Basic Principles of Epidemiology 573A

For whatever reason(s), even those students who have been cruising along just fine in Epi hit a rather nasty bump this past week. It’s not about the grade (I will give credit for HWK 7 even though you had issues as long as you work through this supplementary homework), but you need to understand the concepts in order to succeed in this course (and do well on Exam 2 and the final). Please take some time to: complete the readings if you have not yet done so, download and use the Confounding Algorithm, and work through this Supplementary assignment.

**Supplementary material to Homework 7**

You are not expected to be able to calculate confidence intervals or estimate sample size or study power in 573A. [You’ll see these calculated on homework answers and some students have the background to do these. These calculations are covered in Basic Biostatistics.] However, you should understand the concepts of bias, confounding, and interaction. The term “understand” means that you should know what these terms mean, when to worry about them, and how to test for them when you are looking at study results.

**Readings:**

Found on the D2L Course Content Tree in the Resources section [Confounding Algorithim.jpg] and **Bias Readings** section [Analysis of confounding in Epi Research; Improving long term recall in Epi Studies; Surveillance in Epi Research; Bias in Epi Research].

**\*I have included some additional reading from the RWJ Foundation YES Program in this supplement.**

Part 1 It’s critical that you can take the relevant information stated in a scenario and translate that information into tables that allow you to calculate results.

What you need to know:

* recognize study design (see Study Design Dichotomous Key and Study Design handout1),
* what calculation (measure of association) is appropriate for that study design,
* what information is needed to complete that calculation,
* how to set up your table,
* how to calculate your measure, and
* how to appropriately interpret (say in words) what your results indicate.

HWK 7 question 7.1 was taking information from a paragraph and comparing incidence rates to determine if it appeared that bias may be occurring in the results. Here are two more examples:

7.1a A medical resident attended a seminar in which a visiting physician suggested that coffee drinking may be associated with angina. He noted that the literature indicates that the estimated rate of angina in the general adult population is 25%. As the medical resident was catching up on charting of his patients he noticed that of the 250 angina patients he had examined in the past two weeks, 190 of the patients reported drinking at least 12 ounces of coffee daily. He spoke with a fellow resident in the ulcer clinic and found that 140 of the last 250 patients seen in the ulcer clinic reported drinking at least 12 ounces of coffee daily. Could these results be subject to bias?

**It seems like the doctor might be trying to compare the rate of coffee drinking between those with angina and the general population. Unfortunately, patients in the ulcer clinic are not representative of the general population and may actually be enriched for coffee drinkers (I don’t drink coffee and I don’t know if it is associated with ulcers or not). The results could certainly be subject to bias because they are not from a random sample.**

7.1b How would you expect the medical resident to set up the scenario trying to show that coffee drinking was related to angina?

**I would expect the doctor to set up a case control study, but given that he is looking at incidence rates in the two populations he might just use a cross-sectional study.**

7.1c What would the medical resident initially find as his measure of association?

**RR = 190/250 / 140/250 = 190/140 = 1.36**

7.1d The medical resident excitedly approached his supervisor who suggested examining the EMR of the past 600 patients in the general clinic to see what proportion of those patients drank at least 12 ounces of coffee daily and were being seen for angina. As may be expected, the general medicine clinic patient characteristics were similar to that of the general U.S. population. Reviewing the EMR, the resident noted that 462 of the patients reported drinking 12 ounces or more of coffee daily. Of these 600 patients examined, 150 were visiting the clinic due to experiencing angina and could be considered as angina patients. How should the resident set up his comparison now, using the information from both his angina clinic population (the 250) and the general medicine clinic population?

**The information from the general medicine population would be sufficient for a cross-sectional study without the bias of selecting a biased sample. Since the doctor is interested in assessing the association of coffee and angina in the general population the angina patient sampling frame is not representative of the general population. Including that data in the analysis would not be wise. The most I would suggest would be to compare the incidence of coffee drinking among angina patients to the incidence of coffee drinking in the general population to see if angina patients were more likely to be coffee drinkers.**

7.1e According to the EMR, 340 of the GMC patients that were not being seen for angina, reported drinking at least 12 ounces of coffee daily. Now that you have sufficient information, you should be able to set up the comparison for the resident.

**In the general population:**

|  |  |  |
| --- | --- | --- |
|  | **No Angina** | **Angina** |
| **No coffee** | **110** | **28** |
| **Coffee** | **340** | **122** |

7.1f Now if the medical resident were to complete his comparison, what would he find to be the association of coffee drinking and angina?

**RR = (122/462) / (28/138) = 1.30**

This next section is set up so you can work through the problems while you are completing the readings. Each reading is from the RWJ Foundation and provides straight-forward explanations of bias and confounding while also illustrating the concepts.

**READ: \* Observational Studies and Bias in Epidemiology.pdf**

After completing the bias reading, answer the following questions: (7.2a-3)

7.2 Short Answer

a. In a few sentences, define and contrast bias and sampling error.

**Bias represents a systematic effect of the sampling scheme which leads to an estimated measure of association that is not consistent with the true measure of association in the target population. Sampling error, however, is a random artifact of taking a limited sample from the target population and should not have any systematic bias so long as the sampling error is not systematic.**

b. Define and contrast selection and information bias.

**Selection bias is a shift in the estimate of the measure of association that is due to differences between the sampling frame and the target population. Information bias is a shift in the estimated measure of association which is due to either a systematic measurement error or misclassification of the sample units. Information bias is due to a bias in the data collected by the researchers whereas selection bias is due to who the information is collected from.**

7.2c Tanning and Melanoma (Skin Cancer)



d. Do you agree or disagree with the investigator? Explain your answer in a few sentences (briefly).

e. Briefly explain what types of bias may be present?

In the second RWJ reading, when you get to the bedsores example (pg 10) work through the problem (space provided below) as you continue with your reading.

**READ: \* Confounding in Epidemiology.pdf**

Bedsores and mortality problem (in reading, Table 2)

Qx 7. Exposure: Disease:

8. (Based on Table 2 in reading) complete the following:

9. Calculate the RR:

10. Interpret your RR (above) in words:



<again, the numbers you need for completing Tables 3a and 3b are in your reading>



13. After reviewing your two completed tables (3a and 3b), briefly explain how it helps to see the results displayed in the two separate tables.

14. Calculate the RR in the high medical severity group. State your interpretation in words.

15. Calculate the RR in the low medical severity group. State your interpretation in words.

16. The unadjusted and adjusted relative risks differ by medical severity, indicating the presence of confounding. Can you think why this is so?

17. If the unadjusted and adjusted relative risks were similar, what would you conclude about whether medical severity confounds the association between bedsores and mortality?

18. Using the results in Tables 3a and 3b, determine what proportion of the high medical severity group died. Then determine the proportion of the low medical severity group died.

Proportion of deaths in high medical severity group: ­­­­\_\_\_\_\_

Proportion of deaths in low medical severity group: ­­­­\_\_\_\_\_

19. Is the probability of death in the high severity group similar or different from the probability of death in the low severity group?

What does this suggest about the association between medical severity and death?

To what part of Figure 1 does this conclusion correspond?

20. Using the results in Tables 3a and 3b, determine what proportion of patients with high medical severity had bedsores. Then determine the proportion of patients with low medical severity that had bedsores.

Proportion of patients in high medical severity with bedsores: ­­­­\_\_\_\_\_

Proportion of patients in low medical severity with bedsores: ­­­­\_\_\_\_\_

21. Is the proportion of patients with bedsores in the high medical severity group similar or different from the proportion of patients with bedsores among the low severity group?

What does this suggest about the association between medical severity and bedsores?

To what part of Figure 1 does this conclusion correspond?