Epidemiologic Methods II

PHW250B

Week 8: Cross Sectional & Ecological Studies

October 16, 2025

Recording sessions

- All live sessions will be recorded and available on bCourses.
- If you want remain anonymous during the session, please send a private message to me on Zoom.

Send a private message

If the host has enabled private chat, participants can communicate with each other privately in the meeting. Hosts can't see private chats between participants.

1. While in a meeting, click **Chat** in the meeting controls.
2. In the **To:** drop-down menu, select the participant you want to chat with directly.
3. Enter your message in the chat window.
4. Press **Enter** to send your private message.
   Your message will appear in the chat window indicated by a **(Direct Message)** notification above the message.

Reminders

- Regrade requests for Exam 1 are done!
    - *For questions regarding your regrade or understanding of why you didn't get full credit on a question, please* ***make a private post (or public!) on Ed Discussion*** *in addition to the comment in your Gradescope regrade request. That alerts Andrew and I to respond to your questions directly.*

- Journal Club Assignment #3 Due Monday, 10/21

- Exam 2 in two weeks, week of 10/27-11/2
    - Content is Weeks 5-9

Agenda

- **Refresher problems**
- **Practice problems**
    - 2 from Tab 1, cross-sectional studies
    - 2 from Tab 4, ecological studies
- **Q & A**

Refresher Problems #1

Problem 1. (Fall 2017 250A Problem Set)
Read the following abstract (adapted from Pettifor AE, Rees HV, Kleinschmidt I, Steffenson AE, Macphail C, Hlongwa-Madikizela L, Vermaak K, Padian NS. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. AIDS 2005;19(14):1525-34) and answer the following questions:

Objectives. To determine the prevalence of HIV infection and exposure to national HIV prevention programs in South African youth aged 15–24 years.

Methods. From March to August 2003 we conducted a national survey of HIV prevalence and sexual behavior among 11,904 15–24 year olds.

Results. Young women were significantly more likely to be infected with HIV in comparison with young men (15.5 versus 4.8%). Males and females who reported participation in at least one national HIV prevention program were less likely to be infected with HIV (adjusted OR=0.60; 95% CI, 0.40–0.89; adjusted OR=0.61; 95% CI, 0.43–0.85, respectively).

A. Is the study described in this paper an example of an observational or experimental study? How do you know this?

B. What is the name of this type of study design?

Refresher Problems #1

Problem 1. (Fall 2017 250A Problem Set)
Read the following abstract (adapted from Pettifor AE, Rees HV, Kleinschmidt I, Steffenson AE, Macphail C, Hlongwa-Madikizela L, Vermaak K, Padian NS. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. AIDS 2005;19(14):1525-34) and answer the following questions:

Objectives. To determine the prevalence of HIV infection and exposure to national HIV prevention programs in South African youth aged 15–24 years.

Methods. From March to August 2003 we conducted a national survey of HIV prevalence and sexual behavior among 11,904 15–24 year olds.

Results. Young women were significantly more likely to be infected with HIV in comparison with young men (15.5 versus 4.8%). Males and females who reported participation in at least one national HIV prevention program were less likely to be infected with HIV (adjusted OR=0.60; 95% CI, 0.40–0.89; adjusted OR=0.61; 95% CI, 0.43–0.85, respectively).

A. Is the study described in this paper an example of an observational or experimental study? How do you know this?

This is an observational study design because the investigators had no control over the subjects’ exposure status.

B. What is the name of this type of study design?

Refresher Problems #1

Problem 1. (Fall 2017 250A Problem Set)
Read the following abstract (adapted from Pettifor AE, Rees HV, Kleinschmidt I, Steffenson AE, Macphail C, Hlongwa-Madikizela L, Vermaak K, Padian NS. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. AIDS 2005;19(14):1525-34) and answer the following questions:

Objectives. To determine the prevalence of HIV infection and exposure to national HIV prevention programs in South African youth aged 15–24 years.

Methods. From March to August 2003 we conducted a national survey of HIV prevalence and sexual behavior among 11,904 15–24 year olds.

Results. Young women were significantly more likely to be infected with HIV in comparison with young men (15.5 versus 4.8%). Males and females who reported participation in at least one national HIV prevention program were less likely to be infected with HIV (adjusted OR=0.60; 95% CI, 0.40–0.89; adjusted OR=0.61; 95% CI, 0.43–0.85, respectively).

A. Is the study described in this paper an example of an observational or experimental study? How do you know this?

This is an observational study design because the investigators had no control over the subjects' exposure status.

B. What is the name of this type of study design?
This is a cross-sectional study.

Refresher Problems #1

C. When did assessment of the exposure and outcome occur? Given your answer, do you have any concerns about the directionality of the association between HIV infection status and participation in a HIV prevention program?

D. Calculate and interpret a ratio of disease prevalence among females compared to males (i.e., a prevalence ratio).

Refresher Problems #1

C. When did assessment of the exposure and outcome occur? Given your answer, do you have any concerns about the directionality of the association between HIV infection status and participation in a HIV prevention program?

Assessment of the exposures and outcome occurred simultaneously. At times, this may lead to uncertainty about the directionality of the association between the exposure and outcome in a cross-sectional study. In this example, it's possible that some people learned they had HIV *before* they had the opportunity to join the prevention program, which may have negatively influenced their future willingness to participate in the program. In this case, HIV status influences participation in the program, not vice versa.

D. Calculate and interpret a ratio of disease prevalence among females compared to males (i.e., a prevalence ratio).

Refresher Problems #1

C. When did assessment of the exposure and outcome occur? Given your answer, do you have any concerns about the directionality of the association between HIV infection status and participation in a HIV prevention program?

Assessment of the exposures and outcome occurred simultaneously. At times, this may lead to uncertainty about the directionality of the association between the exposure and outcome in a cross-sectional study. In this example, it's possible that some people learned they had HIV *before* they had the opportunity to join the prevention program, which may have negatively influenced their future willingness to participate in the program. In this case, HIV status influences participation in the program, not vice versa.

D. Calculate and interpret a ratio of disease prevalence among females compared to males (i.e., a prevalence ratio).

Prev ratio = $\text{Prev}_{\text{females}}$ / $\text{Prev}_{\text{males}}$ = 15.5% / 4.8% = 3.23

The prevalence of HIV infection among young women in South Africa in 2003 is 3.23 times the prevalence of HIV infection among young men.

Refresher Problems #1

E. Explain why the authors of this study chose to use this study design to answer their research question. What are the advantages of using such a study design for this research question?

Refresher Problems #1

E. Explain why the authors of this study chose to use this study design to answer their research question. What are the advantages of using such a study design for this research question?

Potential explanations include:

- Cross-sectional study designs are efficient when the exposures and outcomes of interest are relatively common and for outcomes that have long duration. These criteria are met in this study setting where the outcome and exposures of interest are not uncommon among young people.
- Cross-sectional studies are quick and often cost less than other study designs.
- Since this study is nationally representative, it serves to document the nation-national prevalence of HIV and risk factors in South Africa.

Agenda

- **Refresher problems**
- **Practice problems**
    - 2 from Tab 1, cross-sectional studies
    - 2 from Tab 4, ecological studies

- **Q & A**

Tab 1, Problem #2

List the advantages and disadvantages of each type of sampling method.

Sampling method	**Advantage**	**Disadvantage**
Random sampling		
Cluster sampling		
Stratified sampling		

Tab 1, Problem #2

List the advantages and disadvantages of each type of sampling method.

Sampling method	**Advantage**	**Disadvantage**
Random sampling	Ensures study population is representative of target population	Time and cost intensive to generate list of everyone in the population
Cluster sampling	Reduces cost, improves efficiency	May be less representative of population
Stratified sampling	Ensures that the study population includes people in all important categories	Analysis must include weights to account for stratified sampling

Tab 1, Problem #2

List the advantages and disadvantages of each type of sampling method.

Sampling method	Advantage	Disadvantage	When you might use it
Random sampling	Ensures study population is representative of target population	Time and cost intensive to generate list of everyone in the population	*Example:* A statewide health department wants to estimate diabetes prevalence in adults. They have a registry of all licensed drivers, so they randomly select individuals from that list to contact for blood-glucose testing. Use when you have a complete list and need an unbiased population estimate.
Cluster sampling	Reduces cost, improves efficiency	May be less representative of population	*Example:* For a national vaccination survey, instead of randomly selecting households from the entire country, you first sample census tracts, then randomly select households within each tract. Use when people are naturally grouped (schools, neighborhoods, hospitals) or when travel cost is high.
Stratified sampling	Ensures that the study population includes people in all important categories	Analysis must include weights to account for stratified sampling	*Example:* A mental-health survey ensures sufficient numbers of rural, suburban, and urban residents by stratifying the sampling frame by geography and sampling proportionally within each group. Use when subgroup representation matters for your research question or policy relevance.

Tab 1, Problem #3

You are interested in knowing about various health issues in Scottish men. You decide to conduct a study to examine a sample of Scottish men for the presence of peripheral vascular disease (PVD). Other characteristics, such as whether or not a subject had ever smoked, will also be examined. You want to be able to make inference to all Scottish men, so you define a sampling frame of all living Scottish men over the age of 18, and select a random sample of 8,000 of these men for the study. A survey about smoking behaviors, other risk behaviors, and medical diagnoses including PVD is mailed to the 8,000 selected men. 5,000 completed surveys are mailed back to you.

a. What kind of study is this? (1 point)

Tab 1, Problem #3

You are interested in knowing about various health issues in Scottish men. You decide to conduct a study to examine a sample of Scottish men for the presence of peripheral vascular disease (PVD). Other characteristics, such as whether or not a subject had ever smoked, will also be examined. You want to be able to make inference to all Scottish men, so you define a sampling frame of all living Scottish men over the age of 18, and select a random sample of 8,000 of these men for the study. A survey about smoking behaviors, other risk behaviors, and medical diagnoses including PVD is mailed to the 8,000 selected men. 5,000 completed surveys are mailed back to you.

a. What kind of study is this? (1 point)

Cross-sectional

Tab 1, Problem #3

You are interested in knowing about various health issues in Scottish men. You decide to conduct a study to examine a sample of Scottish men for the presence of peripheral vascular disease (PVD). Other characteristics, such as whether or not a subject had ever smoked, will also be examined. You want to be able to make inference to all Scottish men, so you define a sampling frame of all living Scottish men over the age of 18, and select a random sample of 8,000 of these men for the study. A survey about smoking behaviors, other risk behaviors, and medical diagnoses including PVD is mailed to the 8,000 selected men. 5,000 completed surveys are mailed back to you.

b. The results of your study show that the prevalence of PVD in smokers is lower than it is in non-smokers. From your review of the literature, you learn that, in general, other studies have shown that smokers develop PVD and quickly die from it. Knowing this, what type of bias would you be concerned about? Why? (2 points)

Tab 1, Problem #3

You are interested in knowing about various health issues in Scottish men. You decide to conduct a study to examine a sample of Scottish men for the presence of peripheral vascular disease (PVD). Other characteristics, such as whether or not a subject had ever smoked, will also be examined. You want to be able to make inference to all Scottish men, so you define a sampling frame of all living Scottish men over the age of 18, and select a random sample of 8,000 of these men for the study. A survey about smoking behaviors, other risk behaviors, and medical diagnoses including PVD is mailed to the 8,000 selected men. 5,000 completed surveys are mailed back to you.

b. The results of your study show that the prevalence of PVD in smokers is lower than it is in non-smokers. From your review of the literature, you learn that, in general, other studies have shown that smokers develop PVD and quickly die from it. Knowing this, what type of bias would you be concerned about? Why? (2 points)

This is an example of survivor bias or length biased sampling (most smokers with PVD died quickly, so they didn't end up in your study). You could also describe this as a form of selection bias. Your exposed, diseased population (cell "a") is smaller than it should be, so your PR would be smaller than it should be (in this case, it would be biased toward the null).

Tab 1, Problem #3

You are interested in knowing about various health issues in Scottish men. You decide to conduct a study to examine a sample of Scottish men for the presence of peripheral vascular disease (PVD). Other characteristics, such as whether or not a subject had ever smoked, will also be examined. You want to be able to make inference to all Scottish men, so you define a sampling frame of all living Scottish men over the age of 18, and select a random sample of 8,000 of these men for the study. A survey about smoking behaviors, other risk behaviors, and medical diagnoses including PVD is mailed to the 8,000 selected men. 5,000 completed surveys are mailed back to you.

c. Was the sample for this study designed to be a probability or a non-probability sample? Why? (2 points)

Tab 1, Problem #3

You are interested in knowing about various health issues in Scottish men. You decide to conduct a study to examine a sample of Scottish men for the presence of peripheral vascular disease (PVD). Other characteristics, such as whether or not a subject had ever smoked, will also be examined. You want to be able to make inference to all Scottish men, so you define a sampling frame of all living Scottish men over the age of 18, and select a random sample of 8,000 of these men for the study. A survey about smoking behaviors, other risk behaviors, and medical diagnoses including PVD is mailed to the 8,000 selected men. 5,000 completed surveys are mailed back to you.

c. Was the sample for this study designed to be a probability or a non-probability sample? Why? (2 points)

Note the question asked how the study was *designed*, not the final sample. It is a probability sample because it is a random sample. Assuming that the sample is truly random, the probability of being sampled is equal and greater than zero for all individuals in the sample.

We gave partial credit for saying non-probability and justifying it by mentioning the role of non-response.

Agenda

- **Refresher problems**
- **Practice problems**
  - 2 from Tab 1, cross-sectional studies
  - 2 from Tab 4, ecological studies

- **Q & A**

First! Sources of bias in ecological studies

1. Within-group bias:
Bias results from bias within groups due to confounding, selection methods, or misclassification.

- Possible in any study

2. Cross-level bias:

- **Confounding by group:** Bias occurs if the background rate of disease in the unexposed population varies across groups
    - *Unique to ecologic analysis*
- **Effect modification by group (on an additive scale):** Bias results if the rate difference for the exposure effect at the individual level varies across groups
    - *Unique to ecologic analysis*

Example: cross-level bias

Number of New Cases, Person-Years (P-Y) of Follow-up, and Disease Rate (Y, per 100,000), by Group and Exposure Status (x): Summary Parameters for Each Group (top); and Results of Individual-Level and Ecologic Analyses (bottom): Hypothetical Example of Ecologic Bias due to Effect Modification by Group

Exposure Status (x)	Group 1			Group 2			Group 3		
	Cases	P-Y	Rate	Cases	P-Y	Rate	Cases	P-Y	Rate
Exposed ($x = 1$)	20	7,000	286	20	10,000	200	20	13,000	154
Unexposed ($x = 0$)	13	13,000	100	10	10,000	100	7	7,000	100
Total	33	20,000	165	30	20,000	150	27	20,000	135
Percentage of exposed ($100X$)			35			50			65
Rate difference (per 10^5y)			186			100			54
Rate ratio			2.9			2.0			1.5

Individual-level analysis	Ecologic analysis: linear model
Crude rate ratio[a] = 2.0	$\hat{Y} = 200 - 100X$ ($R^2 = 1$)
Adjusted rate ratio (SMR)[b] = 2.0	Rate ratio = 0.50

[a] Rate ratio for the total population, unadjusted for group.
[b] Rate ratio standardized for group, using the exposed population as the standard.

Rothman ME3, Chapter 25

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; *P* < 0.001) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

A. Define the exposure and outcome in this study.

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

A. Define the exposure and outcome in this study.

The exposure is arsenic in drinking water and the outcome is mortality due to pulmonary tuberculosis. Both variables are considered ecologic variables because they were measured at the group/population level (arsenic levels from centralized, aggregate government source data – not individual households – and TB deaths from aggregated rates.

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

B. What observed trend in their data did the authors cite as supportive of the biological plausibility of the arsenic-TB relationship?

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

B. What observed trend in their data did the authors cite as supportive of the biological plausibility of the arsenic-TB relationship?

They cite the observed temporal lag of 10 years between the increased exposure to arsenic and the rise in TB death rates as evidence for a biological relationship. Other arsenic-caused diseases are also observed to have a time lag of around 10 years, making the lagged TB association consistent with similar outcomes.

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

C. Why do the authors believe their study is not very susceptible to the ecologic fallacy?

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

C. Why do the authors believe their study is not very susceptible to the ecologic fallacy?

The authors note both in the Introduction and Discussion that in this particular case, exposure at the group level is likely to be nearly identical to exposure at the individual level. Region II's central water source (the arsenic-contaminated one) was the only water source available, for both business and household use, and everyone in the population drinks water. They note that the only potential source of bias would have been from out-migration, and this did not seem to have a large effect.

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

D. What is the authors' desired level of interest in the exposure in this study?

Tab 4, Problem 2

From Smith et al., 2010: Evidence From Chile That Arsenic in Drinking Water May Increase Mortality From Pulmonary Tuberculosis

Abstract: Arsenic in drinking water causes increased mortality from several cancers, ischemic heart disease, bronchiectasis, and other diseases. This paper presents the first evidence relating arsenic exposure to pulmonary tuberculosis, by estimating mortality rate ratios for Region II of Chile compared with Region V for the years 1958–2000. The authors compared mortality rate ratios with time patterns of arsenic exposure, which increased abruptly in 1958 in Region II and then declined starting in 1971. Tuberculosis mortality rate ratios in men started increasing in 1968, 10 years after high arsenic exposure commenced. The peak male 5-year mortality rate ratio occurred during 1982–1986 (rate ratio = 2.1, 95% confidence interval: 1.7, 2.6; $P < 0.001$) and subsequently declined. Mortality rates in women were also elevated but with fewer excess pulmonary tuberculosis deaths (359 among men and 95 among women). The clear rise and fall of tuberculosis mortality rate ratios in men following high arsenic exposure are consistent with a causal relation. The findings are biologically plausible in view of evidence that arsenic is an immunosuppressant and also a cause of chronic lung disease. Finding weaker associations in women is unsurprising, because this is true of most arsenic-caused health effects. Confirmatory evidence is needed from other arsenic-exposed populations.

D. What is the authors' desired level of interest in the exposure in this study?

The authors' desired level of interest in the exposure seems to be at the individual level since they think their study is not subject to the ecologic fallacy (so the level of interest in the exposure is presumably individual). However, they do not specifically say whether they are interpreting their results at the individual or population level. Because they don't have individual-level data, they cannot determine things like whether arsenic is increasing TB incidence or just mortality among incident TB cases.

Tab 4, Problem 4

Background: There is growing concern that moderate levels of outdoor air pollution may be associated with infant mortality.
Methods: We analyzed daily time-series data on air pollution levels and the infant mortality rate between 1990 and 2000 in 10 major cities of England.
Results: A 10 mg/m^3 increase in sulphur dioxide (SO_2) was associated with a 1.02 times higher infant mortality rate (95% CI 1.01 to 1.04).
Conclusions: Continuing reductions in SO_2 levels in the UK may yield additional health benefits for infants.

Adapted from: Hajat et al. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. J Epidemiol Community Health 2007; 61:719-722.

A. We discussed three types of ecologic measures in class. What type did the authors use in this study for their exposure of interest?

B. Based on this study, the local newspaper publishes a story the next day with the headline: "*An infant's risk of death in the first year of life is increased by 1.02 times when air pollution increases*." You quickly identify something wrong with the level of inference in this statement. Describe the error in 1-2 sentences. ***(3 Points)***

Tab 4, Problem 4

Background: There is growing concern that moderate levels of outdoor air pollution may be associated with infant mortality.
Methods: We analyzed daily time-series data on air pollution levels and the infant mortality rate between 1990 and 2000 in 10 major cities of England.
Results: A 10 mg/m^3 increase in sulphur dioxide (SO_2) was associated with a 1.02 times higher infant mortality rate (95% CI 1.01 to 1.04).
Conclusions: Continuing reductions in SO_2 levels in the UK may yield additional health benefits for infants.

Adapted from: Hajat et al. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. J Epidemiol Community Health 2007; 61:719-722.

A. We discussed three types of ecologic measures in class. What type did the authors use in this study for their exposure of interest?

Environmental – air pollution levels

B. Based on this study, the local newspaper publishes a story the next day with the headline: "*An infant's risk of death in the first year of life is increased by 1.02 times when air pollution increases*." You quickly identify something wrong with the level of inference in this statement. Describe the error in 1-2 sentences. ***(3 Points)***

Tab 4, Problem 4

Background: There is growing concern that moderate levels of outdoor air pollution may be associated with infant mortality.
Methods: We analyzed daily time-series data on air pollution levels and the infant mortality rate between 1990 and 2000 in 10 major cities of England.
Results: A 10 mg/m^3 increase in sulphur dioxide (SO_2) was associated with a 1.02 times higher infant mortality rate (95% CI 1.01 to 1.04).
Conclusions: Continuing reductions in SO_2 levels in the UK may yield additional health benefits for infants.

Adapted from: Hajat et al. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. J Epidemiol Community Health 2007; 61:719-722.

A. We discussed three types of ecologic measures in class. What type did the authors use in this study for their exposure of interest?

Environmental – air pollution levels

B. Based on this study, the local newspaper publishes a story the next day with the headline: "*An infant's risk of death in the first year of life is increased by 1.02 times when air pollution increases*." You quickly identify something wrong with the level of inference in this statement. Describe the error in 1-2 sentences.

Ecologic fallacy, or ecologic bias (Rothman) – to mistakenly infer that the individual level association equals the ecologic association.

Tab 4, Problem 4

Background: There is growing concern that moderate levels of outdoor air pollution may be associated with infant mortality.
Methods: We analyzed daily time-series data on air pollution levels and the infant mortality rate between 1990 and 2000 in 10 major cities of England.
Results: A 10 mg/m^3 increase in sulphur dioxide (SO_2) was associated with a 1.02 times higher infant mortality rate (95% CI 1.01 to 1.04).
Conclusions: Continuing reductions in SO_2 levels in the UK may yield additional health benefits for infants.

Adapted from: Hajat et al. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. J Epidemiol Community Health 2007; 61:719-722.

C. Discuss <u>two</u> reasons why you might choose this study design.

Tab 4, Problem 4

Background: There is growing concern that moderate levels of outdoor air pollution may be associated with infant mortality.
Methods: We analyzed daily time-series data on air pollution levels and the infant mortality rate between 1990 and 2000 in 10 major cities of England.
Results: A 10 mg/m^3 increase in sulphur dioxide (SO_2) was associated with a 1.02 times higher infant mortality rate (95% CI 1.01 to 1.04).
Conclusions: Continuing reductions in SO_2 levels in the UK may yield additional health benefits for infants.

Adapted from: Hajat et al. Outdoor air pollution and infant mortality: analysis of daily time-series data in 10 English cities. J Epidemiol Community Health 2007; 61:719-722.

C. Discuss <u>two</u> reasons why you might choose this study design.

Only have exposure/outcome data at aggregate level
Already have the data and it’s a quick and easy way to generate/test hypothesis.

Agenda

- **Refresher problems**
- **Practice problems**
    - 2 from Tab 1, cross-sectional studies
    - 2 from Tab 4, ecological studies

- **Q & A**