



Confounding Study

Step 1: Learning Objectives

- A. Explain the importance of comparability groups in epidemiological studies
 - 1. Concept of source population
- B. Define confounding
 - 1. Identify three criteria a variable must fulfill to be a confounder in an epidemiological study
 - 2. Diagram the relationship of a confounder with exposure and outcome
- C. Explain the rationale of using various methods to control for confounding
 - 1. Describe ways of handling confounding at the design phase of a study
 - a. Randomization
 - b. Restriction
 - c. Matching
 - 2. Describe ways of handling confounding at the analysis phase of a study
 - a. Stratification
 - b. Multivariate techniques
- D. Describe evaluation of confounding in the data
 - 1. Explain how you would evaluate whether confounding influences an effect estimate
 - 2. Discuss "residual" and "uncontrolled" confounding

Step 2: Introduction to the Study

Dr. Morrissa Zapp is keen on keeping your internship interesting and intellectually fulfilling to you. Now that you have learned quite a bit about bias, you decide to tackle yet another methodological issue in epidemiology, confounding.

Recall that when you worked on the Susser Syndrome outbreak in the cohort exercise, you calculated rate ratios of Susser Syndrome due to exposure to Superclean for four different age categories (refer to question 11 data analysis in step 6).

The reason it is important to consider age in your investigations is because age may potentially confound an association between the exposure of interest and the outcome. Confounding is one of the most important problems in epidemiological studies. Although in this investigation you found that rate ratios for all age groups were the same, (i.e., age did not confound the association between exposure to Superclean and Susser Syndrome) in many other epidemiological studies age may be a major confounder. Dr. Zapp suggests that you take a look at the same two studies you used to learn about the concept of bias.

As you heard in Dr. Stellman's introduction to the study, he and his colleagues were concerned about confounding in their study. In order to reduce confounding, they matched controls to cases on age, sex and hospital. In addition, they also matched controls to cases on hospital room status (either ward, semi-private, or private room) because hospital room status tended to correlate highly with income.

Dr. Shapiro and his colleagues were also concerned about confounding in their study. They collected detailed information on potential confounders such as obesity, prior use of other female hormones and the like.

Step 3: Student Role - Your Plan of Action

You need to first familiarize yourself with these studies.

1. Listen to the introduction about the two studies
2. Read the following synopsis of each study

Questions in steps 4 and 5 require you to demonstrate critical thinking and knowledge of epidemiological concepts. Read carefully through the explanations of both correct and incorrect answers. Finally, answer the discussion questions in Step 6 found at the end of the exercise. Bring your answers to your seminar section and be prepared to discuss them in class.



Dr. Steven Stellman
[TRANSCRIPT]

My name is Steve Stellman, I'm professor of Epidemiology here at the Mailman School of Public Health.

Keeping our food and water supply safe is an important public health function. The Bureau of Foods, which later became the FDA, was created nearly a hundred years ago under President Theodore Roosevelt. Saccharin is a chemical sweetener which was discovered even earlier in the 1870s, and it became one of the Food Bureau's first targets. Unfortunately, Roosevelt was an aficionado of saccharin, and he forbade the FDA from touching it, exclaiming anyone who says saccharin is injurious to health is an idiot. Saccharin has been a political football ever since then.

It became economically important only in the 1960s when the soft drink industry, uh, adopted diet soft drinks as a major product. The future of saccharin was threatened when three separate studies were published in which bladder cancer was induced in rats. This should have triggered The Delaney Amendment to The Food and Drug Act, which forbade any food additive that causes cancer, but under industry pressure, Congress exempted it from regulation. However, epidemiologists responded to the public health challenge by designing a number of studies of saccharin and bladder cancer in humans.

Most of these studies failed to find any association except for one that was co-authored by Dr. Geoffrey Howe in Toronto, now in our own department of epidemiology at Columbia. At that time I was at The American Health Foundation, I was working with Dr. Ernest Wynder the pioneer researcher in tobacco and lung cancer. We already had a case control study of bladder cancer under way at ten United States hospitals, so we simply added several questions to our questionnaire covering use of tabletop artificial sweeteners and diet beverages. Our study was designed to answer the question, is there an association between occurrence of bladder cancer and past consumption of saccharin? We interviewed 302 men and 65 women with bladder cancer, and equal numbers of controls. We knew that socioeconomic status could be an important confounding factor. Bladder cancer was associated with higher socioeconomic status in men. We called this, in fact, the Hubert Humphrey Phenomenon after the former Minnesota senator and vice president who chose to be treated at, for bladder cancer at Memorial Sloan-Kettering Cancer Center, which was regarded as an elite cancer treatment institution, rather than at his home institute at the University Of Minnesota. To reduce confounding, we matched the controls not only to age, sex and hospital, which we always did as a matter of course. But also on hospital room status, either ward, semi-private, or private room. In those days, there were still large wards. This often reflected income. We were also concerned with recall bias. Stories about saccharine and cancer had frequently been reported on the nightly news, and we were concerned that hearing those reports might lead bladder cancer patients to selectively remember saccharine use, which would wrongly inflate our estimates of relative risk.

As it turned out, we found no association between bladder cancer and many different measures of saccharine usage. A few years later, an Institute Of Medicine panel on which I served reached largely the same conclusion, and

that was that. What was once a burning public health issue finally lost most of its importance after Aspartame, or NutraSweet, largely displaced saccharine in soft drinks.

Prior to this study, the relation between saccharin obtained through artificial sweeteners or diet beverages and bladder cancer in humans was a matter of public health and scientific controversy. Dr. Stellman says that "Saccharin has been on the burner 'of epidemiology' for over 125 years." Animal studies demonstrated a statistically significant increase of bladder tumors in male rates while tumor-promoting effects were observed in vitro and in vivo studies. Yet, published epidemiological studies had been negative. Based on the results of the study presented to you in this exercise it was concluded that there was no evidence that the regulated artificial sweeteners on the market in the United States were related to cancer risk in humans. Today, artificial sweeteners are continued to be regulated by the U.S. Food and Drug Administration (FDA).



Dr. Syd Shapiro
[TRANSCRIPT]

My name is Samuel Shapiro. I am a visiting professor of epidemiology in the department of epidemiology at the Mailman School Of Public Health. The background to the study "Risk Of Localized And Widespread Endometrial Cancer In Relation To Recent And Discontinued Use Of Conjugated Estrogens" was this: it had really been documented in several studies that there was an association between the use of unopposed estrogens, conjugated estrogens in particular and endometrial cancer. What was disputed principally by Fienstien and Horowitz was the claim that this association was causal. They thought that it could be due to either diagnostic surveillance of women who use conjugated estrogens, or that it could be due to the unmasking of otherwise silent endometrial cancer.

They even implied, although they didn't explicitly state, that the unmasking of early cancers might be beneficial because estrogens result in their earlier diagnosis, and hence a better prognosis. And support for that argument, they said that most of the studies had shown, a higher risk for early pre-invasive cancer than for later, frankly, evasive cancer. The purpose of the study was to examine those two questions, and firstly it collected sufficient data to look at the risks of both early and late cancer, and showed that the risks were increased for birth, although with the qualification for late cancer that the data were a little on the sparse size. It would be nicer if there'd been more data. With regard to the question of unmarking the cancers, what was done in this study was to look at estrogen use which had been discontinued as much as one year previously, as much as five years previously, and as much as ten years previously. On the underlying assumption that if an increased risk was shown for discontinued use, then the mechanism couldn't be the unmasking of otherwise early cancer. And this was shown, and it was also shown within these periods obviously discontinued, that there was a duration response effect within each of the categories. Which and statistically significant duration response.

So for example, among women in a discontinued use as much as ten years previously, there was a rising monotonic trend of relative risk according to duration reviews of estrogens. So this paper really established for all practical purposes that, the purported explanations proposed by Horowitz and Fienstien did not account for the association. And it substantially strengthened the claim that this association is almost certainly causal and not due to some source of bias.

Another important aspect of this paper is that it collected detailed information on potential confounders, such as obesity, prior use of other female hormones and the like, and none of these explain the increased risk. One bias which is always possible in interview based case control studies, is information bias. This could not be eliminated, but the consistency of the data according to intervals since last use and according to duration within each of those intervals makes this rather implausible.

Selection bias was also a remote possibility, but extremely unlikely because consecutive cases of endometrial cancer were enrolled, and the refusal rate for the enrollment of controls was less than four percent. And there was substantial evidence to suggest that the selection of the controls was independent of the probability of being a conjugated estrogen user. Today it is generally accepted that unopposed estrogens of which conjugated estrogen is the leading example among, post menopausal American women cause endometrial cancer. It isn't often in epidemiology that we can use the term cause, but here this statement appears to be justifiable.

S. Shapiro and colleagues conducted a study of recent and past use of conjugated estrogens in relation to adenocarcinoma of the endometrium. You can learn more Dr. Shapiro's work by listening to his audio clip. What was the controversy all about?

In 1975, a study was published which suggested an association between the use of non-contraceptive estrogen and endometrial cancer. Some argued, however, that the association was due to selection bias of cases because women who used estrogens were more likely to present with symptoms of uterine bleeding and thus, an otherwise undiagnosed asymptomatic tumor was diagnosed because estrogen led to its bleeding whereas women who had asymptomatic tumors but did not take estrogens were less likely to be diagnosed with endometrial cancer.

Dr. Shapiro's study brought this controversy to a resolution because the study showed that uterine bleeding could not be attributed to estrogen use that ceased in the distant past and thus, estrogen use really did have an effect on endometrial cancer.

Synopsis 1: Artificial Sweetener and Bladder Cancer, S. Stellman et al.

Note: These synopses will be used as a background material for homeworks on Bias and Confounding.

Objectives

To assess whether use of artificial sweetener in daily diet increases the risk of bladder cancer.

Hypothesis

Artificial sweetener (AS) and diet beverage (DB) use is associated with bladder cancer.

Design

This is a matched case-control study.

Controls were matched to cases on age (in decades), sex, hospital, and hospital-room status (private, semiprivate, or ward). This was a 1:1 matching with matches found for all but 10 male cases and 14 female cases.



Intellectually curious? Learn more about matching.

In this study, it is necessary to perform a statistical analysis appropriate for a matched case-control study since the authors matched on age (in decades), sex, hospital, and hospital-room status (private, semiprivate, or ward). Since cases and controls are matched on these factors, you can no longer elucidate the effects of these variables. Cases can be matched individually or frequency matched on particular variable. In the first scenario, one or more controls are selected to match a particular case on a set of variables, while in the second, the controls are selected in such a way that their distribution on a set of variables resembles that of the cases.

Population at risk for Disease

Males and females who use artificial sweeteners in their diet.

Source Population

Hospital cases and controls present an ill-defined source population that generally cannot be characterized.

Eligibility criteria for cases and controls

Cases: male and female patients admitted for a first diagnosis of bladder cancer.

Controls: male and female patients admitted for other health conditions, both neoplastic and nonneoplastic.



Intellectually curious? What does "neoplastic" and "nonneoplastic" mean ?

"Neoplastic" means diseases characterized by abnormal new growth of tissue, synonymous to "tumor".

"Nonneoplastic" is synonymous with "noncancer" diseases.

Diagnoses of the male matched controls

Tobacco-related cancers (lung, larynx, mouth, and esophagus): 23%

Other cancers: 38%

Benign neoplastic diseases: 5%

Nonneoplastic conditions: 34%

Diagnoses of the female matched controls

Tobacco-related cancers (lung, larynx, mouth, and esophagus): 14%

Other cancers: 36%

Benign neoplastic diseases: 7%

Nonneoplastic conditions: 43%

Methods of accrual of cases and controls

Cases: Eligible men and women were interviewed between August 1977 and June

Controls: Eligible men and women were interviewed during the same time period as cases.

Data collection

Measurement of Exposure: Artificial Sweetener (AS)

Assessment: information was obtained on demographic variables and on the use of tobacco, alcohol, coffee, tea, and other beverages, including those with artificial sweeteners.

The quantity of regular AS intake was reported in units per day where 1 unit was approximately equal to 20 to 40 mg of saccharin per day.

Measurement of Outcome: Bladder Cancer, verified histopathologically (i.e., cytologic, histologic and pathologic characteristics all showed that this indeed was a bladder cancer)

Data Analysis

Total number of cases: 302 males and 65 females

Total number of controls: 302 males and 65 females

Males and females did not significantly differ in their use of artificial sweeteners. The proportion of males who never used AS, currently used AS and formerly used AS were very similar between male cases and controls. A similar pattern was seen in female use of AS. Please see table 1.

*Table 1. Regular users of artificial sweeteners among bladder cancer patients and matched controls.**

When Regularly Used	Males		Females	
	Cases (%)	Controls (%)	Cases (%)	Controls (%)
Never	74.8	73.5	78.5	70.8
Currently	16.6	18.9	16.9	21.5
Formerly (1 year ago or less)	2.3	3.3		
Formerly (1 year ago or more)	6.3	4.3	4.6	4.6

*Regular use was defined a continued use for at least 1 month.

The proportion of males who never used diet beverages was the same in controls and cases. However, it appears that more female controls used diet beverages currently than female cases. Please, see table 2.

*Table 2. Regular users of diet beverages among bladder cancer patients and matched controls.**

When Regularly Used	Males		Females	
	Cases (%)	Controls (%)	Cases (%)	Controls (%)
Never	85.1	82.8	84.6	75.4
Currently	13.6	16.2	13.8	20.0
Formerly \leq 1 year		0.3	1.5	
Formerly \geq 1 year	1.3	0.7		4.6

*Regular use was defined a continued use for at least 1 month.



Intellectually curious? Learn more on how to obtain adjusted effect estimates

Calculation of the effect estimate (in this case odds ratio) necessarily implies pooling the data together and producing an average estimate of risk. However, there might be situations when crude estimates are misleading. In particular, if there is a strong confounding in the data. There are several ways to deal with confounding in the data.

At the study analysis stage, we can stratify the data into several groups based on the level of the confounding variable, calculate estimates of effect in each strata and then combine them by using a Mantel-Haenszel procedure. This statistic combines information across partial tables and enables you to calculate one common OR, as opposed to many for each strata.

When the crude odds ratio (OR) was adjusted for age, hospital room status, year interview and education, there appeared to be no differences between those males who developed bladder cancer and used artificial sweeteners and those males who developed bladder cancer and did not use artificial sweeteners (see table 3). Similar findings were observed for females. See table 4.

Table 3. Odds Ratio for Bladder Cancer Among Male Artificial Sweetener Users (number of males=402)

Variables included in the Model	Odds Ratio	95% C.I.
None	1.85	1.45-2.36
Age, hospital, hospital room status, Year of interview	1.43	1.10-1.86
All of above plus education	1.13	0.60-2.09

Table 4. Odds Ratio for Bladder Cancer Among Female Artificial Sweetener Users (number of males=122)

Variables included in the Model	Odds Ratio	95% C.I.
None	0.99	0.61-1.59
Age, hospital, hospital room status, Year of interview	0.89	0.48-1.64
All of above plus education	0.80	0.20-2.98

Results: No evidence was found to suggest that artificial sweeteners or diet beverages were associated with bladder cancer

Synopsis 2: Recent and Past Use of Conjugated Estrogens in Relation to Adenocarcinoma of the Endometrium, Shapiro S., Kaufman D.W., et al.

Objectives

To determine whether prior use of estrogen is associated with endometrial cancer

Hypotheses

- Asymptomatic endometrial cancer is not caused by estrogen use.
- Bleeding caused by estrogen use does not cause asymptomatic endometrial cancer.

Design

Matched case-control study

Matching

1:4 (up to 4 controls were matched to each case according to decade of age and geographic areas).

Population at risk for disease

Post-menopausal women aged 50 to 69 years from Eastern Seaboard, Kansas, Arizona, California, and Canada.

Source Population

It is difficult to establish the precise source population for a hospital case-control study; cases might have come from far away to receive specialized treatment, while controls might have lived in the neighborhood surrounding the hospital.

Eligibility Criteria for Cases and Controls

- Cases: postmenopausal women aged 50-69, admitted to the hospitals located on the eastern seaboard, Kansas, Arizona, California, and Canada.
- Controls: women who were admitted for conditions not related to prior estrogen use from the same hospitals as cases and during the same time period.

Diagnoses of Matched Controls

Diagnosis	No. of controls	Use of Conjugated Estrogens No. (%)	Use of Other Estrogen-Containing Hormones Only No. (%)
Nontraumatic orthopedic Conditions	84	17 (18)	11 (13)
Trauma	79	13 (16)	2(1)
Acute infections and Other acute conditions	101	14 (11)	7 (7)
Other disorders	138	23 (14)	10 (9)

Methods of Accrual of Cases and Controls

- Cases: all newly admitted patients with a diagnosis of endometrial cancer were identified and interviewed.
- Controls: female patients admitted to the medical, surgical, and orthopedic wards with diagnoses other than endometrial cancer were sampled in a systemic manner and interviewed.

Data Collection:

Measurement of Exposure: a questionnaire was used with questions pertaining to lifetime histories of regular use of noncontraceptive estrogens for any of the following indications: regulation of periods, menstrual problems, infertility, breast conditions, endometriosis, sexual difficulties, and menopausal symptoms.

Measurement of Outcome: diagnosis of adenocarcinoma of the endometrium recorded either in the discharge summary or the pathology report, within a year of the current admission.

Exclusion criteria:

- 5 cases and 11 controls who first used a noncontraceptive estrogen within two years of the date of diagnosis (for cases) or for whom the date of the first use was unknown were excluded (for controls).
- Use of noncontraceptive estrogens for a total duration of less than three months
- Use of unspecified female hormone only (20cases and 90 controls)

Data Analysis

Total number of cases: 149
 Total number of controls: 453

The proportion of cases who used conjugated estrogens was greater among cases than controls (see Table 1).

Table 1: Relation of Use of Noncontraceptive Estrogens among 149 Cases and 402 Controls

Use of Estrogen	Cases No. (%)	Controls No. (%)
No use	81(54)	305(76)
Conjugated Estrogens	60(40)	67 (17)
Nonconjugated Estrogens only	8 (5)	30 (7)

Odds ratio estimates together with their 95% confidence limits were computed for various categories of estrogen use. Conjugated estrogen use was a statistically significant predictor of endometrial cancer (Table 2).

Table 2. Relation of Use of Noncontraceptive Estrogens to Risk of Endometrial Cancer among 149 Cases and 402 Controls

Use of Estrogen	Cases No. (%)	Controls No. (%)	Rate Ratio*	95% Confidence Limits*/td>
No use	81(54)	305(76)	1.0	---
Conjugated Estrogens	60(40)	67 (17)	3.9	2.5-6.2
Nonconjugated Estrogens only	8 (5)	30 (7)	0.9	0.4-2.3

Conjugated estrogens use played a statistically significant role for all categories of time elapsed since latest use, except for the last time category, ≥ 5 yr (Table 3).

Table 3. Relation of Use of Conjugated Estrogens for Five Years or More to Risk of Endometrial Cancer, According to Time Elapsed since Latest Use

Time Elapsed Since Latest Use	Duration of Use ≥ 5 years		Rate Ratio	95% Confidence Limits
	Cases	Controls		
< 1 year	27	15	8.8	4.4-17
≥ 1 year	14	13	3.6	1.6-8.4
≥ 2 years	11	11	3.3	1.4-8.0
≥ 3 years	8	9	2.7	1.0-7.5
≥ 4 years	7	8	2.7	1.0-7.8
≥ 5 years	6	7	2.6	0.8-7.8

Results: The rate of endometrial cancer was higher in women who used conjugated estrogens, relative to those who did not. There was no evidence of an association for use lasting less than one year but the risk increased with duration of use.

Questions in steps 3, 4, and 5 require you to demonstrate critical thinking and knowledge of epidemiological concepts. Read carefully through the explanations of both correct and incorrect answers. Finally, answer the discussion questions in Step 6 found at the end of the exercise. Bring your answers to you seminar section and be prepared to discuss them in class. Please proceed to Step 4.

Step 4: Questions for Dr. Stellman's Study

1. Hospitalization, a marker of socioeconomic status, was one of the potential confounders considered in the study of artificial sweetener use and bladder cancer. Please explain why hospitalization was considered to be a potential confounder using the three criteria a variable must satisfy to be a confounder.

- a. Hospitalization, a marker of socioeconomic status, is a risk factor for bladder cancer and is associated with artificial sweetener use.
- b. Hospitalization, a marker of socioeconomic status, is in the causal path between artificial sweetener use and bladder cancer.
- c. Hospitalization, a marker of socioeconomic status, is a risk factor for bladder cancer but is not associated with artificial sweetener.

ANSWERS:

A - Correct

For a variable to be a confounder, it must be a marker of disease and be associated with exposure, but not be caused by exposure. Hospitalization, a marker of socioeconomic status (SES), is a confounder because SES is associated with AS use and is a risk factor for bladder cancer. Note, however, that SES is not in the causal pathway between AS and bladder cancer.

B - Incorrect

A confounder cannot be an intermediate step in the causal path. If a risk factor is in the causal pathway, it is no longer a confounder but a mediator.

C - Incorrect

A confounder must be associated with the exposure of interest. It is one of the three criteria that a confounder must meet.

2. How was confounding handled at the design stage of the study?

- a. Randomization of subjects into cases and controls
- b. Restriction of cases and controls within a narrow age category
- c. Matching controls to cases on selected characteristics

ANSWERS:

A - Incorrect:

Subjects were not randomized in this study. Subjects can be randomized in experimental studies, (i.e., Randomized Controlled Studies) but not in observational studies in which the behavior of subjects is not controlled by the investigator.

B - Incorrect:

Cases and controls were not restricted to any specific age category.

C - Correct:

Controls were matched to cases on age (in decades), sex, hospital, and hospital-room status (private, semiprivate, or ward). Recall that there are two types of matching: frequency (group) matching and individual matching. By matching, we impose comparability on certain factor(s). That is we ensure the same proportion of that factor(s) in the cases and controls. As a result, we can attribute the difference in disease between cases and controls to be due to the exposure of interest rather than to a difference caused by a known confounder.

3. Explain how you would evaluate whether confounding affects an effect estimate.

- a. Look at the crude OR
- b. Look at all the adjusted OR's
- c. Compare the crude OR to the adjusted OR

ANSWERS:

A - Incorrect

It is meaningless to look at the adjusted ORs without comparing them to the crude OR because we are unable to determine what happened to the effect estimate after taking into account other risk factors. The adjusted OR is calculated when in addition to the main exposure of interest, other risk factors are taken into the account.

B - Incorrect

It is meaningless to look at the adjusted ORs without comparing them to the crude OR because we are unable to determine what happened to the effect estimate after taking into account other risk factors. The adjusted OR is calculated when in addition to the main exposure of interest, other risk factors are taken into the account.

C - Correct

It is important to compare the adjusted OR with the crude OR to see the change in the effect estimate. In this study, it was important to compare the crude OR, that is when only artificial sweetener was considered, with the adjusted OR's, those OR's which were calculated by taking into account year of interview and education.

4. Would confounding due to socioeconomic status still be a problem if the investigators chose to conduct a cohort study where they enrolled subjects based on whether they used artificial sweetener or not?

- a. confounding would not be a problem in a cohort study
- b. confounding would still be a problem in a cohort study
- c. confounding would be minimal in a cohort study compared to case-control study

ANSWERS:

A - Incorrect

Confounding is a problem in a cohort study just as much as it is a problem in a case-control study because confounding is not a consequence of a study design but a consequence of failing to take into account other variables of interest when attempting to assess a relationship between exposure and disease.

B - Correct

Confounding would still be a problem because regardless of whether investigators chose to conduct a case-control study or a cohort study Socioeconomic status (SES) is still a confounder of the artificial sweetener and bladder cancer association. SES is an independent risk factor of bladder cancer (those in lower SES categories are more likely to develop bladder cancer) even if they are unexposed to artificial sweetener (AS) or diet beverages (DB). On the other hand, SES is always correlated with artificial sweetener or diet beverage use because subjects of higher SES are more likely to consume AS and DB and in larger quantities because they can afford it, regardless of the study design.

C - Incorrect

Confounding can be just as large in a cohort study as it is in a case-control study. Therefore, it is important to consider potential confounders in any study design.

SEE APPENDIX A:
interactive exercise shows the "mixing of effects" when confounding is present in the data.



Intellectually curious? Learn more about the evaluation of confounding .

The rule of thumb in evaluation of confounding is to look at the percent change in the adjusted estimate. If the adjusted estimate differs from the crude by 10% or more, then it is customary to assume that a variable produces substantial confounding and should be adjusted for in future analyses. There are many ways of evaluating confounding in the data analysis, but one of the most intuitive ones is to use the "forward selection" method:

1. calculate crude OR
2. look at available variables (hopefully you thought about confounding beforehand and collected the data on potential confounding variables!) and decide which ones satisfy the three criteria for confounding
3. order potential confounder in the order of importance based on the a priori knowledge from other published studies
4. stratify your data on the n levels of the first potential confounding variable
5. calculate individual OR in each stratum of the potential confounding variable
6. combine individual stratum-specific ORs into one using Mantel-Haenszel procedure or use other methods which calculate ORs pooled across strata
7. compare pooled OR adjusted for the first potential confounding variable with the crude OR
8. if adjusted OR differs from the crude by 10% or more, then this variable is a strong confounder and should be retained in the analysis
9. select the next potential confounding variable and stratify your data on m levels of this variable (if the first variable from the list was a strong confounder, then you will have $(n \times m)$ tables and individual stratum-specific ORs
10. calculate pooled OR adjusted to two variables
11. compare pooled adjusted OR to the OR adjusted for only one confounder
12. if adjusted OR differs from the OR adjusted for only one confounder by 10% or more, then the second variable is a strong confounder and should be retained in the analysis.

Step 5: Questions for Dr. Shapiro's Study

1. **Shapiro, et al. matched controls to cases on two factors: age and geographic area, whereas Stellman and Wynder matched controls to cases on four factors: age, sex, hospital, and hospital room. Which study do you think is best at controlling confounding at the design phase of the study?**
- a. Shapiro et al's study is better because it is best to match controls to cases on fewer factors.
- b. Stellman and Wynder's study is better because it is best to match controls to cases on more factors.
- c. It is not possible to determine which study is better at controlling confounding by looking at the number of matched factors.

ANSWERS:

A - Incorrect

Although, you do not want to match controls to cases on too many factors, matching on fewer factors in itself does not guarantee that confounding is accurately controlled.

B - Incorrect

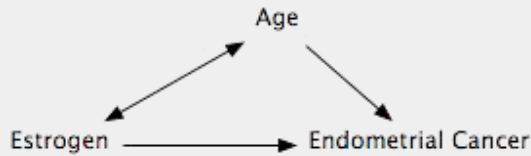
While you want to match controls to cases on several key factors, matching on too many factors may actually be harmful to your study

C - Correct

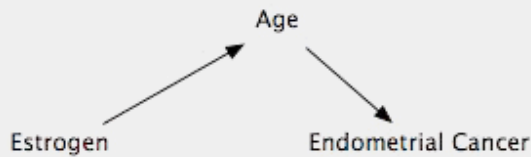
It is not possible to determine whether one study is better than the other at controlling confounding by the number of factors matched. What should be of foremost importance when controlling confounding is whether confounding variables were measured properly and that their effects were removed at the analysis stage.

2. Age was a potential confounder in this study. Choose appropriate diagram representing the relationship of this potential confounder with exposure and outcome.

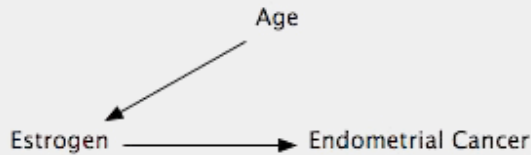
a.



b.



c.



ANSWERS:

A - Correct

In this diagram age meets the requirements to be a confounder because, as depicted in the diagram, age is a risk factor for endometrial cancer and is associated with estrogen use, but it is not a result of estrogen use.

B - Incorrect

This diagram illustrates that age is an intermediate in the pathway between estrogen use and endometrial cancer. If a factor is in the pathway between exposure and outcome, it is called a mediator.

C - Incorrect

In order to be a confounder a factor must be both a known risk factor for disease and be associated with exposure. However, in this diagram age cannot be a confounder because it is not a risk factor for a disease.

3. What if the investigator found during data analysis that the number of abortions performed was associated with estrogen use and was an independent risk factor for endometrial cancer. Should they attempt to control for this confounder?

a. Yes, investigators should still control for this confounder.

b. No, investigators should not control for any confounders which they did not specify a priori.

ANSWERS:

A - Correct

Since attempts to minimize confounding can only be made for known confounders, it is necessary to look for confounders during the analyses as well.

B - Incorrect

It is not always possible to know all the potential confounders at the beginning of the study. This may happen when investigating an exposure-- disease association which has not been studied well or if cost and feasibility may make it impossible to address all potential confounders at the design phase of a study. Therefore, it is necessary to consider confounding at the analysis phase of a study as well.

4. Suppose that investigators wanted to determine if confounding was present during the analysis phase of the study. What can be done at the analysis phase of the study to look at confounding?

- a. Stratified analysis
- b. Restricted analysis
- c. Matched analysis

ANSWERS:

A - Correct

Stratified analysis is one of the means to control for confounding at the analysis phase of a study. Stratification means that the effect of an exposure is evaluated within strata (levels) of the confounder. Once you calculate OR's for each strata you then compare them with crude OR, which represents a collapsed measure. If there is a large difference between stratified OR's and crude OR's, you conclude that confounding is present. If these OR's are similar then it is unlikely that confounding is present.

B - Incorrect

Restricted analysis is one of the means to control for confounding at the design phase of a study but not the analysis phase of a study.

C - Incorrect

Matched analysis is another way to control for confounding at the design phase of a study but not at the analysis phase of a study.

Step 6: Discussion Questions

Please note: this section is structured differently from the other similar sections in Epiville homeworks. You are asked to work on the three types of questions. First, choose the best answer to the three multiple choice questions and click the "Submit" button found on the bottom of the page. Your answers will be sent automatically, without personal identifiers, to your seminar leader who will use them to assess the content areas that are well understood and those that are less well understood so that she may adjust the content on which to focus when your group meets. Second, you are presented with a set of open-ended questions. Write down the answers to them and be prepared to discuss them in class. Finally, if you are looking for an extra challenge, try answering the question for the intellectually curious.

****NOTE: THESE MUST BE SUBMITTED THROUGH THE EPIVILLE WEBSITE****

Who is your section leader?

- Jamie Geier
- Raz Gross
- Beverly Insel
- Tamarra James
- Teresa Janevic
- Elizabeth Kelvin
- Stephen Leeder / Kris Qureshi
- George Loo
- Kellee White
- Lorna Thorpe
- Lina Titievsky
- Lisa Weiss
- Larkin McReynolds
- Emily Leckman
- Regina Zimmerman

Multiple Choice Discussion Questions

1. Imagine that you were going to conduct a case-control study of the role of physical activity and heart disease. While presenting your findings, a colleague asks whether you thought about confounders, such as blood pressure. Under which of the following conditions could this factor have confounded your interpretation of the data?

- it is a risk factor for some other diseases, but not heart disease
- it is a risk factor associated with the physical activity measure and heart disease
- it is part of the causal pathway by which physical activity affects heart disease
- it has caused a lack of follow-up of test subjects
- it may have blinded your study

2. What would happen if we chose to ignore potential confounders in our studies?

- we would not be able to conclude that the exposure outcome relationship is due to the exposure of interest and not an artifact
- external validity would be compromised
- all of the above

3. Which of the following methods used to control confounding in the design phase of the study can be considered for all analytic study designs?

- Randomization
- Restriction
- Matching
- Restriction and Matching

4. Researchers designed a study to investigate the relationship between consumption of oily cold-water fish such as tuna, salmon, and mackerel and coronary heart disease. They also planned to measure several other variables which would be used for adjustment of effect estimates in the data analyses. One of the variables was daily consumption of omega-3 fatty acids. Oily cold-water fish is known to be a primary dietary source of omega-3 fatty acids. In this study variable measuring daily consumption of omega-3 fatty acids is:

- potential confounder
- potential mediator
- matching variable
- neither

Open-Ended Questions

1. Suppose that in the study of Estrogen Use and Endometrial Cancer you wanted to evaluate the hypothesis that endometrial cancer varies by geographic area. Would you match on geographic area? Please, explain your answer.

2. Which study design offers the best opportunity to control confounding-- randomized clinical trial, cohort study, or case-controls study?

Questions for the intellectually curious

1 a) Calculate crude OR

	Cases	Controls
Estrogen	81	305
Estrogen	68	97

1 b) calculate the stratum-specific odds ratios.
Ponderal index <40

	Cases	Controls
Estrogen	48	150
Estrogen	40	150
	98	307

Ponderal index ≥ 40

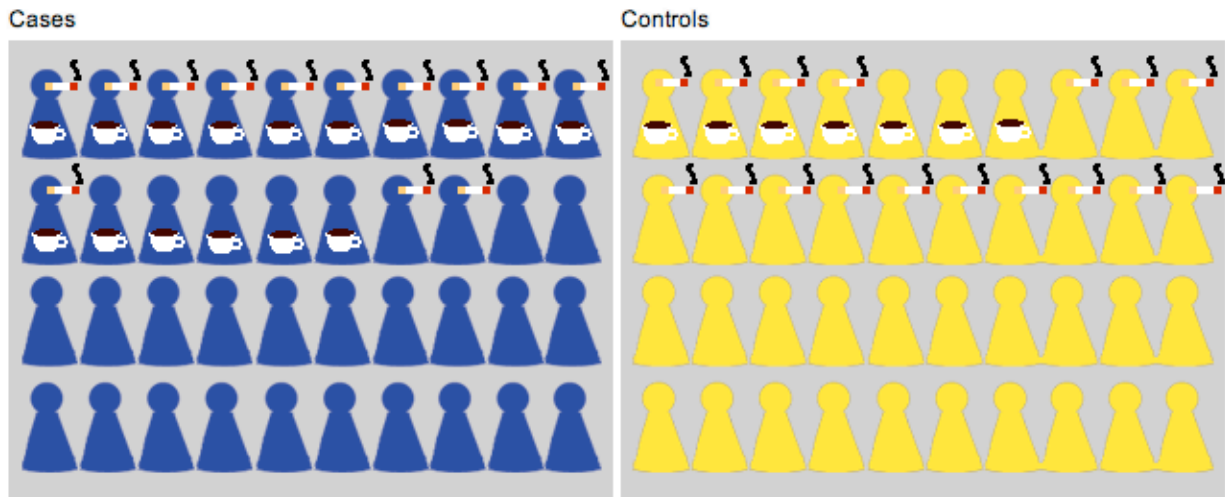
	Cases	Controls
Estrogen	26	43
Estrogen	25	52
	51	95

1 c) Based upon these calculations, is ponderal index a confounder in these data? Justify your answer.

2. Can we achieve perfect control of confounding?

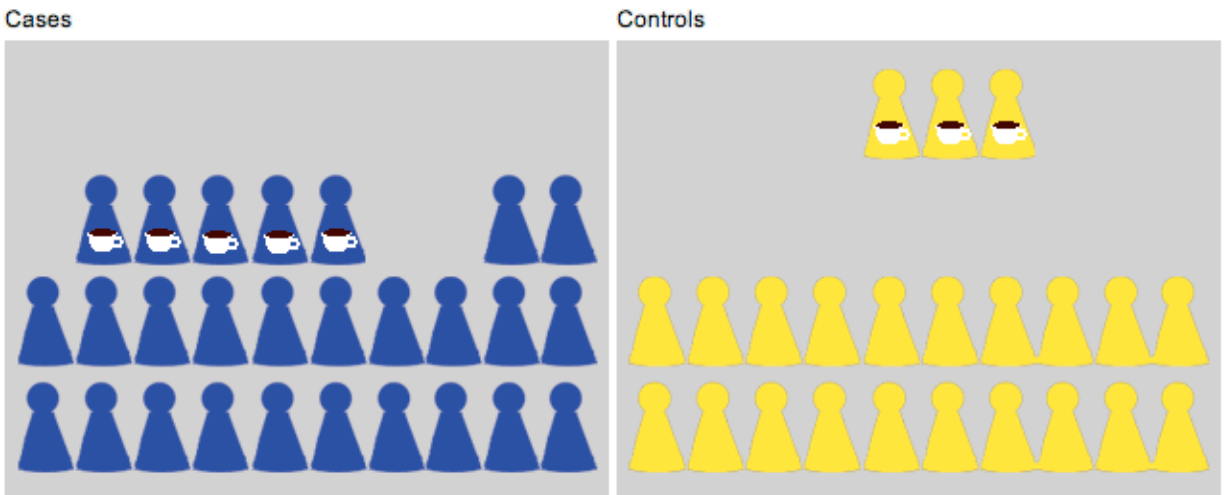
APPENDIX A:

Does drinking coffee have an effect on low-birth weight?



However, we note that many coffee drinkers are also smokers.

Does drinking coffee have an effect on low-birth weight?



2x2 Table

	Cases	Controls
Coffee	5	3
No Coffee	22	20

Odds Ratio = 1.52

After stratifying the data on smoking and looking separately at smokers and non-smokers, it is evident that the effect of coffee consumption is much smaller, with an odds ratio (OR) = 1.52. The data above represent the effect of coffee consumption on non-smokers only (the confounding effects of smoking are removed.)