**EPID600 Lab: Comparisons, Hypothesis Generation, & Team Project Topic**

**Pre-Lab Individual Work**

*Please complete the following* ***three*** *parts and associated questions* ***BEFORE*** *lab.*

**PART 1: Applying the concepts discussed in lecture.**

*Smoking Cessation Example*

1. Last week you calculated the risk of smoking cessation between groups that were exposed to an anti-tobacco advertisement called “Stages” and those who were not. You noted that there were differences in risk of cessation between the two groups, but now you have the tools to quantify it. Calculate the risk for in the table below where those lost to follow up were excluded and answer the following two questions:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Smoking Cessation** | **No Smoking Cessation** | **Total** | **Risk** | **Risk per 100 persons** |  |
| **Advert** | 32 | 138 | 170 | .1882 | 18.82 |  |
| **No advert** | 56 | 454 | 510 | .1098 | 10.98 |  |
| **Total** | 88 | 593 | 681 |  |  |  |
|  |  |  |  | .1294 | 12.94 |  |

1. Calculate and interpret the risk difference between those that were exposed to the anti-tobacco advertisement and those who were not. Use either [OpenEpi](http://www.openepi.com/Menu/OE_Menu.htm) or EpiInfo7 to calculate 95% confidence intervals. Provide an interpretation of the risk difference. Is the 95%CI statistically significant?

*Risk Difference (risk exposed-risk unexposed) = 7.843%*

*95% CI (1.371,14.32)*

*Over repeated sampling 95% of the CIs would be expected to contain the true risk difference.*

*Among those that saw the advertisement, the risk of smoking cessation was 7.843 cases per 100 persons higher than those who did not watch the advertisement.*

*Using an alpha value of .05*

*Our two tailed p-value is .011*

*p-value < alpha which implies statistical significance.*

*Since the 95% CI does not contain 0 (the null value), it is statistically significant.*

1. Calculate and interpret the risk ratio between those that were exposed to the anti-tobacco advertisement and those who were not. Provide an interpretation. Use OpenEpi or EpiInfo7 to calculate 95% confidence intervals. Is the 95%CI statistically significant?

*Risk ratio (risk exposed/risk unexposed)= 1.714*

*Among those that saw the advertisement, the risk of smoking cessation was approximately 1.714 as likely compared with those who did not see the advertisement.*

*95% CI (1.151,2.553)*

*Over repeated sampling 95% of the CIs would be expected to contain the true risk ratio.*

*The CI is statistically significant since it does not contain 1 (the null value)*

*Flu Example*

BACKGROUND

The flu is a contagious respiratory illness caused by influenza viruses which can cause mild to severe illness and can possibly lead to death. Young children, older people, and people with certain health conditions are at high risk for experiencing serious flu complications. Definitive identification of cases are made through laboratory testing of microorganisms which can be labor-intensive and time- consuming and could lead to delayed responses by medical and public health professionals.

Influenza-like illness (ILI) is a medical diagnosis of possible influenza or a number of other

illnesses which cause similar symptoms of fever, shivering, chills, malaise, loss of appetite,

and body aches. Accurate surveillance of ILI is important for timely public health responses to outbreaks and can be done through automated models like those described in Park et al

(Emerging Health Threats Journal, 2011) in the NC Veterans Affairs Medical Center in

Durham.

During October through May, the NC Division of Public Health provides weekly updates on the spread of influenza in North Carolina at [www.flu.nc.gov](http://www.flu.nc.gov). The following questions are adapted from numbers published in their Final Summary 2016-2017 document.

1. As a researcher, you’re working with NC Division of Public Health to study the spread of influenza-like illness. Over a one-week time period, a total of 20,685 people each contributed 1 person-week of time to your study and a total of 251 people met the case definition of incident influenza-like illness during this same period of time. What is the rate of influenza-like illness in these locations (report in person-weeks). What is the interpretation?

*Rate = 251/20685 = .012134 cases per person week*

*This means that there were .012134 new cases of flu like illness per person week during*

*the one week period.*

1. The highest rate during this time period was found during the 8th week of 2017 where a total of 15,227 people each contributed 1 person-week of time and 1,178 people met the case definition of incident influenza-like illness. The lowest rate is seen in the 20th week of 2017 when 5,997 people each contributed 1 person-week of time and only 15 met the case definition for incident influenza-like illness. What is the rate difference between these two weeks? (report in person-weeks) What is the interpretation?

*8th week rate= 1178/15227=.07736 cases per person week*

*20th week rate = 15/5997=.0025 cases per person week*

*Rate difference = 8th week rate-20th week rate= .07736-.0025=.07486 cases per person week*

*The rate of flu like illness was .07486 cases per person week higher in the 8th week than the in the 20th week.*

**PART 2: Hypothesis Generating Exercise.**

Individually, read the two scenarios below and develop potential hypotheses based on the available descriptive data. These will be discussed during lab with your team members. You do not need formally written hypotheses for these questions- only possible explanations (which can be tested) for what is being seen.

1. A recent spike in HIV has been reported among injection drug users in your city. Two years ago, a local NGO instituted a free needle exchange program, allowing drug users to exchange used syringes for new, clean ones. In addition to providing clean needles, the program offers referrals to various social programs and offers free HIV testing to clients. Social workers have reported an increased number of injection drug users congregating in the neighborhood where the needle exchange program is housed. Just over a year ago, the city health departments began collecting HIV data on homeless populations.

*Several things are happening here. First of all, there is increased access to HIV testing for these individuals as well as a motivation to come to the city for free clean needles. In addition HIV data collection on homeless populations (who would likely be at higher risk for using intravenous drugs) is new and thus has few data points. Most likely this is not truly a spike, but a result of an influx of people with higher risk of HIV combined with testing these individuals whose HIV case status was likely previously unknown.*

1. You are working to improve maternal health outcomes in a developing country in Africa. Rates of maternal morbidity and mortality from complications related to birth are 20% higher in your district than anywhere else in the nation. Your district is the poorest region of the country, but dietary evaluations indicate that most citizens do not suffer from malnutrition. Vital statistics indicate that the district experiences a higher rate of unwed teen pregnancies than other districts, with roughly 68% of women experiencing their first pregnancy at an age <18. It is a very rural region with only one hospital and four public health clinics, which operate on unpredictable schedules. Most residents report having to travel 4+ hours to receive medical attention of any sort. Many births are conducted at home with local midwives in attendance.

*Access to medical care (outside of the midwives) could be a significant factor in the maternal health outcomes. Since the hospital is far away and the health clinics have unpredictable schedules and there is a high poverty rate, people are most likely choosing at home births with local midwives as a cheaper more accessible option. Also given that the rate of unwed teen pregnancies is higher, there may be less support for the women giving birth and given there age they may be less knowledgeable about their birthing options than an older woman would be. It would be important to test whether having an at home birth with a local midwife has a higher risk for maternal health complications versus a hospital birth.*

**PART 3: Ideas for team project.**

List 1-2 ideas for your team project below. Think of areas that you’re particularly interested and

questions which could be answered in a survey.

*Exercise and stress levels. Quantify levels of activity by intensity and minutes and look for its impact on reported stress levels and hours of sleep.*

**In-Class Group Work**

*Please complete the following two parts* ***DURING*** *lab.*

1. Review your answers to the calculations and hypotheses generated above in Parts 1 and 2 with your TA and lab section.
2. With your team, review your ideas for a team project from Part 3 above. Discuss and draft a team project hypothesis and answer the questions below. (**see Team Project Part 1 folder in Sakai for details)**. We challenge you to pick a salient public health topic that is of interest to the entire team. You will be able to modify or change your topic as you delve into the literature. Given the short length of the semester, we recommend a cross-sectional study design for your team project. With a cross-sectional study design, you will use measures of prevalence for both your exposure and your outcome.

*Discuss potential topics of interest*.

Consider pressing or potential health related problems in your population. What issues have you noticed on campus or in a local community?

*General advice for selecting a topic.*

Keep it simple. Determine a study sample/study population that you can reach via email with a survey link or interview in person using cellphones.

Consider exposures and outcomes that can be easily and reliably reported. We encourage selecting topics with potential “real world” use (such as assisting a

community to quantify a public health problem).

1. Describe your target population. How will you define it? Think about generalizability. What are some ways you could sample from your target population? Describe advantages and disadvantages of two different ways of sampling.
2. How will you define your main exposure of interest? For example, sleep deprivation would be defined as sleeping less than 6 hours a night for 7 days.
   1. Describe what your numerator and denominator and time frame will be for calculating prevalence of exposure.
   2. Describe what your numerator and denominator and time frame would be if you were calculating risk of exposure.
   3. Describe what your numerator and denominator and time frame would be if you were calculating rate of exposure.
3. How will you define your outcome of interest? For example, high caffeine consumption would be defined as four 8 ounce cups of coffee, tea, or caffeinated beverages per day per 7 days.
   1. Describe what your numerator and denominator and time frame will be for calculating prevalence of outcome.
   2. Describe what your numerator and denominator and time frame would be if you were calculating risk of outcome.
   3. Describe what your numerator and denominator and time frame would be if you were calculating rate of outcome.
4. Work with team members to decide upon and draft 3 standard demographic questions for your team project survey. (Note: You may wish to consult other standardized surveys for comparability, such as CDC’s Behavioral Risk Factor Surveillance Survey: <http://www.cdc.gov/brfss/>, the US Census Bureau: <http://www.census.gov/history/www/through_the_decades/index_of_questions/>, or the National Health Interview Survey: <http://www.cdc.gov/nchs/nhis/nhis_questionnaires.htm)>
5. Each TEAM must submit one Word document under the team assignment folder.
6. One team member will post your team hypothesis draft to the forum board. Include both a definition of both your exposure and your outcome. This does not have to be your final draft.

**Here is some additional clarification from the Team Project part 1 guidelines:**

*Exposure and health outcome*

With your team, decide on both a main exposure of interest and a main health outcome of interest. You need to decide on the definition of your exposure. The exposure could also be a behavior. Be specific. The health outcome can be a behavior, attitude, or a measure of general health, or specific health (such as blood pressure).

Example: “Do public health students who text frequently while driving report more car accidents?” You would need to think through how your team will define the exposure “frequent” texting and the outcome “more car accidents.” Use the peer reviewed literature on the topic to help inform your decisions.

*What is a study sample?*

See Figure 1 below. Examples include randomly select members of a local community organization, or systematically sampled residents of the Northside neighborhood. For those of you who are not familiar with random or systematic population sampling, you may want to view the following clip as an overview: [Survey Design: Sampling](http://media.sph.unc.edu/adobe/cphp/tws/HEP_SDP1/) (slides 5-22, 58). For this project, select a systematic or random sample from your sampling frame and obtain data on at least 100 individuals.

Another option: you can survey all the individuals in your defined “sampling frame” – This is known as a census sampling. Examples include surveying all members of an on-campus fraternity or sorority listserv. If you choose the census survey option, you **MUST** have the total number of members of the listserv (i.e. denominator information), in order to calculate your response rate. Your target sample size is 100 individuals. ***One caveat-you are not allowed to use formal graduate or undergraduate program listservs in the Gillings School of Global Public Health.***

Figure 1. Sampling Terminology  
(from http://www.socialresearchmethods.net/kb/sampterm.php)

