1. Selection bias could occur in a regular case-control study that a nested case control study would be able to prevent. If the cases or controls were selected differentially on the basis of their exposure status (UNC25 Virus), this could affect the study. Since both the exposure, UNC25 Virus and the disease, cancer have occurred by the time the patient is recruited into the study this would be a significant possibility in a regular case-control study.
2. The selection bias of the nested case control is dependent on how biased the cohort is, thus it will not introduce any additional selection bias into the study.
3. DAG1
   1. **Age** confounder
   2. **Sex** confounder
   3. **Depressed immune status** mediator
4. DAG2
   1. **Sex** confounder
   2. **Age** modifier
   3. **Education** none of these
5. Table

|  |  |  |
| --- | --- | --- |
| Table 1. Crude and Sex Stratum-specific Odds Ratio between UNC 25 Virus Exposure and Cancer Diagnosis | | |
| *Measure* | *Estimate* | *95% CI* | |
| *Crude OR* | *4.22* | *1.75 – 10.18* | |
| *Stratum-specific OR* |  |  | |
| *Male* | *2.40* | *.76-7.93* | |
| *Female* | *7.50* | *1.97-28.61* | |

1. Table 2 Interpretation
2. Table 2 Interpretation
3. Table 2 related Interpretation
4. Table 2 related Interpretation

**#6-9** The odds of being diagnosed with cancer is 5.8 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus. The 95% CI is (1.32-12.39), this tell us that 95% of the time, when we calculate a confidence interval in this way, the true parameter will be between the two values.

The odds of being diagnosed with cancer for people Males is 1.50 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus. The 95% CI is (0.5,8.75), this tell us that 95% of the time, when we calculate a confidence interval in this way, the true parameter will be between the two values.

The odds of being diagnosed with cancer Females is 9.20 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus. The 95% CI is (2.48,33.80), this tell us that 95% of the time, when we calculate a confidence interval in this way, the true parameter will be between the two values.

The sex-stratified OR estimate for females is much larger than the crude unadjusted OR estimate. The sex-stratified OR estimate for males is much smaller than the crude unadjusted OR estimate. This would flag that there is evidence that confounding by dichotomized sex is present.

By looking at the odds ratio confidence intervals, the crude OR 95% CI and the sex-stratified OR CI for males are the most precise compared the one for females. (since the range of these two are the narrowest). The largest confidence interval by far is for the Female stratum Odds Ratio.

1. Complete Table 3

|  |  |  |
| --- | --- | --- |
| Table 3.Comparison of Unadjusted OR and Sex-adjusted OR with 95%CIs. | | |
| *Measure* | *Estimate* | *95% CI* |
| *Crude OR* | *4.22* | *1.75-10.18* |
| *Sex-adjusted OR* | *4.13* | *1.75-10.39* |
| *% difference* | *2%* |  |



The odds of being diagnosed with cancer is 4.22 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus.

The odds of being diagnosed with cancer is 4.13 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus when adjusting for Sex.

The odds ratio decreases by 2% when adjusted for Sex. By looking at the difference between the odds ratio, it is unlikely Sex is a confounder (the general cut off is usually greater than 10%).

1. Complete Table 4

|  |  |  |
| --- | --- | --- |
| Table 4.Comparison of Race-adjusted and Unadjusted OR with 95%CIs and p-values using Logistic Regression | | |
| *Measure* | *Estimate* | *95% CI* |
| *Crude OR* | *4.22* | *1.75-10.18* |
| *Race-adjusted OR* | *3.67* | *1.51-9.37* |
| *% difference* | *13%* |  |

1. **And 14**

The odds of being diagnosed with cancer is 4.22 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus.

The odds of being diagnosed with cancer is 3.67 times higher for people exposed to the UNC25 virus compared to people not exposed to the UNC25 virus when adjusting for Race.

The odds ratio increases by 13% when adjusted for Race. Since the percent difference is 13%, it would indicate there is evidence of confounding for Race



PART 2. Comprehensive Review and Integration

1. Time Trend Ecological Study. Unit of measurement: total guns
2. They are weighted by country population, not standardized to the at-risk population.
3. Prevalence is not a rate, it is a proportion. It is the number of people with the outcome at a given time.
4. Most likely a cross sectional study since they are taking a snapshot of a population at a certain time, and generalizing the results to a larger population.
5. Consistency, Specificity, Strength of Association
6. Information Bias
7. Bias away from the null
8. Associations are all greater than the null value, 95%CIs are mostly statistically significant
9. It increases
10. Yes, the pattern of results is the same when researchers used Rapid Eye or Landsat data.
11. Selection Bias
12. 35-59 female since they have the highest RR
13. Non-response bias
14. Greater than or equal to 4 years in the community was the only significant factor given that the null value 1 is not in the confidence interval

OR 2.09 CI 1.31–3.32

1. The odds of being a non-acceptor is .49 times as likely for those who said yes to that question. The CI tells us that 95% of the time, when we calculate a confidence interval in this way, the true adjusted odds ratio will be between .31 and .77.
2. Convenience of having an HIV test at home encouraged me to test since it has the lowest odds ratio, thus indicates the lowest odds of being a non-acceptor.

appendix

etable=function(exposure,disease){  
 t=table(exposure,disease)[2:1,2:1]  
 rownames(t)=c('E+','E-')  
 colnames(t)=c('D+','D-')  
 kable(t)  
}  
  
riskratio=function(a,b,c,d){  
 e=a/(a+b)  
 f=c/(c+d)  
 c(e,f,e/f)  
}  
  
ciriskr=function(ecases,enoncases,uecases,uenoncases){  
 a=ecases  
 b=enoncases  
 c=uecases  
 d=uenoncases  
 e=a/(a+b)  
 f=c/(c+d)  
 rr=e/f  
 z=c(-1.96,1.96)  
 se=sqrt(1/a-1/(a+b)+  
 1/c-1/(c+d))  
 ci=exp(log(rr)+z\*se)  
 c("Risk Ratio"=rr,"95 CI"=ci)  
}  
  
rateratio=function(cases,noncases,pyrcases,pyrnoncases){  
 a=cases  
 b=noncases  
 p=pyrcases  
 q=pyrnoncases  
 (a/p)/(b/q)  
}  
  
cirater=function(cases,noncases,pyrcases,pyrnoncases){  
 a=cases  
 b=noncases  
 p=pyrcases  
 q=pyrnoncases  
 rr=(a/p)/(b/q)  
 se=sqrt(1/a+1/b)  
 z=c(-1.96,1.96)  
 ci=exp(log(rr)+z\*se)  
 c("Rate Ratio"=rr,"95 CI"=ci)  
}  
  
cioddsr=function(ecases,enoncases,uecases,uenoncases){  
 a=ecases  
 b=enoncases  
 c=uecases  
 d=uenoncases  
 or=(a/c)\*(d/b)  
 se=sqrt(1/a+1/b+1/c+1/d)  
 z=c(-1.96,1.96)  
 ci=exp(log(or)+z\*se)  
 c("Odds Ratio"=or,"95 CI"=ci)  
}

cc <- read\_excel("casecontrol.xlsx")

ctab=etable(cc$VIRUS,cc$CANCER)

men=cc%>%filter(SEX==0)  
women=cc%>%filter(SEX==1)  
mtab=etable(men$VIRUS,men$CANCER)  
wtab=etable(women$VIRUS,women$CANCER)  
mtab

|  |  |  |
| --- | --- | --- |
|  | D+ | D- |
| E+ | 12 | 6 |
| E- | 15 | 18 |

wtab

|  |  |  |
| --- | --- | --- |
|  | D+ | D- |
| E+ | 20 | 4 |
| E- | 10 | 15 |

cOR=4.22  
y = glm(CANCER ~ VIRUS+SEX, family=binomial,data = cc)  
(a=round(exp(cbind(coef(y), confint(y))),digits=2))

## 2.5 % 97.5 %  
## (Intercept) 0.70 0.37 1.32  
## VIRUS 4.13 1.75 10.39  
## SEX 1.19 0.51 2.78

aOR=a["VIRUS",1]  
pchange1=(cOR-aOR)/cOR  
round(pchange1,digits=2)

## [1] 0.02

x = glm(CANCER ~ VIRUS+RACE, family=binomial,data = cc)  
round(exp(cbind(coef(x), confint(x))),digits=2)

## 2.5 % 97.5 %  
## (Intercept) 0.49 0.19 1.11  
## VIRUS 3.67 1.51 9.37  
## RACE 1.41 0.86 2.59

pchange2=(cOR-3.67)/cOR  
round(pchange2,digits=2)

## [1] 0.13