EPI 203: *Epidemiologic Methods* **NAME: _____**
Problem Set 10: Confounding and Interaction II

Due: Dec. 1, 2015 at 1:30 pm section
Possible points: 40 (plus 2 extra credit)

*Note: For the questions marked **"FOR DISCUSSION IN SECTION ONLY:",** please familiarize yourself with these items before section such that we can have a rewarding discussion. Also, please be sure to complete all of the required problem set questions.*

1. Consider the following abstract:

> Oral clefts, transforming growth factor alpha gene variants, and maternal smoking: a population-based case-control study in Denmark, 1991-1994.
>
> Studies in the United States have indicated that both maternal first trimester smoking and infant transforming growth factor alpha (TGFA) locus genetic mutations are associated with cleft lip and/or palate (CLP). In a Danish case-control study of CLP, the authors studied the effects of smoking and TGFA alleles in an ethnically homogeneous setting. Interview information was obtained for mothers of 302 CLP cases and for 567 mothers of normal formed children. Multivariate logistic regression analyses revealed that smoking was associated with a moderately increased risk of CLP (odds ratio = 1.40, 95% confidence interval 1.02-2.00). TGFA genotype was not associated with CLP. The "rare" TGFA allele occurred in 25% of both cases and controls compared with an average of 14% in other white populations. Furthermore, the frequency of CLP in Scandinavia is among the highest in the world. Hence, it is possible that the previously reported association between TGFA and CLP to some degree can be attributable to confounding by ethnicity. The role of toxins in the environment as a causal determinant of CLP could not be directly examined.

(a) The authors conclude that the previously reported association between TGFA mutations and cleft lip/palate seen in the U.S. was possibly caused by confounding by ethnicity. What general technique/method did the authors use to attempt to prevent confounding by ethnicity in their study? Briefly explain how the technique/method was used. (1 pt)

(b) Using a DAG, depict how the authors attempted to preclude confounding by ethnicity when evaluating the association between TGFA mutations and cleft lip/palate. (Note: disregard the presence of other variables) (1 pt)

(c) In this study, the authors matched controls to cases (roughly 2:1) based on place of birth of the infant (by selecting controls from the same census tract in which the case mother resided) and date of birth (controls were the two most proximate births in time). Speculate as to why date and place of birth were considered as confounders when evaluating the association between the two exposures of interest (TFGA mutation and maternal smoking) and the outcome of interest (cleft lip/palate) in this study. Draw a DAG to depict this. (1 pt)

(d) Do you believe the use of date and place of birth completely control for the confounding as the authors' intended? Explain your answer. (1 pt)

(e) What are the merits of the use of matching (which the authors performed) versus stratification alone (an alternative strategy used in the analysis phase) to prevent confounding by date and place of birth? (1 pt)

(f) What if place and time of birth were not causally related to the development of cleft lip/palate? Comment on the downside, if any, of matching on place and time of birth. (1 pt)

2. Consider the following abstract. [Clinical note: Child-Pugh score (A, B, or C) is a staging system for the amount of hepatic (liver) reserve in patients who have liver failure. The score is obtained by looking at albumin, bilirubin, prothrombin time, hepatic encephalopathy, and ascites. Class A is least severe; class C is most severe. Each participant has such a score at the beginning of the trial, noted below.]

> Feasibility and potential benefit of maintenance endoscopic variceal ligation in patients with unresectable hepatocellular carcinoma and acute esophageal variceal hemorrhage: a controlled trial.
>
> BACKGROUND:
> Patients with unresectable hepatoma (cancer of the liver) and first-onset acute esophageal variceal bleeding (bleeding from enlarged blood vessesl in the esophagus) have extremely high rates of recurrent bleeding and mortality. This controlled study evaluates the feasibility and potential benefit of maintenance endoscopic variceal ligation in these patients.
>
> METHODS:
> Patients with unresectable hepatoma and acute esophageal variceal bleeding underwent emergent endoscopic variceal ligation. After control of bleeding, patients were randomized to undergo esophageal variceal ligation (EVL) on a maintenance schedule (i.e., on a regular schedule) or as necessary (i.e., as demanded by clinical necessity when the patient bled again).
>
> RESULTS:
> Fifty-four patients underwent maintenance EVL (on a regular schedule) and 55 underwent on-demand EVL (as clinically necessary). Two outcomes -- survival and recurrent bleeding rates -- were similar (i.e., no statistically significant differences) in both groups. A subgroup analysis of patients with Child-Pugh's A and B hepatic reserve was performed. In this subgroup, maintenance ligation reduced the rate of recurrent bleeding compared with on-demand ligation ($p = 0.043$).
>
> CONCLUSION:
> Maintenance ligation in patients with unresectable hepatoma and variceal bleeding is more effective at reducing the rate of recurrent bleeding in patients with good hepatic reserve (Child-Pugh Score A or B) than in patients with poor hepatic reserve (Score C).

(a) Which factor was considered as an effect modifier in this study? (1 pt)

(b) Does a factor have to be a confounder in order to be an effect modifier? Briefly explain your answer. (1 pt)

(c) In evaluating the effectiveness of maintenance esophageal variceal ligation (vs. ligation as clinically necessary) in this study, is it possible that, at baseline, a participant's Child-Pugh class score could be statistically associated with the group to which the participant was randomized? If yes, what would the mechanism be for this? If no, why not? (1 pt)

(d) Does the p value of 0.043 refer to a test of interaction? Briefly explain your answer. (1 pt)

(e) From the data provided, do you agree with the authors' conclusion? Explain your answer. (1 pt)

3. Examine the following abstract:

Sex-Based Differences in the Effect of Digoxin for the Treatment of Heart Failure.
NEJM 347: 1403, 2002

BACKGROUND:
The Digitalis Investigation Group trial reported that treatment with digoxin did not decrease overall mortality among patients with heart failure and depressed left ventricular systolic function, although it did reduce hospitalizations slightly. Even though the epidemiologic features, causes, and prognosis of heart failure vary between men and women, sex-based differences in the effect of digoxin were not evaluated.

METHODS:
We conducted a post hoc subgroup analysis to assess whether there were sex-based differences in the effect of digoxin therapy among the 6800 patients in the Digitalis Investigation Group study. The presence of an interaction between sex and digoxin therapy with respect to the primary end point of death from any cause was evaluated with the use of Mantel–Haenszel tests of heterogeneity.

RESULTS:
There was an absolute difference of 5.8 percent (95 percent confidence interval, 0.5 to 11.1) between men and women in the effect of digoxin on the risk of death from any cause (P=0.034 for the interaction). Specifically, women who were randomly assigned to digoxin had a higher risk (at 37 months) of death than women who were randomly assigned to placebo (33.1 percent vs. 28.9 percent; absolute difference, 4.2 percent, 95 percent confidence interval, –0.5 to 8.8). In contrast, the risk of death was similar among men randomly assigned to digoxin and men randomly assigned to placebo (35.2 percent vs. 36.9 percent; absolute difference, –1.6 percent; 95 percent confidence interval, –4.2 to 1.0). In the multivariable analysis, digoxin was associated with a significantly higher rate of death among women (adjusted hazard ratio for the comparison with placebo, 1.23; 95 percent confidence interval, 1.02 to 1.47), but it had no significant effect among men (adjusted hazard ratio, 0.93; 95 percent confidence interval, 0.85 to 1.02; P=0.014 for the interaction).

CONCLUSION:
The effect of digoxin therapy differs between men and women. Digoxin therapy is associated with an increased risk of death from any cause among women, but not men, with heart failure and depressed left ventricular systolic function.

Note: The manuscript was authored by individuals who were not affiliated with the original Digitalis Investigation Group study. Instead, the authors received a public use dataset from the NIH for free and conducted the analysis.

(a) Based on the data provided, present the data for interaction using the preferred tabular format described in class which features description of the joint effect of both factors with one reference category. Consider men as the reference category for sex, and consider "placebo" use as the reference category for use of digoxin. Assume that your major interest is in additive scale interaction. Note: not all elements of the table can be completed given the data provided; therefore, provide whatever elements you can. (3 pts)

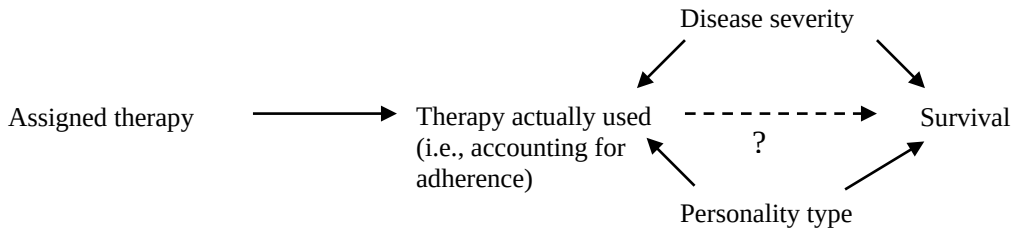
(b) Calculate the metric by which to evaluate the presence of additive interaction. Is additive interaction present? If so, is it sub- or super-additive? (1 pt)

(c) Calculate the metric by which to evaluate the presence of multiplicative interaction. Is multiplicative interaction present? If so, is it sub- or super-multiplicative? (1 pt)

(d) When reporting interaction, is it preferable to do what was performed in the digoxin study or what was done in the variceal bleeding study (in question 2)? (1 pt)

(e) **FOR DISCUSSION IN SECTION ONLY:**
The authors showed data for both additive and multiplicative interaction. In this article, which of the following do you believe is relevant and addressable: statistical, mechanistic, public health, or biologic/physical interaction? More than one answer may be correct.

4. Consider the following DAG. Like all DAGs used in causal inference, all directed edges depict causal relationships and all common causes (for any two variables) are shown in the diagram. The population in this DAG can be considered any population with a particular disease condition. The exposure under study is "therapy actually used" and the outcome is survival. Therapy refers to any form of therapy for the disease.



(a) If the goal is to estimate the causal effect of "therapy actually used" on survival (which is a typical goal of research), what type of variable is "assigned therapy"? In other words, what term is used to describe "assigned therapy"? (1 pt)

(b) Of the different types of study designs used in clinical research, what type of study design do you believe is most likely being depicted in the DAG? Explain your answer. (1 pt)

(c) Assume the DAG that is shown is accurate (i.e., complete). If one performed the study design depicted in the DAG and calculated the relationship between the assigned therapy and survival, could the relationship be biased by confounding? Explain your answer. (1 pt)

5. One day while evaluating the usage of various "statins" (drugs that are conventionally used to treat high cholesterol) in a large administrative database, your research assistant investigates whether statin use has any effect on the development of malignancies. He performs the analysis and finds that the risk ratio (5 year duration) for breast cancer associated with statin use is 0.65 (95% CI: 0.37 to 1.15; p = 0.14). Before going to lunch, he does one more analysis asking whether there is perhaps effect-measure modification by pre- versus post-menopausal status. He performs a test of homogeneity between a stratum limited to pre-menopausal women and a stratum limited to post-menopausal women and finds again that the p value is "not significant". Concluding that "nothing is happening" between statins and breast cancer, he goes to lunch and leaves you with the following hard copy tables:

All Women		
	Breast Cancer	No Breast Cancer
Statins	19	1393
No statins	29	1376

Pre-Menopausal		
	Breast Cancer	No Breast Cancer
Statins	8	205
No statins	5	200

Post-Menopausal		
	Breast Cancer	No Breast Cancer
Statins	11	1188
No statins	24	1176

You don't have the actual data file (you've delegated all of this to your assistant), but you decide to check the accuracy of your assistant's work.

(a) Determine the stratum-specific risk ratios and the p value of the test of homogeneity comparing the multiplicative measures of association (in this case, the risk ratios). (3 pts)

Hint: you can perform these calculations in Stata by creating a dataset in Stata using the information contained in the cells of the two menopause status-specific strata. Do this by considering that there are three variables here that can be coded 0 or 1 (statin use, breast cancer, pre vs. post-menopausal status). Thus, there are 8 possible sets of values that a given person could have. These 8 possible sets of values are actually the 8 cells contained in the two stratum-specific tables. By going into the Data Editor in Stata (click on the Data Editor icon in the row of icons), one can enter these 8 possible sets of values (e.g., 111, 110, 101, 001, etc.) by making 8 rows where each row has 3 variables. Then, use a 4th variable to place the number of persons who possess that particular set of the first three variables. You can then perform most Stata commands with this dataset if you remember to add "`[freq=varname]`" where varname is the variable that holds the number of persons who possess that particular set of first three variables. If you don't change any variable names and just use what Stata defaults to, you would write `[freq=var4]`. This process is known as using frequency weights. For example, the command to look at the crude association between an exposure variable and an outcome variable would look like:

```
cs outcome_variable exposure_variable [freq= frequency_variable]
```

(b) Determine the p value of the test of homogeneity using the additive measures of association (in this case, the risk differences). (3 pts)

(c) What are your conclusions regarding the role statins in the occurrence of breast cancer? (1 pt)

6. Consider the following abstract:

Racial Disparity in Pregnancy-related Mortality Associated with Livebirth:
Can Established Risk Factors Explain It?

The authors conducted a nested case-control study to determine whether the fourfold increased risk of pregnancy-related mortality for US Black women compared with White women can be explained by racial differences in sociodemographic and reproductive factors. Cases were derived from a national surveillance database of pregnancy-related deaths and were restricted to White women *(n* = 840) and Black women *(n* = 448) whose pregnancies resulted in a livebirth and who died of a pregnancy-related cause between 1979 and 1986. Controls were derived from national natality data and were randomly selected White women and Black women who delivered live infants and did not die from a pregnancy-related cause *(n* = 5,437). Simultaneous adjustment for risk factors by using logistic regression did not explain the racial gap in pregnancy-related mortality.

The authors adjusted for the four main known risk factors for pregnancy-related mortality using the following measurements:

- Maternal age at delivery: <20, 20-24, 25-29, 30-34, 35-39, >40 years
- Education: $<12^{th}$ grade, $\underline{\geq}12^{th}$ grade
- Parity: Most prior studies had defined parity as the total number of births given to fetuses 28 weeks or greater, regardless of whether alive or stillborn. Because this was not available for the present study, number of live births was used instead, including the index pregnancy. This variable was obtained from the birth certificate.
- Adequacy of prenatal care: 3 category variable (no care, inadequate, adequate) as determined by number and timing of prenatal care visits and birth weight.

All information was derived from a combination of birth certificates, fetal death certificates, and maternal death certificates.

Despite adjustment for all four of these factors, there still remained a racial difference in pregnancy–related mortality (black women had higher mortality than white women). One formal possible explanation for the findings is that genetic differences that account for race are responsible for the racial gap, operating via some novel biologic mechanism. Can you provide <u>at least two alternative explanations</u> for the finding (i.e., something other than racial genetic differences operating via a novel biologic mechanism)? Assume that the finding is not the result of chance, problems in the selection of the case/controls, or measurement error of the women's race or mortality status. (2 pts)

Hint: we are looking for broad alternative explanations, which do not require detailed understanding of this biological/clinical system

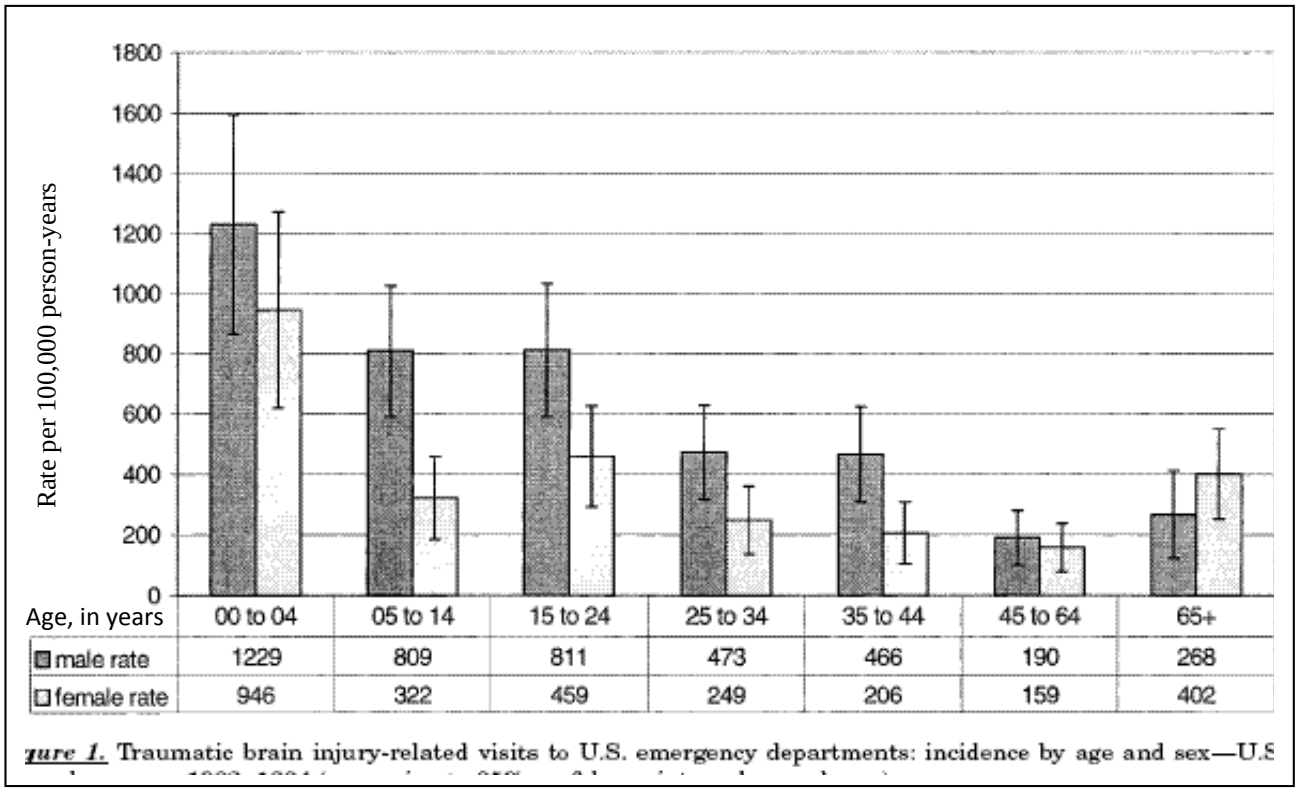
7. Consider the underlined result in the following abstract:

OBJECTIVE:
To describe the influence of sex on the occurrence of traumatic brain injuries (TBIs) treated in U.S. emergency departments (EDs).

METHODS:
A secondary analysis was performed on data from the National Hospital Ambulatory Medical Care Survey administered from 1992 to 1994.

RESULTS:
The rate of new TBI treated in U.S. EDs was 444 per 100,000 person-years (95% CI = 390 to 498). Men were 1.6 times as likely as women to suffer TBI until the age of 65 years, when the female rate exceeded the male.

In this abstract, the figure below was interpreted by the authors as indicating that there was an interaction between age and sex in the occurrence of traumatic brain injuries. Specifically, it was interpreted that men had a greater incidence rate of traumatic brain injury than women, except at ages ≥65 years when the rate for women was greater. Assume that indeed there is, in truth, no interaction between sex and age in the occurrence of traumatic brain injury. Assume also that the observed findings are not the result of chance, selection bias, or mismeasurement of sex, age, or outcome. What else could be causing the apparent interaction? (2 pts)



Age, in years	00 to 04	05 to 14	15 to 24	25 to 34	35 to 44	45 to 64	65+
male rate	1229	809	811	473	466	190	268
female rate	946	322	459	249	206	159	402

ure 1. Traumatic brain injury-related visits to U.S. emergency departments: incidence by age and sex—U.S

Figure 1. Traumatic brain injuries treated in U.S. emergency departments: incidence by age and sex

8. Your colleague is studying the influence of childhood education on the occurrence of diabetes in children. He sends you his DAG, encoded in the following text:

```
Child's_diabetes_status O @1.000,0.004
Child's_education E @0.330,0.003
Family_income_in_childhood 1 @0.076,-0.062
Mother's_diabetes_status 1 @0.454,-0.039
Mother's_genetic_background 1 @0.851,-0.063

Child's_education Child's_diabetes_status
Family_income_in_childhood Child's_education Mother's_diabetes_status
Mother's_genetic_background Mother's_diabetes_status Child's_diabetes_status
```

(a) Create the DAG in dagitty.net by inserting (copying and pasting) the above text in the Model Text Data window (bottom right hand corner of screen) and then clicking "Update DAG". Your colleague mentions that mother's genetic background is theorized, but it has not been measured. Alter the DAG by depicting that "mother's genetic background" was not measured (i.e., it is unobserved). Instructions can be found in the "How to" menu. At this point, show your DAG by pasting it into your answer document or by printing and attaching a separate page. Also show your model text data by copying and pasting it into your answer document. (1 pt)

(b) Your colleague is studying the causal association between childhood education and diabetes with a case-control study (cases = children with incident diabetes; controls = children without diabetes). He now reminds you that during recruitment, he matched on mother's diabetes status because mother's diabetes was associated with childhood education and the child's diabetes status in prior work he had done. Matching is a form of adjustment, and adjusting for a variable can be depicted in dagitty.net; (instructions can be found under the "How to" menu). Assuming that you agree with the original DAG that has been created (and that mother's genetic background is unobserved), do you agree with this matching? If not, is there anything you can do to rescue his study and estimate the causal relationship between education and child's diabetes? Substantiate your answer by showing output/advice from dagitty.net. (2 pts)

9. In the last Journal Club article by *Rookus et al.* (Induced abortion and risk for….*JNCI* 23:1759, 1996), consider the following:

(a) The main methodological point that the authors attempted to prove (i.e., the purpose of the paper) was bolstered by a statistical test of interaction. Describe the numerical data (i.e., measure(s) of association) that the authors used to make their point and the accompanying statistical test of interaction in a few clear sentences. In particular, make sure to explain in detail, in one or more clear and informative sentences, the numerical meaning/interpretation of the statistical test of interaction. (2 pts)

(b) What type of scale was used to assess interaction in this study? What alternative scale (the type not used in this study) to assess interaction could also have been investigated? Explain your answer including how, in general, the alternative scale interaction could have been investigated. (2 pts)

Perform this test of interaction on the alternative scale. Note: do not worry about adjustment for any factors. (2 pts extra credit)

(c) **FOR DISCUSSION IN SECTION ONLY:**
Assuming the authors had intended to demonstrate a specific concept of interaction, do you think they were interested in statistical interaction, mechanistic interaction, public health interaction, or biologic interaction? Explain your answer.

10. **PREPARATION FOR NEXT WEEK'S PROBLEM SET:**
    In the last Problem Set of the year (#11), next week, we will be asking you to use DAGitty.net to create a DAG depicting a research question that you are currently studying. Be sure to make it a causal DAG, meaning that all common causes of any two variables should be shown. If you are not working on a project that fits the spirit of what DAGitty.net does, then create a DAG based upon a research question that you would like to investigate in your field.

11. Please address the following questions from the *Nitz et al.* Journal Club article (to be discussed this week):

(a) While non-differential misclassification is the most likely type of misclassification of exposure that occurred in this work, it will lead to an attenuation of the measure of association and thus cannot explain the effect that was observed by the authors. Can you think of a plausible way that differential misclassification of exposure could have occurred, and what do you want to look for in the Methods section to help you evaluate whether this may have occurred? (1 pt)

(b) Type I statistical error due to multiple statistical comparisons is a common concern in gene association studies. Are you concerned about this in the present study? What have the authors done to convince you one way or another? (1 pt)