Epi 3 Lab – Bias 2022

Each week there will be an in-class lab assignment. It will be an opportunity for students to practice applying that weeks’ concepts with the help of the instructor and TA.

Table of Contents

[Word document usage 2](#_Toc115772378)

[Formatting notes 2](#_Toc115772379)

[Future question ideas: 2](#_Toc115772380)

[Quiz Instructions 3](#_Toc115772381)

[Overview 3](#_Toc115772382)

[Task 1 3](#_Toc115772383)

[Q1. Short Description – Multiple Choice 3](#_Toc115772384)

[Q1. Feedback 3](#_Toc115772385)

[Q2. Short Description – True/False 4](#_Toc115772386)

[Q2. Feedback 4](#_Toc115772387)

[Q3. Short Description – Fill In the Blank 4](#_Toc115772388)

[Q3. Feedback 5](#_Toc115772389)

[Q4. Short Description – Fill In Multiple Blanks 5](#_Toc115772390)

[Q4. Feedback 5](#_Toc115772391)

[Q5. Short Description – Multiple Answers 5](#_Toc115772392)

[Q5. Feedback 6](#_Toc115772393)

[Q6. Short description – Multiple Dropdowns 6](#_Toc115772394)

[Q6. Feedback 7](#_Toc115772395)

[Q7. Short Description - Matching 7](#_Toc115772396)

[Q7. Feedback 8](#_Toc115772397)

[Q8. Short Description – Numerical Answer 8](#_Toc115772398)

[Q8. Feedback 8](#_Toc115772399)

[Q9. Short Description – Formula Question 8](#_Toc115772400)

[Q9. Feedback 9](#_Toc115772401)

[Q10. Short Description – Essay Question 9](#_Toc115772402)

[Q10. Feedback 9](#_Toc115772403)

[Q11. Short Description – File Upload Question 9](#_Toc115772404)

[Q11. Feedback 9](#_Toc115772405)

[Optional Feedback 9](#_Toc115772406)

# Word document usage

*This document should only contain the questions that go on Canvas, and it should not be given to students as a link*. *Each lab will also have a Word version that doesn’t contain questions and answers, which is to be shared with students*. Why put all the questions into a Word document instead of only typing them directly into Canvas? First, who knows if we will always use Canvas? These Word documents should be a good backup even if we switch to a different LMS. Second, we can add comments and questions in progress to Word documents, which isn’t really possible in Canvas.

# Formatting notes

* Code should be formatted as code in the question text, the answer choices, and the question feedback.
  + In canvas, do so by highlighting the code text, click format, click code.
  + In Word, do so by highlighting the code text, change the font to Courier New, and highlight in light gray.
* R package names should be formatted as code. When functions are discussed, they should be formatted as code and empty parentheses added to the end of the function name. For example, “The tibble() function can be used to create data frames…”…” (There should be a “Code” style available in the styles pane)
* Data set/frame names, variable names, and other key words should be bolded. For example, “Please download **chocolate study 1.xls**”.

# Future question ideas:

* Placeholder

# Quiz Instructions

You may work on the lab in groups or on your own. However, you must submit your own answers in Canvas to get credit.

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.

You may complete this lab as many times as you like.

**Links**:

* Click here to view the lab as a Word document (you will still need to submit answers in Canvas).
* Click here to view my R code for this lab as an HTML document.
* Click here to download my R code for this lab as an R markdown document.

# Q1. Type of Misclassification

[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

"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. Non-differential misclassification consequence

[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. Incidence Proportion Ratio calculation

[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 reltionship 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.

# Task 1

[Text (no question]

Load packages: dplyr, freqtables, fmsb

# Task 2

[Text (no question]

Create 2X2 matrix in R: Input values into the matrix. Add margin totals. Add column and row labels. View the table.

# Task 3

[Text (no question]

Calculate incident proportion for each row. Add the column of incidence proportions to the matrix. View the matrix.

# Task 4

[Text (no question]

Calculate the incidence proportion ratio (AKA risk ratio, relative risk).

# Task 5

[Text (no question]

Calculate the incidence proportion ratio (AKA risk ratio, relative risk) using the fmsb package.

**Answer:**

|  |  |
| --- | --- |
| ✅ | 2.667 |

# Q3. Feedback

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.

# Q3.1. Incidence Proportion Ratio CI

[Fill in Multiple Blanks]

1. Using the fmsb package, what is the lower 95% confidence limit for the risk ratio? Round your answer to 3 decimal places.

LCL = [1.483]

1. What is the upper 95% confidence limit for the risk ratio? Round your answer to 3 decimal places.

UCL = [4.796]

**Answer:**

|  |  |
| --- | --- |
| ✅ | 1.483 |
| ✅ | 4.796 |

# Q3.1. Feedback

Output listed under "95 percent confidence interval" in fmsb output.

1. LCL = 1.483

2. UCL = 4.796

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. Biased 2X2 table

[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 |

# Task 6

[Text (no question]

Calculate the solution in R.

**Answer:**

|  |  |
| --- | --- |
| ✅ | 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. Biased Incidence Proportion Ratio calculation

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 |

# Task 7

[Text (no question]

Calculate the solution in R.

**Answer:**

|  |  |
| --- | --- |
| ✅ | 3.143 |

# Q5. Feedback

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. Bias consequences

[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

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. Type of Bias

[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

"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. Selection bias

[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

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. Misclassification

[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

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. Misclassification

[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

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. Type of Bias

[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

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. Selection Bias

[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

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?

# Overview

[Text (no question]

Provide a brief overview about the content of the lab.

# Task 1

[Text (no question]

Enter the tasks from the Word version of the lab that doesn’t contain questions and answers. You may combine multiple tasks into a single question block.

# Q1. Short Description – Multiple Choice

[Multiple Choice]

Question text

|  |  |
| --- | --- |
| ✅ | Correct answer |
|  | Distractor |
|  | Distractor |
|  | Distractor |

**Feedback:**

Explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q1. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q2. Short Description – True/False

[True/False]

Question text

|  |  |
| --- | --- |
| ✅ | True |
|  | False |

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q2. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q3. Short Description – Fill In the Blank

[Fill In the Blank]

For the question text, you must manually put in the blank using underscores (\_\_\_\_\_). Then, the answer choices work in the same way that they would for multiple choice.

|  |  |
| --- | --- |
| ✅ | Correct answer |
|  | Distractor |
|  | Distractor |
|  | Distractor |

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q3. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q4. Short Description – Fill In Multiple Blanks

[Fill In Multiple Blanks]

In the question text box, type a reference word (no spaces) surrounded by brackets in every place you want to show an answer box to the student. For example, "Roses are [color1], violets are [color2]". Then, for each word in brackets, you will add all possible correct answers. Make sure to think about misspellings and other minor errors that you would still count as correct.

For color1

|  |
| --- |
| Red |
| red |

For color2

|  |
| --- |
| Blue |
| blue |

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q4. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q5. Short Description – Multiple Answers

[Multiple Answers]

Question text

|  |  |
| --- | --- |
| ✅ | Correct answer |
| ✅ | Correct answer |
|  | Distractor |
|  | Distractor |

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q5. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q6. Short description – Multiple Dropdowns

[Multiple Dropdowns]

In the question text box, type a reference word (no spaces) surrounded by brackets in every place you want to show an answer box to the student. For example, "Roses are [color1], violets are [color2]". Then, for each word in brackets, you will add all possible correct answers. Make sure to think about misspellings and other minor errors that you would still count as correct.

For color1

|  |  |
| --- | --- |
| ✅ | red |
|  | Distractor |
|  | Distractor |
|  | Distractor |

For color2

|  |  |
| --- | --- |
| ✅ | blue |
|  | Distractor |
|  | Distractor |
|  | Distractor |

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q6. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q7. Short Description - Matching

[Matching]

Question text

|  |  |
| --- | --- |
| Matching left side | Matching right side |
| Matching left side | Matching right side |
| Matching left side | Matching right side |
| Matching left side | Matching right side |
|  | Optional distractor |
|  | Optional distractor |
|  | Optional distractor |

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q7. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q8. Short Description – Numerical Answer

[Numerical Answer]

Question text. If the answer can have a fractional part, it’s typically a good idea to tell students to round to a particular level of precision.

**Answer**: number

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q8. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q9. Short Description – Formula Question

[Formula Question]

Enter your question, build a formula, and generate a set of possible answer combinations. Students will see the question with a randomly selected set of variables filled in and have to type the correct numerical answer. You can define variables by typing variable names surrounded by brackets (i.e. "What is 5 plus [x]?")

**Answer**: number

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q9. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q10. Short Description – Essay Question

[Essay Question]

Enter the essay prompt in the text box. These questions must be graded manually.

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q10. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Q11. Short Description – File Upload Question

[File Upload Question]

Give the students file upload instructions in the text box. These questions must be graded manually. We sometimes use these for HTML notebook files.

**Feedback:**

Explain why the correct answer was correct. Explain why the incorrect answers were incorrect. Give citations if possible.

# Q11. Feedback

[Text (no question]

After each question, there should be a “Text (no question)” question block that gives students the correct answer to the previous question. It should explain why the correct answer was correct, explain why the incorrect answers were incorrect, and give citations if possible.

# Optional Feedback

**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?