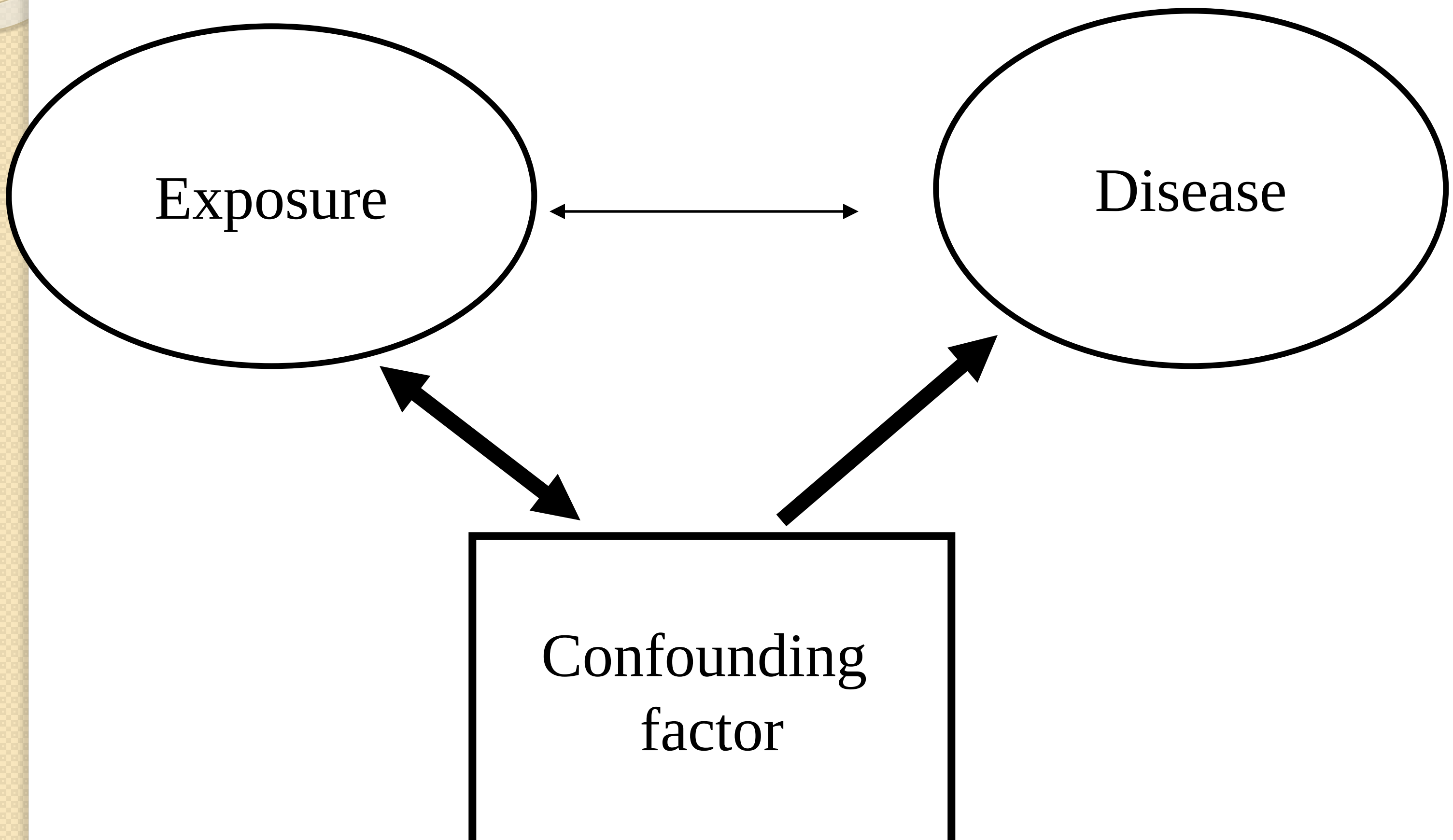
Effect modification and stratification

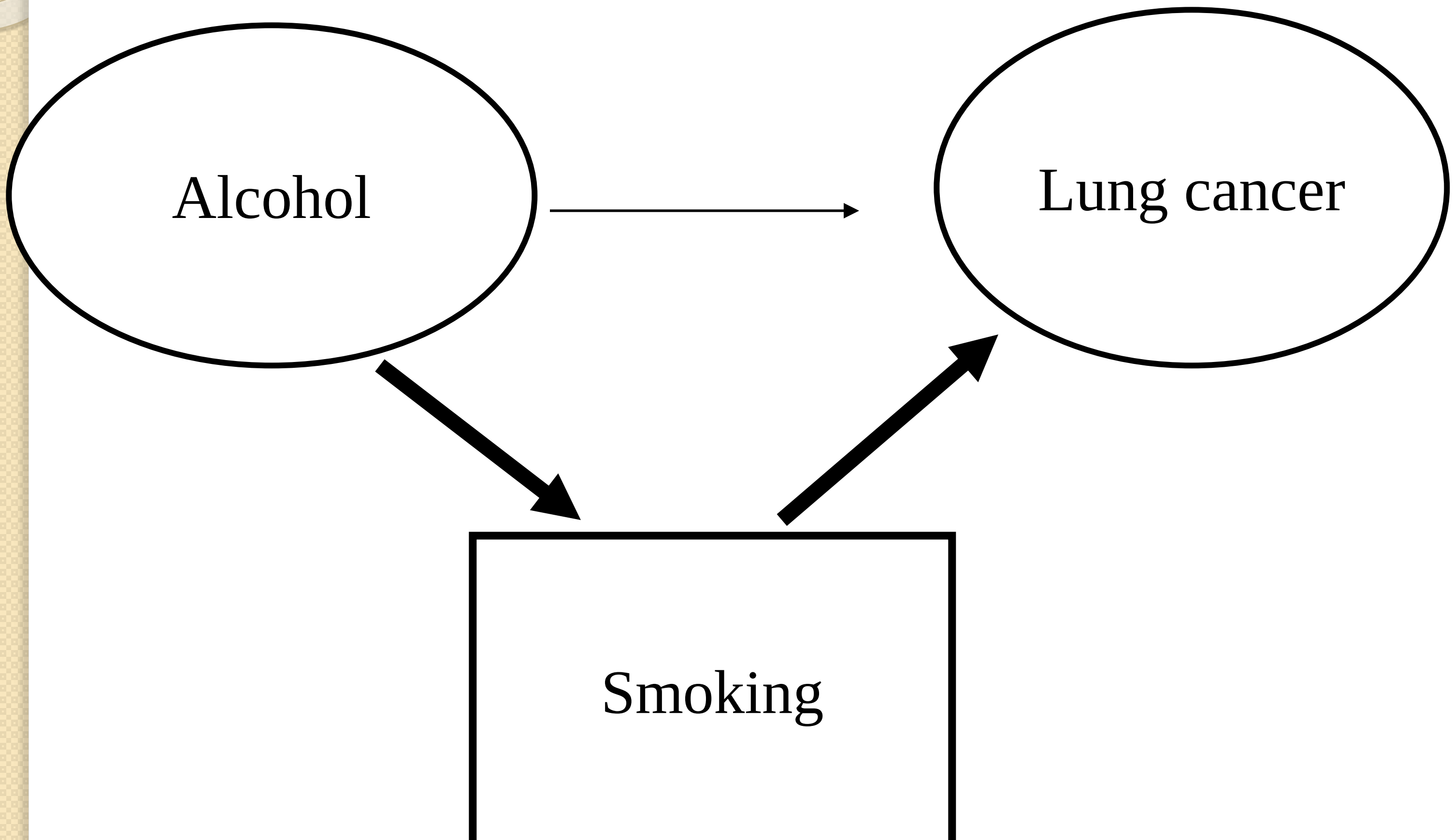
Last session - confounding

- Situation when a third factor is associated with both exposure and disease
- Association between exposure and disease may not be causal; instead, it is due to a third factor which is associated with both exposure and disease.

Confounding



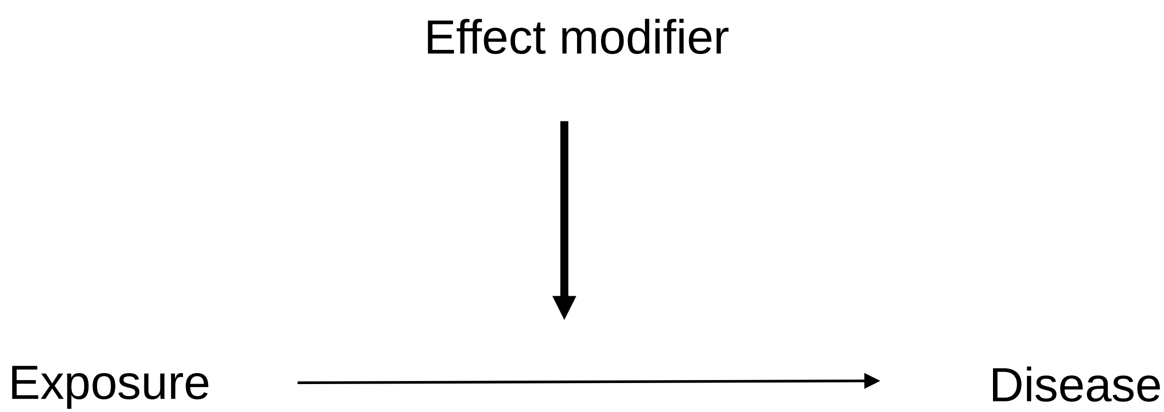
Confounding



Effect modification (interaction)

- the effect of exposure on disease is dependent on the level of a third factor

Effect modification



Biological Interaction

Last's Dictionary of Epidemiology (4th Ed)

Biological interaction is the interdependent operation of two or more causes to produce, prevent or control disease

Factor 1
\+ Factor 2 → Outcome

Examples of biological interaction

1. **Antibiotic tetracycline and tooth discolouration**
   - Tetracycline is associated with discoloration of teeth but mainly among children <8 years
   - effect of antibiotic (exposure) on tooth colour (outcome) is modified by age (effect modifier)

Examples of biological interaction

2. **Measles and vaccination**

- Exposure to measles virus is associated with measles infection if not vaccinated or has not had measles
- Here immune status = effect modifier

Statistical interaction

when the association between exposure and outcome of interest varies according to the level of a third factor (the effect modifier)

Effect modifier (the 3rd factor)

$$\downarrow$$

Exposure $\longrightarrow$ Outcome

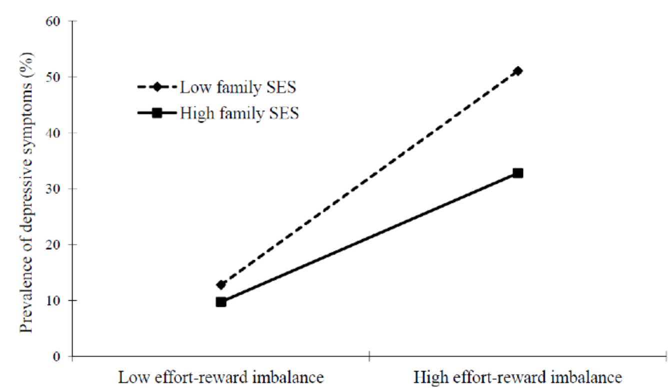
Examples of statistical interaction

1. **Energy from total fat and coronary heart disease (CHD)**

Energy from total fat is associated with CHD among younger women (HR=2.68, 95%CI 1.40,5.12) but not among older women (HR=1.22, 95%CI 0.86,1.71)
(Source: Jakobsen et al. Am J Epidemiol. 2004)

2. Effort Reward Imbalance (ERI) and depressive symptoms among children (China)

School-related stress (ERI school questionnaire) is associated with depressive symptoms among low SES children compared to high SES children
(Source: Guo et al. Int J Environ Res Public Health. 2014)



CHD, smoking and age in British doctors study (rates per 100,000)

	Non-smokers	*Heavy smokers*	
	Rate	*Rate*	*RR*
<45	7	104	14.9
45-54	118	393	3.3
55-64	531	1025	1.9

Positive and negative effect modification

- Positive:
  - “susceptibility factor” or “vulnerability factor”,
  - its presence (or higher values) strengthens the association between exposure and disease.
- Negative:
  - “resiliency factor” or “buffering factor”
  - its presence (or higher values) weakens the association between exposure and disease

CHD, smoking and age in British doctors study (rates per 100,000)

	Non-smokers	*Heavy smokers*	
	Rate	*Rate*	*RR*
<45	7	104	14.9
45-54	118	393	3.3
55-64	531	1025	1.9

Reciprocal nature of effect modification

- For any given outcome and two predictor variables, it is a purely arbitrary decision which predictor variable will be the exposure, and which the potential effect modifier.

- Effect modification is reciprocal. In any of examples, the exposure and other factor (or variable) could have be labelled the other way round, and the same effect would still have been seen.

CHD, smoking and age in British doctors study (rates per 100,000)

	Non-smokers	*Heavy smokers*	
	Rate	*Rate*	*RR*
<45	7	104	14.9
45-54	118	393	3.3
55-64	531	1025	1.9

CHD, smoking and age in British doctors study (rates per 100,000)

	Non-smokers	*Heavy smokers*
	Rate	*Rate*
<45	7	104
45-54	118	393
55-64	531	1025
RR	75.9	9.9

Identification of effect modification

- Stratified analysis
- Compare effect estimates in strata
- Assess differences in effects by significance tests (p-value for heterogeneity)

- Pooled estimates (e.g. standardised) **not appropriate** when there is an interaction

Confounding vs. interaction

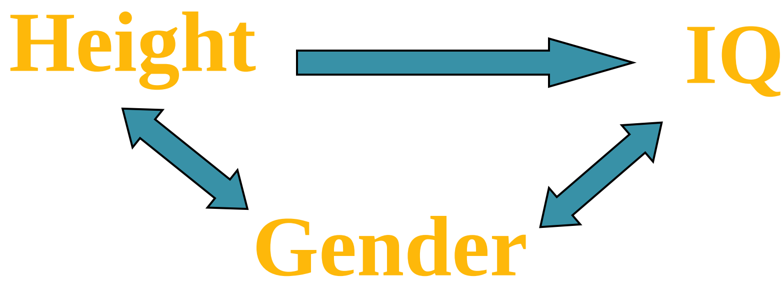
Confounding

- Alternative explanation
- Distorts the "truth"
- Efforts to remove it to get nearer to the "truth"
- When present, stratum specific effects are similar to each other but different from the overall crude effect.

Effect modification

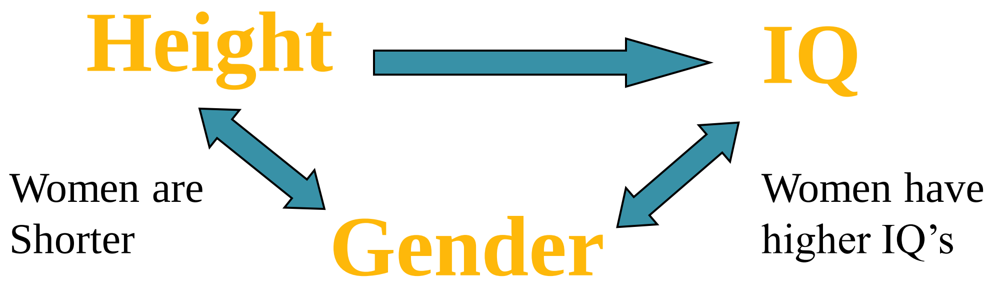
- One factor modifies effect of another factor
- It is genuine, not artefact
- Property of the relationship between factors
- We should detect and describe it but not remove it.

Example: Height and IQ – real association or not?



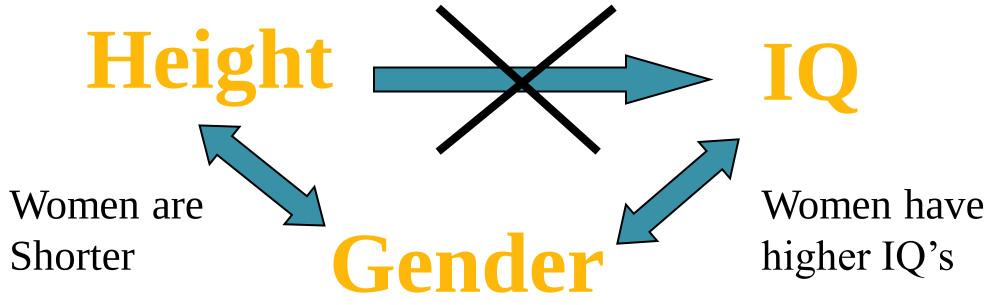
- High negative association between height and IQ

Height and IQ



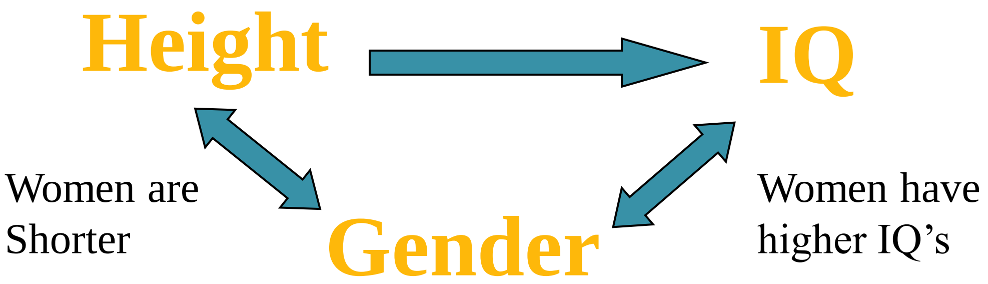
- Find out that Gender is related to Height and that Gender is related to IQ
- Therefore, Gender is a *potential* confounder

Height and IQ



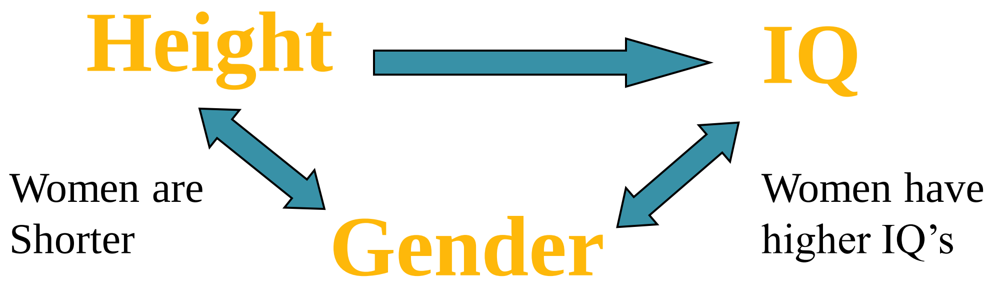
- If after adjustment for Gender there is NO association between height and IQ, then Gender was a confounder

Height and IQ



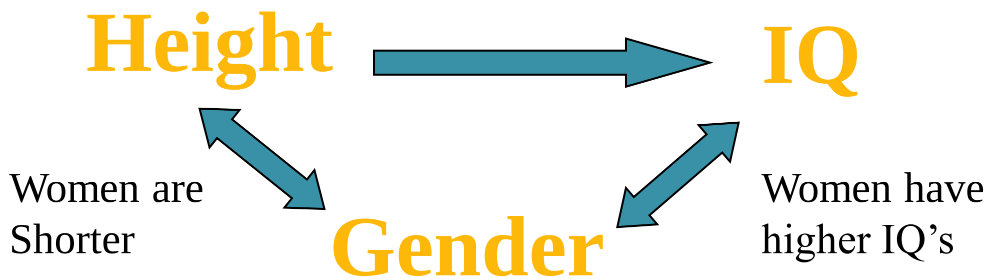
- If after adjustment for Gender there is still a strong negative association between Height and IQ, then Gender is not a confounder

Height and IQ



- If after adjustment for Gender there is still an association between Height and IQ, but the nature and/or strength of the association changes with Gender, then Gender is an **Effect Modifier**.

Height and IQ



- If there is no association between Gender and IQ, then Gender cannot be a confounder
- Likewise, if gender is not associated with height, then Gender cannot be a confounder
- The confounder must be related to both the cause and the effect

Step-by-step guide to the stratified analysis

Example

- A study was undertaken to assess whether smokingh increased risk of stomach cancer. Data were collected from 36,000 individuals

	Stomach cancer		
	Yes	No	Total
Smokers	800 (4.0%)	19200	20000
Non-smokers	400 (2.5%)	15600	16000
Total	1200	34800	36000

Example

- $X^2 = 62.07$ $p < 0.001$

$$OR = \frac{\text{Odds(low)}}{\text{Odds(high)}} = \frac{800/19200}{400/15600} = 1.63$$

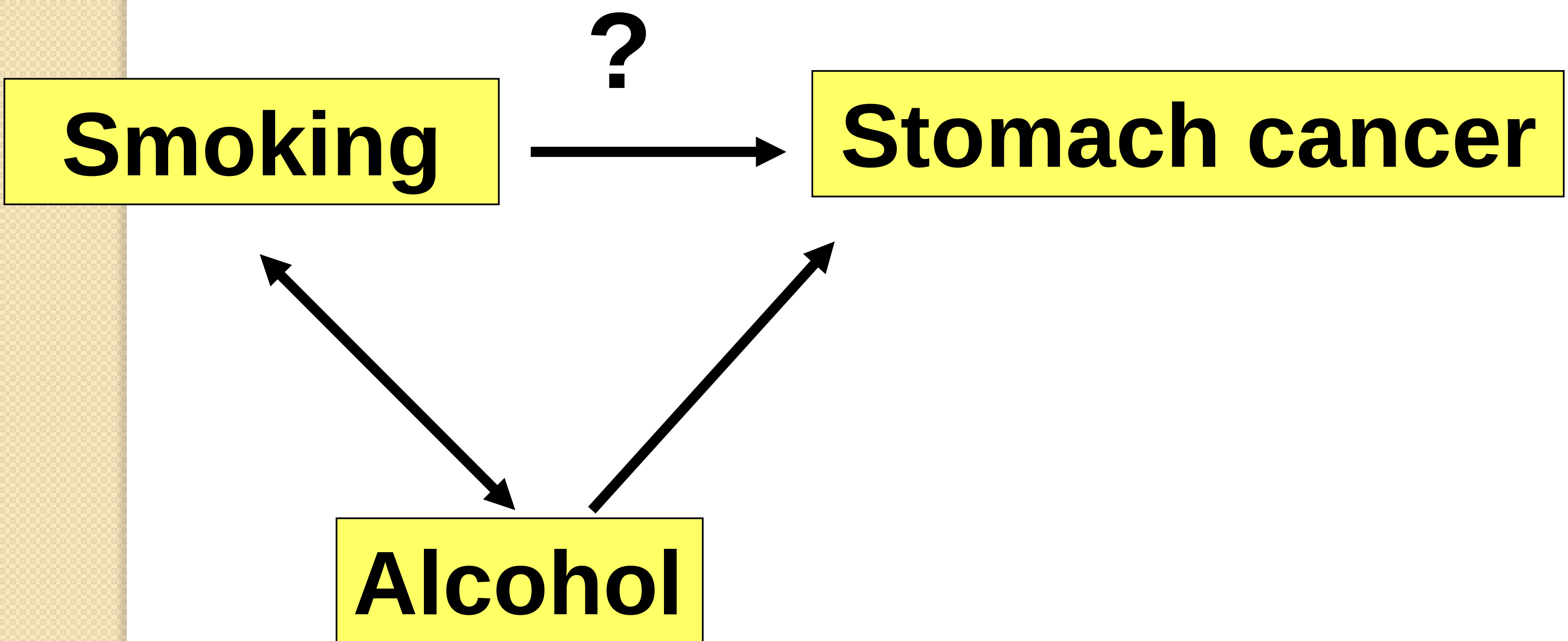
- 95% CI = 1.44-1.84   (Stata)
- The study found a significantly higher odds of cancer in smokers

But is it real association?

- Smokers are more likely to be drinkers
- Drinking doubles the risk of stomach cancer

?

- THEREFORE some of the higher risk in smokers could be because they tend to drink more frequently (and have higher risk because of drinking).



?
Smoking
Stomach cancer
Alcohol

Confounding

- We say that alcohol is a **confounding** variable because it is related both to the outcome variable and to exposure (smoking)

- Ignoring alcohol in the analysis leads to **misleading** results

INDIVIDUALS

Drinkers

Test association between smoking and cancer

X^2 and OR

Non-drinkers

Test association between smoking and cancer

X^2 and OR

Pool these if OR similar across strata = Mantel-Haenszel pooled X^2 and OR

Example

DRINKERS	Stomach cancer		
	Yes	No	Total
Smokers	660	13200	13860
Non-smokers	270	7800	8070
Total	930	21000	21930

DRINKERS	Stomach cancer		
	Yes	No	Total
Smokers	140	6000	6140
Non-smokers	130	7800	7930
Total	270	13800	14070

Example

DRINKERS	Stomach cancer		
	Yes	No	Total
Smokers	660 (4.76%)	13200	13860
Non-smokers	270 (3.35%)	7800	8070
Total	930	21000	21930

NON-DRINKERS	Stomach cancer		
	Yes	No	Total
Smokers	140 (2.28%)	6000	6140
Non-smokers	130 (1.64%)	7800	7930
Total	270	13800	14070

Stratum specific calculations

<u>DRINKERS:</u>

X^2=25.19 p<0.001

OR (95% CI) = 1.44 (1.25-1.67)

<u>NON-DRINKERS</u>

X^2=7.55 p=0.006

OR (95% CI) = 1.40 (1.09-1.79)

- Stratum specific OR are lower than the crude OR (1.44 and 1.40 vs 1.63)
- Stratum specif OR are similar to each other
- This means that it is logical and sensible to pool them
- If they are different (very different) – we should consider drinking to be an EFFECT MODIFIER (the effect of smoking on cancer is modified by drinking status)

Effect modification

- We still need to check one important aspect of M-H analysis – we make the assumption that the association between exposure and the outcome is the same in each level of confounding factor
- If this is NOT true, then you cannot combine stratum specific ORs into one pooled estimate
- If the exposure-outcome association varies in different levels of third variable we say that such third variable modifies the effect of exp on outcome

Steps for dealing with possible confounders

1. Calculate crude X^2 and OR – DONE (X^2 signif. and OR calculated)

2. List possible confounders – we have chosen alcohol in our example

3. Determine whether they are possible confounders
   a. Association with exposure
   b. Association with outcome
   c. Not on causal pathway

Steps for dealing with possible confounders

4. Do stratified analysis by possible confounder

5. Calculate pooled X^2 and OR (= look at the association that is adjusted for confounder)

6. If crude OR and pooled OR different – conclude that variable is a confounder

```
. mhodds cancer smok, by(drink)

Maximum likelihood estimate of the odds ratio
Comparing smok==2 vs. smok==1
by drink

-------------------------------------------------------------------------
    drink | Odds Ratio  chi2(1)          P>chi2       [95% Conf. Interval]
----------+--------------------------------------------------------------
        1 |   1.444444    25.19          0.0000         1.25020    1.66886
        2 |   1.400000     7.55          0.0060         1.10001    1.78181
-------------------------------------------------------------------------
   Mantel-Haenszel estimate controlling for drink
    ----------------------------------------------------------------
     Odds Ratio     chi2(1)          P>chi2         [95% Conf. Interval]
    ----------------------------------------------------------------
       1.433140       32.73          0.0000          1.266074   1.622251
    ----------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(1)    =    0.05
                                     Pr>chi2    =  0.8274
```

```
. mhodds cancer smok, by(drink)

Maximum likelihood estimate of the odds ratio
Comparing smok==2 vs. smok==1
by drink

-------------------------------------------------------------------------
    drink | Odds Ratio  chi2(1)        P>chi2       [95% Conf. Interval]
----------+--------------------------------------------------------------
        1 |   1.444444    25.19        0.0000          1.25020    1.66886
        2 |   1.400000     7.55        0.0060          1.10001    1.78181
-------------------------------------------------------------------------
   Mantel-Haenszel estimate controlling for drink
    ----------------------------------------------------------------
     Odds Ratio     chi2(1)        P>chi2        [95% Conf. Interval]
    ----------------------------------------------------------------
       1.433140       32.73        0.0000         1.266074   1.622251
    ----------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(1)   =    0.05
                                     Pr>chi2   =  0.8274
```

Example

- STATA = test of homogeneity (NULL hypothesis is that stratum specific ORs are homogenous)
- Our example – test of homogeneity: $p=0.83$
- We can assume that stratum specific estimates are same or similar and we can use pooled estimate

Summary of results

- Results are best summarized in the table

Association between smoking and cancer	OR	P-value	Conclusion
Crude assoc.	1.63	<0.001	Odds of cancer 1.63 times higher if smoker
Stratified anal.			
Drinkers	1.44	<0.001	Odds of cancer 1.44 times higher if smoker
Non-drinkers	1.40	0.006	Odds of cancer 1.40 times higher if smoker
Adjusted for drinking	1.43	<0.001	Confounded. Odds of cancer 1.43 times higher rather than 1.63 times higher if smoker

When is effect modification important?

- If we find that stratum specific odds ratios are not homogenous (p-value for test of homogeneity <0.05) we cannot report pooled estiamte
- We need to report stratum specific results!

- Test for homogeneity has low power; → a large p-value does not establish the absence of effect modification. Small p-value however suggest that effect modification is substantial

How to examine effect modification

- Always examine stratum specific odds ratios – how different do they look?
- If there is clear evidence of effect modification, report the exp-outcome association separately for each stratum
- If there is moderate evidence of effect modification, report both M-H OR and stratum specific OR
- If no evidence of effect modification, use M-H OR

Stratification on more than one confounding variable

- Possible
- Combine categories of confounding variables and create strata from all possible combinations
- Problem – number of strata increases fast (for example 3 dichotomous variables = 2x2x2=8 strata)

- We may use other techniques, such as logistic regression