NAME: _____ Student ID #:_____

University of Guelph
Department of Population Medicine

Midterm Examination I – Epidemiology (POPM*3240*01)

Course Coordinator: Dr. Andria Jones-Bitton
February 10, 2018

This exam is out of 37 marks and is worth 25% of the total course grade

EXAM VERSION: **A**

INSTRUCTIONS:

1. Write your name and ID number on the top of the exam AND on the Computer Test Scoring Sheet.

2. Communication with anyone other than the instructors or invigilators during the exam is not permitted.

3. Scientific or desk calculators may be used, but cannot be shared.

4. The use of other electronic devices like computers and cell phones is not permitted. Cell phones must be turned OFF and put away during the exam.

5. All pages of the exam AND the Computer Test Scoring Sheet must be handed in at the end of the exam period.

6. **PART A: MULTIPLE CHOICE QUESTIONS (MCQ)**
   - Fill in the personal information on the Computer Test Scoring sheet
   - Each question is worth one (1) mark
   - There is only one correct answer for each MCQ - choose the most correct answer and,
   - Enter the correct answer for each MCQ on the Computer Test Scoring sheet
   - Use only an HB #2 pencil to fill in the Computer Test Scoring sheet. Make heavy black marks that fill the circle completely. Erase cleanly any answer you wish to change. Make no stray marks on the answer sheet.

7. **PART B: SHORT ANSWER QUESTIONS**
   - Answer ALL questions in the space provided on (the front of) the exam pages ** DO NOT WRITE ON THE BACK OF EXAM PAGES **
   - Please write legibly (we cannot mark things we cannot read)
   - For calculations, show your work. Retain values to 3 decimal places throughout your calculation work, but report end result to one decimal place
   - Answers written in pencil will not be eligible for any re-grade request

NAME: _____ Student ID #:_____

Formulae and Information you may find useful:

$$n = \frac{{Z_\alpha}^2 * \sigma^2}{L^2}$$

$$n = \frac{{Z_\alpha}^2 * p * q}{L^2}$$

$$n = \frac{\left(Z_\alpha \sqrt{2pq} - Z_\beta \sqrt{p_1 q_1 + p_2 q_2}\right)^2}{(p_1 - p_2)^2}$$

$$n = 2 * \left[\frac{\left(Z_\alpha - Z_\beta\right)^2 * \sigma^2}{(\mu_1 - \mu_2)^2}\right]$$

Z-value for when alpha = 5%: 1.96

Z-value for when beta = 20%: -0.84

NAME: _____ Student ID #:_____

PART A: MULTIPLE CHOICE QUESTIONS – 14 marks

The following information is adapted from Public Health Ontario pertains to Questions 1 – 6.

Listeriosis is a serious but uncommon infection caused by the bacterium Listeria monocytogenes. Signs and symptoms of listeriosis include muscle aches, fever, diarrhea, nausea, and vomiting.

1. Which epidemiological term best describes the following statement?
"Signs and symptoms of listeriosis can appear from 3 to 70 days after exposure to the bacteria, with a median time of 3 weeks."

A. Latent period
B. Incubation period
C. Communicability period
D. Shedding period
E. Pathology period

2. Public Health Ontario reports that the annual number of cases of listeriosis for the years 2009 to 2015 ranged between 43 and 65 (average: 53 cases); however, in 2016, 96 cases were reported. Which of the following disease patterns is most consistent with this description?

A. Yearly
B. Sporadic
C. Endemic
D. Epidemic
E. Pandemic

3. You are working for Public Health Ontario, and are helping design a research study to do active surveillance for listeriosis in a high-risk population. You expect the prevalence of listeriosis in this population to be 5%. You want your prevalence estimate to have a precision of +/- 2% and 95% confidence. The minimum required sample size for your study is:

A. 10
B. 91
C. 233
D. 457
E. 4,562

NAME: _____ Student ID #:_____

4. You conduct the active surveillance described above, and you estimate a prevalence of 8% with a 95% confidence interval = (5%, 11%). Which of the following statements best describes this confidence interval?

A. There is a 95% probability that the actual prevalence falls between 5% and 11%
B. There is a 5% probability that the actual prevalence is below 5%
C. There is a 95% probability that the actual prevalence would be lower than 5% or higher than 11%
D. There is a 95% probability that the actual prevalence is 8%
E. More than one of the above is correct

5. You and your research team at Public Health Ontario also want to determine whether people who consume cheeses made from unpasteurized milk have a significantly higher prevalence of listeriosis than people who do not consume such cheese. To calculate an appropriate sample size for this study, which of the following formulae should you use?

A. $$n = \frac{{Z_\alpha}^2 * \sigma^2}{L^2}$$

B. $$n = \frac{{Z_\alpha}^2 * p * q}{L^2}$$

C. $$n = \frac{\left(Z_\alpha\sqrt{2pq} - Z_\beta\sqrt{p_1q_1 + p_2q_2}\right)^2}{(p_1 - p_2)^2}$$

D. $$n = 2 * \left[\frac{\left(Z_\alpha - Z_\beta\right)^2 * \sigma^2}{(\mu_1 - \mu_2)^2}\right]$$

E. None of these formulae should be used

6. You set up a food safety program with Public Health Ontario to test all cheeses and other foods commonly associated with listeriosis to ensure that no Listeria-contaminated food is brought into this high-risk population. Given this objective, and assuming the use of multiple food safety tests in combination, which of the following approaches would be best for you to use?

A. Interpret the tests in series to achieve higher sensitivity
B. Interpret the tests in series to decrease the sensitivity
C. Interpret the test in series to achieve higher specificity
D. Interpret the tests in parallel to achieve a higher sensitivity
E. Interpret the tests in parallel to achieve a higher specificity

7. When epidemiologists judge the evidence to establish possible causes of a health outcome, we consider:

A. The estimated strength of association between an exposure and outcome
B. Evidence that the exposure of interest appeared before the outcome
C. Evidence showing that reductions in the exposure level do not affect the risk of the outcome
D. Two of the above
E. All of the above

8. What is a sufficient cause?

A. Any agent that is required for the development of a given disease
B. A factor that contributes towards disease causation but is not necessary
C. A factor or combination of factors that will inevitably produce disease
D. A component cause that is present in any combination of factors that produce disease
E. None of the above

9. Academic records were used to confirm students' self-reported grade point averages in a research study. By using this validation process instead of just relying on students' provided responses, the researchers are reducing the likelihood of which of the following?

A. Selection bias
B. Information bias
C. Confounding bias
D. Surveillance bias
E. None of the above

10. In a research study, daycares were selected at random, and every child in each daycare was included in the sample. This best describes which type of sampling method?
A. Simple random sampling
B. Systematic random sampling
C. Stratified random sampling
D. Cluster sampling
E. Multistage sampling

11. The local public health department releases a new set of internet memes to help increase the practice of good hand washing in the general public. This represents which level of prevention?

A. Pre-Primary
B. Primary
C. Secondary
D. Tertiary
E. None of the above

12. The use of screening tests to detect diseased individuals has been an integral part of disease control programs. If you are a public health official and your primary concern is to eradicate (i.e. completely eliminate from the population) a specific disease, which of the following would best help to achieve this objective?

A. Selection of a test with a high specificity
B. Selection of a test with a high sensitivity
C. Selection of a test with a low specificity
D. Selection of a test with a low sensitivity
E. Screening tests cannot be used in disease control programs

13. In a study that is designed to compare two groups, the probability of concluding that the outcomes are different, when in fact they are not different, is considered to be:

A. A Type I error
B. A Type II error
C. Confidence of the study
D. Power of the study
E. True state of nature

14. When using 2 tests, classifying an individual as positive if they test positive to one, the other, or both tests, is called "_____ testing", and it increases the _____ of the test:

A. Series testing, sensitivity
B. Series testing, specificity
C. Parallel testing, sensitivity
D. Parallel testing, specificity
E. Long-jump testing, sensitivity

END OF PART A

NAME: _____ Student ID #:_____

PART B: SHORT ANSWER – 23 marks

Question 15 (2 marks)

Very briefly explain the meaning of a "disease carrier". Why are disease carriers particularly important epidemiologically?

A disease carrier is an individual who infected with an infectious disease agent *(0.5 mark)*, but has no signs symptoms *(0.5 mark)*. They are important because they can still act as sources of infection / transmit infections to others *(0.5 mark)*, but because they remain clinically normal, they are more difficult to detect *(0.5 mark)*.

Question 16 (2 marks)

Name one of the core functions of Canada's public health system, and give an example to illustrate.

Any one of the following core functions *(1 mark)*
- *Disease and injury prevention*
- *Health promotion*
- *Health protection*
- *Population Health Assessment*
- *Public Health Surveillance*

And any reasonable example *(1 mark)*

Question 17 (1 mark)

One of Bradford Hill's criteria to consider in establishing causation is temporality ('temporal sequence'). Very briefly, explain this criterion.

To infer causation, one must be able to demonstrate that the suspected risk factor / exposure variable preceded the outcome being investigated *(1 mark)*.

NAME: _____ Student ID #:_____

Question 18 (1 mark)

What does it mean when a researcher states that that their study result is statistically significant?

A statistically significant study result means that the observed result is unlikely to be due to chance *(1 mark)*.

Question 19 (4 marks)

You are reviewing an article that has been submitted to an epidemiological journal for publication. In the article, the authors are suggesting that eating spicy foods is positively associated with longer life expectancy. However, you also know that in the country where the study was conducted, consumption of spicy food is positively associated with household income (i.e. consumption of spicy food increases with increasing household income).

Using the criteria for confounding as a guide, explain why the observed association between spicy food consumption and longer life may be a result of confounding bias.

In this instance, it is likely that the observed association between spicy food consumption and longer life expectancy is due to confounding bias introduced by not adjusting for household income (i.e. that household income is a confounding variable confounding the association between spicy food consumption and life expectancy) *(1 mark for stating the likely confounding bias caused by household income)*. This is because household income:

(1) Is (very likely) associated with life expectancy *(1 mark)*
(2) Is associated with spicy food consumption (as stated in question) *(1 mark)*
(3) Is unlikely a consequence of spicy food consumption *(1 mark)*

As household income meets the 3 criteria for confounding, it is likely a CFV and should be controlled for in the design or analysis phase.

NAME: _____ Student ID #:_____

Question 20 (5 marks)

You are evaluating a new diagnostic test for Leptospirosis (a disease caused by the Leptospira bacterium) in dogs. You and your team use a comprehensive set of diagnostic procedures and techniques and establish a group of 100 dogs infected with Leptospira and a group of 200 dogs that are not infected with Leptospira. You then proceed with your evaluation by applying your new diagnostic test to the dogs in both groups. Your findings are as follows:

Eighty-five of the Leptospira-positive dogs test positive, and 185 of the Leptospira-negative dogs test negative. (Hint: construct a 2x2 table)

	Gold Standard		
New Test	+	-	
+	85	15	100
-	15	185	200
	100	200	300

a. How good is the new test in classifying Leptospira-positive dogs? Name *(1 mark)*, calculate *(1 mark)* and interpret in words *(1 mark)* an appropriate measure to achieve this. Please show your work.

Sensitivity *(1 mark)* = 85 / 100 = 85% *(1 mark)*. This means that the test can correctly classify 85% of truly infected dogs as being Leptospira-positive *(1 mark)*.

1 mark for correct measure (SN), 1 mark for correct calculation, and 1 mark for correct interpretation of their calculation (even if their mathematical calculation was wrong) -- if just gave definition of SN and did not interpret the actual value: 0.5 mark (question asked student to "interpret in words", not just to "define".)

b. A client wants to know the likelihood that her dog is actually free from Leptospira on the basis of her dog's negative test result with the new test. Calculate *(1 mark)* and interpret in words *(1 mark)* the epidemiological term to answer her question. Please show your work.

Negative Predictive Value = 185/200 = 92.5%, which means that 92.5% of new test-negative dogs will actually be non-infected *(or any other correct interpretation)*.

1 mark for correct calculation, 1 mark for correct interpretation of their NPV (even if their mathematical calculation was wrong) -- if just gave definition of NPV and did not interpret the actual value: 0.5 mark (question asked student to "interpret in words the NPV of this test" i.e. in relation to the scenario/question; not just to define NPV.

NAME: _____ Student ID #:_____

Question 21 (8 marks)

Jones, Hagen and Kamleh are practitioners at a highly-specialized (expensive!), private care referral centre located in Toronto. They published a study called: "Imaginosis Disease in People in Ontario: A Referral Centre Survey of the Prevalence of Infection." They asked people who came to their clinic over a 6-month period to participate; 200 people in total, all living in Toronto or the surrounding Greater Toronto Area, agreed. Each participant provided a blood sample, which was analyzed for antibodies to the causative agent of Imaginosis.

(a) Based on the apparent objective as noted in the information provided in the title, what general type of study is this? *(1 mark)*

Descriptive Study *(1 mark)*

(b) What was the apparent target population? *(1 mark)*

Residents of Ontario *(1 mark)*. The title of the study clearly depicts that the authors are attempting to describe / state Imaginosis in "people in Ontario".

(c) What was the source population? *(1 mark)*

Patients who visited the clinic *(1 mark)* over the 6-month period.

(d) What was the study population? *(1 mark)*

Patients who visited the clinic over the 6-month period and agreed to participate *(1 mark)*.

(e) What type of sampling method was used? *(1 mark)*

Convenience sampling *(1 mark)* – the researchers just used the patients who attended their clinic. Note: no cluster level sampling is described here (this would have involved first sampling clinics or sampling at another group-level).

(f) Do you think the study population is representative of the target population? Why? *(1 mark)*

No *(0.5 mark)*, for various reasons *(0.5 mark for <u>one</u> of the following, or anything else reasonable)*: the GTA is unlikely to be representative of all people in Ontario, the highly specialized / expensive nature of the clinic likely means it serves a clientele of high socioeconomic status that is different from the rest of the population, no indication of how patients within the clinic were chosen or what response rate was – non-response may be a further issue.

(g) Based on the sampling described, what major type of bias are you most concerned about? Please explain. *(2 marks)*

NAME: _____ Student ID #:_____

The provided information illustrates a concern with selection bias *(1 mark).* Non-random convenience sampling was used, and the clientele is likely different from the target population indicated in the study title at least in terms of socioeconomic status. Therefore, the people who participated in this study are likely systematically different from those who were potentially eligible for the study (Ontario population) but did not attend the clinic / participate.

1 mark for indicating selection bias, and 1 mark for reasonable explanation that includes the sampling method used (0.5 mark) and the "type" of individual who ended up in the study (0.5 mark).

Any other bias that is reasonable on the basis of the information provided is also acceptable.