NOTE: Answer all questions from section A and four out of five questions from section B.

SECTION A – Multiple Choice Questions (10 marks)

Instructions – Answer this section on the coloured answer sheet provided. Securely attach the answer sheet inside your script book with the tie provided. Answer all questions. Each question is worth 1 mark.

1. The best epidemiological measure of the health status of a population is:
   a. Population attributable risk
   b. Risk difference
   c. Prevalence
   d. Relative risk
   e. Odds ratio

2. The following are all considered part of the Bradford-Hill criteria for causation except one. Which?
   a. Scientific consensus that the exposure causes disease.
   b. Consistency.
   c. Temporality.
   d. Experimental evidence.
   e. Strength of association.
3. Recall bias is **most often** a cause of error in which epidemiological study design?
   a. Cohort studies
   b. Cross-sectional studies
   c. Randomised controlled trials
   d. Case-series.
   e. Case-control studies

4. A 95% confidence interval accounts for what sort of error, when associated with an effect estimate in an epidemiological study?
   a. Information bias
   b. Random (sampling) error
   c. Measurement error
   d. Loss to follow up
   e. Confounding

5. Which one of the following epidemiological measures estimates the proportion of cases that will be prevented once a causal exposure is identified and removed from a population?
   a. Relative risk
   b. Risk difference
   c. Odds ratio
   d. Population proportion attributable risk
   e. Prevalence

6. A randomised-controlled trial:
   a. Has the least ethical constraints of all epidemiological studies.
   b. Is usually carried out retrospectively.
   c. Is usually the cheapest epidemiological study design.
   d. The best study design for assessing causation.
   e. Often inaccurate due to the presence of unmeasured confounders.

7. Which of the following methods is most likely to reduce information bias in a cohort study?
   a. Calibrating equipment and standardising measurement methods used to gather clinical information during the study
   b. Adjusting for confounding during the analysis.
   c. Increasing the sample size.
   d. Checking effect estimates for interaction with other baseline exposures.
   e. Using regression modelling rather than stratification methods.

CONTINUED
8. You are asked by a physician for advice to conduct an epidemiological study to investigate risk factors for multiple myeloma, a type of cancer, which is extremely rare. He is principally interested in investigating the effect of cigarette smoking on the onset of myeloma. Which of the following study designs would be most appropriate, given only this information?
   a. Cohort
   b. Randomised controlled trial
   c. Cross-sectional
   d. Case-control
   e. Case-series.

9. During the analysis of an epidemiological study, if an adjusted (or multivariate) effect estimate differs by more than 10%, compared to the crude (univariate), this is sometimes considered evidence of what type of phenomenon?
   a. Effect modification
   b. Stratification
   c. Confounding
   d. Information bias
   e. Loss to follow up

10. When planning a cohort study, in which the outcome is measured as cumulative incidence, all of the following information are necessary to estimate the sample size, except one. Which?
    a. The expected variation in outcomes measured.
    b. The predetermined probabilities of avoiding type I and II errors.
    c. The population attributable risk of exposure.
    d. The expected proportion of disease in the unexposed.
    e. The proportion of disease in the exposed.
Section B – Short Answers

Instructions – Each question is worth 10 marks
Answer 4 out of 5 questions

11. The evidence linking vitamin D with cardiovascular disease has changed considerably over the last 50 to 60 years.

a. In the 60s and 70s, vitamin D was thought to cause cardiovascular disease. The evidence for this theory initially came from case-series and animal studies, and an assay for measuring vitamin D status in humans had not yet been developed. Outline in a maximum of 5 sentences how such studies could lead to erroneous conclusions.

b. Cardiovascular disease (CVD) event rates increase sharply during winter months. Two explanations were proposed – (1) that reduced sun exposure is the cause, mediated by lower vitamin D levels and (2) that lower temperatures cause the increased risk. Many observational studies show that risk of developing CVD lowers with increasing altitude. How does this information help discriminate between the competing theories? What other explanations are there for these results?

c. A randomised controlled trial is proposed to explore the question of whether vitamin D supplementation reduces the incidence of cardiovascular disease. Which ethical issue associated with such a design will be important to address in planning the study? What information would you seek to resolve this issue?
12. In the late 1980s a case control study was undertaken to determine whether the drug fenoterol was causing an excess of asthma deaths. Cases were derived from a mortality survey of asthma deaths, and controls were ethnicity, age and district matched with cases who had been admitted to hospital with asthma in the past year. Drug exposure information was derived from GP records for cases and hospital records for controls. The study results are printed in the summary table below:

<table>
<thead>
<tr>
<th>Taking Fenoterol?</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>34</td>
<td>76</td>
</tr>
<tr>
<td>No</td>
<td>18</td>
<td>87</td>
</tr>
</tbody>
</table>

a. With the following set of formulae, derive the odds ratio and 95% confidence interval. (Show your working).

Standard error (log_e(OR)) = √[1/a+1/b+1/c+1/d]
Error factor= exp[1.96 x s.e.(log_e(OR))]
95% CI = OR / EF to OR x EF (4 marks)

b. Describe the meaning of the odds ratio and 95% confidence interval in words to a news reporter interested in the study. (2 marks)

c. A representative from the drug company that manufactures fenoterol criticises the study because the exposure is measured differently between cases and controls. His criticism is that fenoterol use is more likely to be recorded in general practice records compared to those entered in hospital.

What effect is this likely to have on the effect measure you reported above? Describe an alternative case-control study design that results in recording of exposure status from the same data source. What is the advantage of such a design? (4 marks)

13.

a. Explain the meaning of a P-value, when used to interpret the effect measure from an observational study. (2 marks)

b. Which factors influence the size of the P-value? (2 marks)

c. When associated with an effect estimate, discuss the meaning of a 95% confidence interval. (2 marks)

d. Compare and contrast the use of P-values and 95% confidence intervals, when reporting epidemiological study results. (4 marks)
14. In the following studies, E=exposure, D=disease, and C= is a third variable. In which of the following circumstances would you expect C to be a confounder? Give reasons for your selection. If you are uncertain, what information would you seek to make up your mind?
   a. E=dental decay; D=cardiovascular disease; C=sugar consumption (3 marks)
   b. E=indoor air pollution; D=tuberculosis; C=socioeconomic status (3 marks)
   c. E=obesity; D=cardiovascular disease; C=diagnosis of diabetes (4 marks)

15. Outline seven of the nine Bradford-Hill criteria for causation. Then, critique the criteria, offering alternative methods which may help distinguish whether an association is causal. (10 marks)

16. You have just conducted a study which investigates the effect of protective helmets on the risk of head injury for skiers and snowboarders, using a case-control design.

<table>
<thead>
<tr>
<th>Helmet use?</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>96</td>
<td>656</td>
</tr>
<tr>
<td>No</td>
<td>480</td>
<td>2330</td>
</tr>
</tbody>
</table>

   a. Calculate the appropriate effect estimate for the study. (4 marks)
      Standard error \((\log_e(OR)) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c + 1/d}}\)
      Error factor= \(\exp[1.96 \times \text{s.e.}(\log_e OR)]\)
      95% CI = \(\frac{OR}{\text{EF}}\) to \(\frac{OR \times \text{EF}}{\text{EF}}\)

   b. A reporter asks you to explain the results of the study, what do you tell him? (2 marks)

   c. Assuming the study sample size was calculated using and alpha level of 0.05, and the sample size from the calculation was finally included in the study. What is the probability of making a type I error, of the effect estimate? Explain what a type I error is. (2 marks)

   d. Outline other possible explanations for the observed effect estimate. Does the 95% confidence interval you calculated in (a) include the effects of such threats to the validity of your study results? (2 marks)
SECTION A - ANSWER SHEET

First Name: ___________________________  Surname: ______________________________

AUID: ____________________________________________________________

Instructions – Make sure you have completed the section above with your name and ID. Select your answer by circling the appropriate letter in the grid below. If you wish to change your answer please carefully erase your previous answer. Tie this answer sheet inside your script book with the tie provided.

<table>
<thead>
<tr>
<th>Question</th>
<th>a.</th>
<th>b.</th>
<th>c.</th>
<th>d.</th>
<th>e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>