## 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: Frequency matching**

[Multiple Choice]

A group of researchers wants to investigate the association between exhaust fumes (exposure) and asthma (outcome). They enrolled 1,000 participants who had asthma into the study. They want to frequency-match controls to the cases on a 4 to 1 basis (i.e. 4 controls to 1 case) by race-ethnicity. The following table presents the race-ethnicity distribution of the cases.

|  |  |
| --- | --- |
| **Race and Ethnicity** | **Cases** |
| Non-Hispanic White | 450 |
| Non-Hispanic Black | 225 |
| Hispanic | 250 |
| Others | 75 |

How many Non-Hispanic White controls can they enroll in the study?

|  |  |
| --- | --- |
| ✅ | 1,800 |
|  | 900 |
|  | 1,000 |
|  | 300 |

**Q1: Feedback**

The correct answer choice is 1,800.

First, we obtain the percentage of each of the race-ethnicity categories.

|  |  |  |
| --- | --- | --- |
| **Race and Ethnicity** | **Cases** | **Proportion** |
| Non-Hispanic White | 450 | 0.45 |
| Non-Hispanic Black | 225 | 0.225 |
| Hispanic | 250 | 0.250 |
| Others | 75 | 0.075 |

Then we apply that to the total number of expected controls (4,000) given the matching ratio (4:1). The obtained table is shown below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Race and Ethnicity** | **Cases** | **Proportion** | **Controls** |
| Non-Hispanic White | 450 | 0.45 | 1,800 |
| Non-Hispanic Black | 225 | 0.225 | 900 |
| Hispanic | 250 | 0.250 | 1,000 |
| Others | 75 | 0.075 | 300 |

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: Frequency matching**

[Multiple Choice]

An investigator decided to perform frequency matching in a case-control study to investigate the association between maternal pesticide exposure and pediatric leukemia. Pediatric leukemia includes patients age 1 to 20 years old.

The ratio of controls to cases is 3 to 1, and the total number of cases was 2,000. The case population was 41% non-Hispanic white, 38% Hispanic, 5% non-Hispanic black, and remaining were categorized as other.

What is the maximum number of non-Hispanic white controls that the investigators can recruit for the study?

|  |  |
| --- | --- |
|  | 300 |
|  | 960 |
|  | 2,280 |
| ✅ | 2,460 |

**Q2: Feedback**

The ratio of cases to controls is 3 to 1. There are 2000 cases. Therefore, there will be 3 \* 2,000 = 6,000 controls.

In frequency matching, the proportion of cases in each level of the matching variable should be identical to the proportion of controls in each level of the matching variable.

Therefore, the number of controls who are non-Hispanic white should be 6,000 \* 0.41 = 2,460.

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: Shellfish OR**

[Multiple Choice]

A state epidemiologist wanted to investigate stomach cancer (outcome) and shellfish consumption (exposure).

She conducted a study with 125 stomach cancer cases and 125 controls, pair-matched on age and gender.  Shellfish consumption was dichotomized as more than 3 times per week (exposed) vs. less than once per month (unexposed).  After conducting the study, it was found that there were 35 pairs in which both cases and controls ate shellfish more than 3 times per week (exposed), and 15 pairs in which both cases and controls ate shellfish less than once per month (unexposed).  There were 28 pairs where the control ate shellfish more than 3 times per week and the case rarely at shellfish and 47 pairs where the case at shellfish more than 3 times per week and the control at shellfish rarely.

Calculate the ratio of the odds of exposure to eating shellfish regularly versus to eating shellfish rarely among participants with and without stomach cancer.

|  |  |
| --- | --- |
|  | 1.2 |
| ✅ | 1.7 |
|  | 2.1 |
|  | 3.2 |

**Q3: Feedback**

The correct answer is 1.7

|  |  |  |
| --- | --- | --- |
|  | **Controls: Exposed** | **Controls: Unexposed** |
| **Cases: Exposed** | 35 | 47 |
| **Cases: Unexposed** | 28 | 15 |

To calculate the matched OR, divide case-exposed discordant pairs over control-exposed discordant pairs:

b / c = 47 / 28 = 1.7

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: Shellfish 95% CI**

[Multiple Choice]

A state epidemiologist wanted to investigate stomach cancer (outcome) and shellfish consumption (exposure).

She conducted a study with 125 stomach cancer cases and 125 controls, pair-matched on age and gender.  Shellfish consumption was dichotomized as more than 3 times per week (exposed) vs. less than once per month (unexposed).  After conducting the study, it was found that there were 35 pairs in which both cases and controls ate shellfish more than 3 times per week (exposed), and 15 pairs in which both cases and controls ate shellfish less than once per month (unexposed).  There were 28 pairs where the control ate shellfish more than 3 times per week and the case rarely at shellfish and 47 pairs where the case at shellfish more than 3 times per week and the control at shellfish rarely.

Calculate the 95% CI of the OR (1.7) you calculated in the previous question.

|  |  |
| --- | --- |
|  | (0.0467, 0.8954) |
|  | (0.0627, 0.9985) |
|  | (0.2387, 0.3487) |
| ✅ | (1.064, 2.714) |

**Q4: Feedback**

The correct answer is (1.064, 2.714).

SE = sqrt (1/47+1/28)

SE = sqrt(0.057)

SE = 0.2387

95%CI = ln(1.7) +/- (1.96\*0.2387)

95%CI = (0.5306-0.4679, 0.5306+0.4679)

95%CI = (0.0627, 0.9985)

Exp(0.0627, 0.9985) = (1.064, 2.714)

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: Shellfish chi-square**

[Multiple Choice]

A state epidemiologist wanted to investigate stomach cancer (outcome) and shellfish consumption (exposure).

She conducted a study with 125 stomach cancer cases and 125 controls, pair-matched on age and gender.  Shellfish consumption was dichotomized as more than 3 times per week (exposed) vs. less than once per month (unexposed).  After conducting the study, it was found that there were 35 pairs in which both cases and controls ate shellfish more than 3 times per week (exposed), and 15 pairs in which both cases and controls ate shellfish less than once per month (unexposed).  There were 28 pairs where the control ate shellfish more than 3 times per week and the case rarely at shellfish and 47 pairs where the case at shellfish more than 3 times per week and the control at shellfish rarely.

Calculate the chi-square statistic for the matched paired data provided.

|  |  |
| --- | --- |
|  | 0.26 |
|  | 1.25 |
|  | 2.65 |
| ✅ | 4.32 |

**Q5: Feedback**

The correct answer is: 4.32

McNemar's Chi-square =  Diagram

Description automatically generated with medium confidence

McNemar's Chi-square = Diagram

Description automatically generated with medium confidence

McNemar's Chi-square = A picture containing text

Description automatically generated

McNemar's Chi-square = Text

Description automatically generated with medium confidence

McNemar's Chi-square = A picture containing text

Description automatically generated

McNemar's Chi-square = 4.32

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: Shellfish p-value**

[Multiple Choice]

A state epidemiologist wanted to investigate stomach cancer (outcome) and shellfish consumption (exposure).

She conducted a study with 125 stomach cancer cases and 125 controls, pair-matched on age and gender.  Shellfish consumption was dichotomized as more than 3 times per week (exposed) vs. less than once per month (unexposed).  After conducting the study, it was found that there were 35 pairs in which both cases and controls ate shellfish more than 3 times per week (exposed), and 15 pairs in which both cases and controls ate shellfish less than once per month (unexposed).  There were 28 pairs where the control ate shellfish more than 3 times per week and the case rarely at shellfish and 47 pairs where the case at shellfish more than 3 times per week and the control at shellfish rarely.

Compare the value of the chi-squared statistic obtained in the previous question (4.32) to the appropriate critical value (significance level of 0.05) on the chi-squared distribution table (Google it). What is the appropriate conclusion?

|  |  |
| --- | --- |
|  | There is no statistically significant association between stomach cancer and shellfish consumption. |
| ✅ | There is a statistically significant association between stomach cancer and shellfish consumption. |
|  | There is insufficient information to develop any conclusion. |

**Q6: Feedback**

The correct answer is: There is a statistically significant association between stomach cancer and shellfish consumption.

Since 4.32 (our test statistic) > 3.84 (critical value from the chi-square distribution)

We reject the null hypothesis and conclude that our data are consistent with a statistically significant association between stomach cancer and shellfish consumption.

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: Matched OR**

[Multiple Choice]

A case-control study examined the association between smoking (exposure) and lung cancer (outcome).

The controls were matched to the cases by gender on a 1 to 1 basis. The matched paired data obtained from the study are shown below.

|  |  |  |
| --- | --- | --- |
|  | Controls Exposed  (smoking) | Controls Unexposed  (No smoking) |
| Cases Exposed  (smoking) | 64 | 72 |
| Cases Unexposed  (no smoking) | 54 | 108 |

Calculate the odds ratio of the matched study

|  |  |
| --- | --- |
|  | 1.778 |
| ✅ | 1.333 |
|  | 0.582 |
|  | 2.624 |

**Q7: Feedback**

A case-control study examined the association between smoking (exposure) and lung cancer (outcome).

The controls were matched to the cases by gender on a 1 to 1 basis. The matched paired data obtained from the study are shown below.

|  |  |  |
| --- | --- | --- |
|  | Controls Exposed  (smoking) | Controls Unexposed  (No smoking) |
| Cases Exposed  (smoking) | 64 | 72 |
| Cases Unexposed  (no smoking) | 54 | 108 |

Calculate the odds ratio of the matched study

The correct answer is 1.333.

The discordant pairs (b and c) are the only values used for calculating OR for a matched study.

OR = B/C = 72/54 = 1.33

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: Interpret matched OR**

[Multiple Choice]

A case-control study examined the association between smoking (exposure) and lung cancer (outcome). The controls were matched to the cases by gender on a 1 to 1 basis. The matched paired data obtained from the study are shown below.

|  |  |  |
| --- | --- | --- |
|  | Controls Exposed  (smoking) | Controls Unexposed  (No smoking) |
| Cases Exposed  (smoking) | 64 | 72 |
| Cases Unexposed  (no smoking) | 54 | 108 |

Interpret the odds ratio obtained as an exposure odds ratio

|  |  |
| --- | --- |
| ✅ | The odds of smoking among people with lung cancer is 1.333 times that of people without lung cancer after adjusting for gender. |
|  | The risk of smoking among people with lung cancer is 1.333 times that of people without lung cancer after adjusting for gender. |
|  | The odds of smoking among people with lung cancer is 133 percent higher than that of people without lung cancer after adjusting for gender. |
|  | None of the above |

**Q8: Feedback**

A case-control study examined the association between smoking (exposure) and lung cancer (outcome).

The controls were matched to the cases by gender on a 1 to 1 basis. The matched paired data obtained from the study are shown below.

|  |  |  |
| --- | --- | --- |
|  | Controls Exposed  (smoking) | Controls Unexposed  (No smoking) |
| Cases Exposed  (smoking) | 64 | 72 |
| Cases Unexposed  (no smoking) | 54 | 108 |

Interpret the odds ratio obtained as an exposure odds ratio

The correct answer choice is: The odds of smoking among people with lung cancer is 1.333 times that of people without lung cancer or 33% percent higher.

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: Calculate matched OR**

[Multiple Choice]

A matched case-control study was conducted to examine the association between residential levels of benzene (exposure) and acute myeloid leukemia (AML) (outcome. The controls were matched to the cases based on butadiene exposure on a 1 to 1 basis. Both the exposure (benzene) and the matched variable (butadiene) were dichotomized as ‘High’ or ‘Low’. Among the matched pairs, there were 45 in which both the case and the control were categorized as ‘High’ for benzene exposure; 25 pairs where the case was categorized as ‘high’ benzene exposure and the control as ‘low’ benzene exposure; 21 pairs where the control was categorized as ‘high’ benzene exposure and the case was categorized as ‘low’ benzene exposure; 20 where the case and the control were categorized as ‘low’ benzene exposure.

Calculate the odds ratio association. Please round to 3 decimal places.

|  |  |
| --- | --- |
|  | 0.583 |
|  | 0.857 |
| ✅ | 1.190 |
|  | 1.219 |

**Q9: Feedback**

A matched case-control study was conducted to examine the association between residential levels of benzene (exposure) and acute myeloid leukemia (AML) (outcome. The controls were matched to the cases based on butadiene exposure on a 1 to 1 basis. Both the exposure (benzene) and the matched variable (butadiene) were dichotomized as ‘High’ or ‘Low’. Among the matched pairs, there were 45 in which both the case and the control were categorized as ‘High’ for benzene exposure; 25 pairs where the case was categorized as ‘high’ benzene exposure and the control as ‘low’ benzene exposure; 21 pairs where the control was categorized as ‘high’ benzene exposure and the case was categorized as ‘low’ benzene exposure; 20 where the case and the control were categorized as ‘low’ benzene exposure.

Calculate the odds ratio association. Please round to 3 decimal places.

The correct answer choice is 1.190. The table obtained is shown below

|  |  |  |
| --- | --- | --- |
|  | Exposed Control  (High Benzene) | Unexposed control  (Low Benzene) |
| Exposed case  (High Benzene) | 45 | 25 |
| Unexposed case  (Low Benzene) | 21 | 20 |

OR = B/C =25/21 = 1.190

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: Calculate matched CI**

[Multiple Choice]

A matched case-control study was conducted to examine the association between residential levels of benzene (exposure) and acute myeloid leukemia (AML) (outcome. The controls were matched to the cases based on butadiene exposure on a 1 to 1 basis. Both the exposure (benzene) and the matched variable (butadiene) were dichotomized as ‘High’ or ‘Low’. Among the matched pairs, there were 45 in which both the case and the control were categorized as ‘High’ for benzene exposure; 25 pairs where the case was categorized as ‘high’ benzene exposure and the control as ‘low’ benzene exposure; 21 pairs where the control was categorized as ‘high’ benzene exposure and the case was categorized as ‘low’ benzene exposure; 20 where the case and the control were categorized as ‘low’ benzene exposure.

Calculate the 95% confidence interval for the odds ratio. Please round to 3 decimal places at each step.

|  |  |
| --- | --- |
|  | (0.866, 1.572) |
|  | (0.433, 0.733) |
|  | (1.008, 1.430) |
| ✅ | (0.666, 2.132) |

**Q10: Feedback**

A matched case-control study was conducted to examine the association between residential levels of benzene (exposure) and acute myeloid leukemia (AML) (outcome. The controls were matched to the cases based on butadiene exposure on a 1 to 1 basis. Both the exposure (benzene) and the matched variable (butadiene) were dichotomized as ‘High’ or ‘Low’. Among the matched pairs, there were 45 in which both the case and the control were categorized as ‘High’ for benzene exposure; 25 pairs where the case was categorized as ‘high’ benzene exposure and the control as ‘low’ benzene exposure; 21 pairs where the control was categorized as ‘high’ benzene exposure and the case was categorized as ‘low’ benzene exposure; 20 where the case and the control were categorized as ‘low’ benzene exposure.

Calculate the 95% confidence interval for the odds ratio. Please round to 3 decimal places at each step.

The correct answer is (0.666, 2.132).

|  |  |  |
| --- | --- | --- |
|  | Exposed Control  (High Benzene) | Unexposed control  (Low Benzene) |
| Exposed case  (High Benzene) | 45 | 25 |
| Unexposed case  (Low Benzene) | 21 | 20 |

Text, letter

Description automatically generated

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: Calculate chi-square**

[Multiple Choice]

A matched case-control study was conducted to examine the association between residential levels of benzene (exposure) and acute myeloid leukemia (AML) (outcome. The controls were matched to the cases based on butadiene exposure on a 1 to 1 basis. Both the exposure (benzene) and the matched variable (butadiene) were dichotomized as ‘High’ or ‘Low’. Among the matched pairs, there were 45 in which both the case and the control were categorized as ‘High’ for benzene exposure; 25 pairs where the case was categorized as ‘high’ benzene exposure and the control as ‘low’ benzene exposure; 21 pairs where the control was categorized as ‘high’ benzene exposure and the case was categorized as ‘low’ benzene exposure; 20 where the case and the control were categorized as ‘low’ benzene exposure.

Calculate the 95% confidence interval for the odds ratio. Please round to 3 decimal places at each step.

Calculate the Chi-Square Statistic for the matched pair data. Please round to 3 decimal places.

|  |  |
| --- | --- |
| ✅ | 0.196 |
|  | 0.235 |
|  | 1.434 |
|  | 4.320 |

**Q11: Feedback**

A matched case-control study was conducted to examine the association between residential levels of benzene (exposure) and acute myeloid leukemia (AML) (outcome. The controls were matched to the cases based on butadiene exposure on a 1 to 1 basis. Both the exposure (benzene) and the matched variable (butadiene) were dichotomized as ‘High’ or ‘Low’. Among the matched pairs, there were 45 in which both the case and the control were categorized as ‘High’ for benzene exposure; 25 pairs where the case was categorized as ‘high’ benzene exposure and the control as ‘low’ benzene exposure; 21 pairs where the control was categorized as ‘high’ benzene exposure and the case was categorized as ‘low’ benzene exposure; 20 where the case and the control were categorized as ‘low’ benzene exposure.

Calculate the 95% confidence interval for the odds ratio. Please round to 3 decimal places at each step.

Calculate the Chi-Square Statistic for the matched pair data. Please round to 3 decimal places.

The correct answer is 0.196.

|  |  |  |
| --- | --- | --- |
|  | Exposed Control  (High Benzene) | Unexposed control  (Low Benzene) |
| Exposed case  (High Benzene) | 45 | 25 |
| Unexposed case  (Low Benzene) | 21 | 20 |

McNemar's Chi-square =  Diagram

Description automatically generated

McNemar's Chi-square = A picture containing text

Description automatically generated

McNemar's Chi-square = Text

Description automatically generated with low confidence

McNemar's Chi-square = A picture containing text

Description automatically generated

McNemar's Chi-square = A picture containing text

Description automatically generated

McNemar's Chi-square = 0.196

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: 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?