



Bias Study

Step 1: Learning Objectives

- A. Define bias
- B. Describe factors that contribute to random errors
 - 1. Poor precision
 - 2. Sampling error
 - 3. Measurement error
- C. Describe factors that contribute to each type of bias and give an example of each:
 - 1. Selection bias (self-selection, i.e. volunteer bias, healthy worker effect, sample selection bias)
 - 2. Information bias (interviewer, recall, misclassification)
- D. Identify types of biases specific for different study designs:
 - 1. RCT
 - 2. Case-control
 - 3. Cohort studies
- E. Describe simple ways to minimize bias at the design phase of a study
 - 1. Develop an explicit case definition
 - 2. Enroll all cases in a defined time and region from a defined source population
 - 3. Strive for high participation rate (minimize loss to follow-up and drop-out rate)
 - 4. Take precautions to ensure representativeness of the sample
- F. Describe simple ways to minimize bias at the analysis phase of a study
 - 1. Handling of misclassification (observer bias, measurement bias)
 - 2. Other methods

Step 2: Introduction to the Study

The internship at the Epiville Department of Health has kept you very busy. After having completed investigative work on the two outbreaks (SARS and Susser Syndrome), you feel that in addition to learning "shoe leather epidemiology" you would like to gain a deeper understanding of some of the complexities of epidemiologic methods.

One problem in particular - how to make sure that you have drawn appropriate study samples - is worrying you. In the cohort and case-controls studies you completed over the last few weeks you learned how to properly choose exposed and unexposed groups and cases and controls. For instance, recall Question 3 from Step 4 of the cohort study where you were asked to define eligibility criteria for study participants. The correct answer was to choose those who have worked at the factory for at least two years and who were shown to be healthy at their initial or annual health check-ups (these people had to be employed at the factory from September 2000 to September 2002).

Dr. Zapp explained to you that if you decided to choose everyone at the time point when you began your study (September 2002), you would have introduced selection bias because "everyone" could have included persons who have been on the job for less than two years. She also said that because the minimal induction period is at least 6 months if you had included in your sample people who were employed for less than the induction period, the association between exposure and disease would have been diluted since these new employees would not have had the same opportunity as the long-term employees to develop the disease. This made sense to you, but you begin to worry that there must be many other possible methodologic errors just waiting to derail your march from wide-eyed intern to senior investigator.

You begin to lose sleep over this. Your friends suggest that you "get a life" but, instead, you decide, once again, to turn to Dr. Morissa Zapp. She tells you that you must be very careful in interpreting study results and pay particular attention to the choice of your comparison groups to prevent selection bias. Furthermore, she says that if you are not careful in your interviewing techniques then you can introduce information bias. Dr. Zapp recommends that you

undertake careful study of the concept of bias by exploring two studies that were conducted by her friends, prominent scientists at the Mailman School of Public Health at Columbia University.

The first study which you will look at to learn about biases is "Artificial Sweetener Use and Bladder Cancer" by E.L. Wynder and Steven D. Stellman. This study was extremely influential in resolving the debate about the safety of saccharin. The second study, conducted by Dr. Samuel Shapiro and colleagues, looked at the relationship between estrogen and cancer of the endometrium in women, and is another classic study that resolved a different critical debate in the scientific community.

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.

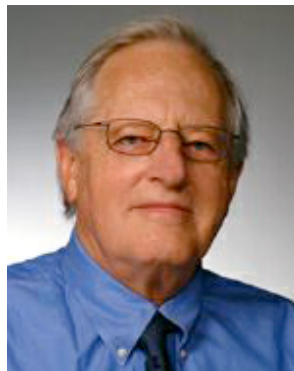
It became economically important only in the 1960s when the soft drink industry, uh, adopted diet soft drinks as a major product. The future of saccharin was threatened when three separate studies were published in which bladder cancer was induced in rats. This should have triggered The Delaney Amendment 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 your seminar section and be prepared to discuss them in class. Please proceed to Step 4.

Step 4: Questions for Dr. Stellman's Study

1. Suppose that you were told that the clinical setting from which researchers got their cases and controls was a tertiary care facility, which is a specialized hospital equipped with diagnostic and treatment facilities to treat bladder cancer. What type of bias might be introduced in the selection of controls and why? Please, note that different tertiary care facilities may be specialized to treat different diseases and therefore people may travel from far away places to get to the facility. They may also be admitted to it simply because they happen to live nearby.

- ☐ a. selection bias
- ☐ b. recall bias
- ☐ c. surveillance bias

ANSWERS:

A - Correct

This is an example of selection bias because your cases do not come from the same source population as your controls. Your cases are probably better off economically since they can afford to choose to be admitted to a specialized facility whereas controls may be admitted simply because they happen to live nearby perhaps in a poor neighborhood). As a result, your cases and controls may differ systematically from each other in ways unrelated to the disease or the exposure of interest. Hypothetical selection of cases and controls for the hospital-based case-control study to show selection procedures and their effects.

B - Incorrect

This has nothing to do with recall bias because recall bias is a type of information bias. The problem in this situation is not with obtaining information from cases and controls but with selecting cases and controls.

C - Incorrect

Surveillance bias deals with differential measurement of outcome but we are concerned not so much how to measure outcome but how to get our cases and controls.

2. Which situation do you think is the best:

- ☐ a. situation 1
- ☐ b. situation 2
- ☐ c. situation 3

ANSWERS:

A - Incorrect

Controls are subjects who were drawn from the tattoo removal clinic. They come from 3 areas (see areas circled by the orange lines) of Epiville. At the same time, cases come from all over the Epiville. Since various variables are distributed differently throughout Epiville (look for income distribution, health insurance and prevalence of smoking), it is likely that cases and controls would differ on many important variables. Thus, cases and controls are drawn from different source populations which makes this sampling biased.

B - Correct

Controls come from many different areas of Epiville as do cases. Because Epiville General is a tertiary care facility, some of the cases come from outside of Epiville. This may introduce some bias, but we will be collecting information on many variables and adjust for their possible confounding effects in the analysis.

C - Incorrect

Both cases and controls come only from Epiville. While this situation would be preferable, it is highly unlikely. It is not likely that there will be many cases of disease under study just in Epiville. Researchers are more likely to recruit a sufficient number of cases from a large tertiary care facility such as Epiville General Hospital.

3. If the researchers had a choice of how and where they could obtain subjects with bladder cancer, what do you think would be the best source for the cases?

- ☐ a. hospital
- ☐ b. tumor registry
- ☐ c. death certificates

ANSWERS:

A - Incorrect

Given a choice, hospital cases are not the best source to get your cases because you are capturing only those cases who were well enough to travel to the hospital and probably had a slower progressing tumor. As a result, you are likely to miss patients which are too sick to travel and who have very fast growing tumors.

B - Correct

A tumor registry is the best source for obtaining cases but only if reporting of all cancer cases is mandatory. In that case it would be possible to ascertain all bladder cancer cases.

C - Incorrect

Death certificates are notoriously inaccurate. Often, the cause of death recorded on the death certificate is only a contributory cause and not the underlying cause. For instance, if a patient had skin cancer and then developed liver cancer due to metastasis, liver cancer may be recorded as a cause instead of skin cancer.



Intellectually curious? Learn more about tumor registries .

Tumor registries are established with the sole purpose of gathering and disseminating current epidemiologic data on all primary tumors, usually malignant, but sometimes benign and malignant, for the purposes of accurately describing the incidence and survival patterns, evaluating diagnosis and treatment, facilitating etiologic studies, establishing awareness of the disease, and ultimately, for the prevention of all tumors. In the U.S. there is no one unified tumor registry as is the case in some European countries. An example of tumor registry is the Connecticut Tumor Registry, which has provided the data for many population-based studies.

4. What potential problem could have been introduced if you found out that interviews with cases took 30 minutes longer than interviews with controls?

- ☐ a. selection bias
- ☐ b. information bias
- ☐ c. volunteer bias

ANSWERS:

A - Incorrect

Selection bias refers to the way cases and controls were selected. In this situation the problem is not with selection but with interviewing.

B - Correct

Information bias occurs when the means for obtaining information from cases and controls differ. This is potentially problematic since it suggests that there may be systematic differences in the way in which exposures are reported and classified.

C - Incorrect

Volunteer bias is a type of selection bias but the problem here is not with selection of cases and controls but with interviewing.

5. In the study of artificial sweetener use and bladder cancer, authors collected information regarding artificial sweetener use and diet beverages. The duration of artificial sweetener use and or diet beverage differed for both cases and controls. For instance, some persons consumed these products as far back as 10 years ago whereas others continued to consume these products during the time that the study was conducted. What potential problems could arise when trying to measure exposures that happened over different time periods?

- ☐ a. Misclassification
- ☐ b. Reporting bias
- ☐ c. Surveillance bias

ANSWERS:

A - Correct

Misclassification may be a problem because it is very difficult for subjects to accurately recall what they ate or drank going back so many years. This will result in improper exposure classification. If misclassification is random (not associated with exposure), then its effect will be to decrease the estimate of effect, however if misclassification is related to exposure, its effect would be hard to estimate.

B - Incorrect

Reporting bias may occur when subjects are reluctant to report an exposure because of negative attitude, beliefs, or perceptions. This type of bias is common in studies evaluating smoking or alcohol. However, artificial sweetener use or diet beverage use is not commonly considered to be a negative exposure.

C - Incorrect

Surveillance bias has nothing to do with the question because it pertains to disease ascertainment but the question is concerned with exposure ascertainment.

6. What would have happened if interviewers were aware of the disease status of the study subjects?

- ☐ a. It would have benefited the study as the interviewers would have collected better data.
- ☐ b. It would not change the results of the study.
- ☐ c. It would irreparably damage the data by introducing the interviewer bias.

ANSWERS:

A - Incorrect

One of the principles of conducting a case-control study is not to inform the person collecting the data and the participant about the hypotheses that are being investigated in the study. The best way to conduct data collection is to gather information as similarly as possible from cases and controls.

B - Incorrect

If interviewers were aware of the disease status they would conduct interviews differently with cases and controls and their results would be different from the situation when interviewer was not aware of disease status.

C - Correct

If interviewer is aware of disease status this might introduce interviewer bias which is characterized by more vigilant data collection (inappropriate probing, leading the subject to "correct" answers, etc.) from cases than from controls.

Step 5: Questions for Dr. Shapiro's Study

Table A. Distribution of 149 Cases of Endometrial Cancer and 402 Controls, According to Various Characteristics

Characteristic		Cases No. (%)	Controls No. (%)
Race	White	140 (94)	370 (92)
	Nonwhite	9 (6)	32 (8)
Religion	Christian	125 (84)	340 (85)
	Jewish	15 (10)	38 (10)
	Non-Christian	9 (6)	24 (5)
	Non-Jewish		
Marital status	Single	14 (9)	51 (13)
	Not Single	135 (91)	351 (87)
Years of Education	< 12	45 (30)	129 (32)
	≥ 12	104 (70)	273 (68)
Parity	< 3	108 (72)	235 (58)
	≥ 3	41 (28)	167 (42)
Ponderal Index	< 40	98 (66)	307 (76)
	≥ 40	51 (34)	95 (24)
Age at Menopause	< 50 years old	64 (43)	217 (54)
	≥ 50 years old	85 (57)	185 (46)
Prevalence of Diabetes Mellitus	Affected	7 (5)	10 (2)
	Unaffected	142 (95)	392 (98)

1. Based on the table above you notice that cases and controls are similar in terms of several factors including race, religion, marital status, and years of education. However, cases and controls differ in their parity, 72% of cases have less than 3 children whereas only 58% of controls have less than 3 children. It's likely that women with more children see their doctor more often than others because to receive prenatal care. On the other hand, women who see their doctor more often would be more likely to be treated for precancerous lesions and less likely to progress to full-blown cancer of the endometrium. If this were the case in this study, what type of bias could have been introduced?

- ☐ a. interviewer bias
- ☐ b. selection bias
- ☐ c. volunteer bias

ANSWERS:

A - Incorrect

Although interviewer bias is a type of information bias, it is not our primary concern in this situation. We are more concerned here with the way cases and controls had been monitored and disease ascertainment.

B - Correct

This is a correct answer because women with a greater number of pregnancies would self-select themselves to see their doctors more often. As a result, their physicians are more likely to notice any abnormalities and treat them before they progress to cancer. This self-selection would have introduced selection bias and made cases and controls different on many characteristics (coming from the different source population).

C - Incorrect

We are not concerned with volunteer bias here because we know that controls were not selected from volunteers but through a systematic review of patients from the hospital, thus they are representative of the population from which they were drawn. Volunteers, on the other hand, frequently differ from the population from which they come, i.e., they are usually more health-conscious, have less bad habits, etc.

2. In this study nothing is mentioned regarding the way cases and controls were interviewed. We do not know whether cases and controls were interviewed during the same time or cases were interviewed first and then controls were interviewed later. Which is the best way to interview cases and controls?

- ☐ a. simultaneously
- ☐ b. interview cases first and some time later interview controls
- ☐ c. interview controls first and some time later interview cases

ANSWERS:

A - Correct

In a case-control study the best way to go about interviewing your cases and controls is to do so at the same time in order to avoid introducing bias due to changes in diagnostic procedures.

B - Incorrect

If you were to interview your cases first and then wait a while to interview your cases you may end up introducing a diagnostic bias.

C - Incorrect

You may introduce recall bias if your controls come into contact with cases and cases learn from them about the study. As a result, they will be more likely to exaggerate their exposures.

3. The authors of the study estimated effect of estrogen use on endometrial cancer by categorizing exposure according to the time elapsed since latest use and duration of use. Refer to Table 3 in the synopsis. What potential bias would have been introduced if the authors also asked women to recall doses of estrogens?

- ☐ a. measurement error
- ☐ b. recall bias
- ☐ c. no bias

ANSWERS:

A - Correct

Measurement error occurs when either the subjects report incorrect amounts or when the tool to measure doses is faulty.

B - Incorrect

Recall bias occurs when information about exposure is recalled by a case but forgotten by a control.

C - Incorrect

When subjects in the study are asked to quantify their exposures, everyone will do so with the some degree of inaccuracy, and as a result, some bias is generally introduced.

4. Errors in the recall of exposures is a common problem in case-control studies. In this study estrogen use for five years and more was evaluated according to the time elapsed since latest use. Refer to Table 3 from the synopsis for detailed results. How could limitations in the recall of estrogen use have affected this study?

- ☐ a. Recall would had been the same regardless of time elapsed since latest use and no recall limitations took place.
- ☐ b. Recall would had been better amongst those women for whom the time elapsed since latest use was shorter than for those women for whom the time elapsed since latest use was longer and thus recall limitations occurred.
- ☐ c. Recall would had been worse amongst those women for whom the time elapsed since latest use was shorter than for those women for whom the time elapsed since latest use was longer and thus recall limitations occurred was.

ANSWERS:

A - Incorrect

This is incorrect answer because for some women time elapsed since latest estrogen use was more than 5 years whereas for others women the time elapsed was less than 1 year thus recall could not had been the same for women with differing time elapsed.

B - Correct

This is a correct answer since women who used estrogen close to the time of the study would had been more likely to remember such use whereas women who used estrogen many years prior to the study would had been more limited in their ability to recall estrogen use.

C - Incorrect

This answer does not make any sense.

5. Authors indicated that cases and controls were matched on decade of age and geographic area. What do you think were the effects of controlling or adjusting for these variables on the effect estimates (ORs)?

- ☐ a. It was possible to estimate effects of age and geographic area as predictors of endometrial cancer.
- ☐ b. It was possible to estimate partial effect of age and geographic area on the endometrial cancer occurrence.
- ☐ c. It was not possible to estimate effect of age and geographic area on the endometrial cancer occurrence.

ANSWERS:

A - Incorrect

If you match your controls to cases according to various characteristics, you can no longer study those characteristics. This is so because matching artificially ensures that the proportion of women in different age categories and from varying geographic areas is identical in cases and controls. Consequently, the prevalence of these two factors is the same in the cases and the controls and thus it no longer makes sense to ask whether cases and controls differed in the prevalence of age or geographic area.

B - Incorrect

Similarly to a, it is not possible to estimate full or partial effect of variables on which the cases and the controls were matched.

C - Correct

Since in this study controls are matched to cases according to age and geographic area, it is no longer possible to explore effects of these two variables.

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: THIS INFORMATION 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

Comment on the effects of the following biases on the odds ratio (OR) in case-control studies. Would they cause an over or under estimation or no change in the OR?

1. Exposed individuals are more likely to participate than non-exposed individuals.

- ☐ overestimate
- ☐ underestimate
- ☐ no change

2. Controls are selected with conditions that are positively associated with exposure.

- ☐ overestimate
- ☐ underestimate
- ☐ no change

3. All cases and controls came from the hospital settings, which implies that they were all insured. How do you think this affects internal validity and external validity (generalizability)?

- ☐ both internal validity and external validity are affected
- ☐ internal validity is affected but external validity is not affected
- ☐ internal validity is not affect but external validity is affected
- ☐ neither internal validity nor external validity is affected

4. In case-control studies, one of the potential biases, which effects validity of a study, is recall bias. What do you think is a potential bias that affects validity the most in cohort and randomized controlled studies?

- ☐ Recall bias
- ☐ Selection bias
- ☐ Interviewing bias

Open-Ended Questions

1. In both studies presented to you, hospital controls were used. Discuss some of the major limitations of using hospital controls in case-control studies.

2. What would had happened if researchers decided to use friend controls (i.e., asking the cases to nominate a friend to serve as a control) instead of hospital controls?

Questions for the intellectually curious

1. In the study of artificial sweetener and bladder cancer, researchers evaluated artificial sweetener use during the previous 10 years. How do you think the result of the study would have been affected if the researchers attempted to evaluate artificial sweetener use during the previous 20 years?

2. Is it better to select controls from the same source population that gives rise to cases and sample them in such a way that level of exposure amongst controls is similar to that of general population or is it better to select cases and controls in such a way that they are similar in all respects except for the disease in question?