Lecture 13: Bias

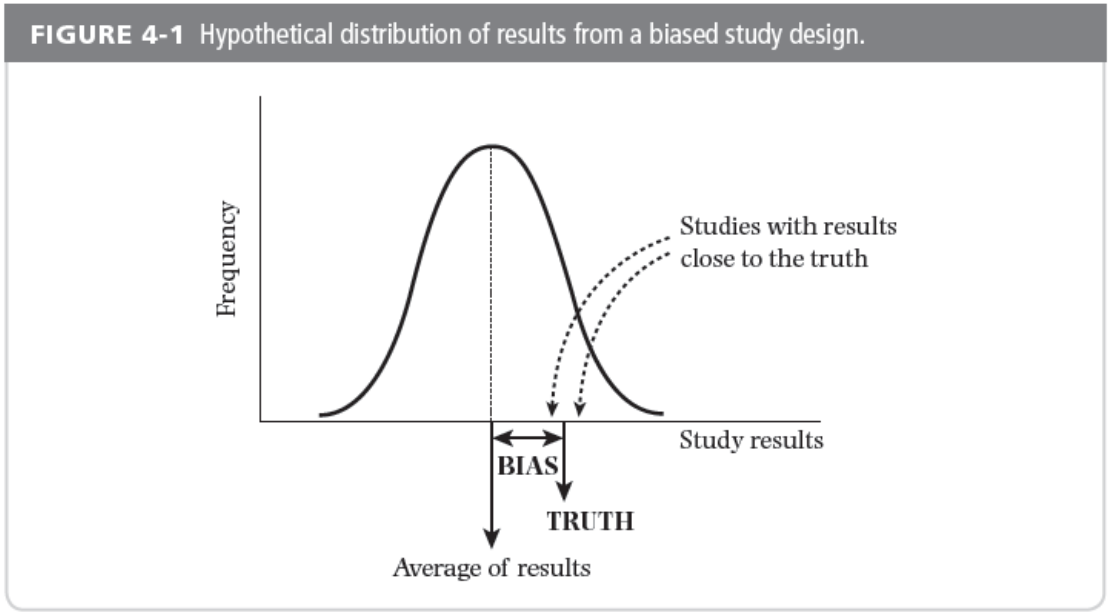
Lecture prepared by Dr. Hailey Banack, PhD

What is bias?

Systematic deviation of results or inferences from the truth [Porta, 2008]

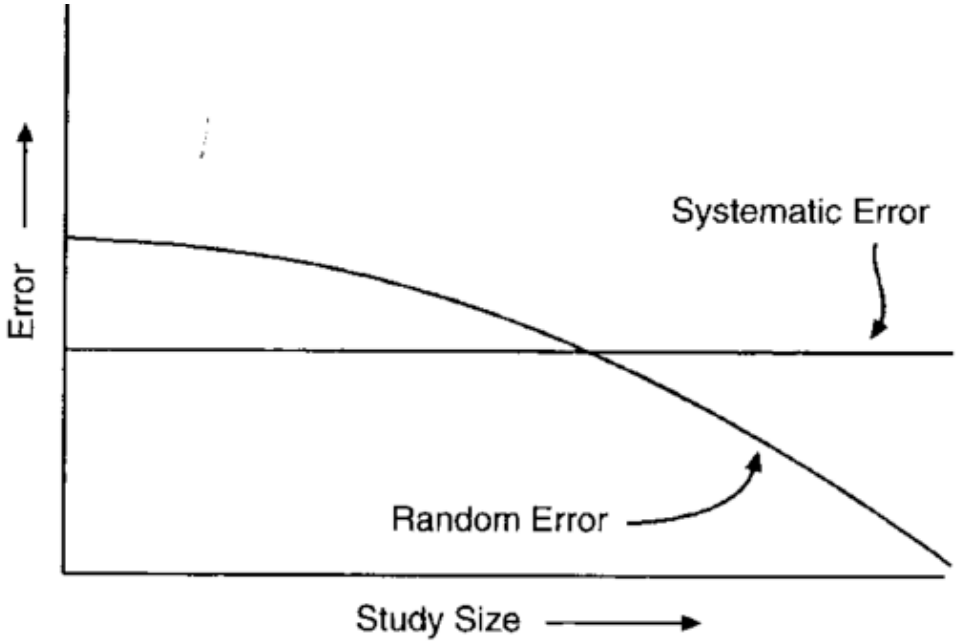
- Bias = lack of validity

Can occur at the study design stage, during the conduct of a study, or while analysis is being done



Systematic vs. Random Error

Systematic error (=bias/lack of validity) is not the same as random error (=lack of precision)



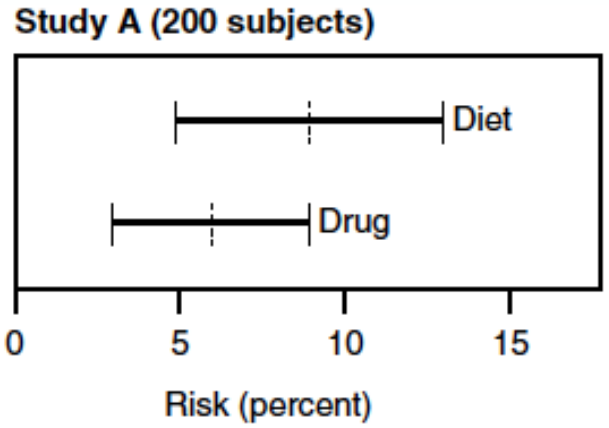
Systematic error is present regardless of study size, even in an infinitely large study

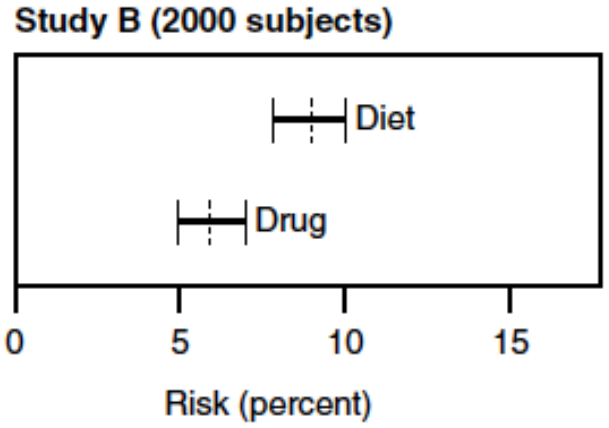
Random error decreases as study size increases

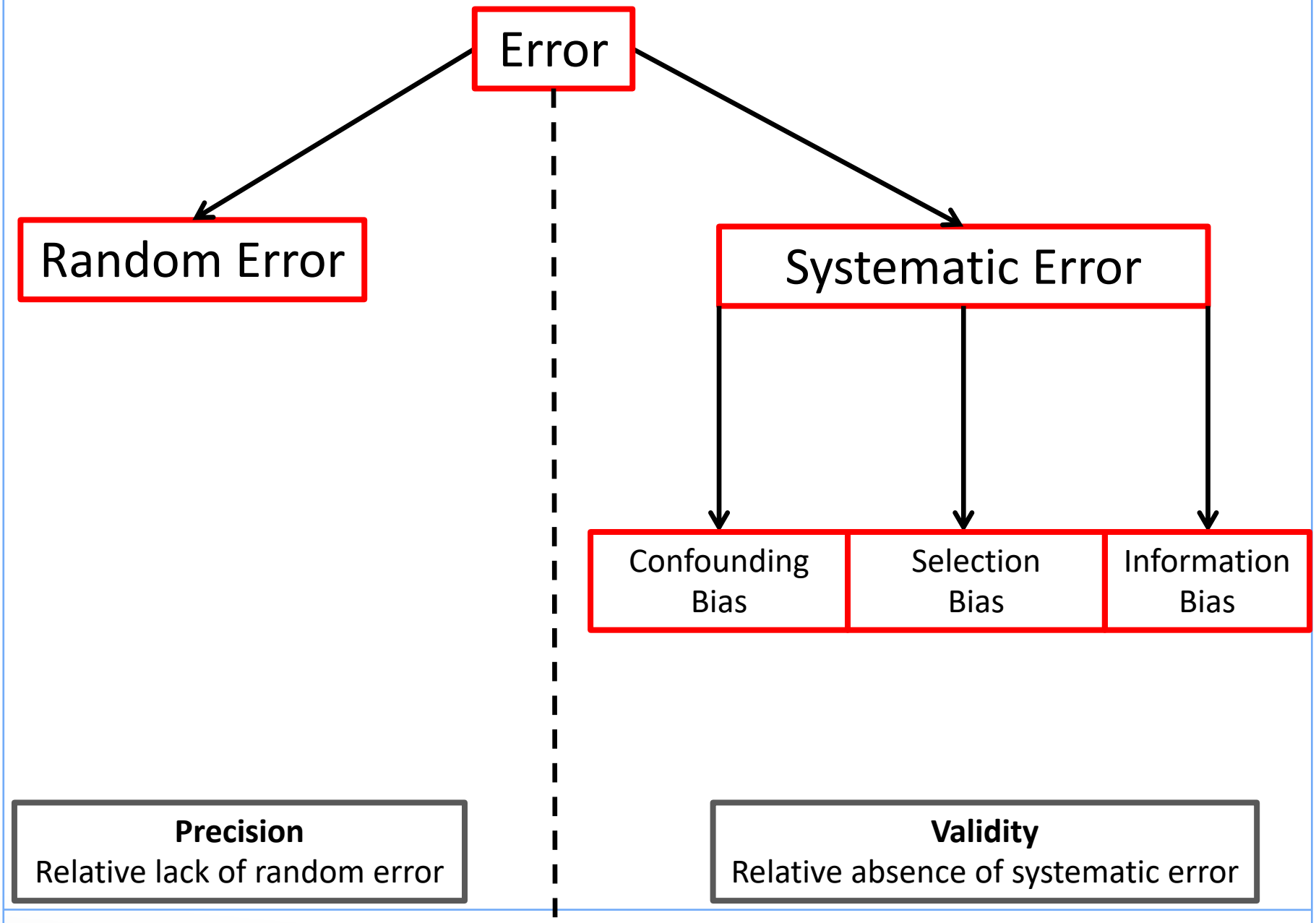
Here we have two hypothetical RCTs where patients were randomly allocated to either receive a cholesterol-lowering drug or a dietary Intervention

In both studies, the risk of myocardial infarction (MI) was 9% for the diet group and 6% for the drug group

However, study A (200 people) would conclude that there is no difference in risk of developing MI between the two treatments, while Study B (2000 people) would conclude that the drug is more effective than the dietary intervention at reducing the risk of MI






Error
Random Error
Systematic Error
Confounding Bias
Selection Bias
Information Bias
Precision
Relative lack of random error
Validity
Relative absence of systematic error

Reliability and Validity

Consistent but wrong	Right answer, on average	Wrong and variable	Consistent and correct

Types of Bias in Epidemiologic Research

- Selection Bias
  - Results from procedures used to select study subjects and factors that influence study participation
- Information Bias
  - Results from either imperfect definitions of study variables or flawed data collection procedures
- Confounding
  - Results when the effect of the exposure of interest is mixed with the effect of another variable (“mixing of effects”)

Direction of Bias

Positive bias (upward bias)— observed effect is higher than the true causal effect

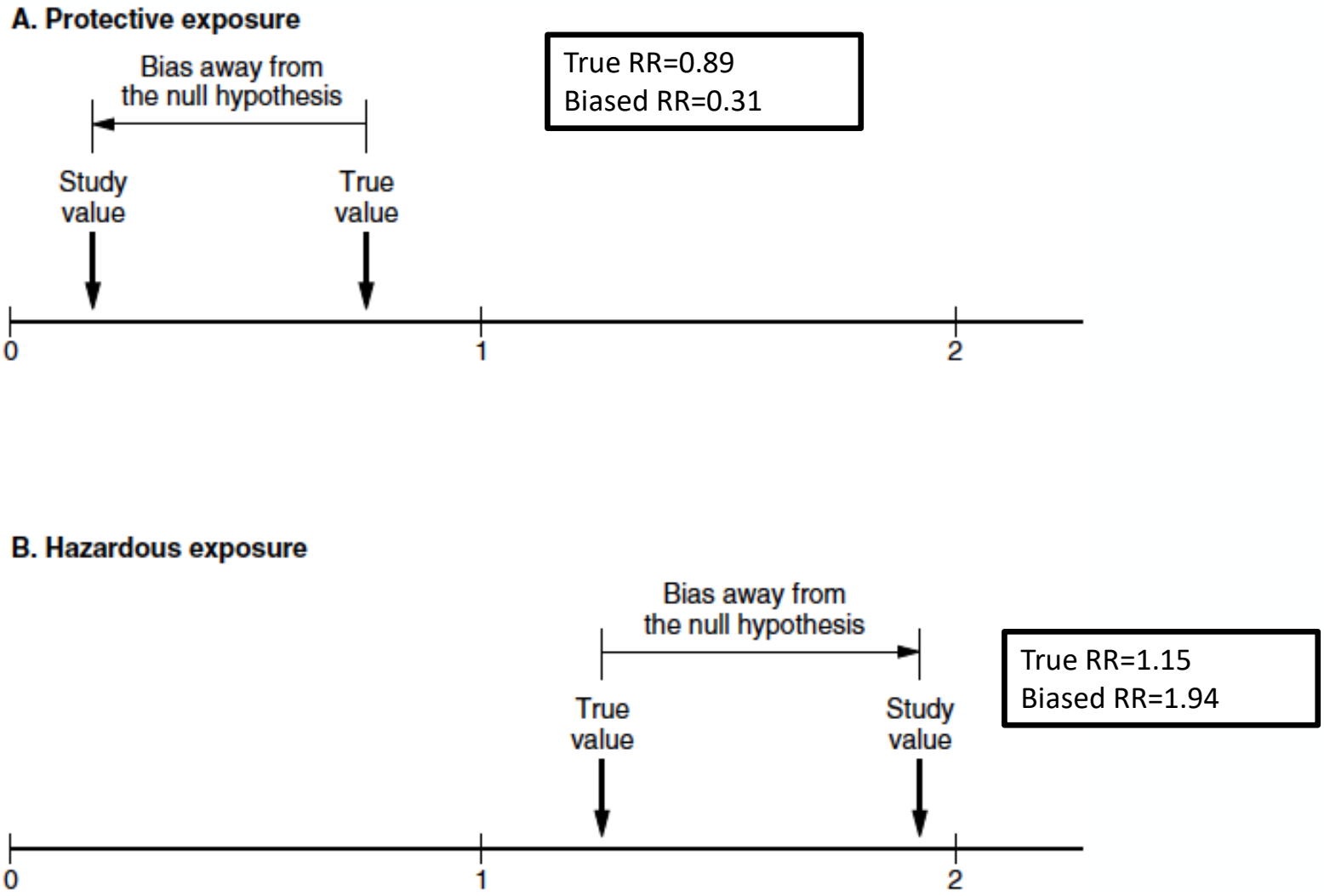
Negative bias (downward bias)— observed effect is lower than the true causal effect

Often we refer to:

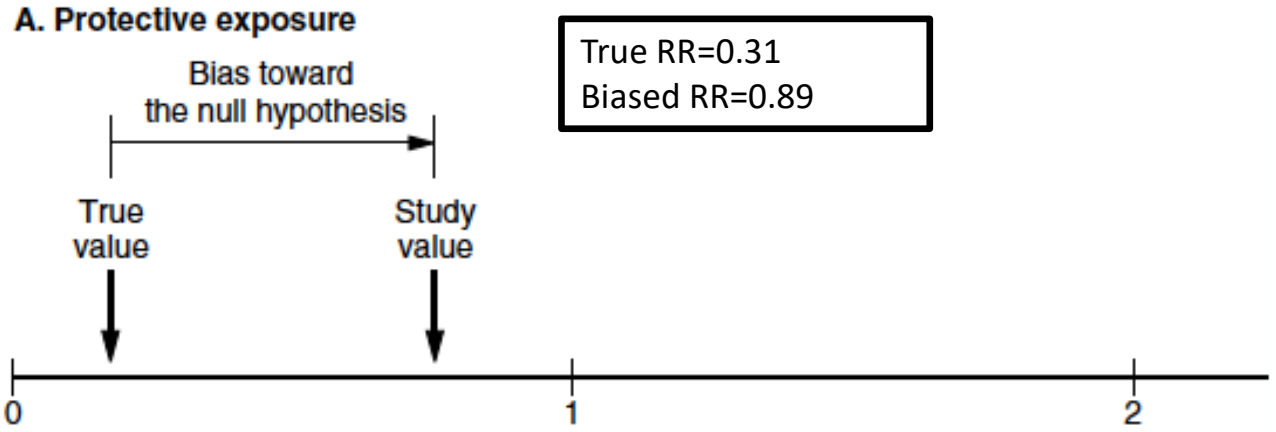
- Bias towards the null— observed value closer to 1.0* (for a ratio measure) than is the true causal effect
- Bias away from the null— observed value farther from 1.0* than is the true causal effect

**closer or further from 0 for a difference measure*

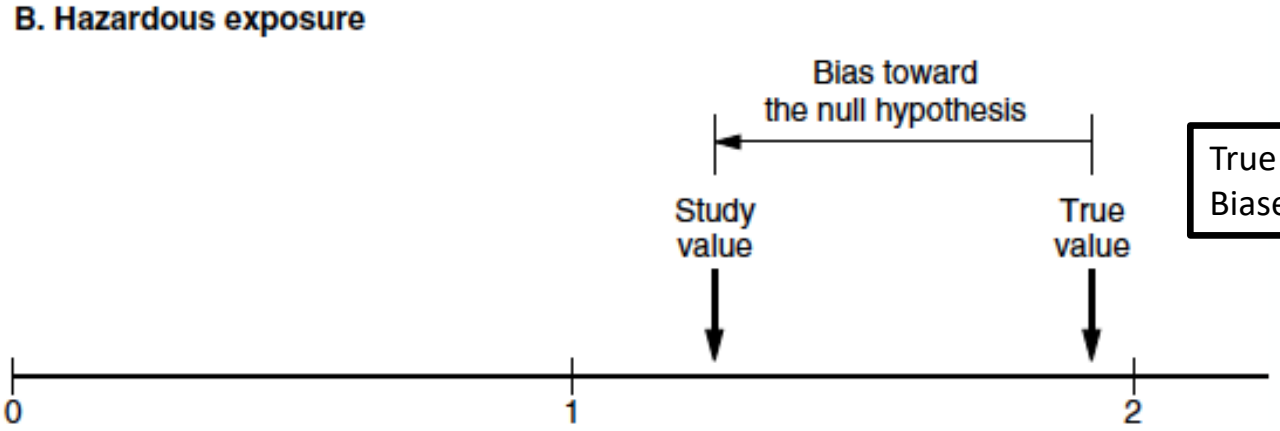
Overestimation of a risk ratio for (**A**) a protective exposure and (**B**) a hazardous exposure.



Underestimation of a risk ratio for (**A**) a protective exposure and (**B**) a hazardous exposure.



True RR=0.31
Biased RR=0.89



True RR=1.94
Biased RR=1.15

Selection Bias

Definition

- Systematic error in the ascertainment of study subjects that occurs when the association between exposure and disease differs in those who participate and those who do not participate in a study
- Can occur:
  - At the the stage of recruitment of participants
  - And/or during the process of retaining them in the study

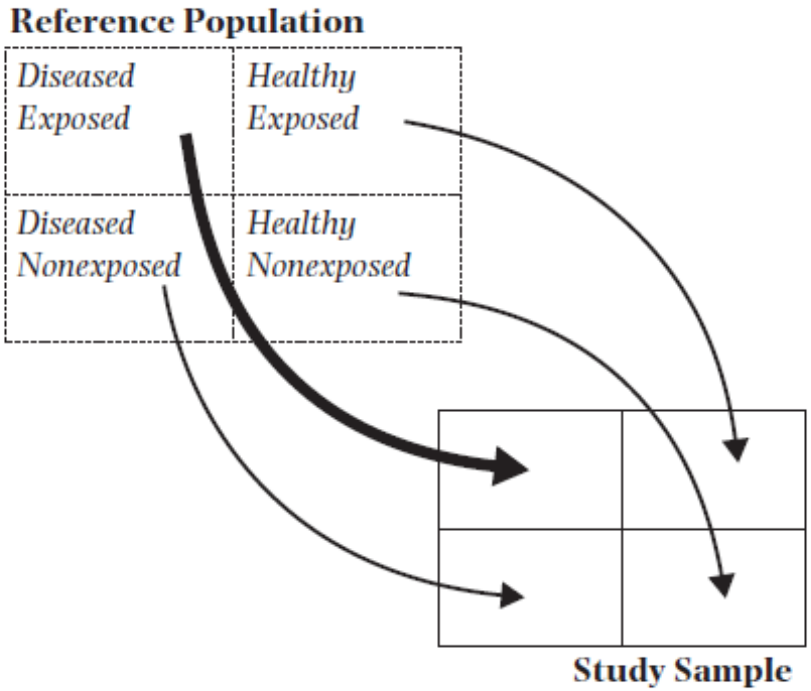
Selecting Participants & Selection Bias

- May affect external validity
  - E.g., Nurses' Healthy Study or RCTs -- highly selected group of people (volunteers, exclusion criteria, etc.)

- May threaten the internal validity of the study, leading to biased measures of effect and invalid inferences about the exposure-disease relationship
  - E.g., Improper control selection in a case-control study

Selection bias arises when the cells of the 2 x 2 table in your study population are sampled with different probabilities from the 2 x 2 table in the source population

FIGURE 4-2 Selection bias: one relevant group in the population (exposed cases in the example) has a higher probability of being included in the study sample.



Selection Probabilities

	E	$\bar{E}$
D	A	B
$\bar{D}$	C	D

Target

	E	$\bar{E}$
D	A^o	B^o
$\bar{D}$	C^o	D^o

Actual

alpha (α) = (A^o/A) the probability that a person in cell A (in the target population) will be selected into the actual population from which the study population is a random sample

beta (β) = (B^o/B) the probability that a person in cell B (in the target population) will be selected into the actual population

gamma (γ) = (C^o/C) the probability that a person in cell C (in the target population) will be selected into the actual population

delta (δ) = (D^o/D) the probability that a person in cell D (in the target population) will be selected into the actual population

Cross Products

α	β
γ	δ

There is no selection bias present if the cross-product ($\alpha * \delta$)/ ($\beta * \gamma$) of the selection probabilities equals 1

Example 1

.5	.2
.5	.2

(.5*.2) / (.5*.2)= 1

Example 2

.5	.2
.8	.1

(.5*.1) / (.8*.2)= 0.31

Selection Bias in Cohort Studies

- Differential loss to follow-up
  - Also called ‘informative censoring’
  - Study subjects leave the study before the end of follow up for reasons that are related to the exposure and disease
- Restrictions on cohort entry
  - Must be affected by exposure and associated with the outcome (e.g., volunteer bias, survivor bias, healthy worker effect)

Selection Bias in Cohort Studies

- Healthy worker effect
  - Bias occurs when comparing outcomes between a worker cohort and the general population
  - Lower expected mortality for exposed workers
- Healthy user bias
- Survivor Bias
  - Must survive to cohort entry
- Losses to follow-up
  - E.g., exposure → side effects → drop out
  - E.g., sicker people → drop out

Selection Bias in Case-Control Studies

- When cases and controls are not drawn from the same source population
  - E.g., Exposure distribution in controls is not representative of source population that the cases came from

Coffee and pancreatic cancer:

- The exposure distribution in the control group did not represent the exposure distribution in the source population

COFFEE AND CANCER OF THE PANCREAS

BRIAN MACMAHON, M.D., STELLA YEN, M.D., DIMITRIOS TRICHOPOULOS, M.D., KENNETH WARREN, M.D., AND GEORGE NARDI, M.D.

Abstract We questioned 369 patients with histologically proved cancer of the pancreas and 644 control patients about their use of tobacco, alcohol, tea, and coffee. There was a weak positive association between pancreatic cancer and cigarette smoking, but we found no association with use of cigars, pipe tobacco, alcoholic beverages, or tea. A strong association between coffee consumption and pancreatic cancer was evident in both sexes. The association was not affected by controlling for cigarette use. For the sexes combined, there was a significant dose-response relation ($P \sim 0.001$); after adjustment for cigarette smoking, the relative risk associated with drinking up to two cups of coffee per day was 1.8 (95 per cent confidence limits, 1.0 to 3.0), and that with three or more cups per day was 2.7 (1.6 to 4.7). This association should be evaluated with other data; if it reflects a causal relation between coffee drinking and pancreatic cancer, coffee use might account for a substantial proportion of the cases of this disease in the United States. (N Engl J Med. 1981; 304:630-3.)

"Relative risk associated with drinking up to two cups of coffee per day was 1.8 (95% CI 1.0, 3.0)"

Coffee and Pancreatic Cancer

	Case	Control
Coffee: ≥ 1 cup day	84	82
No coffee	10	14

OR= (84/10) / (82/14) = 1.4 (95% CI, 0.55 - 3.8)

So, when population-based controls were used, there was no strong association between coffee and pancreatic cancer

Selection Bias in RCTs

- **Due to lack of allocation concealment**
  - Major benefit of RCT = participants do not choose their exposure group
  - Lack of allocation concealment completely eliminates this benefit (participants select their groups)
- **Due to attrition**
  - During the course of the trial, individuals are going to drop out
  - Using an intention to treat analysis will avoid attrition related selection bias

Control of Selection Bias

- Best avoided at the design stage
  - Appropriate control selection (case-control)
  - Very thorough follow-up procedures
- Can collect data to estimate the magnitude/direction of selection bias
  - E.g., collect data from non-responders or censored participants
- Bias analysis: effect estimates can be ‘adjusted’ if selection probabilities are known

Collider stratification bias

- A specific type of selection bias
- Occurs when you condition (stratify on, adjust for, restrict) on a variable that is affected by exposure and outcome

FIGURE 3. Conditioning on a common effect C of exposure E and outcome D.

- Conditioning induces an association within levels of C even if there is no true relationship between E and D

You attend a meeting with 99 of your colleagues

10 people are pre-infected with influenza upon arrival

At the meeting, everyone has either an egg sandwich or chicken sandwich for lunch.

The sandwiches are distributed at random

Cole et al., 2009

What is the expected number of influenza cases among the 50 people who ate chicken sandwiches?

$$Risk_{chicken} = 5/50 = 0.1$$

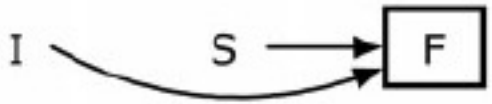
What is the expected number of influenza cases among the 50 people who ate egg sandwiches?

$$Risk_{egg} = 5/50 = 0.1$$

Cole et al., 2009

That evening, you and 54 colleagues develop a fever of 102°F

Everyone with a fever were exposed to either influenza or an egg sandwich (or both)



Cole et al., 2009

What is the risk of influenza among all people at the meeting?

What is the risk of influenza among individuals with a fever?

Table 1 Data illustrative of selection bias, due to conditioning on a collider

	Influenza				
	Yes	**No**	**Total**	**Risk**	**Risk difference**
Panel A					
Sandwich					
Chicken	5	45	50	0.1	0.0
Egg salad	5	45	50	0.1	
Panel B					
Fever					
Sandwich					
Chicken	5	0	5	1.0	0.9
Egg salad	5	45	50	0.1	

Conditioning on fever status induces an association between sandwich type and influenza

Cole et al., 2009

Information Bias

Definition

- Bias in an estimate arising from measurement error
- Systematic error arising from the collection of erroneous information about exposure, outcome, or other covariates
- Results in different quality (accuracy) of information between comparison groups
- Arising during data collection
- Also called misclassification

Types of Variables

Categorical	Continuous
-Two or more mutually exclusive categories Binary/dichotomous = two categories Nominal= K+ unordered categories Ordinal= K+ ordered categories	-Potentially infinite number of possible values along a continuum

Classification

Exposures

- Exposed/unexposed
- Level (dose) of exposure
- Duration of exposure
- Time since exposure

Outcomes

- Dead/not dead
- Diseased/not diseased
- Level of abnormality

Accurate measurement is important to ensure study results are valid

Bias in self-reported estimates of obesity in Canadian health surveys: An update on correction equations for adults

by Margot Shields, Sarah Connor Gorber, Ian Janssen and Mark S. Tremblay

August 2011

Statistics Canada Statistique Canada

Canada

The Canadian Community Health Survey (2005, 2008) and the Canadian Health Measures Survey (2007-8) collected both self-reported and measured height and weight for a subsample of respondents.

Mean height, weight, body mass index (BMI) and prevalence of obesity, by collection method and sex, household population aged 18 to 79, Canada, 2008, 2007 to 2009, and 2005

	Self-reported			Measured			Bias		
		95% confidence interval			95% confidence interval		Self-reported minus measured	95% confidence interval	
	Estimate	from	to	Estimate	from	to		from	to
2008 Canadian Community Health Survey									
Men									
Mean height (cm)	175.8*	175.3	176.3	174.6	174.1	175.1	1.2	0.9	1.5
Mean weight (kg)	81.6*	80.7	82.5	83.8	82.8	84.7	-2.2	-2.4	-1.9
Mean BMI (kg/m^2)	26.4*	26.1	26.6	27.5	27.2	27.9	-1.2	-1.4	-1.0
% obese (BMI 30.0 kg/m^2 or more)	18.5*	16.0	21.2	26.1	23.4	28.9	-7.6	-9.5	-5.7
Women									
Mean height (cm)	162.1*	161.7	162.5	161.2	160.7	161.6	0.9	0.6	1.2
Mean weight (kg)	66.8*	66.0	67.7	69.5	68.6	70.4	-2.7	-2.9	-2.4
Mean BMI (kg/m^2)	25.4*	25.1	25.7	26.9	26.5	27.3	-1.5	-1.7	-1.2
% obese (BMI 30.0 kg/m^2 or more)	16.1*	14.2	18.2	23.3	20.8	25.9	-7.2	-9.2	-5.2

Both men and women were found to under-report their weight

Misclassification

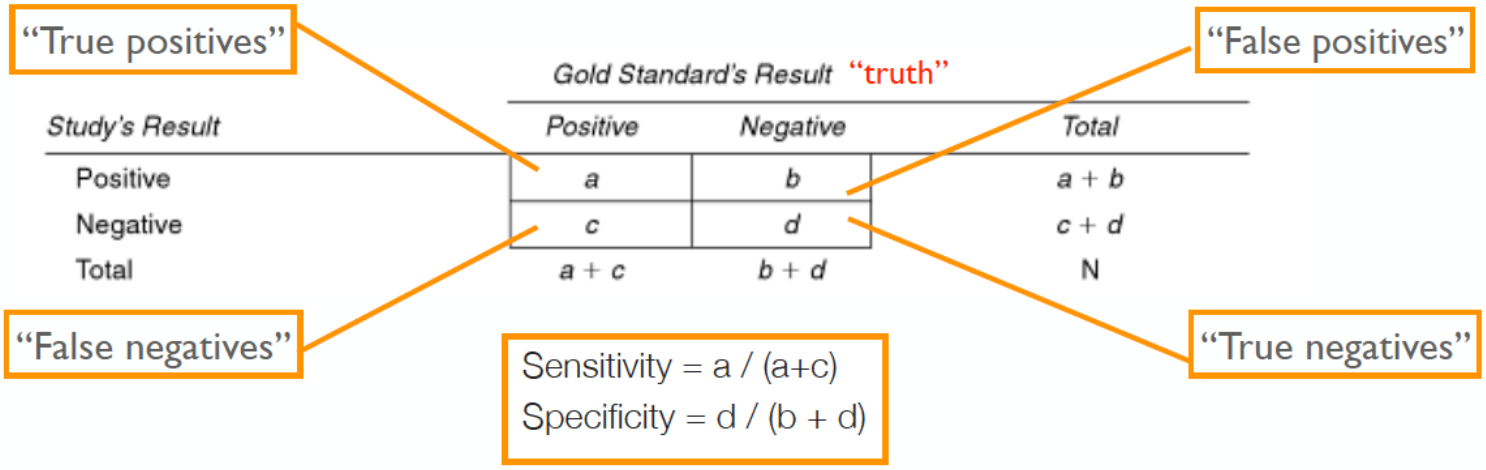
- **Non-differential misclassification:** Errors in classification are not related to other study variables
  - E.g., poor recall
  - E.g., poor data collection tools/instruments
- **Differential misclassification:** Errors in classifying the variable are related to other study variables

How good is the measurement tool?

- Misclassification occurs when sensitivity and/or specificity of the procedure used to detect exposure/outcome is not perfect
  - Ideal measurement tool would perfectly identify E+/E-/D+/D-

- We use sensitivity and specificity in this context as well
  - Sensitivity: ability of a test or measure to correctly identify those who have the exposure/outcome of interest
  - Specificity: ability of a test or measure to correctly identify those who do not have the exposure or outcome of interest

(Delgado-Rodriguez et al., 2004)



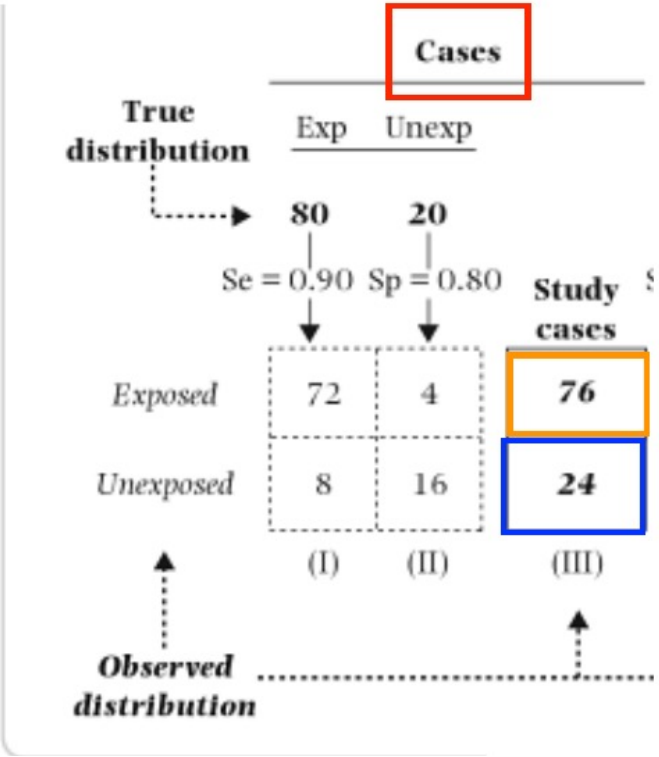
	Gold Standard (Measured)		
Self-reported	+	-	Total
Overweight	3234	134	3368
Not overweight	507	3580	4087
Total	3741	3714	7455

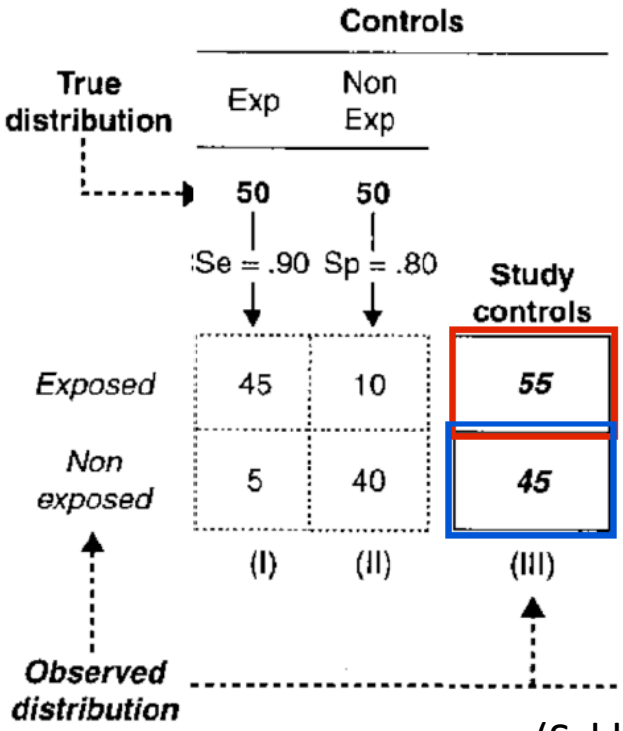
- Sensitivity = 3234/3741 = 86.4%
- Specificity = 3580/3714 = 96.4%

Credit: Sam Harper, McGill

Non-Differential Misclassification of Exposure

In a case control study, the investigators measured exposure using an instrument that has 90% sensitivity and 80% specificity





(Szklo & Nieto, 2007)

Effect of Non-Differential Misclassification

	TRUTH	
	Cases	Controls
E+	80	50
E-	20	50

OR=(50*80) / (50*20)= 4.0

	STUDY	
	Cases	Controls
E+	76	55
E-	24	45

OR=(76*45) / (55*24)= 2.6

Typically, non-differential misclassification results in bias toward the null

Non-differential because sensitivity and specificity of exposure ascertainment is equal for cases and controls.

Non-differential misclassification and prevalence

- The magnitude of bias due to non-differential misclassification depends on sensitivity, specificity, and the prevalence of exposure
- Exposure prevalence from the previous example was 50% among the controls
- If you work through the same example using an exposure prevalence of 2.5% among controls...
  - Leads to a greater degree of bias in the study estimates
  - OR_{biased}= 1.3 vs. OR_{true}=4.0

	TRUTH	
	Cases	Controls
E+	50	20
E-	500	800

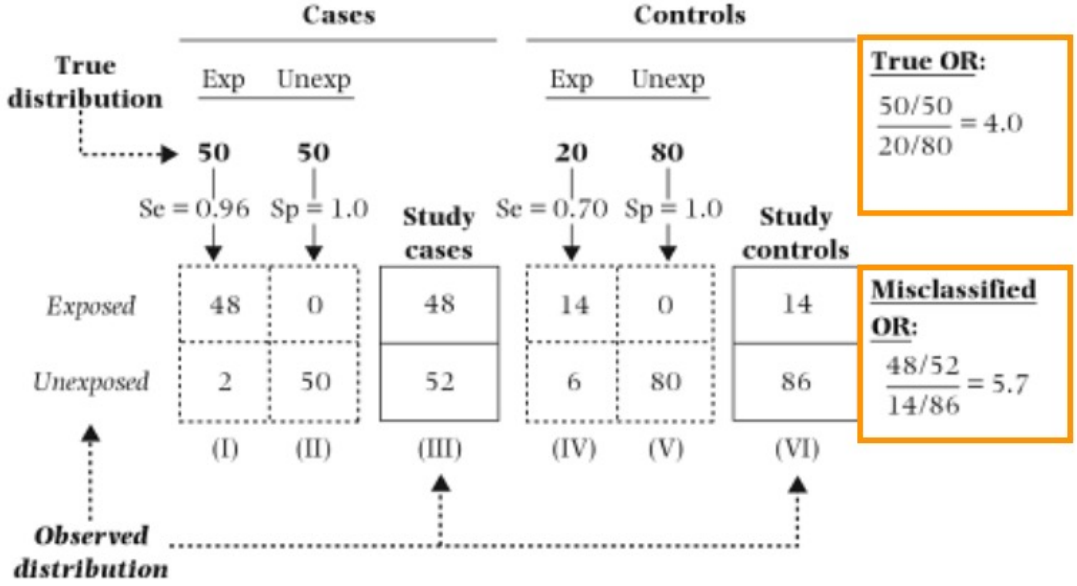
Differential Misclassification

- A situation in which either:
  - Sensitivity and/or specificity of exposure differs by disease status

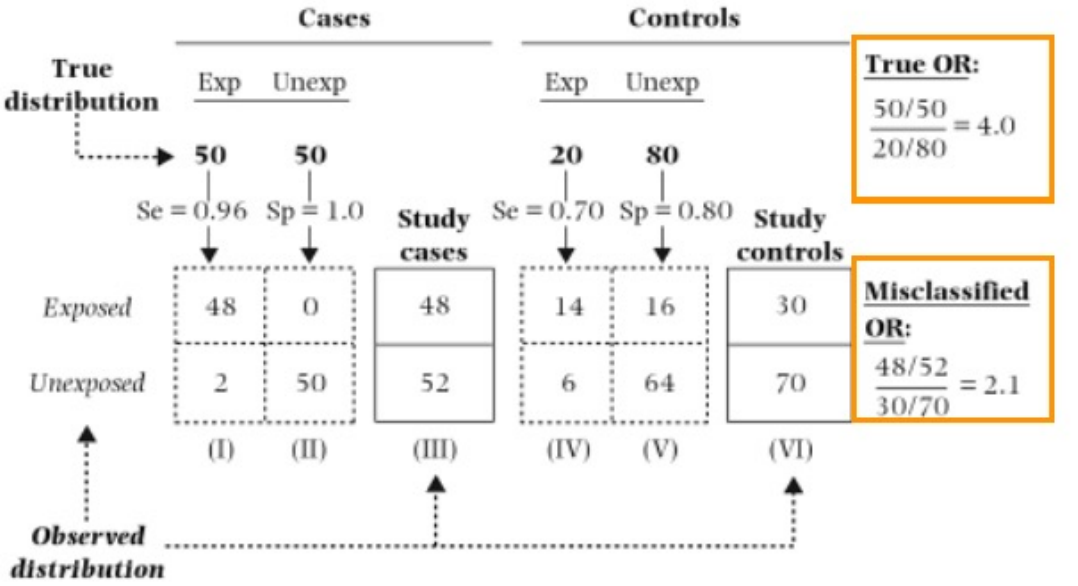
  OR
  - Sensitivity and/or specificity of disease differ by exposure status

- The direction of differential misclassification is not predictable like with non-differential misclassification
  - Could lead to bias in either direction

Differential Misclassification



1) Sensitivity differs in cases and controls



2)Sensitivity and specificity differ in cases and controls

Misclassification of Exposure

- **Recall Bias**
  - Exposure information is misclassified differentially for those with or without the disease (e.g., cases might exaggerate exposure history)
- **Interviewer Bias**
  - Interviewer aware of the study subject's outcome status may ask prompting questions or emphasize certain questions when ascertaining exposure info
- **Changes in exposure status over time**

Cohort studies: misclassification of exposure tends to be non-differential (because usually outcome hasn't happened yet)

Case-control studies: misclassification of exposure could be differential or non-differential

Recall Bias

SHORT REPORT

Recall bias, MMR, and autism

N Andrews, E Miller, B Taylor, R Lingam, A Simmons, J Stowe, P Waight

Arch Dis Child 2002;**87**:493–494

Parents of autistic children with regressive symptoms who were diagnosed after the publicity alleging a link with measles, mumps, and rubella (MMR) vaccine tended to recall the onset as shortly after MMR more often than parents of similar children who were diagnosed prior to the publicity. This is consistent with the recall bias expected under such circumstances.

Minimizing Recall Bias

- Collect objective measures of exposure where possible (e.g., check vaccination records rather than relying on parental report)
- Verify exposure information obtained from all or some participants medical or pharmaceutical records, physician reports
- Because recall bias can be caused by the rumination of cases regarding the cause of their disease, another approach to minimize recall bias is to use a control group composed of subjects with similar diseases

Minimizing Interviewer Bias

- Blind interviewers to case-control status if possible
  - Important when interviewers are aware of study hypothesis
- Use standardized questionnaires consisting of closed-end, easy to understand questions with appropriate response options
  - Training all interviewers to adhere to the question and answer format, with similar questioning for both cases and controls

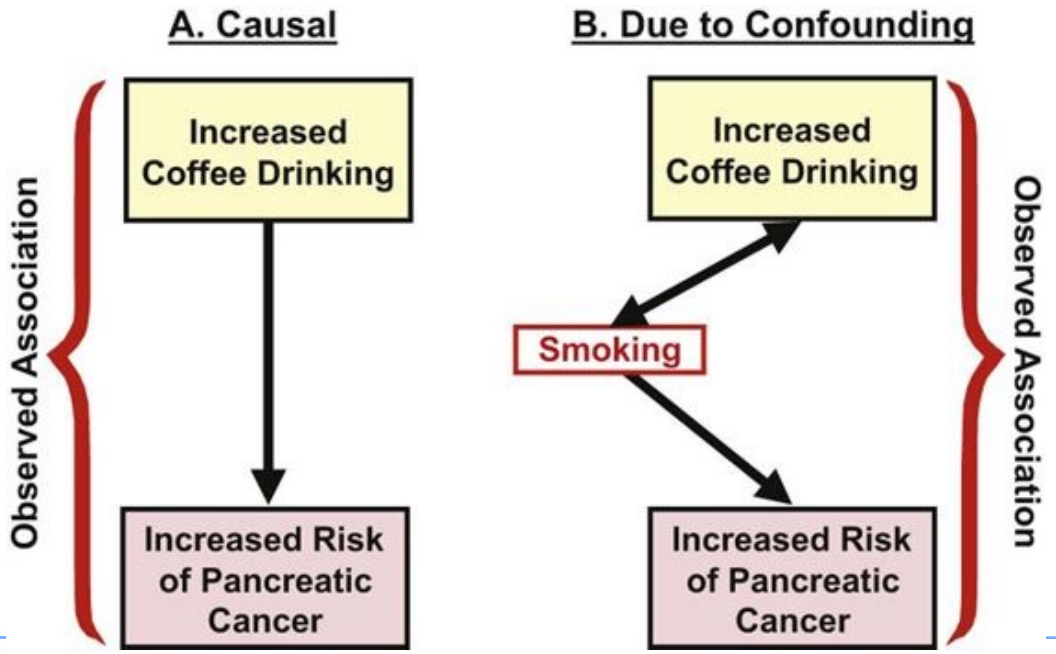
Misclassification of Outcome

- Observer Bias – different quality of info about outcome collected from exposed and unexposed groups
- Surveillance Bias – when a medically relevant exposure leads to closer surveillance for study outcome (increased probability of detection in E+)
- Respondent Bias – participants more likely to report outcomes they believe support study hypotheses/ underreport socially unacceptable outcomes

Confounding

Definition

The term **confounding** refers to a situation in which a non-causal association between a given exposure and an outcome is observed as a result of the influence of a third variable, usually referred to as a **confounder**.



(Szklo & Nieto, 2007)

Confounding

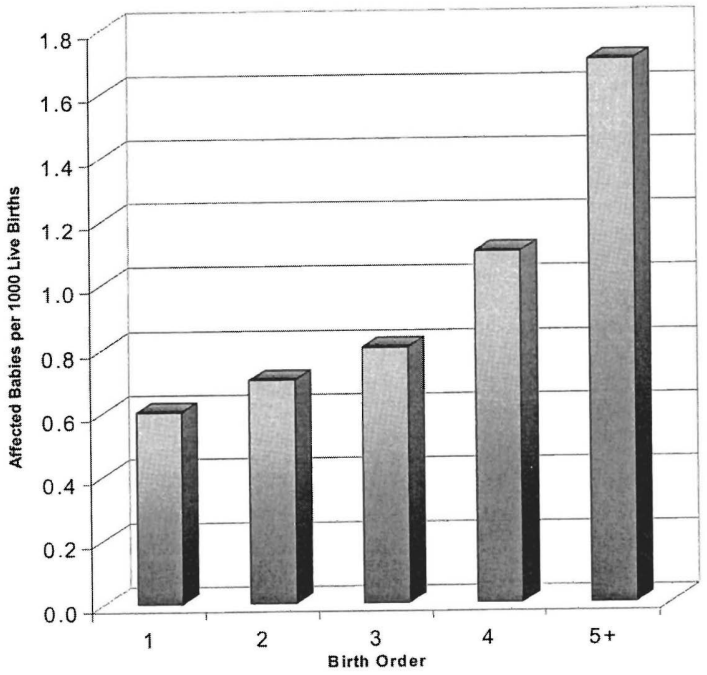
- Latin “confundere” is to mix together
- One way to understand confounding:

“Confounding is confusion, or mixing, of effects; the effect of the exposure is mixed together with the effect of another variable, leading to bias”

-Rothman, 2002

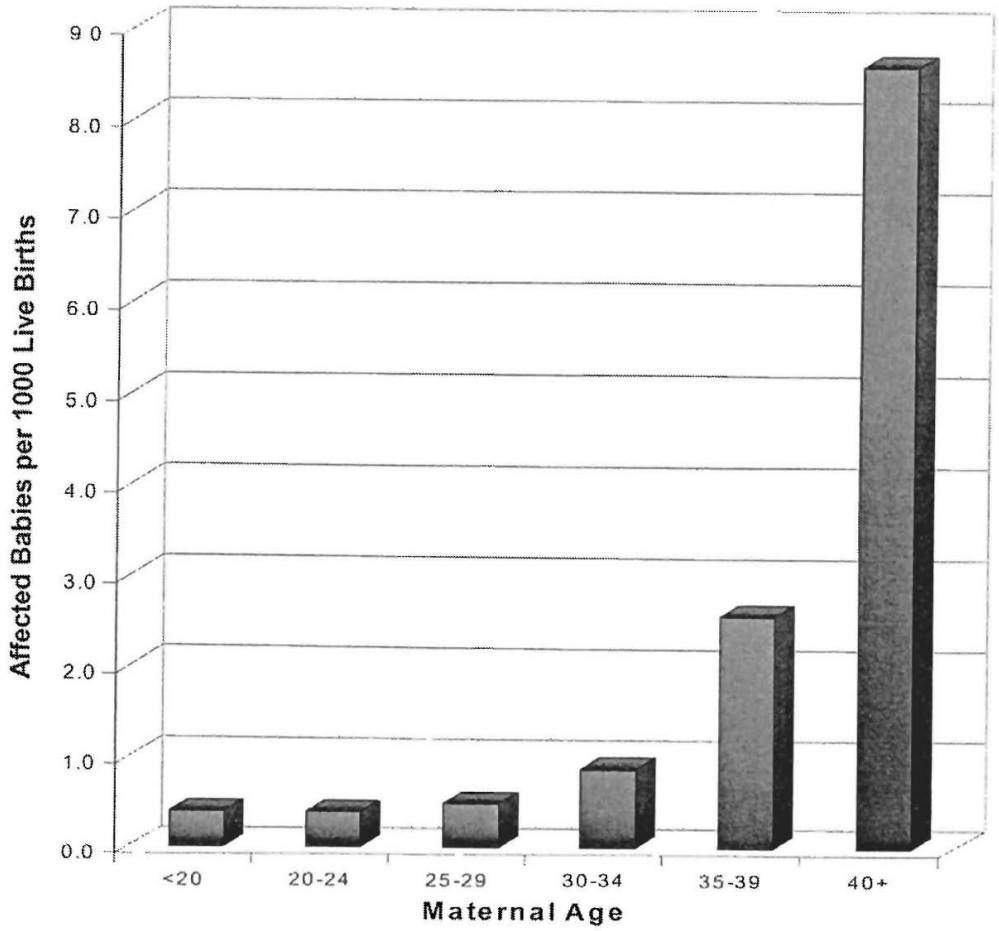
Birth Order and Down Syndrome

Is there a relationship between birth order and down syndrome?

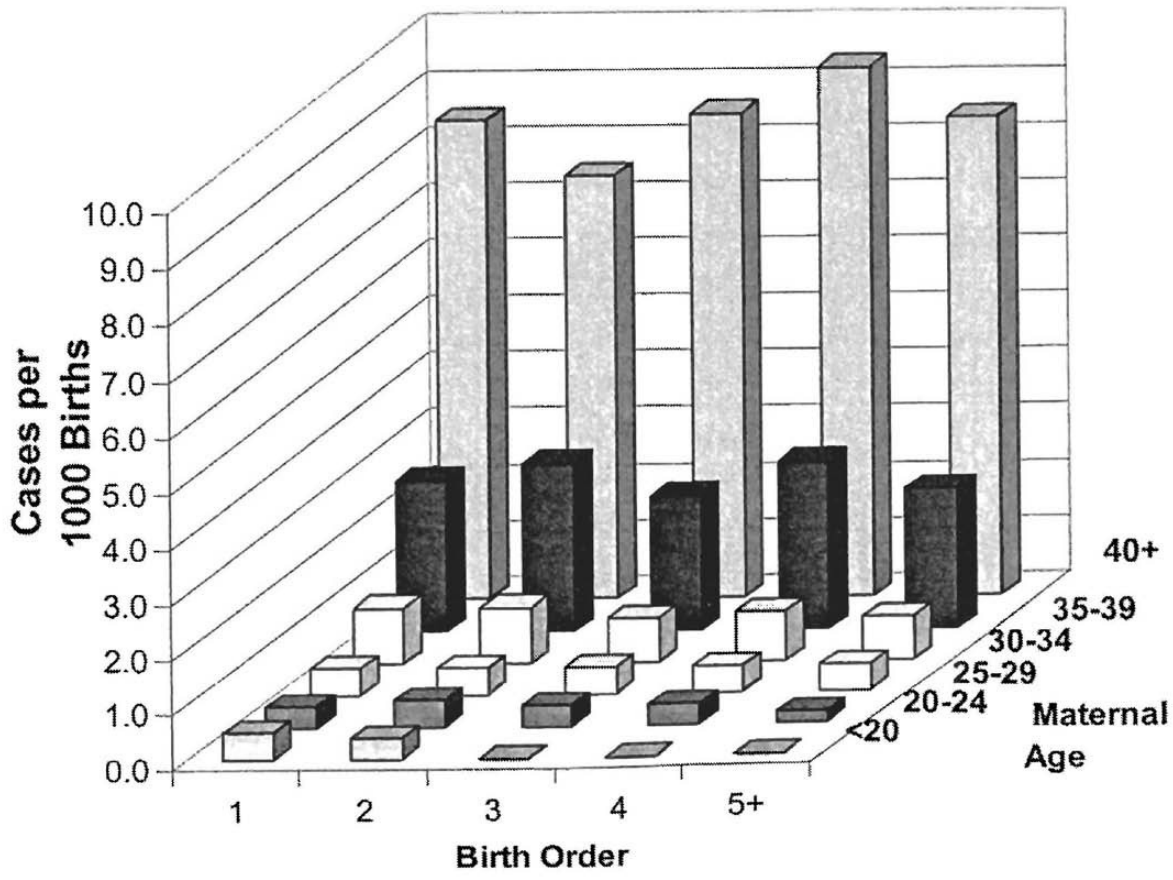


Data from Stark and Mantel (1966)

Maternal Age and Down Syndrome



Confounding?



Maternal age confounds the relationship between birth order and Down syndrome

Data from Stark and Mantel (1966)

Confounding Criteria

Three criteria for identifying a confounding variable:

1. Must be associated with the exposure

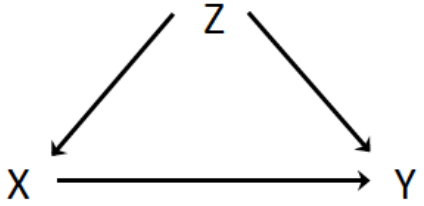
- Maternal age is associated with birth order

2. Must be associated with the outcome

- Maternal age is a known risk factor for Down Syndrome

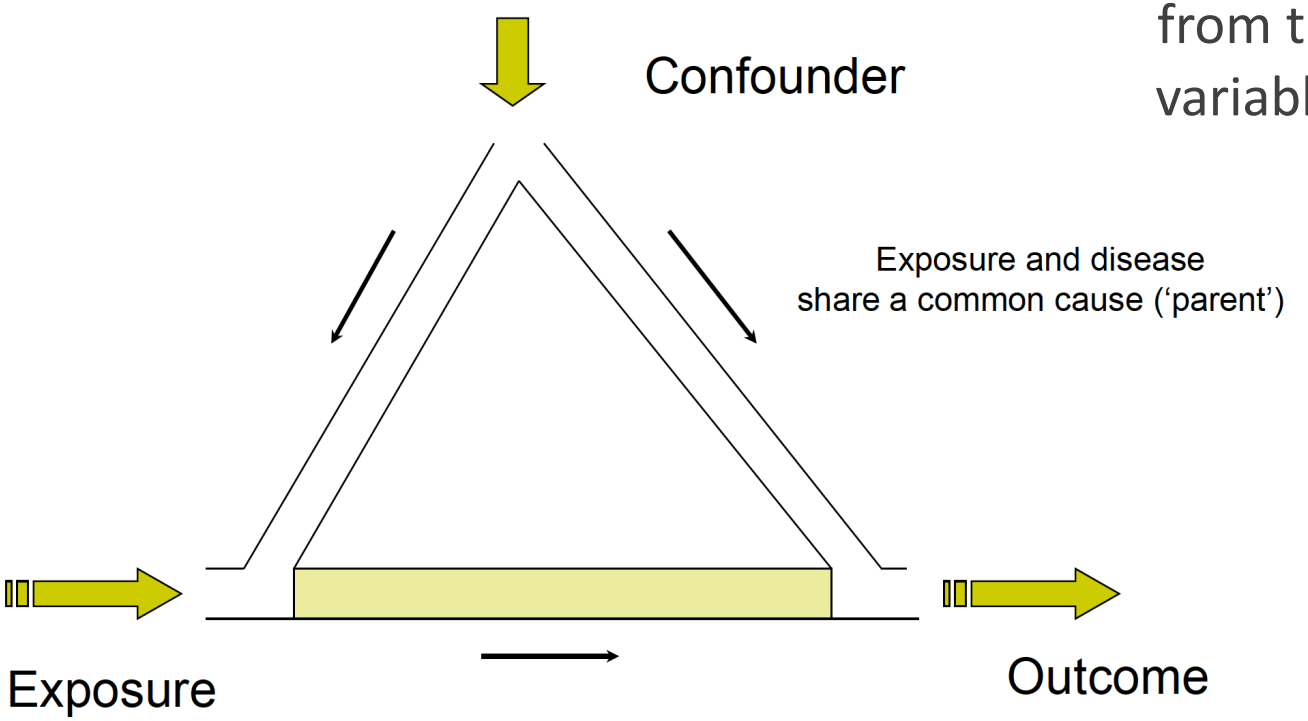
3. Must not be on the causal pathway between exposure and outcome

- Birth order does not cause maternal age



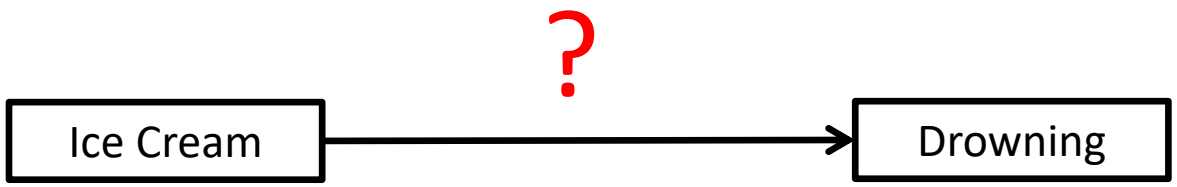
Mixing of Effects: Water Pipes

Cannot separate the effect of exposure from that of the third variable (confounder)



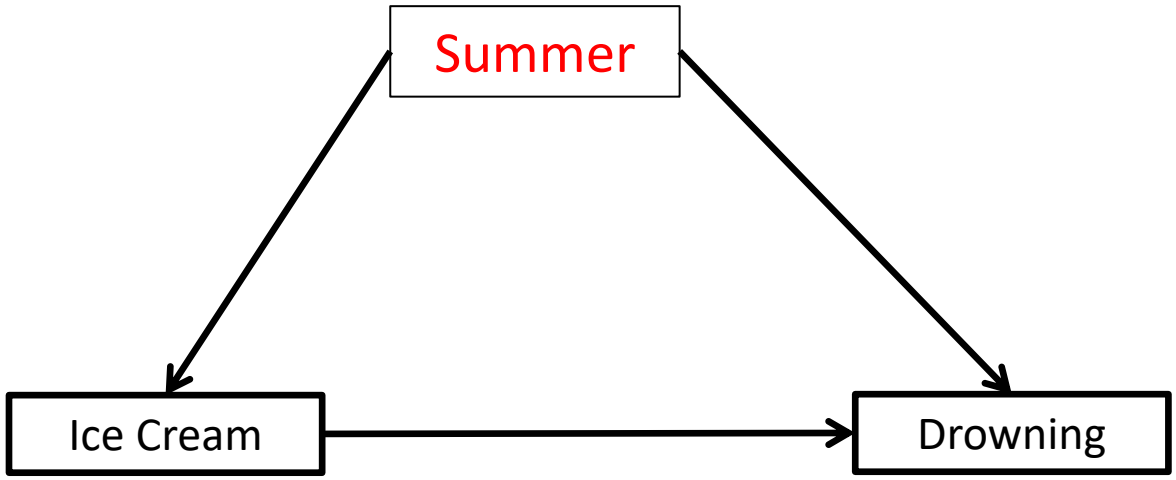
Confounding Example

- There is an observed association between ice cream eating and drowning deaths
- Do you think this is the result of a causal relationship or could it be due to confounding?



Confounding Example

The relationship is not causal, it is confounded by seasonality



Confounding Criteria

Three criteria for identifying a confounding variable:

1. Must be associated with the exposure

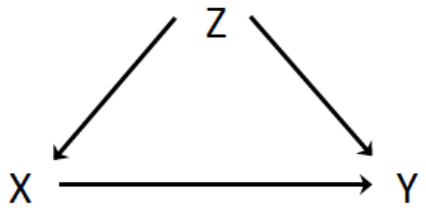
- People eat more ice cream in the summer

2. Must be associated with the outcome

- Drownings are more likely to happen in the summer

3. Must not be on the causal pathway between exposure and outcome

- Ice cream consumption does not cause the season to change



Control of Confounding

- **Preventing confounding at the study design phase:**
  - Randomization
  - Restriction
  - Matching
- **Control of measured confounding at the data analysis phase:**
  - Adjustment (via regression models)
  - Stratification

Crude vs. Adjusted

Crude effect estimate

Does not take into account any confounding variable(s)

Adjusted effect estimate

Accounts for confounding variable(s)

Empirical assessment of confounding:

Crude effect estimate ≠ Adjusted effect estimate

Randomization

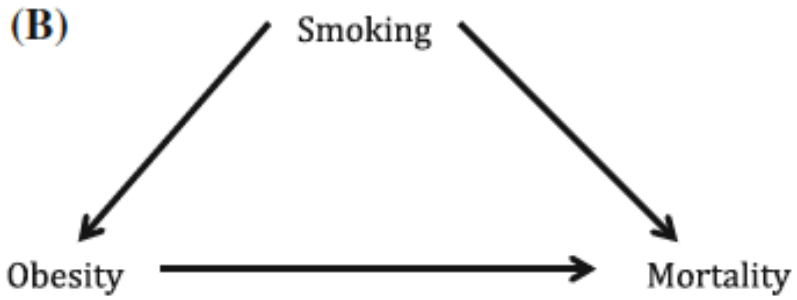
- Successful randomization produces groups that are exchangeable with regard to both measured and unmeasured variables
- Exchangeability= no confounding

	Nicotine e-cigarettes (n=289)	Patches (n=295)
Age (years)	43·6 (12·7)	40·4 (13·0)
Women	178 (62%)	182 (62%)
Ethnicity*		
New Zealand Māori	95 (33%)	95 (32%)
Non-Māori	194 (67%)	200 (68%)
Education below year 12† or no qualification	150 (52%)	123 (42%)
Average number of cigarettes (including RYO) smoked per day	18·4 (7·2)	17·6 (6·0)
Age started smoking (years)	15·6 (4·7)	15·2 (3·8)
Number of years smoking continuously	25·9 (13·1)	23·5 (12·9)
Type of tobacco usually smoked		
Factory made only	167 (58%)	167 (57%)
RYO only	92 (32%)	92 (31%)
Both	30 (10%)	35 (12%)
Lives with other smokers	151 (52%)	149 (51%)
At least 1 quit attempt in past 12 months	158 (55%)	169 (57%)
FTND score	5·6 (2·0)	5·5 (2·0)
FTND >5 (high dependence)	157 (54%)	162 (55%)
GN-SBQ score	20·1 (7·9)	20·1 (8·4)
Self-efficacy to quit‡	3·7 (1·0)	3·7 (0·9)
AUTOS total score	22·6 (7·2)	23·1 (7·6)

Restriction

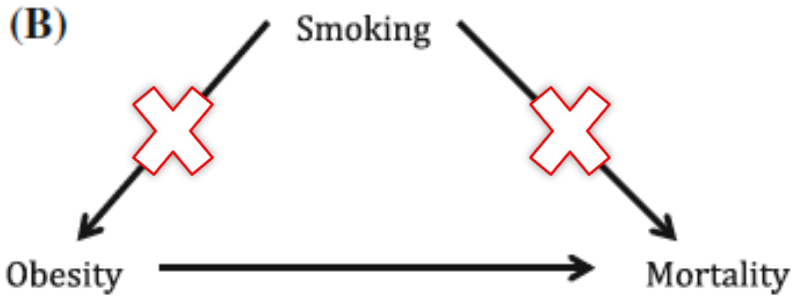
- Restricting entry into the study to individuals who have the same value for a particular variable
  - E.g., Restricting study entry to non-smokers
  - E.g., Restricting study entry to women only

- Very effective method for preventing confounding in any type of study design, though has important implications for generalizability of results.



Smoking may confound the relationship between obesity and mortality

However, if the study were confined to non-smokers, smoking cannot be a confounder of the obesity-mortality relationship



Control of Confounding: Data Analysis

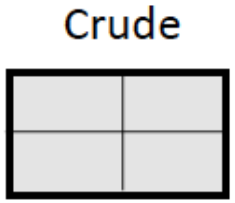
1. **Stratification**:

- The objective of stratified analysis is to set the level of the confounding variable and produce groups within which the confounder does not vary

- Then, we evaluate the exposure-disease relationship within each stratum of the confounder

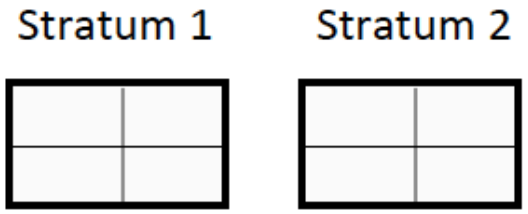
Example: obesity-smoking-mortality

If you stratify on smoking status (smokers vs. non-smokers), we can assess whether the obesity mortality relationship different between the strata



Calculate crude measure of association

Stratify by confounder



Calculate stratum-specific associations

Are stratum-specific associations similar?

If yes, calculate pooled estimate

If not, do not pool estimates*

**Evidence of effect measure modification*

If crude association > or < adjusted association then confounding is likely

If crude association similar to adjusted association then confounding is **not** likely

Limits to Stratification

- Can only stratify on categorical variables
- Numerous strata can be problematic
  - Sparse data and imprecise estimates
- Impractical to adjust for multiple confounding variables
  - Controlling for age and gender, if gender is measured with 2 categories and age is measured with 5 , end up with 10 strata

Control of Confounding: Data Analysis

2. Adjustment

- If the number of potential confounders is large, multivariate analyses (regression analysis) offer the only real solution
- Can handle a large number of confounders simultaneously
- Uses statistical regression models
- Always done with statistical software
  - SAS, Stata, R

Example: Multivariate Adjustment

Table 3: Relative and absolute differences in preterm and very preterm births among non-Hispanic black women relative to non-Hispanic white women in Canada and the United States, 2004–2006

	Crude			Adjusted*			Adjusted†		
Measure; outcome	Canada	US	p_H value‡	Canada	US§	p_H value‡	Canada	US§	p_H value‡
Risk ratio (95% CI)									
Preterm birth (< 37 v. 37–41 wk)	1.49 (1.32 to 1.66)	1.57 (1.56 to 1.58)	0.3	1.46 (1.29 to 1.63)	1.41 (1.40 to 1.42)	0.5	1.60 (1.39 to 1.81)	1.45 (1.44 to 1.47)	0.1
Very preterm birth (< 32 v. 32–41 wk)	2.70 (1.95 to 3.44)	2.81 (2.77 to 2.86)	0.8	2.61 (1.88 to 3.35)	2.36 (2.31 to 2.40)	0.5	2.62 (1.83 to 3.41)	2.43 (2.36 to 2.52)	0.6
Risk difference (95% CI)									
Preterm birth (< 37 v. 37–41 wk)	2.94 (1.91 to 3.96)	4.63 (4.56 to 4.70)	0.003	2.76 (1.74 to 3.78)	3.41 (3.33 to 3.48)	0.2	3.59 (2.32 to 4.85)	3.57 (3.43 to 3.70)	1.0
Very preterm birth (< 32 v. 32–41 wk)	1.22 (0.71 to 1.73)	1.67 (1.64 to 1.70)	0.08	1.16 (0.66 to 1.67)	1.31 (1.28 to 1.35)	0.6	1.17 (0.62 to 1.71)	1.32 (1.25 to 1.38)	0.6

Note: CI = confidence interval.
*Term birth = 37–41 wk, preterm = < 37 wk, very preterm = < 32 wk.
*Adjusted for maternal age, maternal education, marital status, birth order, sex of child and missing paternal information.
†Adjusted for maternal age, maternal education, marital status, birth order, sex of child, missing paternal information and maternal nativity. For the US sample, the estimates include only births in 2004 (n = 1 493 259).
‡p_H = p value for χ^2 test for heterogeneity of the risk ratios or risk differences.
§Adjusted US estimates are standardized to the covariate distribution of the Canadian study population.