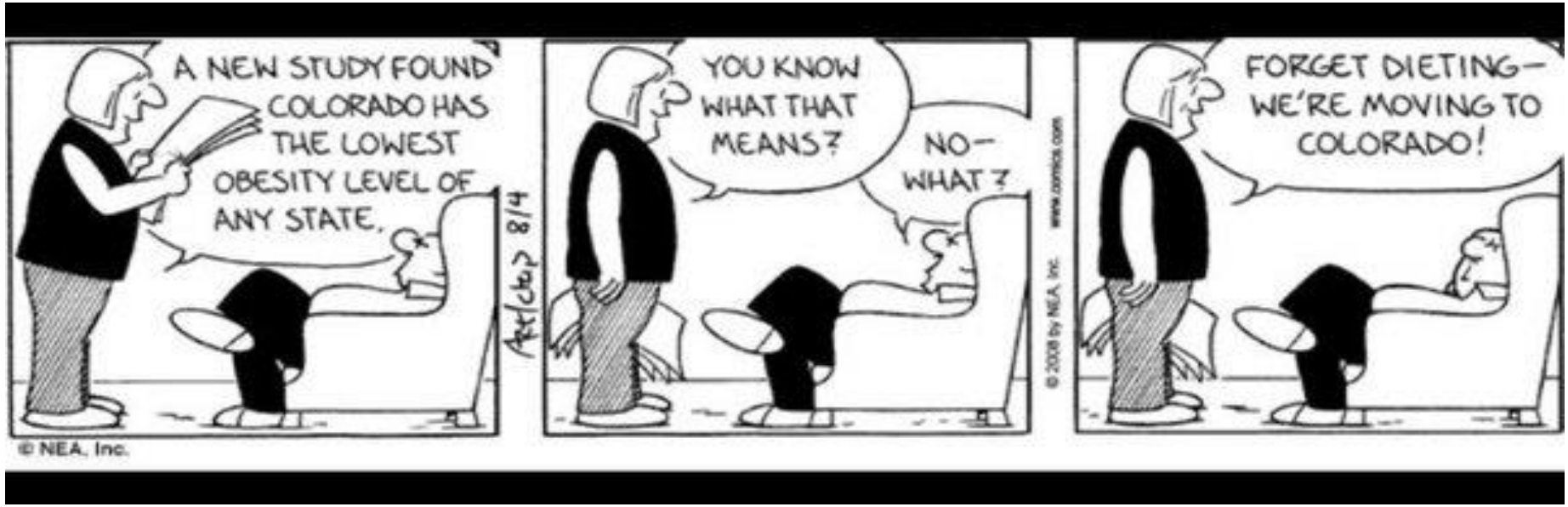
Chance findings



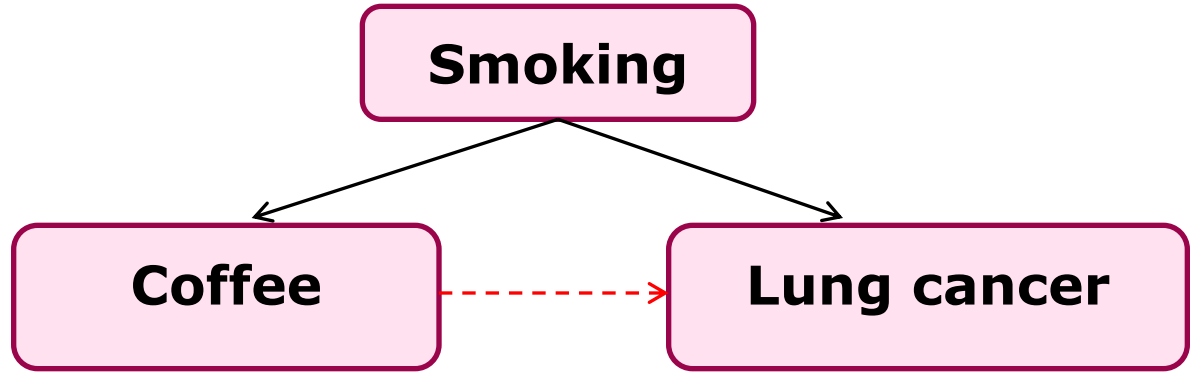
Reverse causality



Confounding

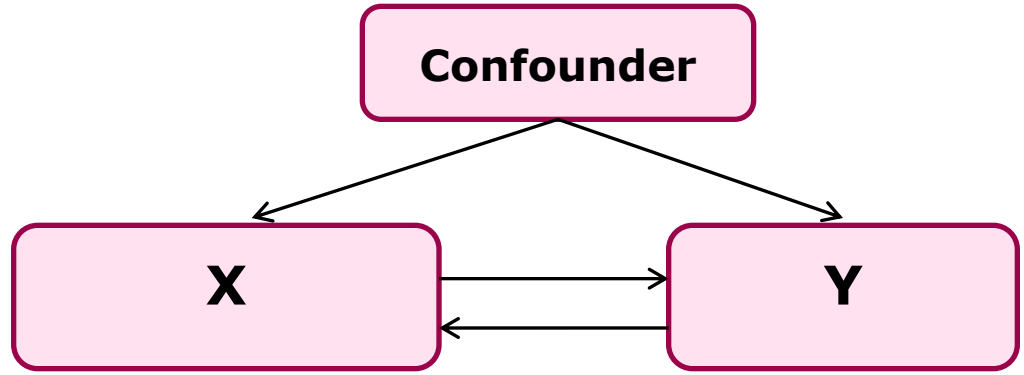


Confounding



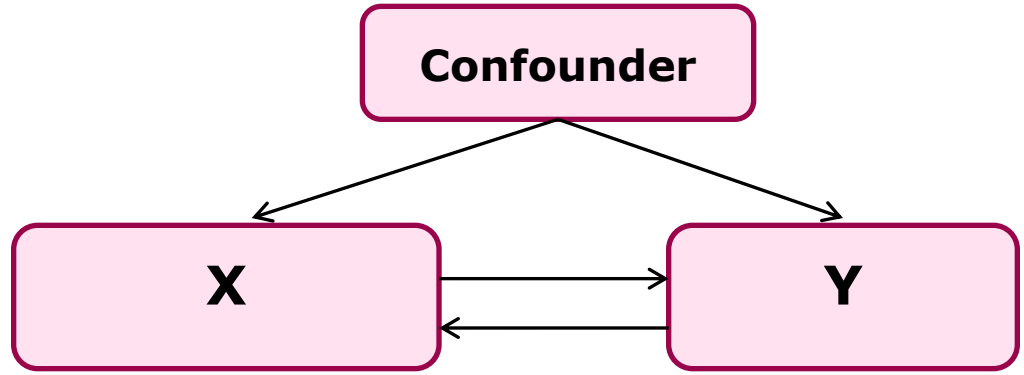
- ***Household smoking***
- ***Low SES***
- ***Poor air quality***
- ***…..***

Observed Associations – Possible explanations?



- ***X causes Y***
- Chance finding (*False Positive*)
- Y causes X (*Reverse causality*)
- Both X & Y are downstream of *Z* (‘*confounder’*)

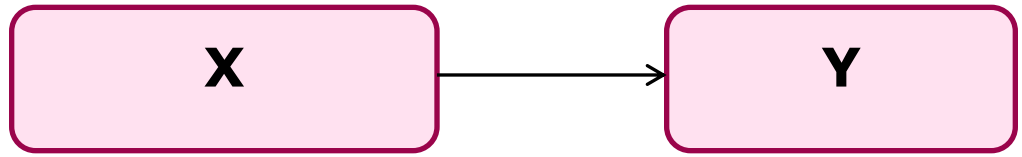
Observed Associations – How to infer causality



- Chance finding
  - Use a Robust P-value threshold; Seek Confirmation
- Y causes X
  - Temporal Study design; Adjust for Baseline differences
- Confounding
  - Measure & Control (e.g. Adjust) for Confounders

Observed Associations – why aren't they all causal?

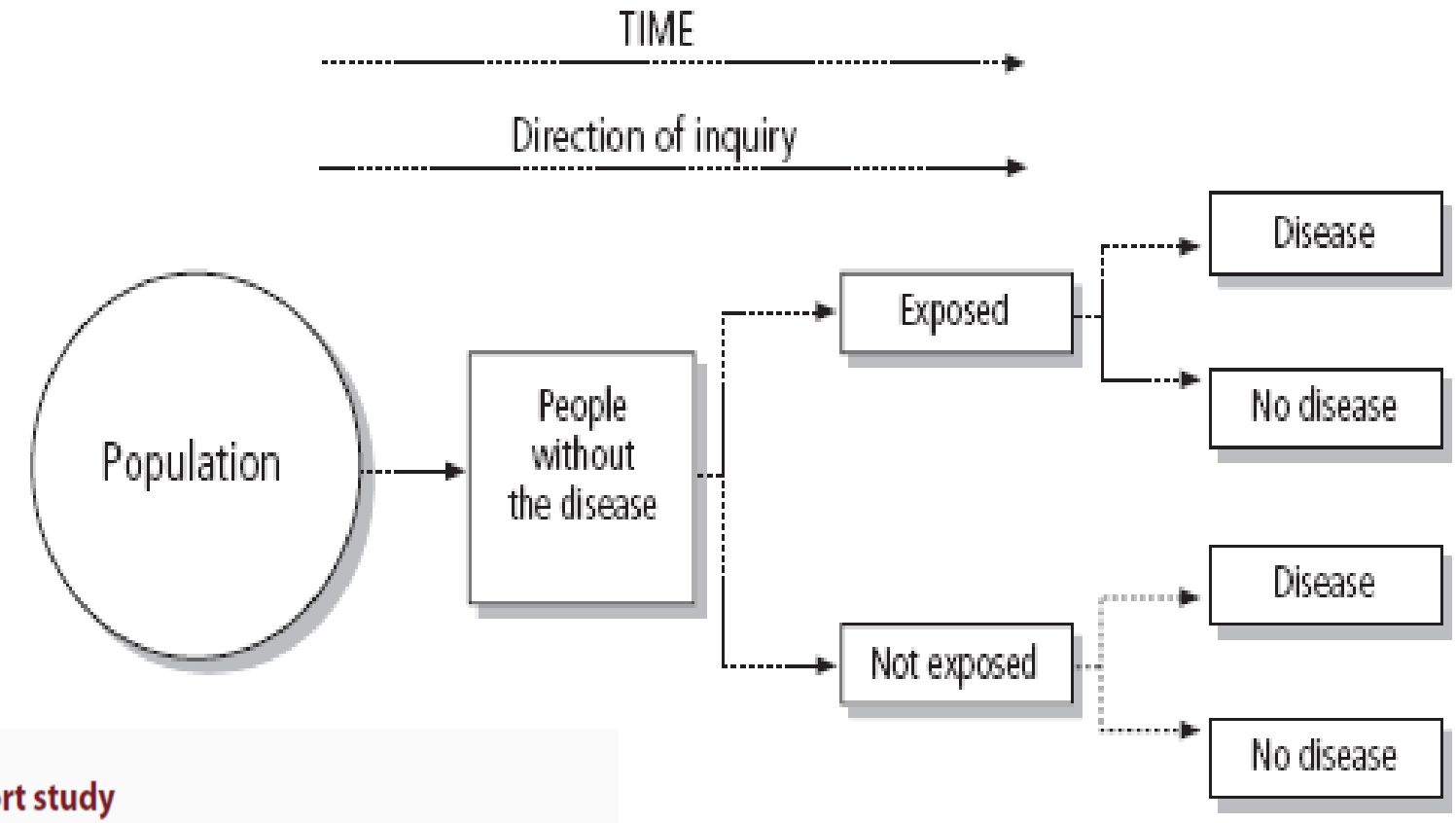
- ***X causes Y***
- Chance finding (*False Positive*)
- Y causes X (*Reverse causality*)
- Both X & Y are downstream of *Z* ('*confounder'*)



BIAS

- **Selection** of study population / study groups
- **Measurement** of X or Y

Cohort Study



- **Prospective cohort study**
  - Concurrent cohort study or longitudinal study
- **Retrospective cohort study**
  - Non-concurrent cohort or historical cohort study

Measures of Risk Difference

Absolute Risk difference = $I_e - I_{ue}$

- What is the additional risk of disease following exposure, over and above that experienced by people who are not exposed?

Relative Risk difference (relative risk) = I_e / I_{ue}

- How many times more likely are exposed persons to get the disease relative to nonexposed persons?

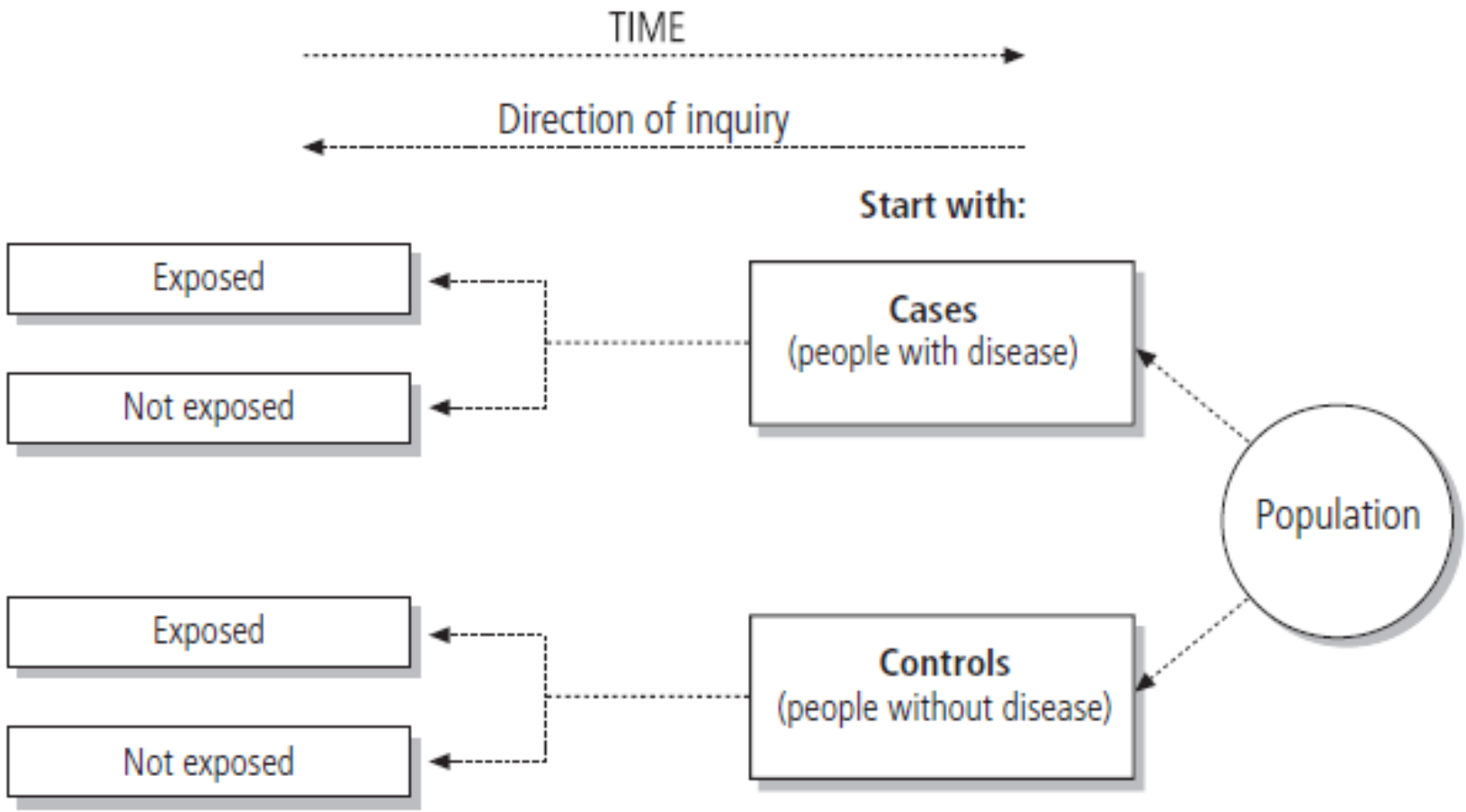
Advantages of Cohort Studies

- Provides estimates of incidence
- Can study natural history
- Can deal with exposure and disease as a continuum
- Rare exposures can be studied
- Multiple disease outcomes from a given exposure can be studied
- Effect of change in exposure status can be assessed
- Information on exposure precedes disease
- Can do nested case-control studies

Disadvantages of Cohort Studies

- Expense – money, manpower
- Time, loss to follow-up
- Difficult to maintain consistent measurements
- Can test only fairly specific hypotheses
- Rarely topical by the time completed
- Needs large numbers
- Unsuitable for rare diseases
- Organizational complexities
- Needs fixed and stable populations
- Not easy to know induction and latent periods
- Defining exposure, effect, and subject not easy

Case Control Study



Advantages of Case-control Studies

- Quick and cheap
- Topical
- Can study many causes
- Can study rare diseases
- Smaller numbers
- Can use elaborate tests
- No problem with losses to follow up
- Can evaluate preventive measures

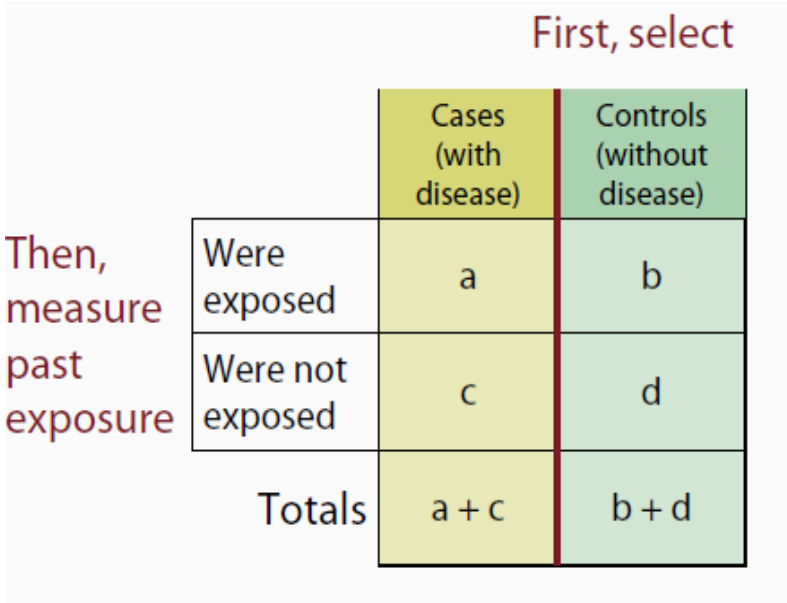
Disadvantages of Case-control Studies

- Biases
- Cannot deal with disease as a continuum
- Cannot estimate incidence
- Difficulty defining controls

Controls should represent people who would have been eligible to be included as cases had they developed the disease

- *They should be from the same specified population as the cases and be representative of the exposures being studied within that population*

Measures of association in case control studies



First, select

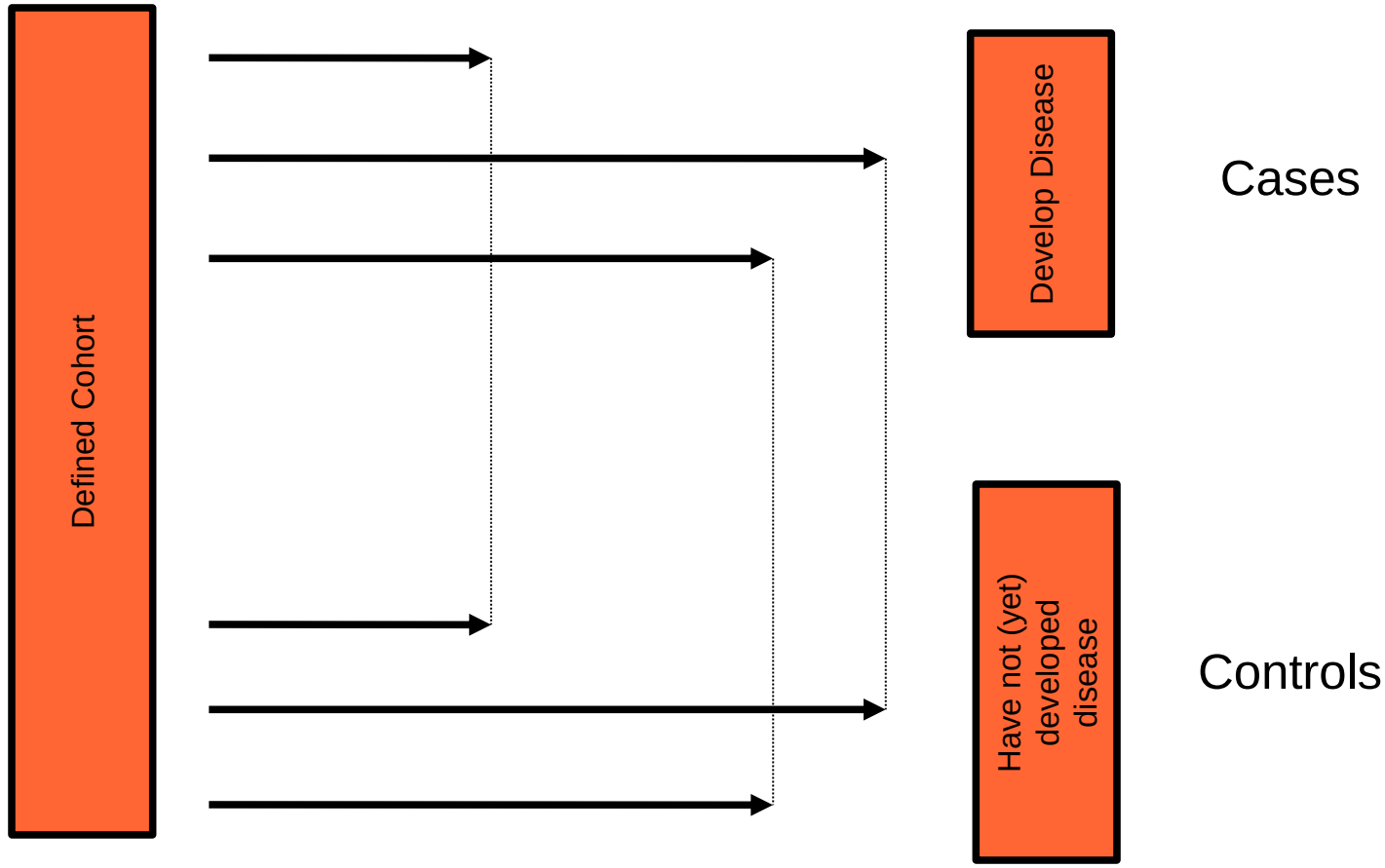
Then, measure past exposure	Cases (with disease)	Controls (without disease)
Were exposed	a	b
Were not exposed	c	d
Totals	a + c	b + d

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

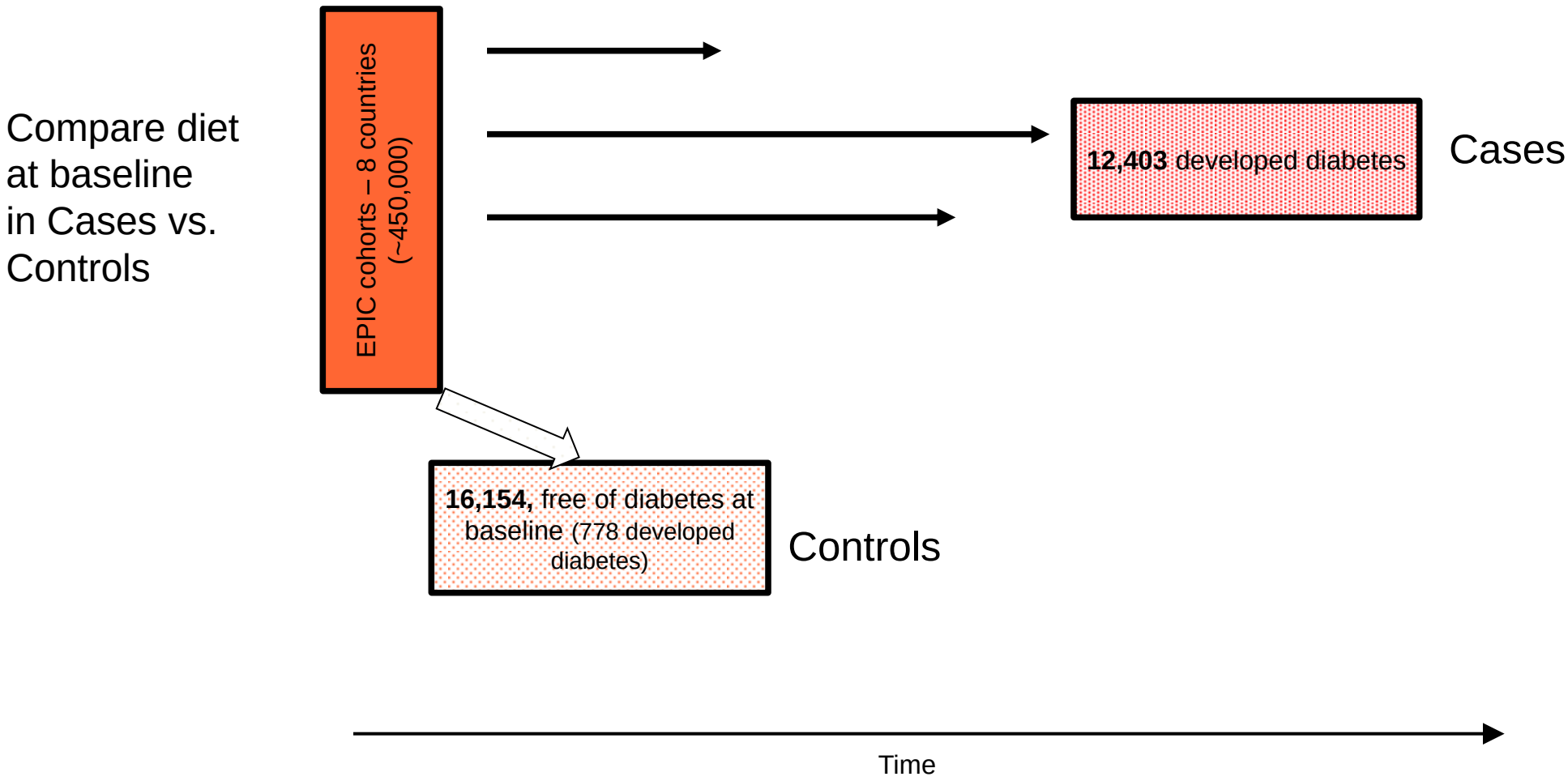
Odds ratio is an approximation of relative risk if:

1. Outcome is rare
2. Cases are a random sample of incident cases in the population
3. Controls are random sample of **non-diseased** in the population

Nested case control study



Case-cohort design, the Epic-Interact Study



Cause and effect

Bradford Hill's criteria

Sir Austin Bradford Hill J Roy Soc Med 1965;58:295-300

- Strength of the association
  - relative risk
- Consistency
  - lots of well performed studies
- Specificity
  - one cause/effect
- Temporal relationship
  - cause precedes the effect
- Dose response
  - higher exposure more disease
- Biological plausibility
- Independence
  - of **confounders and bias**
- Reversibility
  - reduced exposure has opposite effect

Cause & effect: Bradford Hill criteria

Strength of association

- Relative risk for smokers
  - Lung cancer = 4-6x
  - Renal cancer = 1.1-1.6x
- Relative risk of hepatitis B
  - Hepatocellular carcinoma = nearly 300x

Cause & effect: Bradford Hill criteria

Consistency

- Several studies
- Different times
- Different settings
- Different types of patient

But several studies may ALL make the same mistake!

- Therefore, consistency across different study designs is reassuring

Cause & effect: Bradford Hill criteria

Dose-response relationship

Lung cancer deaths per 100,000/year

- Non-smoker: 10
- 1-14 Cigs/day 76
- 15-24 Cigs/day 127
- 25+ Cigs/day 251

Cause & effect: Bradford Hill criteria

Biological plausibility

- A good biological explanation for **X → Y**
- BUT: Lack of biological plausibility may indicate limitations of scientific knowledge

Specificity

- One cause, one effect

Cause & effect: Bradford Hill criteria

Reversibility

Years since stopped smoking	Relative risk in ex-smokers vs never smokers
0	15.8
<5	10.7
5-9	5.9
10-14	4.7
15+	2.0

Control for confounding

- **Study design**
  - (Randomization)
  - Restriction
  - Matching
  - **Use of instrumental variable**

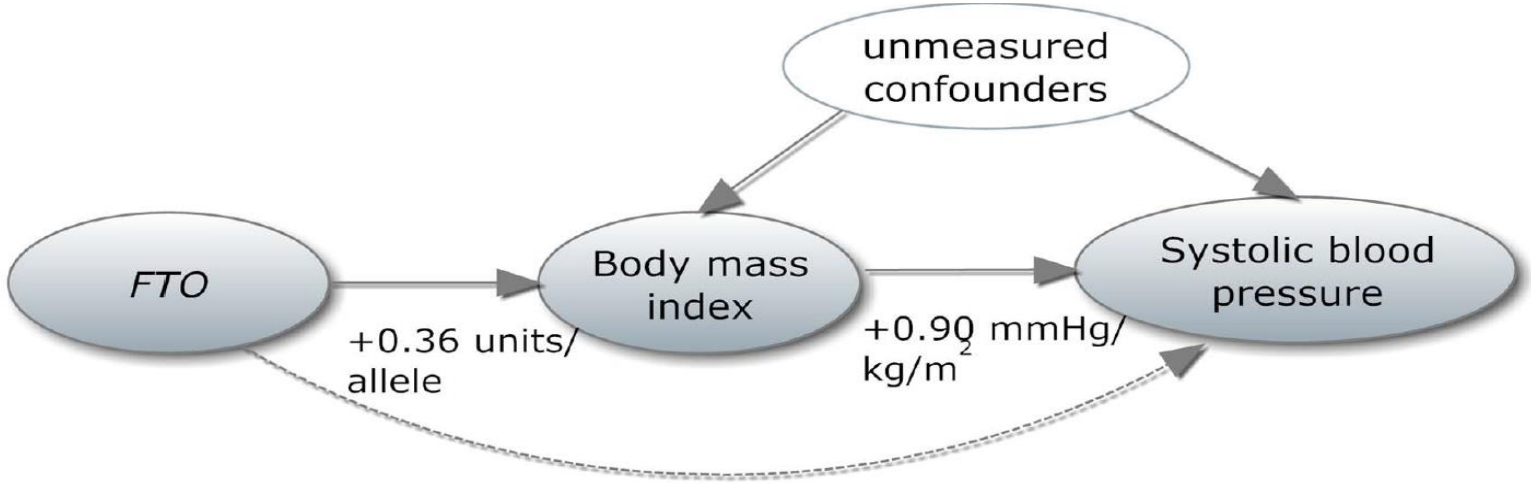
  Analysis
  - Stratification
  - Multi-variable analysis

Instrumental variable

Allows un-confounded estimation of causal effects under certain conditions/assumptions

1. It is associated with the exposure
2. ***It affects the outcome only through the exposure***
3. It does not share any common cause with the outcome

Using the *FTO* genotype as an instrumental variable to assess the unconfounded effect of BMI on systolic blood pressure



IV estimator for causal relationship of BMI and systolic blood pressure: +0.89 mmHg/(kg/m^2)

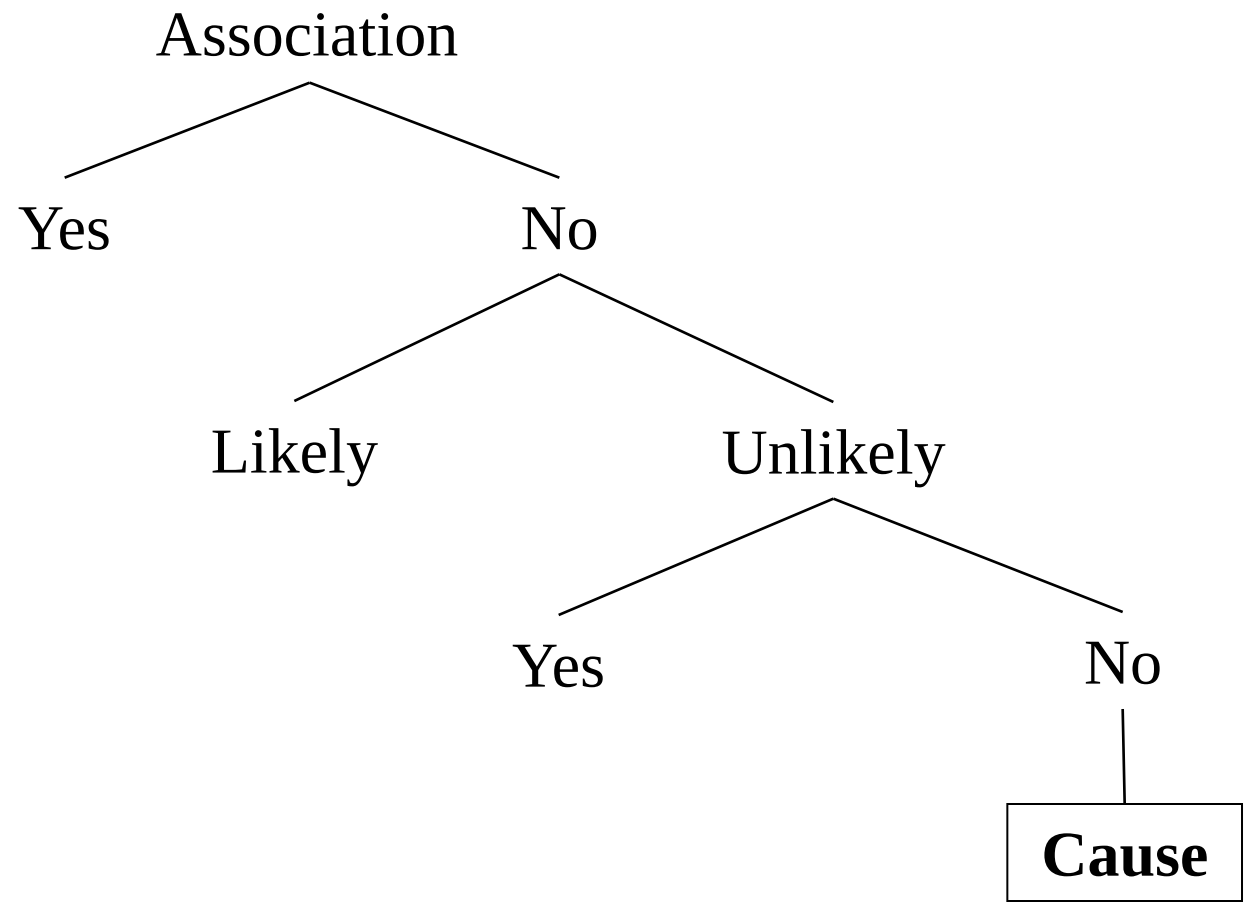
Fall et al, *PLoS Medicine*, 2013

Weighing Causal Evidence

Strength of evidence	Study Designs	Finding
Strongest	Clinical trial Cohort study Case control study Cross-sectional study Ecological study Case series Case reports	Temporality
		Strength of association
		Reversibility
		Dose response
		Consistency
		Biologic plausibility
		Specificity
Weakest		Analogy

Establishing cause

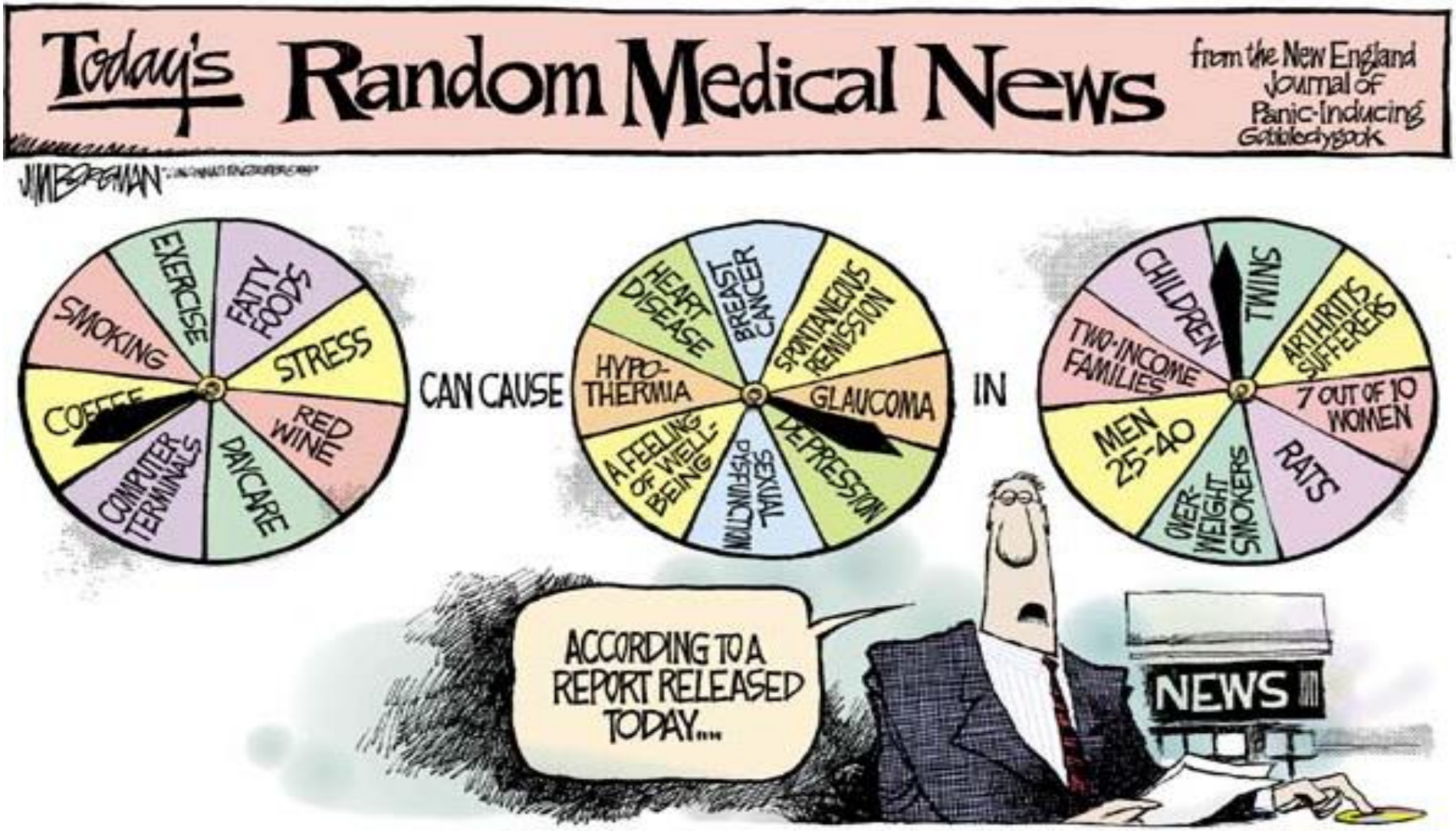
Explanation	Finding



Bias in selection or measurement

Chance

Confounding


Today's Random Medical News
from the New England Journal of Panic-Inducing Gobbledygook
JIM BORGMAN
SMOKING
EXERCISE
FATTY FOODS
STRESS
RED WINE
DAYCARE
COMPUTER TERMINALS
COFFEE
CAN CAUSE
HEART DISEASE
BREAST CANCER
SPONTANEOUS REMISSION
HYPO-THERMIA
GLAUCOMA
A FEELING OF WELL-BEING
SEXUAL DYSFUNCTION
DEPRESSION
IN
CHILDREN
TWINS
ARTHRITIS SUFFERERS
TWO-INCOME FAMILIES
7 OUT OF 10 WOMEN
MEN 25-40
RATS
OVER-WEIGHT SMOKERS
ACCORDING TO A REPORT RELEASED TODAY...
NEWS