Preview Assessment: Post-class Quiz #4 Random Error

Question 1  5 points
Which of the following are the correct values for the null hypothesis?

- A. Risk ratio=0; Odds ratio=0; Risk difference=0; Attributable proportion=0
- B. Risk ratio=1; Odds ratio=1; Risk difference=0; Attributable proportion=0
- C. Risk ratio=1; Odds ratio=1; Risk difference=1; Attributable proportion=1
- D. Risk ratio=0; Odds ratio=0; Risk difference=1; Attributable proportion=1

Question 2  5 points
Which of the following statements are true regarding p-values? (Select all that apply).

- A. P-values depend upon both the magnitude of association and the precision of the estimate (the sample size).
- B. P-values provide a way of evaluating the precision of a point estimate.
- C. P-values measure the compatibility of the data with the null hypothesis, not the probability that the null hypothesis is correct.
- D. In addition to evaluating random error, p-values take into account the errors introduced by bias and confounding.

Question 3  5 points
Richard Doll and Bradford Hill were British researchers who conducted one of the earliest analytic studies of the association between smoking and lung cancer. The study was conducted in 20 hospitals in the London area, in which they identified patients who had been admitted for lung cancer. The comparison group consisted of age and sex-matched patients who had been admitted to the same hospital for problems other than cancer. All patients were interviewed about past smoking and other exposures. Most of the subjects were men, but they were able to identify 60 women with lung cancer; 41 of these stated that they were smokers. Among the 60 women who did not have cancer, 28 stated that they were smokers.

Use Epi_Tools.XLS to conduct an appropriate analysis of these finding and compute the p-value for this association. Round off the p-value to 3 decimal places and enter the p-value. For example, if the p-value is 0.23435, enter 0.234.

[ ]

Question 4  5 points
Mannino et al. used the National Health Interview Survey (a cross-sectional study) to examine the association between exposure to environmental tobacco smoke (ETS) and adverse health effects in children. The results section stated:

“We compared data from 9632 children with no ETS exposure in the home with data from 5047 children with daily ETS exposure in the home. In the two weeks
before the survey, children who were exposed to ETS in the home had a
greater incidence of respiratory illness (7.9% vs. 6.8%, p=0.07) and chronic
respiratory disease exacerbations (2.5% vs. 2.3%, p=0.45) than children who
were not exposed, although chance cannot be excluded as a reason for these
findings. Children who were exposed to ETS had a higher prevalence of
chronic respiratory disease (18.0% vs. 15.7%, p=0.19) than children who were
not exposed, but again, chance cannot be excluded as a reason for these
findings."

Which of the following statements are true? (Select all that apply.)

A. The sample size was ridiculously small.
B. The p-values clearly indicate that the null hypothesis can be rejected
C. Despite the large sample size, the differences were small and the p-
values suggest that random error could account for the differences.
D. If the sample size had been larger, the results would have been

Question 5
5 points
Suppose that at the completion of a study, an appropriate statistical test was
performed and the p-value was 0.02. Which of the following statements would
be appropriate conclusions regarding this study?

A. The null hypothesis can be rejected, because the p value was less
   than 0.05.
B. The null hypothesis should be accepted, because the p value is less
   than 0.05.
C. The risk of developing the outcome is >0.02.
D. The study was underpowered (i.e. the sample size was too small).

Question 6
5 points
Which of the following statements are true about the 95% confidence interval
for a RR or an OR?

A. It indicates whether the observed data are compatible with the null
   hypothesis (whether p <0 .05),
B. It indicates a range of values for the magnitude of association that are
   compatible with the observed data.
C. It indicates the precision of the estimate.
D. All of the above.

Question 7
5 points
A case-control study was done to determine whether high levels of lead in the
soil around a child's home is associated with an increased risk of lead
poisoning. What is the null hypothesis for this study?
A. Children whose yard has high levels of lead have an increased risk of lead poisoning.
B. That lead can cause health problems.
C. That lead can be absorbed through the lungs.
D. That the risk of lead poisoning is the same for children who live in yards with high lead levels and children who live in yards with low lead levels, i.e. that the OR=1.0.

Question 8
5 points
For the case-control study investigating whether high levels of lead in the soil around a child's home were associated with an increased risk of lead poisoning the odds ratio was 1.9, and the 95% confidence interval was 0.8 - 3.0. What do these results indicate? (Select all that apply).

☐ A. The odds ratio from this study suggests that children living in yards with high lead levels have 1.9 times the risk of developing lead poisoning compared to children whose yards have low levels of lead, but the increase was not statistically significant.
☐ B. We can be 95% confident that the true odds ratio lies between 0.8 and 3.0.
☐ C. The p-value must be greater than 0.05.
☐ D. Lead in the soil is absorbed through the skin.

Question 9
5 points
Antibiotic resistant bacteria

We have known for years that bacteria are evolving in ways that make them resistant to antibiotics. Staphylococci, for example, were very sensitive to penicillin when it was first introduced, but they rapidly became resistant to it. Newer synthetic antibiotics, such as methicillin have been useful in treating Staph. infections, but we have seen more and more cases of methicillin-resistant Staph. aureus (commonly known as MRSA). Most of the problems with antibiotic resistant bacteria were initially seen in hospitalized patients, but the problem is now being identified in community-acquired infections as well. Moran et al. studied the bacteriology of patients seen in emergency rooms for acute, purulent (with pus) skin and soft tissue infections in 11 US cities (N Engl J Med 2006;355:666-74). They were specifically interested in the prevalence of infections with MRSA. For example, in Atlanta they enrolled 32 patients who were seen in the emergency room for acute, purulent skin or soft tissue infections and found that 23 of these had infections with MRSA. Use the Epi_Tools.XLS spreadsheet to compute the prevalence of MRSA in skin/soft tissue infections in Atlanta, and compute the 95% confidence interval for this proportion. Which of the following is closest to the 95% confidence interval?

☐ A. 72% (54.6-84.4)
☐ B. 72% (57.5-82.8)
Avian Influenza A (H5N1) Infection in Eastern Turkey in 2006. Ahmet Faik Oner, et al N Engl J Med 2006;355:2179-85. Excerpt from Introduction: The first documented case of human infection with the avian influenza A (H5N1) virus occurred in Hong Kong in 1997. As of July 26, 2006, a total of 232 H5N1 virus infections in humans had been documented, with a mortality rate of 58% among hospitalized patients. An outbreak of H5N1 virus had been detected in poultry in eastern Turkey from December 27, 2005, through January 26, 2006. This outbreak in poultry was followed by infection in humans. In contrast to the previously reported cases of H5N1 in southern and eastern Asia, the human outbreak in Turkey was concentrated in a small geographic area during a short period. H5N1 infection was diagnosed in 12 patients in Turkey and confirmed by the World Health Organization (WHO). Eight of these patients were followed at our hospital in Van. The four other patients were admitted to other health centers and survived without complications. We report the clinical, epidemiologic, and radiologic features and history of exposure of the eight patients with H5N1 virus infection cared for at our center.

In the report by Ahmet Faik Oner, et al. four out of eight patients with avian flu died in Turkey, for a mortality rate of 50%. What was the 95% confidence interval for this mortality rate?

A. p<0.05  
B. p>0.05  
C. 24.8% - 75.1%  
D. 21.5% - 78.5%  
E. None of the above

In the report by Ahmet Faik Oner, et al on 8 eight cases of avian flu in Turkey, the authors noted in the introduction that as of their report there had been a total of 232 cases of avian flu in humans and 135 of these had died (mortality rate= 58%). How would the width of the confidence interval for this mortality rate compare to that in the report by Oner et al.?

A. The overall mortality rate would have a narrower 95% confidence interval because of the larger sample size.  
B. The overall mortality rate would have a wider 95% confidence interval because of the larger sample size.  
C. The overall mortality rate would have the same 95% confidence interval.  
D. It is impossible to say for sure.
Influenza-Associated Deaths among Children in the United States, 2003-2004. Niranjan Bhat, et al. N Engl J Med 2005;353:2559-67. "**Methods:** During the 2003-2004 influenza season, we requested that state health departments report any death associated with laboratory-confirmed influenza in a U.S. resident younger than 18 years of age. Case reports, medical records, and autopsy reports were reviewed, and available influenza-virus isolates were analyzed at the Centers for Disease Control and Prevention. **Results:** One hundred fifty-three influenza-associated deaths among children were reported by 40 state health departments. The median age of the children was three years, and 96 of them (63 percent) were younger than five years old." There were roughly 2,045,000 infants less than six months old during this time span, and 18 of these died. What was the age-specific mortality rate from influenza during this flu season?

- A. 8.8 per 1,000,000
- B. 8.8 per 100,000
- C. 8.8 per 10,000
- D. 8.8 per 1,000

**Question 13**

For the previous question the authors reported that the 95% confidence for the mortality rate in infants under 6 months of age was 0.52-1.39 per 100,000. The best interpretation of this confidence interval would be:

- A. With 95% confidence, the true proportion of deaths was between 0.52 per 100,000 and 1.39 per 100,000 among infants less than six months old.
- B. With 95% confidence the mortality rate was between 0.52-1.39% in infants less than six months old.
- C. With 95% confidence the mortality rate was between 52-139% in infants less than six months old.
- D. None of the above

**Question 14**

A Study of Human Papillomavirus and Oropharyngeal Cancer by D'Souza et al., N Engl J Med 2007;356:1944-56. In this article the authors were primarily interested in determining whether there was an association between infection with human papillomavirus and development of oropharyngeal cancer. As part of their study they examined a number of possible risk factors for oropharyngeal cancer. One interesting finding was that subjects who did not brush their teeth every day had an increased risk of oropharyngeal cancer with an odds ratio of 5.4 and a 95% confidence interval of 1.7 to 17.8. Which of the following is the best interpretation of this confidence interval?

- A. Not brushing daily was a statistically significant risk factor for oropharyngeal cancer, because the 95% confidence interval did not
include the null value.

B. Not brushing daily was not a statistically significant risk factor for oropharyngeal cancer, because the 95% confidence interval did not include the null value.

C. Not brushing daily was a statistically significant risk factor for oropharyngeal cancer, because the 95% confidence interval included the odds ratio of 5.4.

D. Not brushing daily was not a statistically significant risk factor for oropharyngeal cancer, because the 95% confidence interval included the odds ratio of 5.4.

Question 15  
5 points

The 95% confidence interval for failure to brush (1.7-17.8) is wider than the 95% confidence interval for tooth loss (1.7-6.8). Why is this?

A. Not brushing daily is a stronger risk factor for oropharyngeal cancer.

B. The estimated measure of association is more precise for not brushing daily.

C. The number of subjects who admitted to not brushing was relatively small, so the estimated odds ratio was less precise.

Question 16  
5 points


Previous studies had found that persons living in poorer residential neighborhoods have an increased prevalence of risk factors for coronary heart disease (CHD) and increased risk of CHD mortality and all-cause mortality. Hypertension (high blood pressure) is a common and important form of cardiovascular disease that contributes significantly to CHD. These authors used data from the Black Women's Health Study to examine the influence of neighborhood socioeconomic level, measured by median housing value from census data, on the risk for developing hypertension. The study enrolled about 59,000 black women in 1995, and they have been followed at intervals using questionnaires to obtain information on risk factors and various outcomes. The current study was limited to 42,168 subjects who did not have hypertension at the beginning of the study.

The primary thrust of the study by Cozier et al. was to compare the risk of developing hypertension among black subjects in 5 different socioeconomic groups based on median housing value of their neighborhoods. The table below shows how certain other characteristics compared across these 5 groups.

Using the group with the HIGHEST median housing value as the reference (least exposed) group, compute the incidence rates, the incidence rate ratio (relative risk), and the 95% confidence interval for the relative risk of the
LOWEST housing value group using the Stat Tools spreadsheet.

CozierTable.jpg

A. Subjects with the lowest median housing value had 0.69 times the risk of developing hypertension compared to subjects with the highest median housing value.
B. Subjects with the lowest median housing value had 0.69 times more risk of developing hypertension compared to subjects with the highest median housing value.
C. Subjects with the lowest median housing value had 1.44 times the risk of developing hypertension compared to subjects with the highest median housing value.
D. Subjects with the lowest median housing value had 1.44 times less risk of developing hypertension compared to subjects with the highest median housing value.

Question 17
Which of the following are correct statements regarding the statistical significance of the comparison in the previous question? (Select all that apply).

☐ A. The increased risk of hypertension in those with the lowest housing values was statistically significant because the 95% confidence interval for the relative risk did not contain the null value.
☐ B. The increased risk of hypertension in those with the lowest housing values was NOT statistically significant because the 95% confidence interval for the relative risk included the rate ratio.
☐ C. The increased risk of hypertension in those with the lowest housing values was statistically significant because the p-value from the chi-square test was <0.05.
☐ D. The increased risk of hypertension in those with the lowest housing values was NOT statistically significant because the p-value from the chi-square test was <0.05.

Question 18
5 points


In 1976 the Nurse's Health Study enrolled 121,700 female registered nurses who were 30 to 55 years old. Subjects completed a mailed questionnaire regarding their medical history and lifestyle. The women have returned follow up information every two years. For the current study women who reported cardiovascular disease and cancer in the baseline questionnaire were excluded, leaving 116,564 women in the analysis of obesity and mortality. Many studies indicate that obesity and lack of exercise are significant risk factors for cardiovascular disease and death, but obesity and exercise are strongly related to one another. The goal of this study was to determine the
extent to which obesity and lack of activity were independent risk factors. The illustration from this article shows a portion of a table summarizing deaths and person-years of observation time among nine categories of BMI.

Compute the incidence rate of death for all nine categories. Then, using the women with BMI<21.0 as the reference group, compute the rate ratios (relative risk) for all nine categories. Which of the following is the best summary of what the rate ratios indicate?

A. There is a progressive decline in mortality rate as BMI increases
B. There is a progressive increase in mortality rate as BMI increases.
C. The mortality rate was similar for the first and second BMI categories but increased progressively after that.
D. The mortality rate was similar for the first and second BMI categories but decreased progressively after that.

Question 19 5 points

The table in the article by Hu et al. summarized the effects of obesity and physical inactivity on mortality after adjusting for a number of other factors, such as age, smoking, drinking, menopausal status, etc. They used lean (BMI<25), active (>3.5 hr/week) women as the reference group. For each of the three BMI categories examine the relative risks and 95% confidence intervals as activity decreases. Which of the following are true? (Select all that apply.)

A. For women with BMI<25 the risk of death increased significantly as their exercise time diminished.
B. Among the women who exercised >3.5 hr/week the risk of death increased significantly as their BMI increased.
C. Lean women (BMI<25) who exercised 1.0-3.4 hr/week had a statistically significant increased risk of death compared to lean women who exercised > 3.5 hr/week.
D. Women with BMI >30 who exercised < 1.0 hr/week had 2.42 times the risk of dying compared to lean women who exercised >3.5 hr/week.

Question 20 5 points

Refer to the table in question 19. Does moderate exercise carry a higher risk of death than high exercise for the women in the leanest group and why?

A. No, because the risk ratio was only 1.18.
B. No, because the risk ratio was within the confidence interval.
C. Yes, because the confidence interval excluded the null value.
D. You can't tell because they didn't report the p-value.