Week 3 – Thursday session

Matched Study Designs & Analysis

EPI202 – Epidemiologic Methods II
Murray A. Mittleman, MD, DrPH
Department of Epidemiology, Harvard TH Chan School of Public Health

HARVARD T.H. CHAN
SCHOOL OF PUBLIC HEALTH

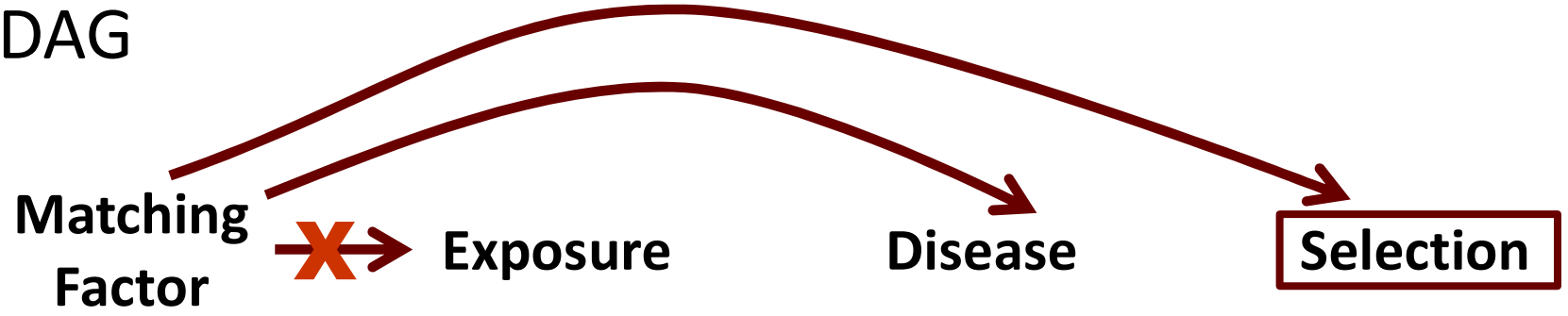
Week 3: Discussion Topics

1. Stratified case-control data data
   - Hypothesis tests
   - Point and interval estimates – weighted averages
   - Assumption of homogeneity
   - Effect measure modification
2. Matching
   - Matched cohort design
   - Matched case-control design
   - Appropriate, unnecessary and overmatching
   - Matching ratio – impact on precision
   - Matched case-control analysis
     - McNemar test and odds ratio estimator
     - Relationship to Mantel-Haenszel methods

Matched Cohort Studies
Design

- In a cohort study, the exposed group is matched to the unexposed group on the matching factor(s)

- DAG



- Matching in a cohort study prevents an association between the exposure and the matching factor(s)

After matching in a cohort study, the investigator can assess:

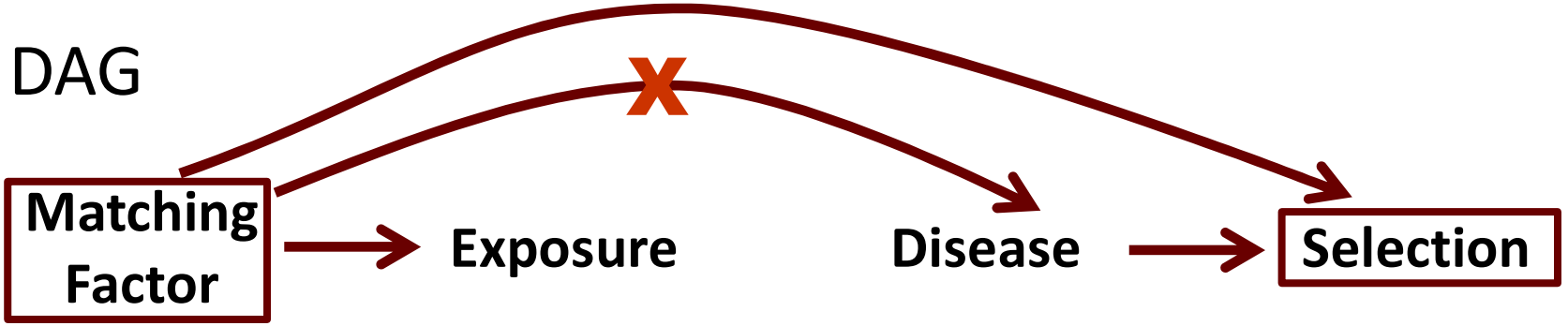
- The association between the matching factor on the outcome of the study.
- Whether the association between the primary exposure and the outcome of the study is modified on the additive scale.
- Whether the association between the primary exposure and the outcome of the study is modified on the multiplicative scale.
- None of the above

Total Results: 0

Matched Case-control Studies
Introduction of Selection Bias By Design

- In a case-control study, the case group is matched to the control (referent) group on the matching factor(s)

- DAG



- Matching in a case-control study abolishes any association between the matching factor(s) and the outcome of interest

Matched Case-control Studies
Validity and Precision

In each stratum, matching enables to have the same balance between cases & controls (and the whole)

- Matching can improve precision (efficiency) but it is not, in itself, a measure to achieve (or improve) validity
- If the matching factor(s) are associated with the exposure of interest, matching will cause the exposure distribution among the control group to be more similar to the cases than the distribution of exposure in the study base
- This is a form of selection bias (collider-stratification bias)
- The bias which is introduced by matching has a tendency towards the null value
- This bias can be eliminated by accounting for the matching factor(s) in the analysis e.g. by conditioning

In a case-control study of the association between fluoride in public drinking water sources and the incidence of osteosarcoma in children, the investigators matched each case to a child of the same age living closest to the case. Matching on "nearest neighbor" in this study is an example of

When poll is active, respond at **PollEv.com/epi2017**

Matching factor relates to the exposure

- Appropriate matching
- Useless matching
- Overmatching
- Matching on an intermediate

Total Results: 0

To show this poll

1

Install the app from
pollev.com/app

2

Start the presentation

Still not working? Get help at pollev.com/app/help
or
Open poll in your web browser

In a case-control study, after matching on order of pregnancy, maternal age and paternal age, can the investigators statistically evaluate whether the OR for the association between induced abortion and ectopic pregnancy varies according to maternal age?

- Yes
- No
- Don't know

Total Results: 0

EMM

Answer

A. Yes

After matching, the main effect of age on the ectopic pregnancy is not estimable (by design it will be null in the data). However, it is possible to assess whether there is effect modification of the odds ratio. This can be done by stratifying the data by age, and then comparing the stratum-specific McNemar estimates.

In a pair-matched case-control study, to statistically evaluate whether the OR for the association between induced abortion and ectopic pregnancy varies according to mother's age, the form of the test statistic is:

Equation 1:

$$\frac{[f_{10} - f_{01}]^2}{f_{10} + f_{01}}$$

Equation 2:

$$\sum_{i=1}^{I} \frac{[ln(\hat{OR}_i) - ln\,\hat{OR}_{MH}]^2}{\hat{Var}_i[ln(\hat{OR}_i)]}$$

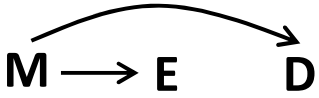
Equation 3:

$$\sum \frac{a_i d_i}{T_i} \Big/ \sum \frac{b_i c_i}{T_i}$$

- Equation 1
- Equation 2
- Equation 3
- None of the above

Total Results: 0

Effect of Matching on Validity and Precision in Case-control Studies

		Design	Analysis	
			Stratified	Not Stratified
	**Appropriate Matching**	Match Do not match	**V PPP** **V P**	**BIAS** **BIAS**
	**Unnecessary Matching**	Match Do not match	**V PP** **V PP**	**V PPP** **V PPP**
	**Over-matching**	Match Do not match	**V P** **V P**	**BIAS** **V PPP**
	**Match on Intermediate**	Match Do not match	**BIAS** **BIAS**	**BIAS** **V PPP**

V=Valid; PPP=maximum precision; PP=Slightly reduced precision; P=reduced precision

Abortion and Ectopic Pregnancy
Notation for Matched Case-Control Data

- 2x2 table for matched case-control data

		Controls: E	Controls: $\bar{E}$
Cases: E		f_{11}	f_{10}
Cases: $\bar{E}$		f_{01}	f_{00}

- The data for this study can be displayed with this layout

		Controls: E	Controls: $\bar{E}$
Cases: E		4	8
Cases: $\bar{E}$		1	5

Notation for Matched Case-Control Data
MH, McNemar, and variance

- The Mantel-Haenszel χ^2 test statistic has the form:

$$Z^2 = \frac{[f_{10} - f_{01}]^2}{f_{10} + f_{01}} \sim \chi_1^2$$

  - This test is also known as McNemar's test.

- The Mantel-Haenszel odds ratio is estimated as:

$$\hat{OR}_{MH} = \frac{f_{10}}{f_{01}}$$

- The variance for the $ln\hat{OR}_{MH}$ has the form:

$$\text{Var}[ln(\hat{OR}_{MH})] = \frac{1}{f_{10}} + \frac{1}{f_{01}}$$

HAVE A GOOD WEEKEND