BIOS668 HW10
Sara O'Brien
4/26/23

Honor Code: On my honor, I have neither given nor received unauthorized aid on this assignment." *Sara O'Brien*

**Q1.**

a. Since this is a case-control study, an odds ratio is the more appropriate measure because we are studying the association between an exposure and outcome with the outcome status already available to us.

**Table of Obesity Occurrence Among Children w/ or w/o Breastmilk Exposure**

| | **Case (Obese)** | **Control (Not obsese)** | |
|---|---|---|---|
| **Breastmilk** | 205 | 1503 | 1708 |
| **No Breastmilk** | 79 | 526 | 605 |
| | 284 | 2029 | |

OR = (205*526)/(1503*79) = 0.908

The estimated odds ratio for the association between breastmilk intake and obesity was 0.908, which suggests that participants who received breastmilk as infants were .908 times as likely to be obese compared to those who did not receive breastmilk and breastmilk consumption during infancy may have a very small protective effect against childhood obesity.

SAS code and output:

```
1  libname data "/home/u49497589/BIOS668";
2
3  data question1;
4      set data.hw10_casecont;
5  run;
6
7  proc freq data=question1;
8      table bfany*case / norow nocol nopercent;
9  run;
```

The FREQ Procedure

Frequency

Table of BFany by case

| BFany | case 0 | case 1 | Total |
|---|---|---|---|
| 0 | 526 | 79 | 605 |
| 1 | 1503 | 205 | 1708 |
| Total | 2029 | 284 | 2313 |

b. Using logistic regression, the estimated odds ratio of obesity was 0.923 with a 95% CI of (0.698, 1.220) when controlling for age and daily caloric intake. A child who received breastmilk as a child was 0.923 times as likely to be obese compared to a child who did not receive breastmilk, similarly suggesting that there is a very slightly decreased odds of obesity for breastfed children even when adjusting for age and daily caloric intake.

SAS code and output:

```
11 proc logistic data=question1 descending;
12     model case=bfany age kcal;
13 run;
```

**Odds Ratio Estimates**

| Effect | Point Estimate | 95% Wald Confidence Limits | |
|---|---|---|---|
| BFany | 0.923 | 0.698 | 1.220 |
| age | 0.955 | 0.913 | 0.998 |
| kcal | 1.086 | 0.991 | 1.191 |

c. If the obese or non-obese students who chose to not participate in the study have different characteristics compared to obese and non-obese study participants, respectively, the sample of cases is not representative of all students in the population. As such, the study may not be generalizable to the entire population of students in the given school because the external validity of the study is compromised. Differing participation rates between obese and non-obese students could also have affected the internal validity of the study due to potential selection bias.

d. Yes, I would recommend that future case-control studies examining the relationship between breastfeeding and obesity match on age and/or daily caloric intake. Both age and caloric intake are established risk factors for obesity and could result in confounding if not controlled for. Matching on age and caloric intake makes it so that cases and controls are more similar across these variables, improving the precision of the produced estimates, reducing potential confounding, and increasing internal validity.

**Q2.**

a. Incident rate ratio = IR in men / IR in women

IR men: 358/2314540 = 0.155 per 1,000 person-years
IR women: 230/2461498 = 0.093 per 1,000 person-years

IRR = 0.154/0.093 = 1.656

SE(lnIR) = sqrt(1/358 + 1/230)
95% CI = exp(ln(IR) $\pm$ 1.96*SE) = (1.403, 1.954)

The incidence rate of Parkinson's disease is 1.656 times higher in men than in women and the corresponding 95% CI for this incidence rate ratio is (1.403, 1.954), suggesting that the difference in rates is statistically significant and men develop Parkinson's at a higher rate than women.

b. A cohort follow-up study design would be the most appropriate to address the research question of whether Parkinson's disease is associated with the use of insect repellant in women. In our study of 2000 women, a cross-sectional design would not be ideal since the exposure and outcome must be captured at a single point in time. It also is not ideal to use a case-control design because if we select cases and controls of Alzheimer's, there may be an issue of recall bias for women with Alzheimer's on their use of insect repellant. Given the nature of the condition being studied, a cohort follow-up design is most appropriate since the women can be followed over time to determine if their exposure to insect repellant is associated with the development of Alzheimer's.