

Confounding in health research

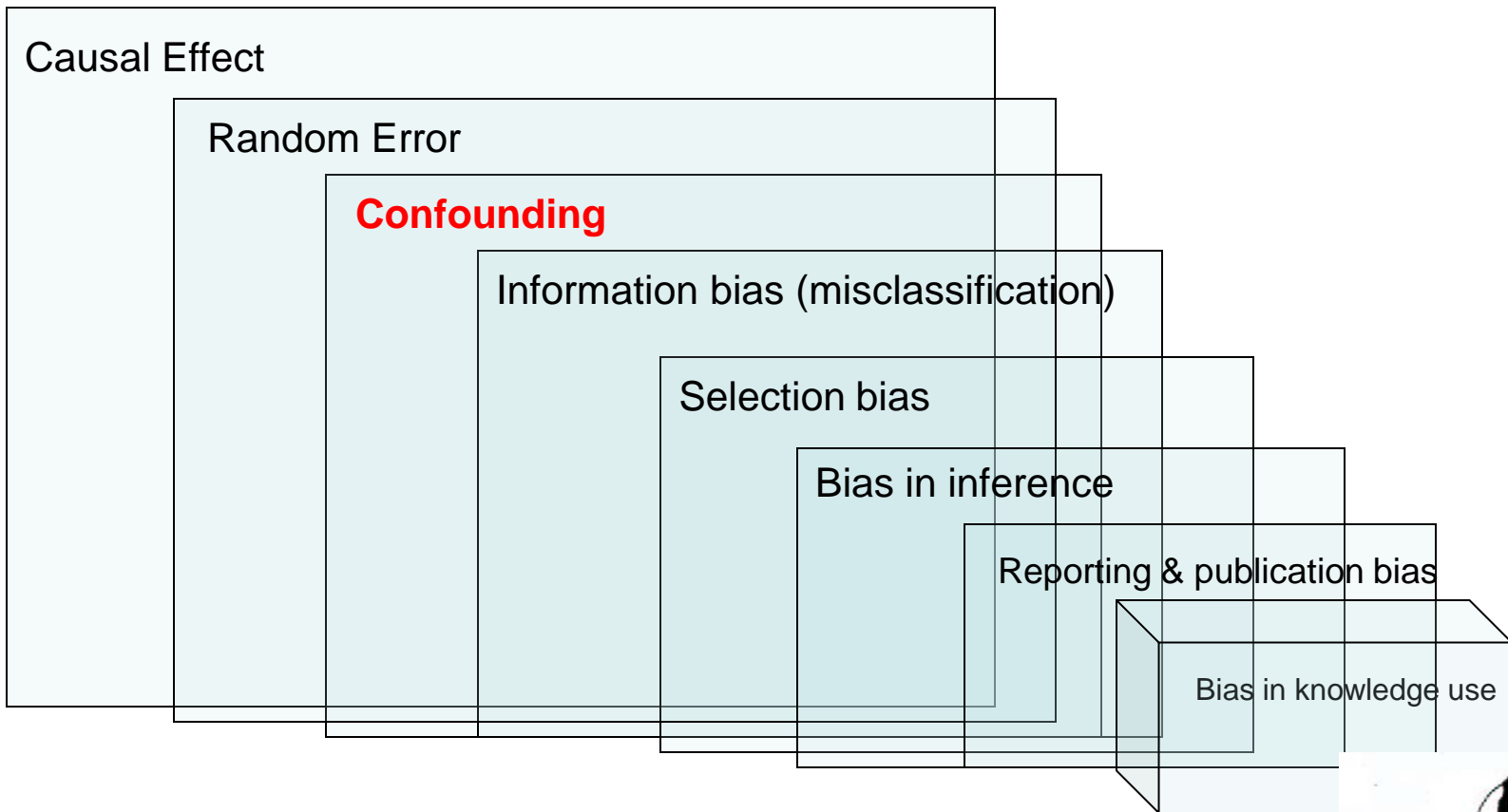
Part 2: adjustment and control

Madhukar Pai, MD, PhD
Assistant Professor of Epidemiology
McGill University
madhukar.pai@mcgill.ca



Quick Recap

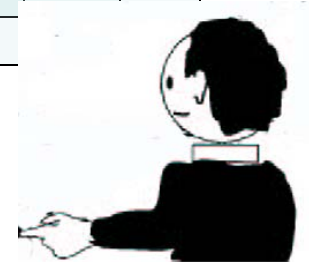
Confounding is one of the key biases in identifying causal effects



RR_{causal}
"truth"



$RR_{\text{association}}$



Confounding: 4 ways to understand it!

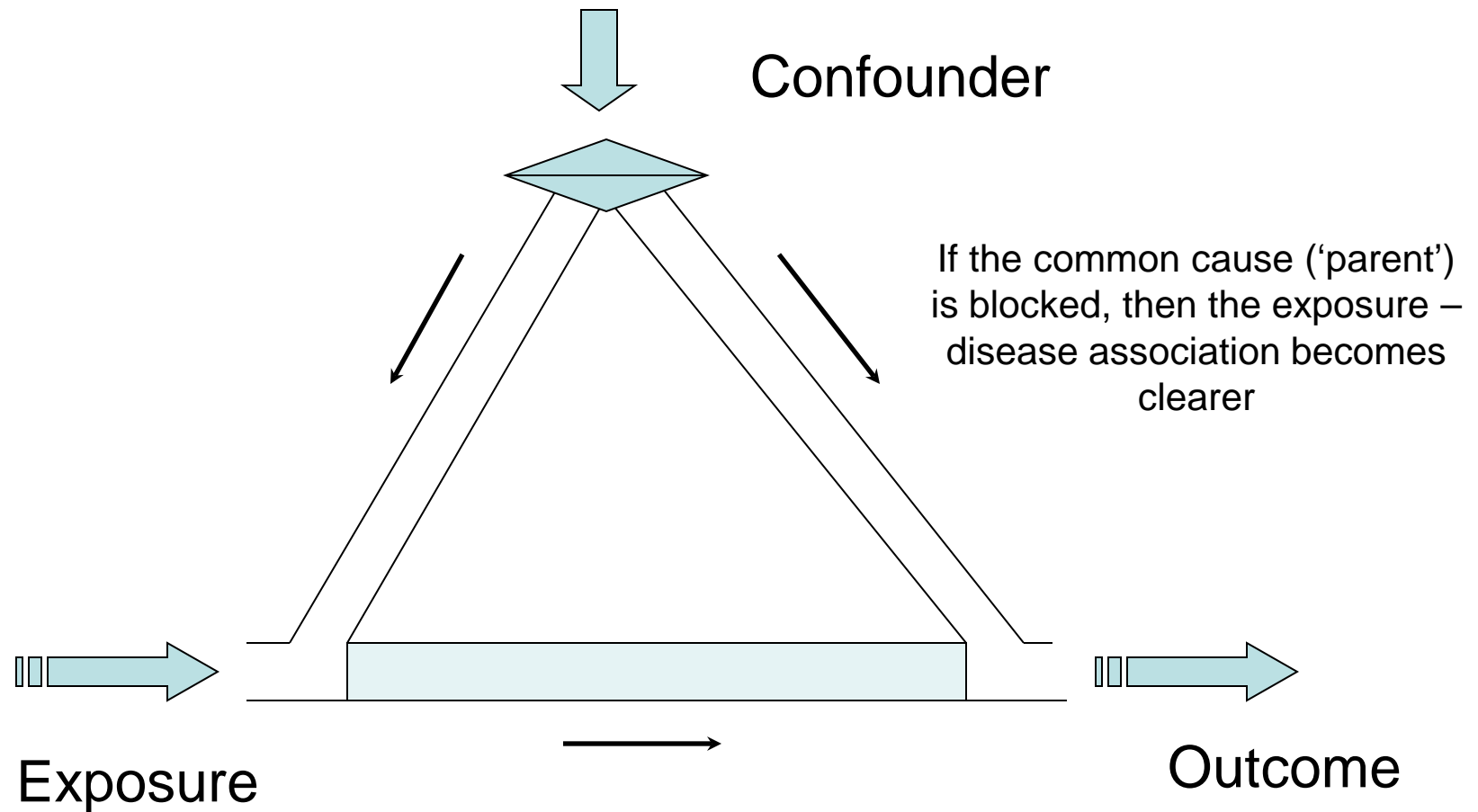
1. “Mixing of effects”
2. “Classical” approach based on *a priori* criteria
3. Collapsibility and data-based criteria
4. “Counterfactual” and non-comparability approaches

First approach: Confounding: mixing of effects

- “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

Latin: “confundere” is to mix together

Mixing of Effects: “control” of the confounder



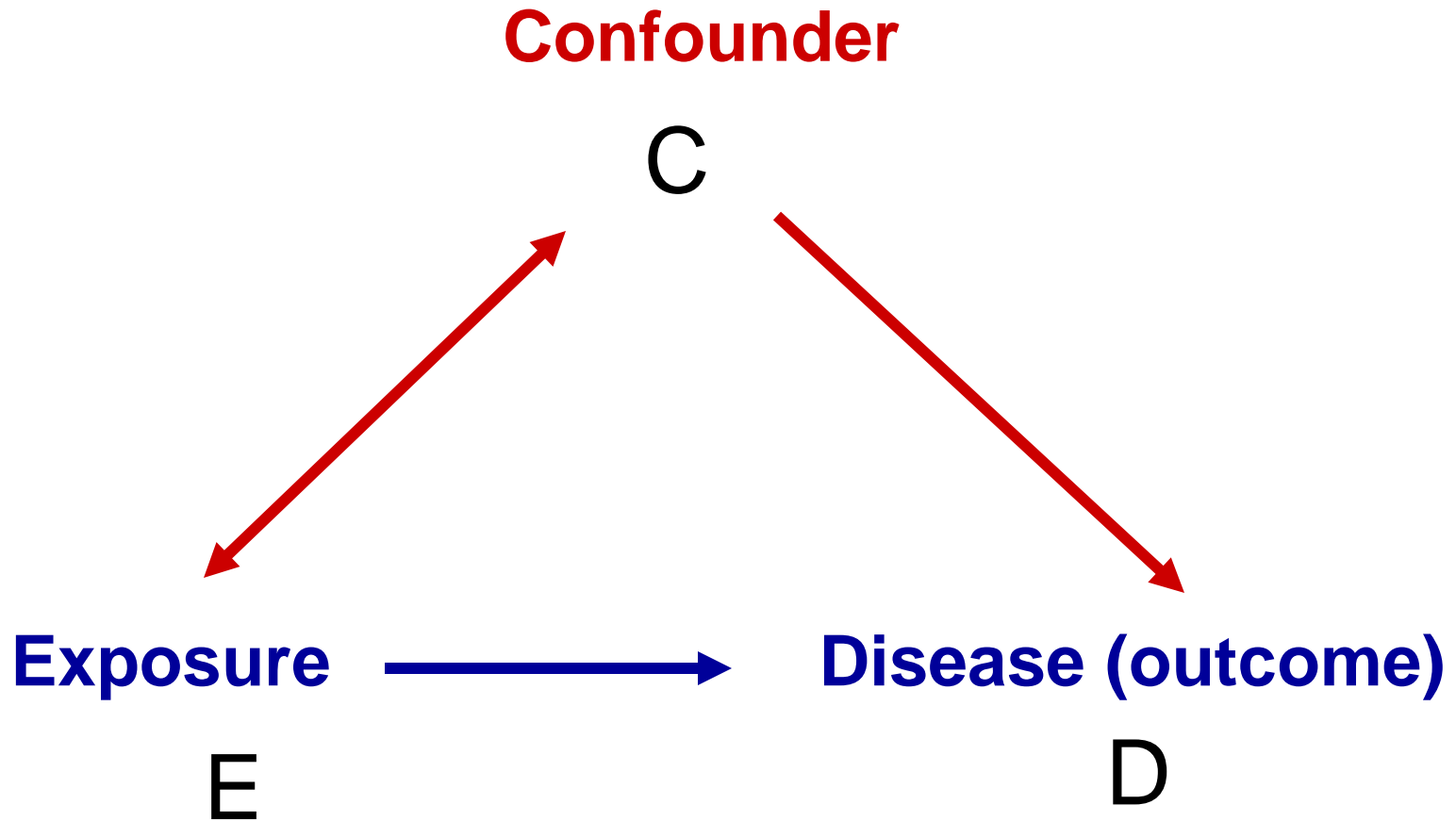
Successful “control” of confounding (adjustment)

Second approach: “Classical” approach based on *a priori* criteria

“Bias of the estimated effect of an exposure on an outcome due to the presence of a common cause of the exposure and the outcome” – Porta 2008

- A factor is a confounder if 3 criteria are met:
 - a) a confounder must be causally or noncausally associated with the exposure in the source population (study base) being studied;
 - b) a confounder must be a causal risk factor (or a surrogate measure of a cause) for the disease in the unexposed cohort; and
 - c) a confounder must not be an intermediate cause (in other words, a confounder must not be an intermediate step in the causal pathway between the exposure and the disease)

Confounding Schematic



Third approach: Collapsibility and data-based approaches

- According to this definition, a factor is a confounding variable if
 - a) the effect measure is homogeneous across the strata defined by the confounder and
 - b) the crude and common stratum-specific (adjusted) effect measures are unequal (this is called “lack of collapsibility”)
- Usually evaluated using 2x2 tables, and simple stratified analyses to compare crude effects with adjusted effects

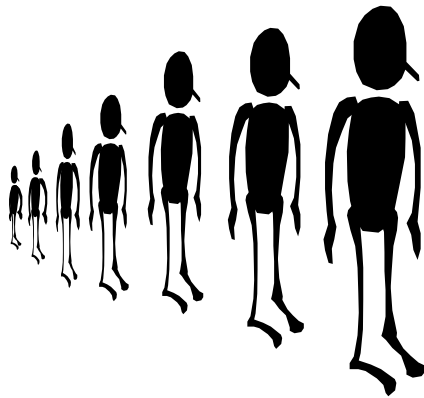
“Collapsibility is equality of stratum-specific measures of effect with the crude (collapsed), unstratified measure” Porta, 2008, Dictionary

Crude vs. Adjusted Effects

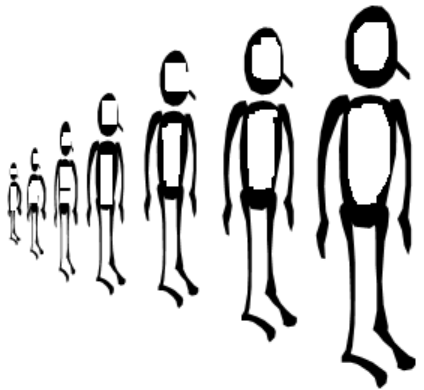
- **Crude**: does not take into account the effect of the confounding variable
- **Adjusted**: accounts for the confounding variable(s) (what we get by pooling stratum-specific effect estimates)
 - Generating using methods such as Mantel-Haenszel estimator
 - Also generated using multivariate analyses (e.g. logistic regression)
- Confounding is likely when:
 - $RR_{\text{crude}} \neq RR_{\text{adjusted}}$
 - $OR_{\text{crude}} \neq OR_{\text{adjusted}}$

Fourth approach: Causality: counterfactual model

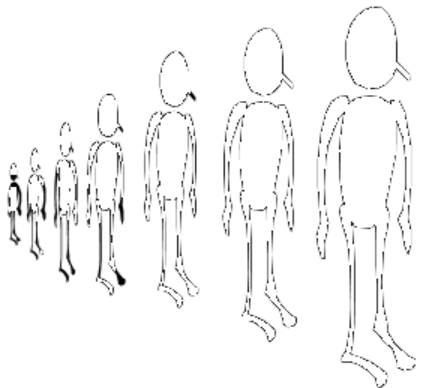
- Ideal “causal contrast” between exposed and unexposed groups:
 - “A causal contrast compares disease frequency under *two* exposure distributions, but in *one* target population during *one* etiologic time period”
 - If the ideal causal contrast is met, the observed effect is the “causal effect”



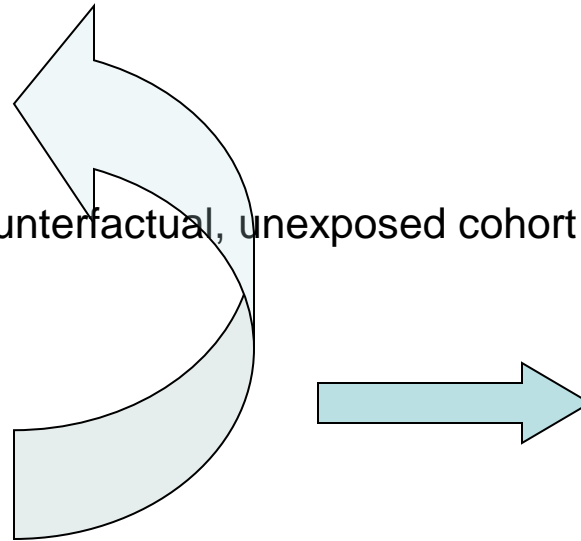
Exposed cohort



Counterfactual, unexposed cohort



Substitute, unexposed cohort



“Confounding is present if the substitute population imperfectly represents what the target would have been like under the counterfactual condition”

Control of confounding: Outline

- Control at the design stage
 - Randomization
 - Restriction
 - Matching
- Control at the analysis stage
 - Conventional approaches
 - Stratified analyses
 - Multivariate analyses
 - Newer approaches
 - Graphical approaches using DAGs
 - Propensity scores
 - Instrumental variables
 - Marginal structural models

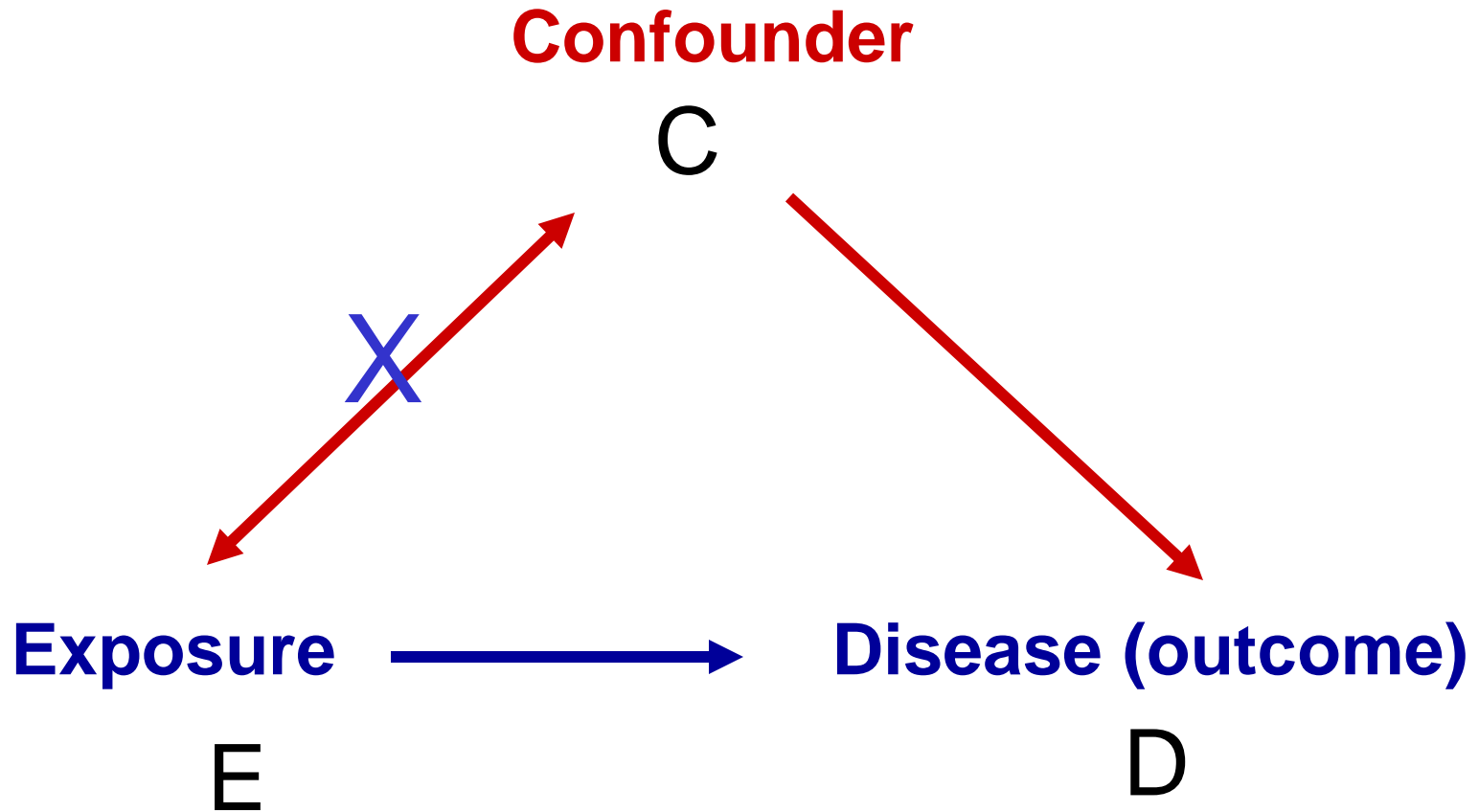
Control of confounding: at the design stage

- Options at the design stage:
 - **Randomization**
 - Reduces potential for confounding by generating groups that are fairly comparable with respect to known and unknown confounding variables (attempts to simulate a counterfactual contrast)
 - **Restriction**
 - Eliminates variation in the confounder (e.g. only recruiting males into the study will eliminate confounding by sex)
 - **Matching**
 - Involves selection of a comparison group that is forced to resemble the index group with respect to the distribution of one or more potential confounders

Randomization

- Randomization
 - Useful only for intervention studies
 - Definition: random assignment of study subjects to exposure categories (this breaks any links between exposure and confounders)
 - The special strength of randomization is its ability to control/reduce the effect of confounding variables about which the investigator is unaware (i.e. both known and unknown confounders get distributed evenly because of randomization)
 - Randomization does not always eliminate confounding
 - Covariate imbalance after randomization can occur in small trials
 - If there is “maldistribution” of potentially confounding variables after randomization (the reason for the classic “Table I: Baseline characteristics” in the randomized trial) then other confounding control options (see below) are applied

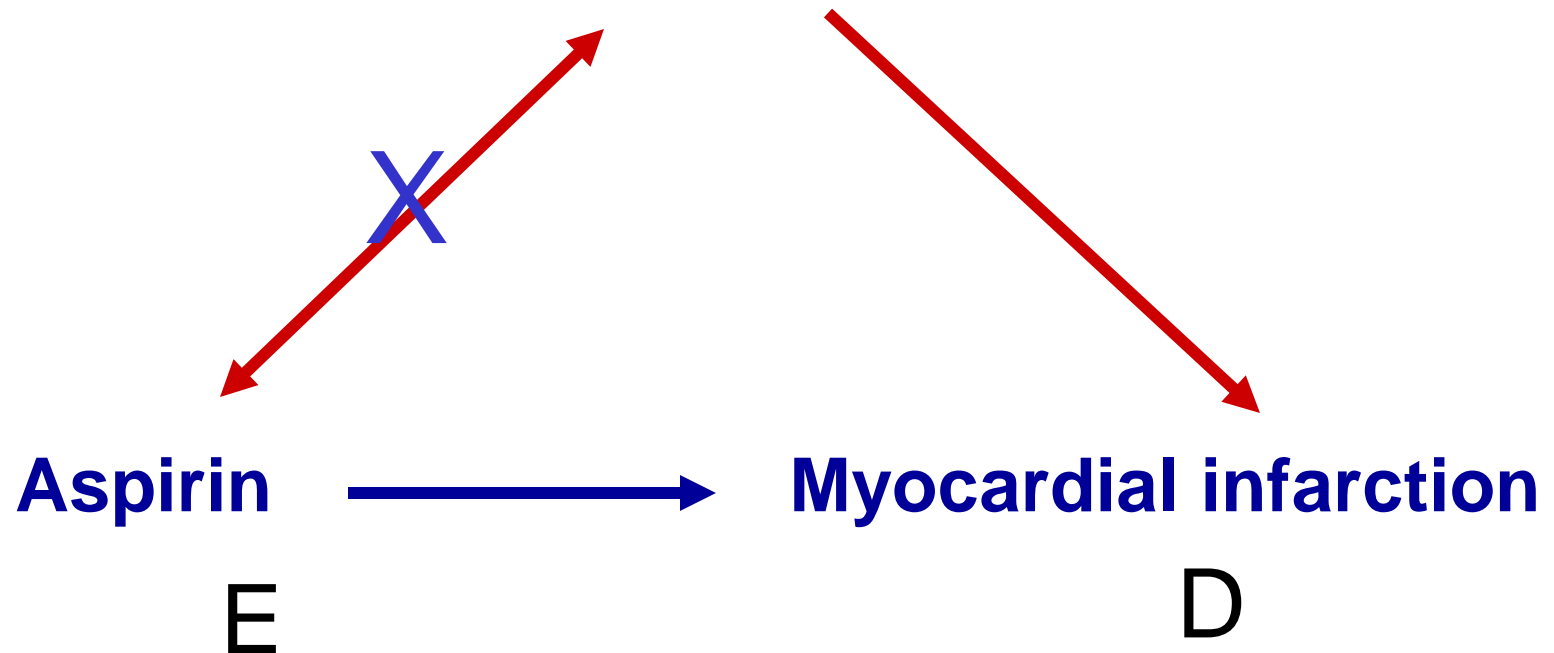
Randomization breaks any links
between treatment and prognostic factors



There can be no association between exposure and confounders (because intervention is randomly allocated)

Example: aspirin to prevent MI

Confounders: Obesity, smoking, family history of heart disease, etc



In a RCT, there can be no association between aspirin and confounders (because intervention is randomly allocated)

Randomization results in groups with comparable distribution of confounders

Example: First RCT on circumcision and HIV

Table 2. Baseline Characteristics of HIV-Negative Men Enrolled in the Trial

Background Characteristics		Control <i>n</i> = 1,582	Intervention <i>n</i> = 1,546
Age	Less than or equal to 21 y	52.4%	48.6%
	More than 21 y	47.6%	51.4%
Primary level of education completed		98.4%	98.3%
Religion	African traditional	47.0%	51.6%
	Protestant or Catholic	11.1%	11.9%
	Other religion	41.8%	36.5%
Ethnic group	Sotho	47.3%	49.0%
	Zulu	38.1%	32.8%
	Other	14.6%	18.2%
Drank alcohol in the past month		41.9%	42.2%
Reported sexual behaviour			
Have had first sexual experience		90.5%	91.8%
Median (IQR) age at first sex (years) ^a		16.6 (15.2–18.4)	16.8 (15.4–18.5)
Median (IQR) number of lifetime sex partners ^b		4 (2–7)	4 (3–7)
Used a condom at first sex ^b		13.4%	15.2%
Ever used a condom ^b		81.2%	82.3%
At-risk behaviour ^{c,d}		46.7%	46.8%
Married or living as married ^d		1.8%	1.8%
Mean (IQR) number of non-spousal partners ^e		1.4 (0–2)	1.4 (0–2)
At least one sexual partnership with only one sexual contact ^e		29.8%	30.7%
Mean (IQR) number of sexual contacts ^e		8.0 (0–8)	8.7 (1–8)
Attended a clinic for a health problem related to the genital area ^e		10.0%	9.6%

Randomization resulted in highly comparable
distribution of potential confounders; so
confounding is not an issue!

Restriction

- Confounding cannot occur if the distribution of the potential confounding factors do not vary across exposure or disease categories
 - Implication of this is that an investigator may restrict study subjects to only those falling with specific level(s) of a confounding variable
 - Extreme example: an investigator only selects subjects of exactly the same age or same sex.
- Advantages of restriction
 - straightforward, convenient, inexpensive (but, reduces recruitment!)

Restriction

- Disadvantages
 - Will limit number of eligible subjects
 - Will limit ability to generalize the study findings
 - Residual confounding may persist if restriction categories not sufficiently narrow (e.g. “decade of age” might be too broad)
 - Not possible to evaluate the relationship of interest at different levels of the confounder

Matching

- Matching is commonly used in case-control studies (where each case is matched to a control)
- For example, if age and sex are the matching variables, then a 35 year old male case is matched to a 35 year old male control
- Types:
 - Pair matching (one to one individual matching)
 - Frequency matching
- The use of matching usually requires special analysis techniques (e.g. matched pair analyses and conditional logistic regression)

Matching

Table 1. Advantages and Disadvantages of Matching in Epidemiologic Studies

Advantages	Disadvantages
Control of unmeasured confounders	Potential additional costs and time
Sufficient controls in sub-analyses	Exclusion of cases without suitable controls
Time comparability	Complex data analysis
Direct control of confounders	Confounder effect cannot be estimated
	Overmatching

Matching

Reasons For Matching

Gain precision

Control for variables difficult to measure

Practical aspects

Control for confounding of both matched and unmatched variables

Reasons Against Matching

Matching on weak risk factors may lose precision

Matching may be costly in terms of time and money

You may have difficulty finding matches

You may overmatch

Can control for confounding even without matching

Matching

- Disadvantages of matching
 - Finding appropriate control subjects may be difficult and expensive and limit sample size
 - Confounding factor used to match subjects cannot be itself evaluated with respect to the outcome/disease
 - Obviously, matching does not control for confounding by factors other than that used to match
 - The use of matching makes the use of stratified analysis (for the control of other potential but non-matched factors) very difficult

Matching

- Disadvantages of matching
 - Matching is most often used in case- control studies because in a large cohort study the cost of matching may be prohibitive
 - In a case-control study, matching may introduce confounding rather than eliminate confounding!
 - *Très compliqué* and will be covered briefly here, but more in intermediate epi (EPIB603)

Matching in a case control study

- Why can matching introduce confounding in a case-control study?
 - The purpose of the control series in a case-control study is to permit an estimate of the person-time distribution of exposure in the source population (study base) from which cases arose
 - In other words, controls selected for a case-control study are supposed to reflect the distribution of exposure in the source population.
 - If controls are selected to match the cases for a factor that is correlated with exposure, this will change the distribution of the control population away from the distribution in the source population
 - then the crude exposure odds in controls is distorted in the direction of similarity to that of the cases (e.g. friend controls of the same age and sex).
 - This will introduce a selection bias that is very similar to confounding.
 - If the matching factor were perfectly correlated with the exposure, the exposure distribution of controls would be identical to cases, and the crude OR estimate would be 1.0

Matching in case control studies

- So, matching (whose well-meaning goal is to reduce confounding) does not attain that objective in case-control studies
 - It substitutes a new confounding structure for the old one
 - At times, it will introduce new confounding where none previously existed
 - So, use matching with caution
 - If you must match, then match on a strong confounder
- The confounding introduced by matching into a case-control study is not irremediable
 - The proper form of analysis is needed in matched studies in order to reduce the confounding introduced by matching
 - Matched Odds Ratios
 - Conditional logistic regression
 - General rule: paired data must be treated as such in the analysis (do not break up the pairs!)

Matched odds ratio

		Control	
		Exposed	Not exposed
Case	Exposed	a	b
	Not exposed	c	d

$$\text{Matched OR (mOR)} = b/c$$

Definition: OR in a matched case-control study is defined as the ratio of the number of pairs a case was exposed and the control was not to the number of ways the control was exposed and the case was not

The pairs in cells A and D do not contribute any information since they are concordant

Control of confounding: at the analysis stage

- Unlike selection and information bias, confounding is one type of bias that can be, to some extent, adjusted in the analysis
- Options at the analysis stage:
 - Stratification
 - Multivariate methods
- To control for confounding in the analyses, one must measure the confounders in the study!
 - Investigators usually do this by collecting data on all known, previously identified confounders
 - A sound knowledge of disease biology will help decide what to measure

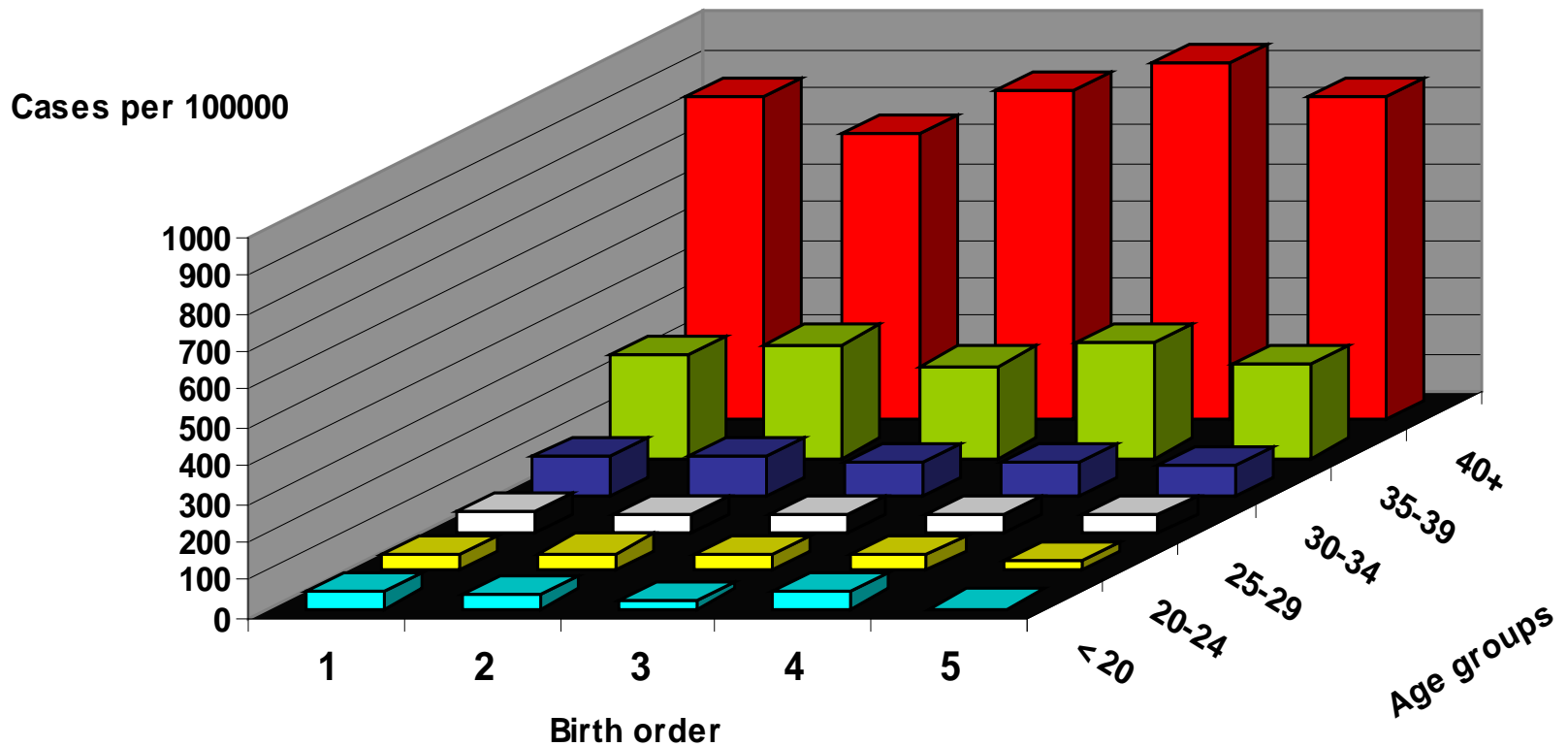
Stratification

- Objective of stratified analysis is to “fix” the level of the confounding variable and produce groups within which the confounder does not vary
- Then evaluate the exposure-disease association within each stratum of the confounder
- Within each stratum, the confounder cannot confound because it does not vary

- Question: Who is an epidemiologist?
- Answer: A physician broken down by age and sex!

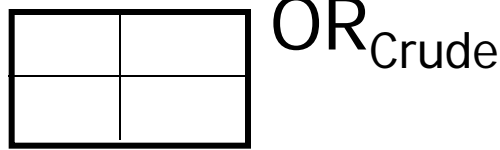
Stratified Analysis: graphic approach

Cases of Down syndrom by birth order and mother's age



Stratified Analysis

Crude



Crude 2 x 2 table

Calculate Crude OR (or RR)

Stratify by Confounder

Stratum 1

Stratum 2



OR_1

OR_2

Calculate OR's
for each stratum

If stratum-specific OR's are similar,
calculate adjusted OR (e.g. M-H)

If Crude OR \neq Adjusted OR,
confounding is likely

If Crude OR = Adjusted
OR, confounding is
unlikely

Mantel-Haenszel (M-H) estimators of adjusted effects from stratified data

Case-Control Study:

$$OR_{MH} = \frac{\sum(ad / T)_i}{\sum(bc / T)_i}$$

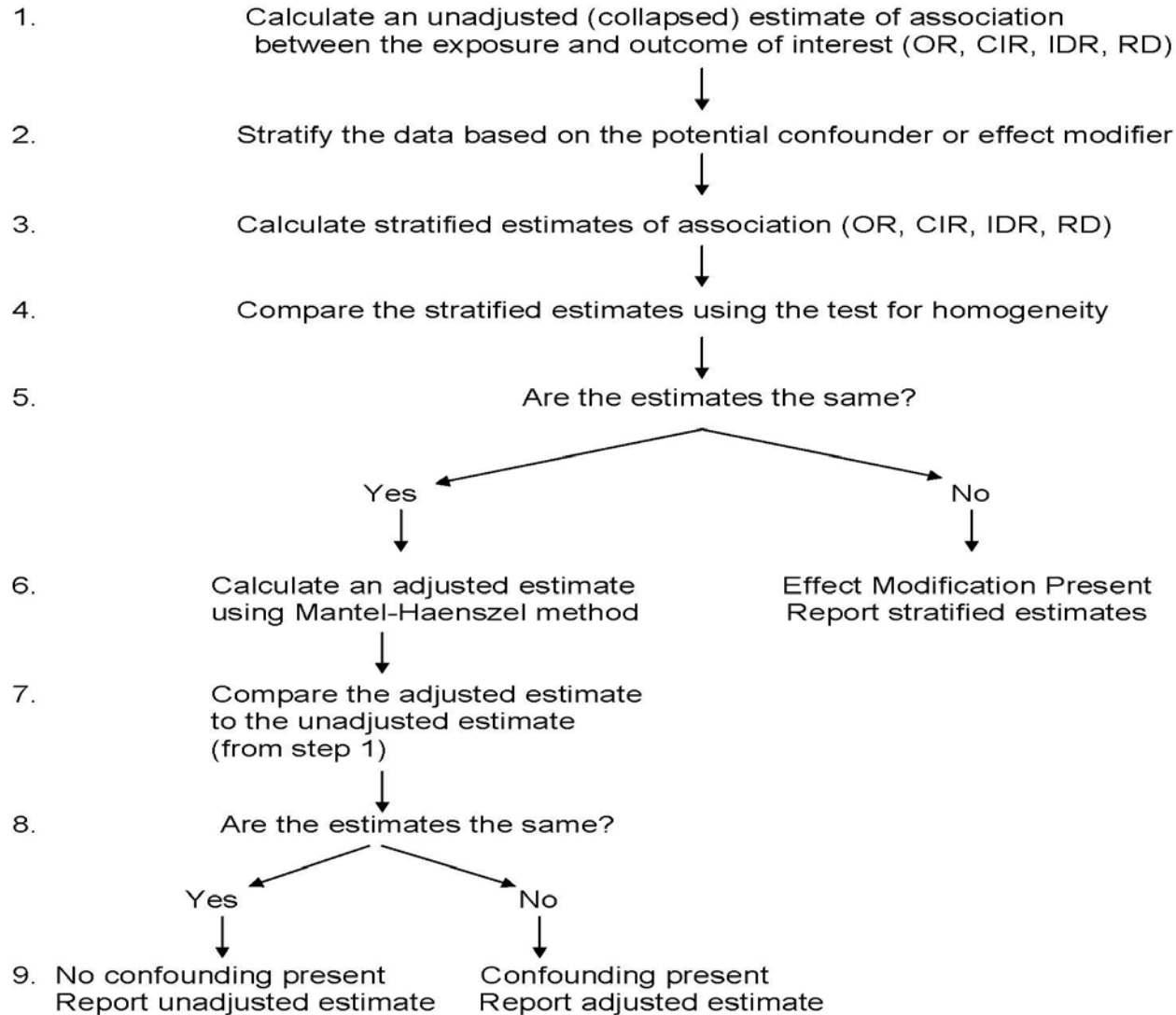
Cohort Study with Count Denominators:

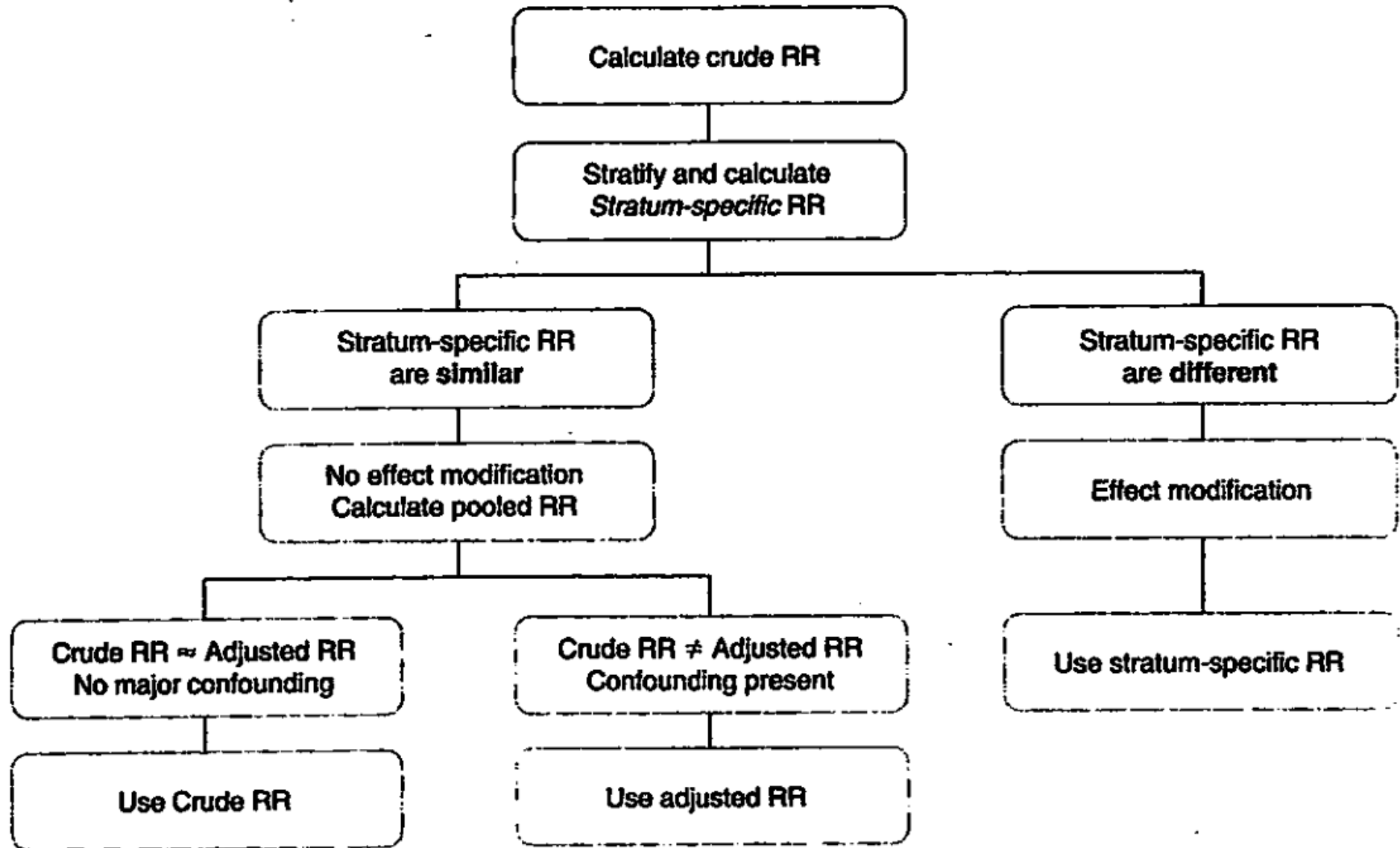
$$RR_{MH} = \frac{\sum\{a(c + d) / T\}_i}{\sum\{b(a + b) / T\}_i}$$

Cohort Study with Person-years Denominators:

$$RR_{MH} = \frac{\sum\{a(PY_0) / T\}_i}{\sum\{b(PY_1) / T\}_i}$$

Decision tree for evaluating confounding and effect modification





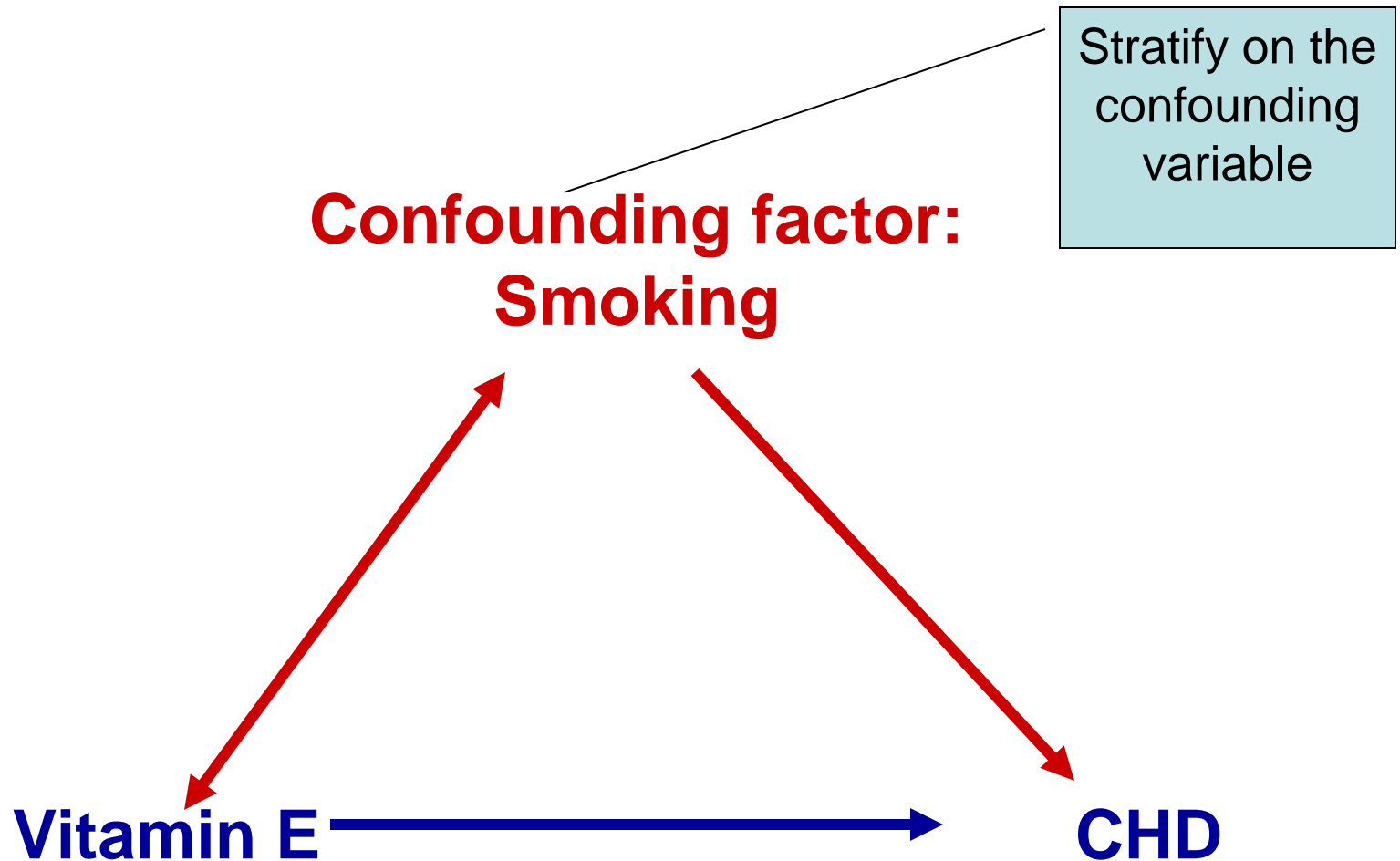
Observational Study on Vit E and Coronary Heart Disease

		CHD	
		Present	Absent
Vitamin E Supplement	Yes	50	501
	No	65	384

$$\text{Crude OR} = (50)(384)/(501)(65) = 0.59$$

Are there potential confounders that can explain this crude OR?

Could reduced smoking among Vit E users partly explain the observed protective effect?



Stratified Analyses (by smoking status)

Smokers:

		CHD	
		Present	Absent
Vitamin E Supplement	Yes	11	40
	No	49	200

Stratum 1

$$\text{OR (smokers)} = (11)(200)/(40)(49) = 1.12$$

Non-Smokers:

		CHD	
		Present	Absent
Vitamin E Supplement	Yes	39	461
	No	16	184

Stratum 2

$$\text{OR (non-smokers)} = (39)(184)/(461)(16) = 0.97$$

Inference

- Before adjusting for smoking, Vit E had a strong protective effect (Crude OR=0.59)
- After stratification, there is little evidence of an association between vitamin E and CHD after controlling for smoking (Stratum-specific ORs 1.1 & 0.97)
- Why is there an apparent contradiction with the results obtained from the crude data?
- Vitamin E group contains considerably fewer smokers (9.3% versus 55.5%), thereby decreasing this group's apparent risk of CHD

Direction of Confounding Bias

- Confounding “pulls” the observed association away from the true association
 - It can either **exaggerate/over-estimate** the true association (positive confounding)
 - Example
 - $RR_{\text{causal}} = 1.0$
 - $RR_{\text{observed}} = 3.0$
 - or
 - It can **hide/under-estimate** the true association (negative confounding)
 - Example
 - $RR_{\text{causal}} = 3.0$
 - $RR_{\text{observed}} = 1.0$

Limitations of Stratified Analyses

Crude

OR_{Crude}

Stratum 1

OR_1

Stratum 2

OR_2

Confounder 1

Stratum 1

Stratum 2

Stratum 1

Stratum 2

Confounder 2

“Sparseness of data” problem

Computational and interpretational difficulties

Multivariate Analysis

- Stratified analysis works best when there are few strata (i.e. if only 1 or 2 confounders have to be controlled)
- If the number of potential confounders is large, multivariate analyses offer the only real solution
 - Can handle large numbers of confounders (covariates) simultaneously (e.g. in the Vit E example, one could control for smoking, alcohol, physical activity, diet, in the same analysis)
 - Based on statistical regression “models”
 - E.g. logistic regression, multiple linear regression
 - Always done with statistical software packages (e.g. SAS, Stata)
 - Multivariate analysis was a nightmare before the advent of microcomputers and software options!



Computing has come a long way!

Multivariate Analysis: Demo

- Example: Association between preterm labor and low birth weight
- N = 189 subjects
- Outcome: low birth weight [**low**] (yes=1; no=0)
- Main exposure of interest: History of preterm labor [**ptl**] (yes=1; no=0)
- Covariates:
 - Smoking status of mother [**smoke**] (yes=1; no=0)
 - **Age** (mother's age in years)
 - **Race** (0=white; 1=black/others)

	ID	low	age	smoke	ptl	race	
1	4	1	28	1	1	1	
2	10	1	29	0	0	0	
3	11	1	34	1	0	1	
4	13	1	25	0	1	1	
5	15	1	25	0	0	1	

Bivariate analysis: Preterm labor and low birth weight

- Outcome: low birth weight [**low**] (yes=1; no=0)
- Main exposure of interest: History of premature labor [**ptl**] (yes=1; no=0)
- 59/189 (31%) were low birth weight
- Crude association: OR = 4.31

```
. cc low ptl
```

	Exposed	Unexposed	Total	Proportion Exposed
Cases	18	41	59	0.3051
Controls	12	118	130	0.0923
Total	30	159	189	0.1587
	Point estimate		[95% Conf. Interval]	
Odds ratio	4.317073		1.777511	10.65589 (exact)
Attr. frac. ex.	.7683616		.4374155	.9061552 (exact)
Attr. frac. pop	.2344154			

chi2(1) = 13.76 Pr>chi2 = 0.0002

Stratified analysis (by smoking status)

Preterm labor and low birth weight

Mantel-Haenszel adjusted OR:

```
. cc low pt1, by(smoke)
```

smoke	OR	[95% Conf. Interval]		M-H weight	
0	3.478261	.8320418	14.22408	1.2	(exact)
1	4.222222	1.20591	15.76105	1.459459	(exact)
Crude	4.317073	1.777511	10.65589		(exact)
M-H combined	3.886532	1.697176	8.900158		

Test of homogeneity (M-H) $\chi^2(1) = 0.05$ $Pr>\chi^2 = 0.8191$

Test that combined OR = 1:
Mantel-Haenszel $\chi^2(1) = 10.97$
 $Pr>\chi^2 = 0.0009$

Crude OR = 4.31
OR (smokers) = 4.22
OR (non-smokers) = 3.47
M-H (adjusted) OR = 3.88

Multivariate analysis: Preterm labor and low birth weight

Model 1: Logistic regression model with no covariates (unadjusted)

```
. logistic low pt1
Logistic regression
Log likelihood = -110.94888
Number of obs = 189
LR chi2(1) = 12.77
Prob > chi2 = 0.0004
Pseudo R2 = 0.0544
```

low	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
pt1	4.317073	1.789132	3.53	0.000	1.916128	9.72645

OR is the same as we got from a simple 2 x 2 table

Multivariate analysis: Preterm labor and low birth weight

Model 2: Logistic regression model with covariates (adjusted)

- ORs are adjusted for only smoking

```
. logistic low pt1 smoke
```

```
Logistic regression
```

```
Number of obs   =      189  
LR chi2(2)      =      15.34  
Prob > chi2     =      0.0005  
Pseudo R2      =      0.0654
```

```
Log likelihood = -109.66445
```

	low	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
pt1		3.863837	1.628601	3.21	0.001	1.69137 8.826711
smoke		1.709971	.5702925	1.61	0.108	.8894148 3.287554

After adjusting for smoking, the OR is significant and strong, but less than the crude OR

Multivariate analysis: Preterm labor and low birth weight

Model 3: Logistic regression model with all covariates (adjusted)

- ORs are adjusted for age, race and smoking

```

. logistic low pt1 age race smoke

Logistic regression
Log likelihood = -104.21339
Number of obs = 189
LR chi2(4) = 26.25
Prob > chi2 = 0.0000
Pseudo R2 = 0.1118

```

	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
low						
pt1	4.099114	1.817752	3.18	0.001	1.718798	9.775868
age	.9459476	.0340036	-1.55	0.122	.8815953	1.014997
race	2.620422	1.006873	2.51	0.012	1.233971	5.564648
smoke	2.477854	.9507154	2.36	0.018	1.168107	5.256165

After adjusting for age, race and smoking, the OR is significant and strong, but less than the crude OR; Smoking and race also appear to be independent risk factors for low birth weight

Logistic regression equation

- General:
 - $\ln(\text{odds of disease}) = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4$
 - Where, a is the intercept; b_1 b_2 etc are the beta coefficients; and X_1 , X_2 , etc are covariates in the model
- Low birth weight study:
 - $\ln(\text{odds of low birth wt}) = a + b_1(\text{preterm labor}) + b_2(\text{age}) + b_3(\text{race}) + b_4(\text{smoking})$
 - From the final model, we can get ORs for each covariate by generating the exponential of the beta coefficients (i.e. e^b)

More on this in EPIB621
and computer lab on Oct 21

Residual confounding

- Confounding can persist, even after adjustment
- Why?
 - All confounders were not adjusted for (unmeasured confounding)
 - Some variables were actually not confounders!
 - Confounders were measured with error (misclassification of confounders)
 - Categories of the confounding variable are improperly defined (e.g. age categories were too broad)

Residual confounding: a case study

ORIGINAL ARTICLE

Tobacco smoking and pulmonary tuberculosis

C Kolappan, P G Gopi

Thorax 2002;**57**:964–966

See end of article for authors' affiliations

Correspondence to:
Dr C Kolappan,
Epidemiology Unit,
Tuberculosis Research
Centre, Mayor V R
Ramanathan Road,
Chetput, Chennai 600
031, Tamil Nadu, India;
kola15@sify.com

Received
21 November 2001
Accepted for publication
31 July 2002

Background: The prevalence of tuberculosis in adult men in India is 2–4 times higher than in women. Tobacco smoking is prevalent almost exclusively among men, so it is possible that tobacco smoking may be a risk factor for developing pulmonary tuberculosis. A nested case control study was carried out to study the association between tobacco smoking and pulmonary tuberculosis.

Methods: A tuberculosis disease survey was carried out in two Panchayat unions in the Tiruvallur district of Tamil Nadu in India. Eighty five men aged 20–50 years with bacteriological tuberculosis (smear and/or culture positive) were selected as cases and 459 age matched men without tuberculosis were selected randomly as controls. Information on smoking status, type of tobacco smoked, quantity of tobacco smoked, and duration of tobacco smoking was collected from cases and controls using a questionnaire.

Results: The estimated crude odds ratio (OR) of the association between tobacco smoking and bacillary tuberculosis was 2.48 (95% confidence interval (CI) 1.42 to 4.37), $p < 0.001$. The age adjusted OR (Mantel-Hanszel estimate) was 2.24 (95% CI 1.27 to 3.94), $p < 0.05$. The ORs for mild (1–10 cigarettes/day), moderate (11–20/day), and heavy (>20/day) smokers were 1.75, 3.17, and 3.68, respectively ($p < 0.0001$ test for linear trend). The ORs for smokers with <10 years, 11–20 years, and >20 years of smoking were 1.72, 2.45, and 3.23, respectively ($p < 0.0001$ test for linear trend).

Conclusion: There is a positive association between tobacco smoking and pulmonary (bacillary) tuberculosis (OR 2.5). The association also shows a strong dose-response relationship.

Residual confounding: a case study

Smoking and Tuberculosis among the Elderly in Hong Kong

Chi C. Leung, Teresa Li, Tai H. Lam, Wing W. Yew, Wing S. Law, Cheuk M. Tam, Wai M. Chan, Chi K. Chan, Kin S. Ho, and Kwok C. Chang

TB and Chest Service; Elderly Health Service, Department of Health; Department of Community Medicine, The University of Hong Kong; and the TB and Chest Unit, Grantham Hospital, Hong Kong, China

TABLE 4. ADJUSTED HAZARD RATIOS OF TUBERCULOSIS-RELATED OUTCOMES BY SMOKING STATUS IN COX PROPORTIONAL HAZARDS ANALYSIS

Outcome	Ex-smokers Versus Never-smokers			Current Smokers Versus Never-smokers			p Value for Trend
	HR*	95% CI	p Value	HR*	95% CI	p Value	
Active TB	1.41	1.02–1.95	0.04	2.63	1.87–3.70	< 0.001	< 0.001
Culture-confirmed TB	1.68	1.13–2.50	0.01	2.80	1.82–4.31	< 0.001	< 0.001
TB > 3 mo	1.48	1.04–2.10	0.03	2.39	1.63–3.50	< 0.001	< 0.001
Culture-confirmed TB > 3 mo	1.73	1.13–2.67	0.01	2.36	1.45–3.86	0.001	< 0.001
New TB	1.44	1.01–2.04	0.04	2.61	1.80–3.80	< 0.001	< 0.001
Retreatment TB cases	1.16	0.49–2.73	0.74	2.48	1.04–5.89	0.04	< 0.05
Pulmonary TB	1.39	0.98–1.97	0.06	2.87	2.00–4.11	< 0.001	< 0.001
Extrapulmonary TB	1.32	0.57–3.05	0.52	1.04	0.33–3.30	0.95	0.79
Extrapulmonary TB only	1.77	0.72–4.35	0.22	0.73	0.16–3.46	0.70	0.86

Definition of abbreviations: CI = confidence interval; HR = hazard ratio; TB = tuberculosis.

* Adjusted hazard ratio, adjusted for sex, age, alcohol use, language, marital status, education, housing, working status, public financial assistance status, monthly expenditure, participation in social activities, self-rated health status, hospital admission within 12 months, diabetes mellitus, chronic obstructive pulmonary disease, hypertension, heart disease, and cerebrovascular disease, with categories as listed in Table 1.

Residual confounding: a case study

- In the observational component of the Multiple Risk Factor Intervention Trial, subjects were followed to determine whether changes in risk behaviour (e.g. smoking cessation) reduces the risk of coronary heart disease
- A coincidental finding was an increased risk of suicide among smokers.
- After adjustment for a large number of confounding factors, compared with non-smokers, the RRs of suicide for those who smoked 1–19, 20–39, 40–59 and ≥ 60 cigarettes per day were 1.4, 1.9, 2.3 and 3.4, respectively (significant trend).
- One explanation is that depressed people tend to smoke and also to commit suicide.
- That is, depression confounded the association; it produced a statistically significant, but nevertheless spurious, dose-response relationship
- The confounding effect of depression could not be eliminated because it was not adequately defined or measured.

Overadjustment and unnecessary adjustment

Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies

Enrique F. Schisterman,^a Stephen R. Cole,^b and Robert W. Platt^c

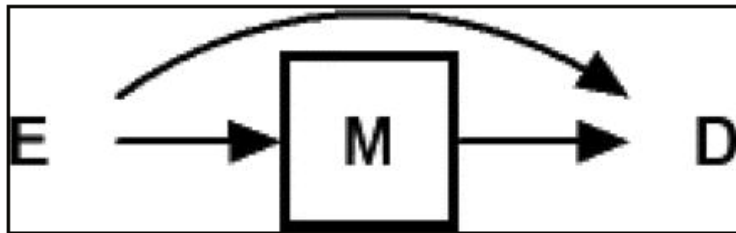
Abstract: Overadjustment is defined inconsistently. This term is meant to describe control (eg, by regression adjustment, stratification, or restriction) for a variable that either increases net bias or decreases precision without affecting bias. We define overadjustment bias as control for an intermediate variable (or a descending proxy for an intermediate variable) on a causal path from exposure to outcome. We define unnecessary adjustment as control for a variable that does not affect bias of the causal relation between exposure and outcome but may affect its precision. We use causal diagrams and an empirical example (the effect of maternal smoking on neonatal mortality) to illustrate and clarify the definition of overadjustment bias, and to distinguish overadjustment bias from unnecessary adjustment. Using simulations, we quantify the amount of bias associated with overadjustment. Moreover, we show that this bias is based on a different causal structure from confounding or selection biases. Overadjustment bias is not a finite sample bias, while inefficiencies due to control for unnecessary variables are a function of sample size.

(Epidemiology 2009;20: 488–495)

confounding¹ and selection biases^{2,3} have been discussed extensively in the epidemiologic literature, the concept of “overadjustment” has had relatively little attention. The definition of overadjustment remains vague and the causal structure of this concept has not been well described.

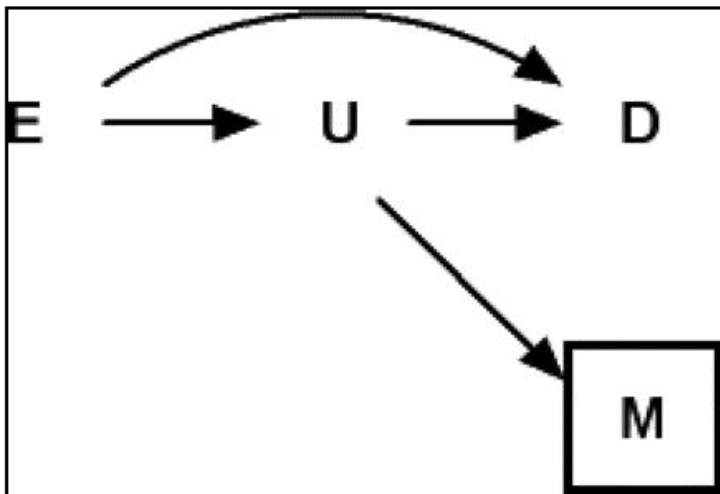
The Dictionary of Epidemiology⁴ cites a seminal paper by Breslow⁵ in broadly defining overadjustment as “Statistical adjustment by an excessive number of variables or parameters, uninformed by substantive knowledge (eg, lacking coherence with biologic, clinical, epidemiological, or social knowledge). It can obscure a true effect or create an apparent effect when none exists.” Rothman and Greenland⁶ discuss overadjustment in the context of intermediate variables: “Intermediate variables, if controlled in an analysis, would usually bias results towards the null. . . . Such control of an intermediate may be viewed as a form of overadjustment.” One also finds reference to the term overadjustment in settings with unnecessary control for variables.⁷ In summary, overadjustment sometimes means control (eg, by regression

Overadjustment bias is control for an intermediate variable (or a descending proxy for an intermediate variable) on a causal path from exposure to outcome



Example: mediating role of triglycerides (M) in the association between prepregnancy body mass index (E) and preeclampsia (D)

If M is adjusted for, then the causal effect will be biased towards null



Example: adjusting for prior history of spontaneous abortion (M); an underlying abnormality in the endometrium (U) is the unmeasured intermediate caused by smoking (E), and is a cause of prior (M) and current (D) spontaneous abortion. M is a “descending” proxy for the intermediate variable U.

If U is adjusted for, the observed association between the exposure E and outcome D will typically be biased toward the null with respect to the total causal effect

Content area expertise is important for evaluation of confounding



American Journal of Epidemiology
Copyright © 2002 by the Johns Hopkins Bloomberg School of Public Health
All rights reserved

Vol. 155, No. 2
Printed in U.S.A.

Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology

Miguel A. Hernán,¹ Sonia Hernández-Díaz,² Martha M. Werler,² and Allen A. Mitchell²

Common strategies to decide whether a variable is a confounder that should be adjusted for in the analysis rely mostly on statistical criteria. The authors present findings from the Slone Epidemiology Unit Birth Defects Study, 1992–1997, a case-control study on folic acid supplementation and risk of neural tube defects. When statistical strategies for confounding evaluation are used, the adjusted odds ratio is 0.80 (95% confidence interval: 0.62, 1.21). However, the consideration of a priori causal knowledge suggests that the crude odds ratio of 0.65 (95% confidence interval: 0.46, 0.94) should be used because the adjusted odds ratio is invalid. Causal diagrams are used to encode qualitative a priori subject matter knowledge. *Am J Epidemiol* 2002;155:176–84.

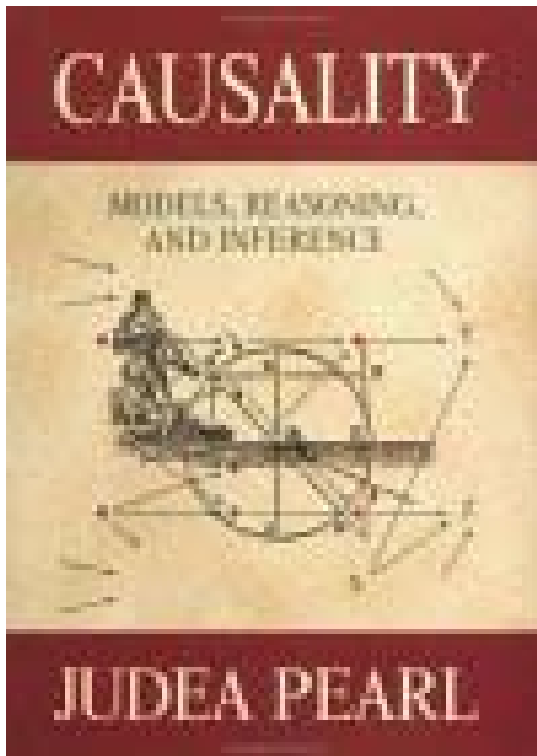
Know your field!!

Newer approaches to address confounding

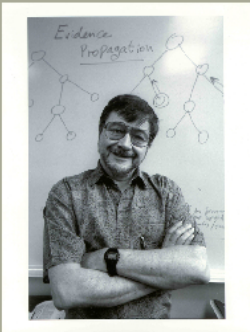
- Mostly based on the counterfactual causality model
 - Causal graphs (Directed Acyclic Graphs) [Pearl 2000]
 - Marginal structural models (Robins 2000)
 - Propensity score analysis (Rubin 1974; Rosenbaum & Rubin 1983)

You are not responsible for this material

Graphical solutions to the adjustment problem (based on directed acyclic graphs [DAG])



JUDEA PEARL



UCLA
Computer Science Department

Cognitive Systems Lab
1532 Boelter Hall
Tel: (310) 825-3243
Fax: (310) 825-2273
judea@cs.ucla.edu

JUDEA PEARL - HOME

HOME BIOGRAPHICAL PUBLICATIONS CAUSALITY CSL
DANIEL PEARL FNDN

Welcome to my homepage.

To view my slides from the Joint Statistical Meetings (JSM-07) held in Salt Lake City, Utah, [click here](#).

For a gentle introduction to my current research on causality, [click 1 or 2].

- 1 Transcript and slides of 1996 Faculty Research Lecture: [The Art and Science of Cause and Effect](#)
- 2 Transcript and slides of 1999 IJCAI Award Lecture: [Reasoning with Cause and Effect](#)

[For a recent video on the state of causality in economics click here.](#)

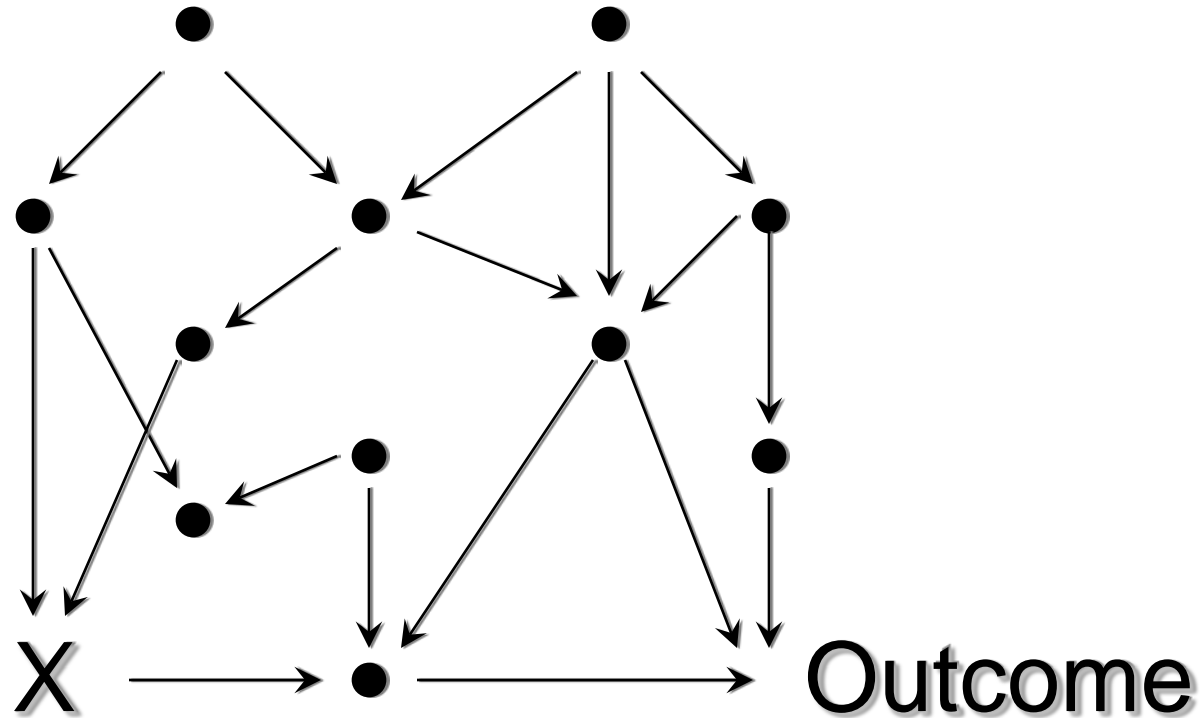
Technical material supporting the story in 1-2, can be found [postscript] or [pdf] in:

- (R-299): [\[pdf\]](#)
J. Pearl, "Statistics and Causal Inference: A Review," *Test Journal*, 12(2):281-345, December 2003 (with discussions).
- (R-264): [\[postscript\]](#) [\[pdf\]](#)
J. Pearl, "Simpson's paradox: An anatomy" Extracted from Chapter 6 of [CAUSALITY](#).
- (R-218-B): [\[postscript\]](#) [\[pdf\]](#)
J. Pearl, "Causal Diagrams for Empirical Research" *Biometrika*, 82(4), 669-710, December 1995.

Pearl J. Causality: models, reasoning and inference.
Cambridge: Cambridge University Press, 2000

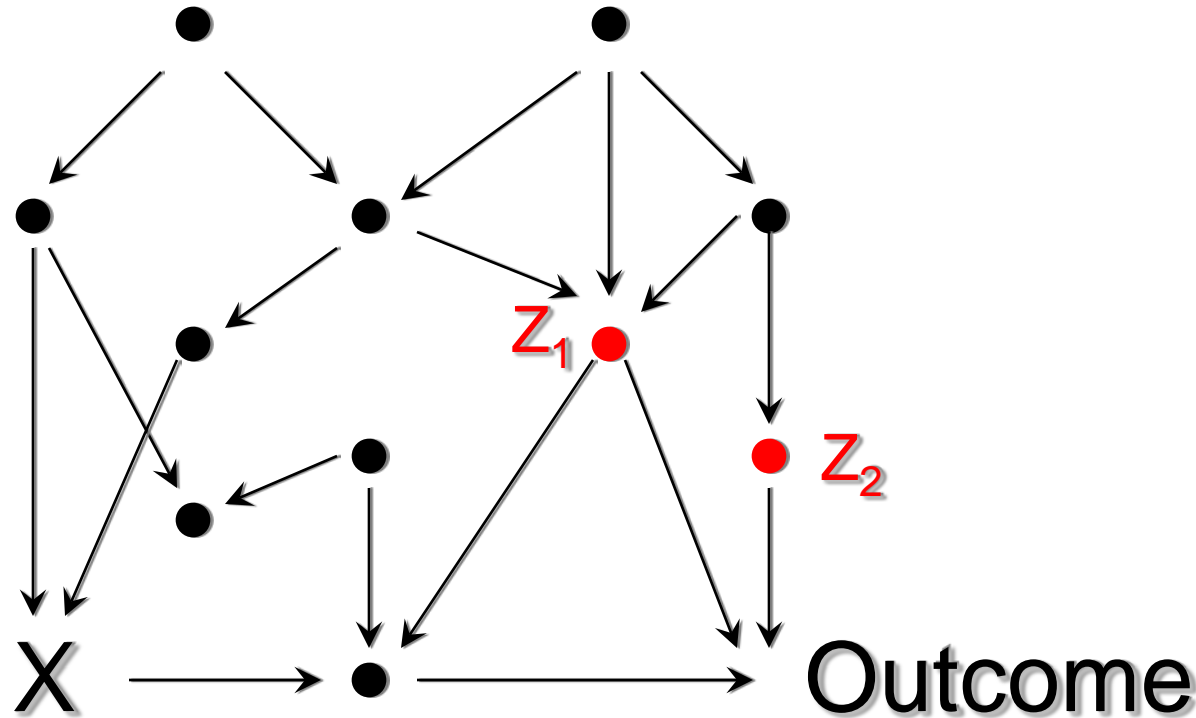
Greenland S, Pearl J, Robins JM.
Causal diagrams for epidemiologic
research. *Epidemiology*
1999;10(1):37-48.

Graphical solutions to the adjustment problem



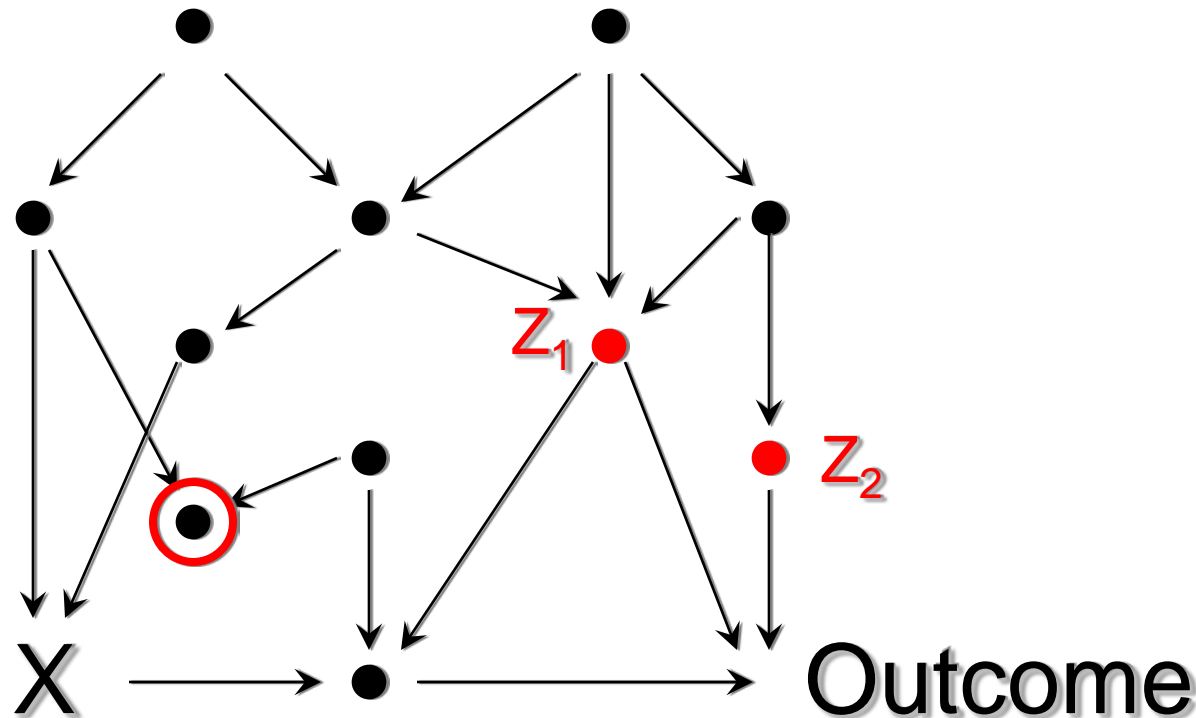
Which measurements should be included in the model if we are interested in the relation between X and Outcome?

Graphical solutions to the adjustment problem



Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z_1 and Z_2 remove confounding?

Graphical solutions to the adjustment problem

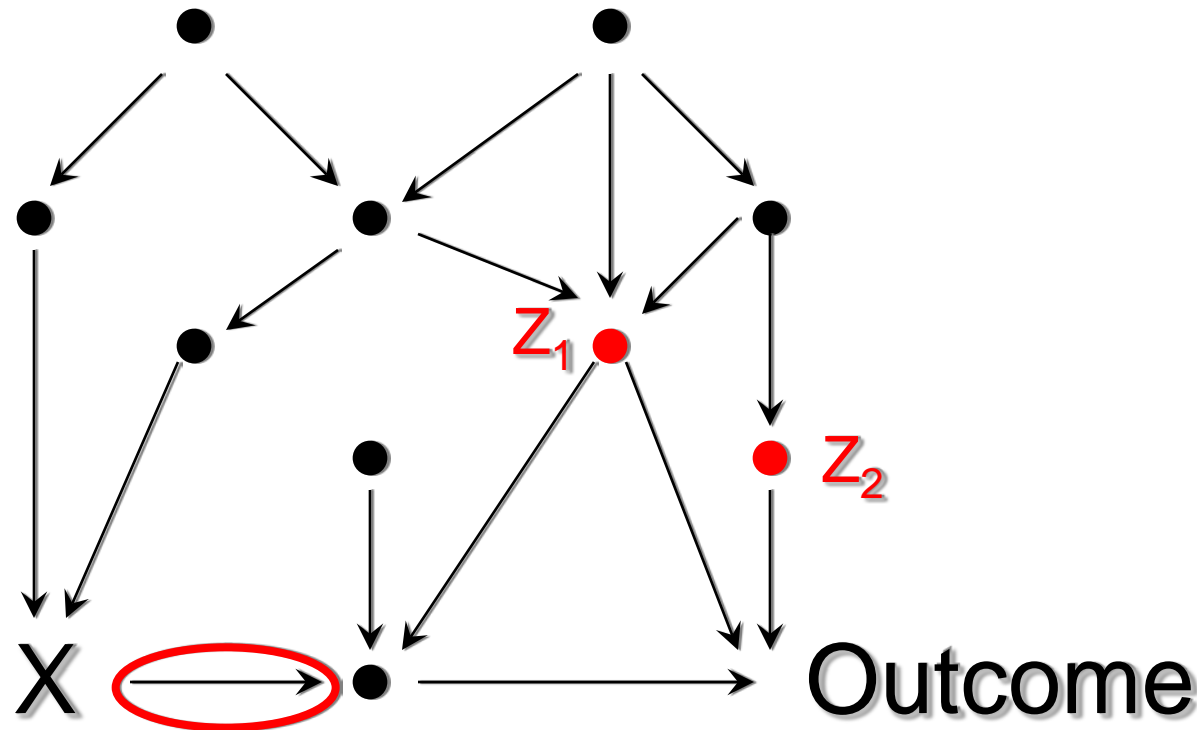


Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z_1 and Z_2 remove confounding?

Step 1: Z_1 and Z_2 should not be descendants of X .

Step 2: Delete all non-ancestors of $\{X, \text{Outcome}, Z_1 \text{ and } Z_2\}$.

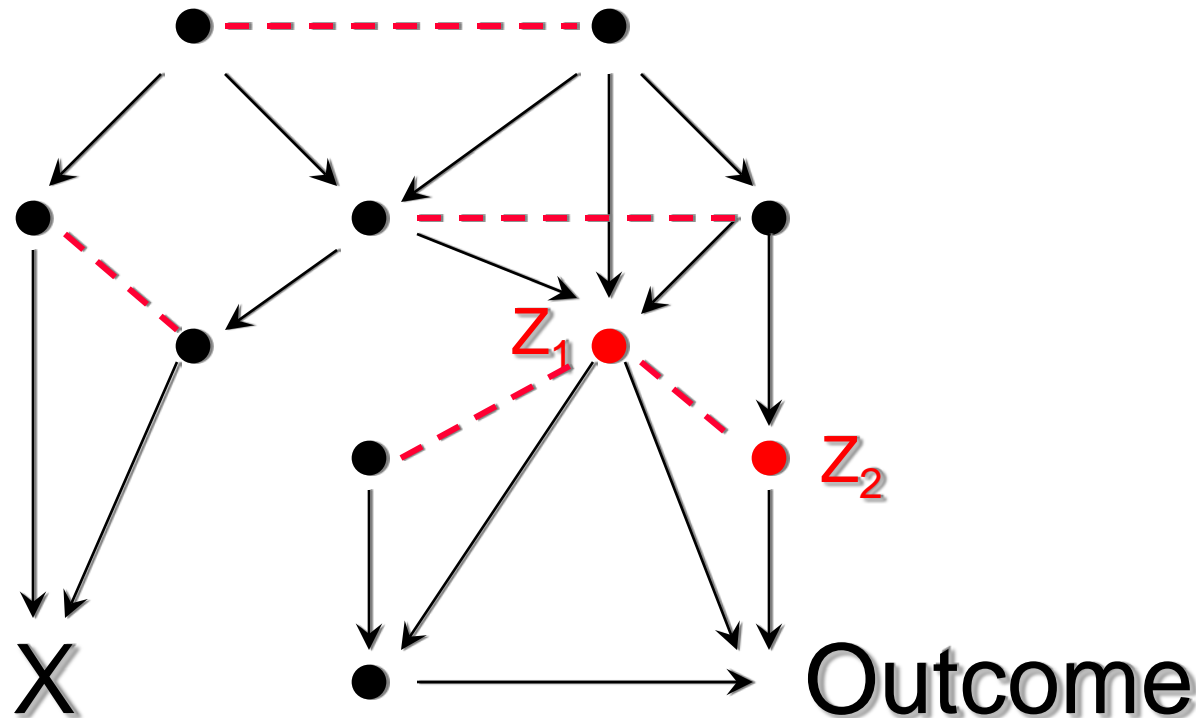
Graphical solutions to the adjustment problem



Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z_1 and Z_2 remove confounding?

Step 3: Delete all arcs emanating from X .

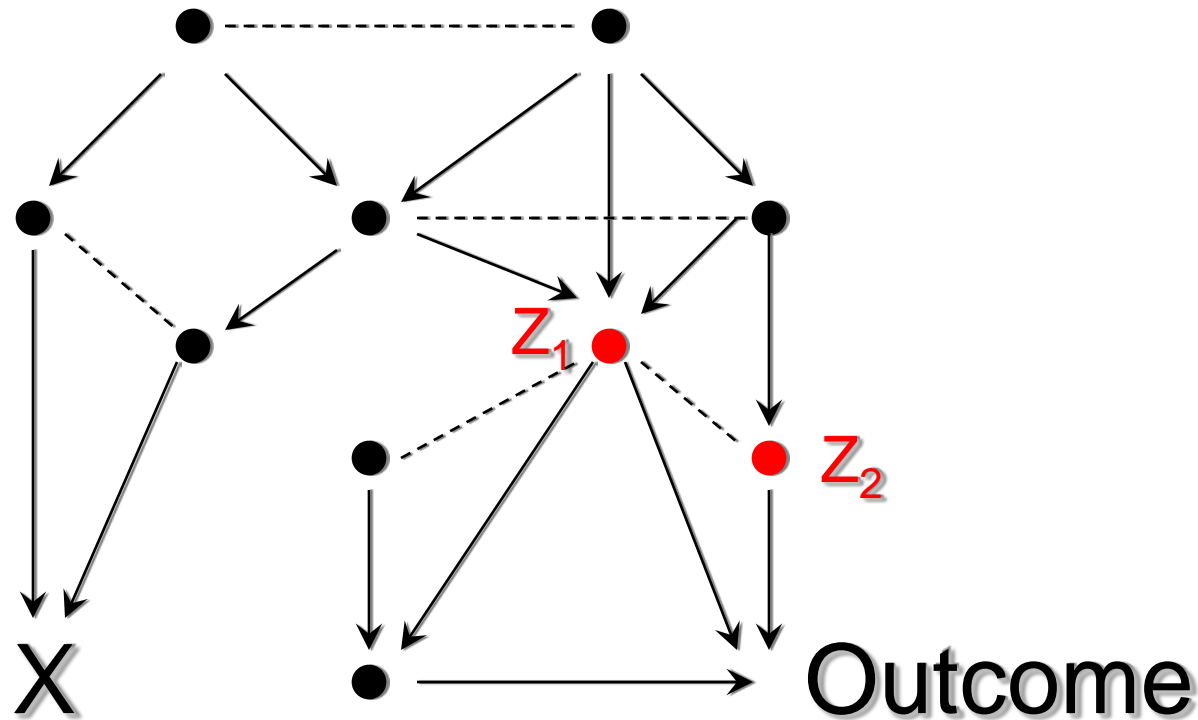
Graphical solutions to the adjustment problem



Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z_1 and Z_2 remove confounding?

Step 4: Connect any two parents sharing a common child.

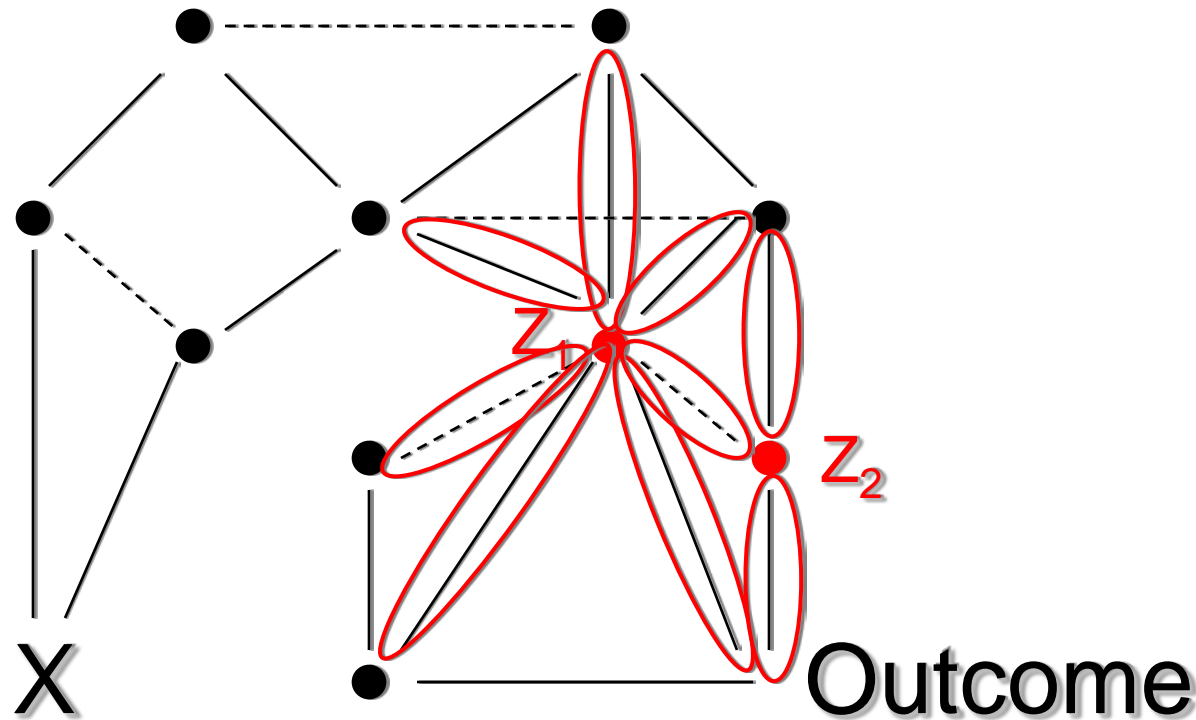
Graphical solutions to the adjustment problem



Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z_1 and Z_2 remove confounding?

Step 5: Strip arrow-heads from all edges.

Graphical solutions to the adjustment problem

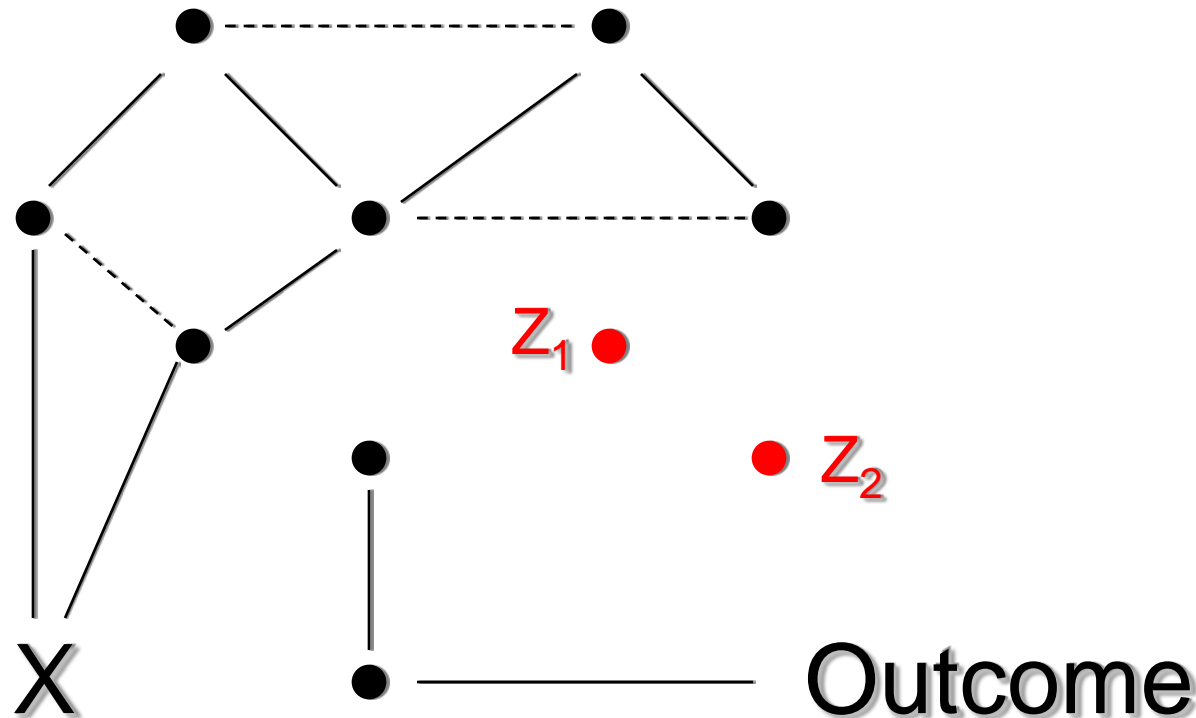


Which measurements should be included in the model if we are interested in the relation between X and Outcome? **Do Z_1 and Z_2 remove confounding?**

Step 5: Strip arrow-heads from all edges.

Step 6: Delete all lines from Z_1 and Z_2 .

Graphical solutions to the adjustment problem

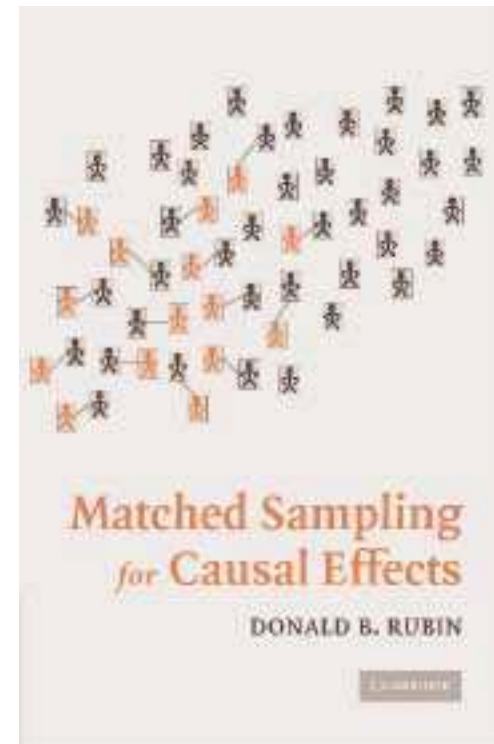


Which measurements should be included in the model if we are interested in the relation between X and Outcome? Do Z_1 and Z_2 remove confounding?

If X is disconnected from Outcome, then there is no confounding

Newer analytical methods to address confounding

- Marginal structural models & propensity scores
 - Build on previously used methods such as “multivariate confounder scores” (Miettinen 1976) and “propensity scores” (Rosenbaum & Rubin)
 - Use a new class of estimators, the inverse-probability-of-treatment weighted estimators
 - Particularly helpful for observational studies with exposures (or treatments) that vary over time and are affected by time-dependent confounders



Robins et al. Epidemiology. 2000 Sep;11(5):550-60

Rosenbaum PR, Rubin DB. Biometrika (1983) 70:41–55⁶⁸



TEN COMMANDMENTS FOR DEALING WITH CONFOUNDING



- I. Always worry about confounding in your research, especially at the design/protocol stage. Try to use design elements (e.g. randomization) that will help reduce potential confounding.
- II. Prior to the study, review the literature and consider the underlying causal mechanisms (e.g. draw causal diagrams such as directed acyclic graphs [DAGs]). Then make sure you collect data on all potential confounders; otherwise you will not be able to adjust for them in your analyses.
- III. Know your field or collaborate with an expert who does! Subject-matter knowledge is important to recognize (e.g. draw causal diagrams) and adjust for confounding.
- IV. Use *a priori* and data-based methods to check if the potential confounders are indeed confounders that should be adjusted for.
- V. Use stratified analyses and multivariable methods to handle confounding at the analysis stage. Choose the multivariate model that best suits the type of data (e.g. dichotomous vs. continuous) you collected and the design you employed (e.g. case-control vs. cohort).
- VI. Do not adjust for covariates that may be intermediate causes (on the causal pathway between the exposure and disease). Do not adjust for covariates that may not be genuine confounders. And beware of time-varying covariates that will need special approaches.
- VII. Use matching with great caution. Use analytic methods that are appropriate for the design used; for example, if matching was done, use methods that take matching into account (e.g. conditional logistic regression, matched pairs analyses).
- VIII. Always consider effect measure modification, but perform and interpret subgroup analyses with caution. The subgroup analysis should be one of a small number of hypotheses tested, and the hypothesis should precede rather than follow the analysis (i.e. subgroups must be pre-specified).
- IX. Always remember that adjustment for confounding can be inadequate due to residual confounding because of unmeasured confounders, misclassification of confounders, and inadequate adjustment procedures (e.g. model misspecification, categorization of continuous covariates).
- X. If conventional methods prove to be inadequate, consider using newer approaches such as propensity scores, matched sampling, instrumental variables and marginal structural models. However, make sure you work with statisticians who understand these new methods (not many do).

When all else fails, pray! If prayer fails, consider changing professions!!

Readings for this lecture

- Rothman text:
Chapter 8

- Gordis text:
Chapters 15

Grades Don't Matter, Sources Say

Palo Alto, CA (AP) - Documents obtained by the Associated Press indicate that grades achieved in post-graduate classes have no effect on future prospects for students enrolled in academic institutions.

According to interviews with several current and past graduate students, "grades don't count," said former grad student and now billionaire Jerry Yang, co-founder of Yahoo! Inc. "I got mostly B's in grad school, which at Stanford was really really bad."

A poll conducted by the Los Angeles Times showed that over 85% of first year grads believe getting high marks "is worth the effort" and "a valuable way to spend my time". Fewer than 10% of fifth year students felt the same way.

In reality, neither employers nor your parents appear to care if you get an A or a B in your advanced Nonlinear Optimization class. "I'm just glad I don't have to pay for tuition any more," said a mother who wished to remain anonymous.

Reaction among graduate TA's was mixed, with some expressing shock that their late hours grading amount to nothing, while others showed visible relief that losing a student's final exam will not really ruin their life.

Sources close to academic faculty reveal that this fact is well known among professors. "Of course grades don't matter," said Prof. Smith, "we only care about the lab work." Grades only serve to "feed the ego of the smart students, and break the spirit of the mediocre ones."



NOW you tell me?? A grad student expresses frustration over the revelation

Continued on page A23

*copyright 2004 Jorge Cham
www.phdcomics.com*