## Instructions

Please complete the following lab assignment. You may work on the assignment in groups or on your own. However, to get credit, you must submit your own answers in Canvas. This lab is open note and open book. You may also ask the instructor and the TA questions. Please note that in most cases we will try to guide you towards answering your own question rather than directly providing you with an answer.

**Calculations**

**Q1.**

[Multiple Choice]

In a case-control study, recall bias is most likely to result in what type of misclassification?

|  |  |
| --- | --- |
| ✅ | Differential |
|  | Non-differential |
|  | Neither |

**Q1. Feedback**

In a case-control study, recall bias is most likely to result in differential misclassification?

"Because, on occasion in case-control studies, recall bias may be caused by “rumination” by cases regarding the causes of their disease," it is often assumed to result in differential misclassification.

Szklo, Moyses,Nieto, F. Javier. Epidemiology (Kindle Locations 3580-3581). Jones & Bartlett Learning. Kindle Edition.

Said another way, cases in a case-control study are often assumed to more accurately remember their past exposures than controls because they spend more time thinking about why they become sick (when the outcome is a disease).

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q2.**

[Multiple Choice]

Non-differential misclassification, when there are two exposure (exposed & non-exposed) categories, is expected to typically lead to:

|  |  |
| --- | --- |
|  | An overestimation of the true association |
| ✅ | An underestimation of the true association |
|  | Either an under- or overestimation of the true association |

**Q2. Feedback**

Non-differential misclassification, when there are two exposure (exposed & non-exposed) categories, is expected to typically lead to an underestimation of the true association.

"In this simple situation when there are only two exposure categories (for instance, “yes” or “no”), nondifferential misclassification tends to bias the association toward the null hypothesis."

Szklo, Moyses, Nieto, F. Javier. Epidemiology (Kindle Locations 3702-3703). Jones & Bartlett Learning. Kindle Edition.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q3**

[Numerical Answer]

In a hypothetical retrospective cohort study that is investigating an occupational exposure to an organic solvent that occurred 15-20 years ago in a factory and its relationship to adverse health events of employees. The study was conducted using the employee health records. If all records had been retained the results may have looked like those shown in the contingency table below

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Diseased** | **Non-diseased** | Total |
| **Solvent Exposure** | 40 | 960 | 1,000 |
| **No Solvent Exposure** | 15 | 985 | 1,000 |
| Total | 55 | 1,945 | 2,000 |

Using the table provided, calculate the risk ratio. Round your answer to 3 decimal places.

|  |  |
| --- | --- |
| ✅ | 2.667 |

**Q3. Feedback**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Diseased** | **Non-diseased** | Total |
| **Solvent Exposure** | 40 | 960 | 1,000 |
| **No Solvent Exposure** | 15 | 985 | 1,000 |
| Total | 55 | 1,945 | 2,000 |

Risk in Exposed = 40 / 1,000 = 0.04

Risk in Unexposed = 15 / 1,000 = 0.015

RR = 0.04 / 0.015 = 2.667

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q4**

[Fill in Multiple Blanks]

Suppose that many of the old records had been lost or discarded, but given the suspicion about the effects of the solvent, the records of employees who had worked with the solvent and subsequently had health problems were more likely to be retained. Record retention was **95%** among workers who were **exposed and developed health problems**, but record retention was only **80**% for **all other worker**s. Create a new contingency table to match the data. Assume no confounding effects and no information bias.

*For this answer, round to the nearest integer, if necessary.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Disease | Non-diseased | Total |
| Solvent Exposure | [38] | [768] | 806 |
| No Solvent Exposure | 12 | [788] | [800] |
| Total | [50] | 1556 | 1606 |

|  |  |
| --- | --- |
| ✅ | 38 |
| ✅ | 768 |
| ✅ | 788 |
| ✅ | 800 |
| ✅ | 50 |

**Q4. Feedback**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Diseased** | **Non-diseased** | Total |
| **Solvent Exposure** | 40 \* 0.95 = 38 | 960 \* 0.8 = 768 | 806 |
| **No Solvent Exposure** | 15 \* 0.8 = 12 | 985 \* 0.8 = 788 | 800 |
| Total | 50 | 1,556 | 1,606 |

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q5**

[Numerical Answer]

Using the data from the new contingency table from the last question, calculate the biased risk ratio. Please round your answer to 3 decimal places at each step.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Diseased** | **Non-diseased** | Total |
| **Solvent Exposure** | 38 | 768 | 806 |
| **No Solvent Exposure** | 12 | 788 | 800 |
| Total | 50 | 1,556 | 1,606 |

|  |  |
| --- | --- |
| ✅ | 3.133 |

**Q5. Feedback**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Diseased** | **Non-diseased** | Total |
| **Solvent Exposure** | 38 | 768 | 806 |
| **No Solvent Exposure** | 12 | 788 | 800 |
| Total | 50 | 1,556 | 1,606 |

Risk in Exposed = 38 / 806 = 0.047

Risk in Unexposed = 12 / 800 = 0.015

RR = 0.047 / 0.015 = 3.133

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q6**

[Multiple Dropdowns]

Compare the true RR (2.667) with the biased RR (3.133). The [Type] loss of records resulted in selection bias and an [Direction] of the association.

Type

|  |  |
| --- | --- |
| ✅ | Differential |
|  | Non-differential |

Direction

|  |  |
| --- | --- |
| ✅ | Overestimation |
|  | Underestimation |

**Q6. Feedback**

Compare the true RR (2.667) with the biased RR (3.133). The biased RR is larger than the true RR.  Differential loss of records resulted in selection bias and an overestimation of the association.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q7**

[Multiple Choice]

What potential bias could have been introduced if you found out that those who interviewed cases took 30 minutes longer on average than those who interviewed controls?

|  |  |
| --- | --- |
|  | Loss to follow-up |
|  | Volunteer bias |
|  | Selection bias |
| ✅ | Interviewer bias |

**Q7. Feedback**

What potential bias could have been introduced if you found out that those who interviewed cases took 30 minutes longer on average than those who interviewed controls?

"When data collection in a case-control study is not masked with regard to the disease status of study participants, observer bias in ascertaining exposure, such as **interviewer bias**, may occur. Interviewer bias may be a consequence of trying to “clarify” questions when such clarifications are not part of the study protocol and failing to follow either the protocol-determined probing or skipping rules of questionnaires."

Szklo, Moyses, Nieto, F. Javier. Epidemiology (Kindle Locations 3620-3622). Jones & Bartlett Learning. Kindle Edition.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q8**

[Multiple Choice]

Researchers conducted a prospective cohort study of the association between air pollution exposure and asthma. Some study participants were lost to follow-up (dropped out of the study) over time. The researchers were able to obtain data on the exposure and the health outcome for participants who remained in the study as well as for participants who dropped out of the study. The researchers discovered that the rate of loss to follow-up did not differ when comparing exposed and unexposed groups. The researchers also found that the rate of loss to follow-up did not differ when comparing people who developed asthma and people who did not develop asthma. Based on this information, which one of the following statements is most likely to be true?

|  |  |
| --- | --- |
|  | Selection bias likely did not occur in this study because people cannot choose if they are exposed to air pollution or not exposed to air pollution |
|  | Selection bias likely occurred in this study because both of the outcome groups (people with asthma and people without asthma) experienced loss to follow-up |
| ✅ | Selection bias likely did not occur in this study because exposure status and health outcome status did not influence whether or not people dropped out of the study |
|  | Selection bias likely occurred in this study because both exposure groups experienced loss to follow-up |

**Q8. Feedback**

Researchers conducted a prospective cohort study of the association between air pollution exposure and asthma. Some study participants were lost to follow-up (dropped out of the study) over time. The researchers were able to obtain data on the exposure and the health outcome for participants who remained in the study as well as for participants who dropped out of the study. The researchers discovered that the rate of loss to follow-up did not differ when comparing exposed and unexposed groups. The researchers also found that the rate of loss to follow-up did not differ when comparing people who developed asthma and people who did not develop asthma. Based on this information, which one of the following statements is most likely to be true?

Selection bias likely did not occur in this study because exposure status and health outcome status did not influence whether or not people dropped out of the study.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q9**

[Multiple Choice]

During the sample selection for this study, if researcher A misclassified the exposure of both the cases and controls by a sensitivity of 0.8 and a specificity of 0.8, what type of misclassification would possibly occur?

|  |  |
| --- | --- |
|  | Differential misclassification |
| ✅ | Non-differential misclassification |

**Q9. Feedback**

During the sample selection for this study, if researcher A misclassified the exposure of both the cases and controls by a sensitivity of 0.8 and a specificity of 0.8, what type of misclassification would possibly occur?

In this case, we have non-differential misclassification of the exposure. The sensitivity and specificity of the exposure are not dependent on outcome status. Said another way, there is equal misclassification of the exposure between participants with the outcome and participants without the outcome.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q10**

[Multiple Choice]

During the sample selection for this, if the researcher B misclassified the exposure of cases by a sensitivity of 0.4 and a specificity of 0.6, and the exposure of controls by a sensitivity of 0.6 and a specificity of 0.8. What type of misclassification would possibly occur?

|  |  |
| --- | --- |
| ✅ | Differential misclassification |
|  | Non-differential misclassification |

**Q10. Feedback**

During the sample selection for this, if the researcher B misclassified the exposure of cases by a sensitivity of 0.4 and a specificity of 0.6, and the exposure of controls by a sensitivity of 0.6 and a specificity of 0.8. What type of misclassification would possibly occur?

In this case, we have differential misclassification of the exposure. The sensitivity and specificity of the exposure are dependent on outcome status. Said another way, there is unequal misclassification of the exposure between participants with the outcome and participants without the outcome.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q11**

[Multiple Choice]

A cohort study was conducted among 2000 subjects ages 15 – 75, to examine if age was a risk factor for cardiovascular disease. The cohort was categorized into 2 groups (<40 yrs and 40 yrs and above) of equal number. The participants were followed for 10 years. A total of 400 subjects dropped out of the study for several reasons before the study ended. Out of the 400 that dropped out, 300 of them were over 50 years of age.

**What type of bias will most likely affect the result of this study?**

|  |  |
| --- | --- |
|  | Information bias |
| ✅ | Selection bias |
|  | Recall bias |
|  | The study is not biased. |

**Q11. Feedback**

A cohort study was conducted among 2000 subjects ages 15 – 75, to examine if age was a risk factor for cardiovascular disease. The cohort was categorized into 2 groups (<40 yrs and 40 yrs and above) of equal number. The participants were followed for 10 years. A total of 400 subjects dropped out of the study for several reasons before the study ended. Out of the 400 that dropped out, 300 of them were over 50 years of age.

What type of bias will most likely affect the result of this study?

The correct answer is selection bias. The age of the study participants is associated with the outcome of the study. Since the dropouts are older, those left in the study are less likely to have the outcome in comparison to those that dropped out of the study. This biases the original selection (exposed-unexposed ratio) and the true estimate between the exposure of interest and the outcome.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q12**

[Multiple Choice]

Even if the investigators are careful in the selection of cases and controls,  selection bias can make interpretation of results difficult. Which of the following is NOT a situation that can produce selection bias? Choose one best answer.

|  |  |
| --- | --- |
|  | The exposure has some influence on the process by which controls are selected |
|  | The exposure has some influence on the process of case ascertainment. |
|  | The exposed cases are reported to registries more than unexposed. |
| ✅ | All of the above will produce selection bias |

**Q12. Feedback**

Even if the investigators are careful in the selection of cases and controls,  selection bias can make interpretation of results difficult. Which of the following is NOT a situation that can produce selection bias? *Choose one best answer.*

The best answer is that all of the above will produce selection bias.

Please make sure you understand why this is the correct answer. You may use the "Previous" button below to update your answer if your original answer was incorrect. Click the "Next" button below to move on to the next question.

**Q13. Optional Feedback**

[Essay Question]

**Optional**: Please feel free to leave any comments below about the usefulness of this lab. Which parts were helpful? What could I do to improve it? What is still unclear?