NAME: ________Diego Liang__________

EPI202: Fall 2022
Homework 2
To be uploaded as PDF to course website by 9:30AM on November 10, 2022

Please provide concise, precise answers.

We encourage collaborative learning in this course. You may discuss homework assignments with other students. However, all written work that you submit for grading must be your own, in your own words, reflecting your understanding of the homework assignment. Homework assignments should not be prepared by copying, paraphrasing, or summarizing someone else's work.

To complete this assignment, you can use the EPI202 calculator, a hand calculator, Excel or any statistical package (SAS, Stata, R etc.).

- Please round all final estimates for measures of association to 2 digits and p-values to 3 digits.
- No output will be considered as part of a response, although you may choose to attach output to demonstrate how you came to your answer in a separate appendix at the end of your work (which will not be graded).

Note: for all questions, please include all relevant formulas and define any variables that you use and have not already defined in previous questions to receive full credit.

NAME: ________Diego Liang__________

Part I. Adjusted analyses.

The following questions are based on the paper: Sponholtz TR, Palmer JR, Rosenberg L, Hatch EE, Adams-Campbell LL, Wise LA. Reproductive factors and incidence of endometrial cancer in U.S. black women. Cancer Causes Control. 2017;28:579-588.
https://link-springer-com.ezp-prod1.hul.harvard.edu/article/10.1007/s10552-017-0880-4

1. Use the data in Table 3 of the paper to construct a new table showing the data stratified by menopausal status to examine the association between age at first birth (categorized as <20 years versus ≥ 20 years) and the incidence of endometrial cancer.

	Age at first birth	
	<20 years	≥ 20 years
Premenopausal		
Cases	31	36
Person-time (years)	79081.63	236532.30
Postmenopausal		
Cases	51	92
Person-time (years)	43330.50	91838.04

2. Use the data in the table you created in question 1 to calculate the Mantel-Haenszel incidence rate ratio (IRR_{MH}) for the association between age at first birth (categorized as <20 years versus ≥ 20 years) and the incidence of endometrial cancer accounting for potential confounding by menopausal status.
    a. Calculate the $IRR_{MH.}$
    b. Interpret the numerical result in words.
    c. Construct a 95% confidence interval for the IRR_{MH} from question 2. Interpret the numerical result in words.

	a_i	b_i	M_i	N_{0i} (person-years)	N_{1i} (person-years)	T_i (person-years)
$i = 1$ (premenopausal)	31	36	67	79082.63	236532.30	315614
$i = 2$ (postmenopausal)	51	92	143	43331.50	91838.04	135169

$$\widehat{IRR}_{MH} = \frac{\sum_{i=1}^{I} \frac{a_i N_{oi}}{T_i}}{\sum_{i=1}^{I} \frac{b_i N_{1i}}{T_i}} = \frac{\frac{31*79082}{315614} + \frac{51*91838}{135169}}{\frac{36*79082}{315614} + \frac{92*43331}{135169}} = 1.5030$$

NAME: ________Diego Liang__________

After adjusting for menopausal status, the incidence rate of endometrial cancer for those whose age at first birth is less than 20 years is 1.5030 times the incidence rate of endometrial cancer for those whose age at first birth is equal or above 20 years, assuming no EMM, residual confounding, confounding by other variables, selection bias, or information bias.

The 95% confidence interval for the IRR_{MH}:

$$X \pm 1.96\sqrt{\widehat{Var}(X)} = ln(\widehat{IRR}_{MH}) \pm 1.96\sqrt{\widehat{Var}(ln(\widehat{IRR}_{MH}))}$$

$$= ln\left(\frac{\sum_{i=1}^{I}\frac{a_i N_{oi}}{T_i}}{\sum_{i=1}^{I}\frac{b_i N_{1i}}{T_i}}\right) \pm 1.96\sqrt{\frac{\sum_{i=1}^{I}(M_{1i} N_{1i} N_{0i})/T_i^2}{\left(\sum_{i=1}^{I}\frac{a_i N_{oi}}{T_i}\right)\left(\sum_{i=1}^{I}\frac{b_i N_{1i}}{T_i}\right)}}$$

$$= ln(1.5030) \pm 1.96\sqrt{\frac{((67*79082*236532)/315614)+((143*43331*91838)/135169)}{\left(\frac{31*79082}{315614}+\frac{51*43331}{135169}\right)\left(\frac{36*236532}{315614}+\frac{92*91838}{135169}\right)}}$$

$$= (0.1330,\ 0.6820)$$

$$e^{(0.1330,\ 0.6820)} = (1.1422,\ 1.9778)$$

After adjusting for menopausal status, these data are consistent with incidence rate ratios ranging from 1.1422 to 1.9778 with 95% confidence for the association between age at first birth (< 20 years compared to ≥ 20 years) and the incidence rate of endometrial cancer, assuming no EMM, residual confounding, confounding by other variables, selection bias, or information bias.

NAME: ________Diego Liang__________

Part II. Adjusted analyses in closed cohort data

In this section, you will analyze data from the Myocardial Infarction Onset Study. This study was a multicenter cohort study of myocardial infarction patients enrolled between 1989 to 1996 at the time of their myocardial infarction. All participants were followed through the National Death Index until death or a minimum of 10 years after enrollment. There was no loss to follow-up. In this analysis, you will evaluate the relationship between reporting ever using marijuana on the baseline interview and the cumulative incidence of death from any cause over the following 10 years. You will account for potential confounding by age. [Note: In this study, conducted in the late 1980s, sex was recorded and categorized as female or male based on NIH reporting requirements in place at the time. No data on gender was recorded.]

The variables in this dataset are described below:

Variable Name	Description
id	ID number
age	Age (continuous, years)
age_cat	Age Category (1: <50yrs, 2: 50-64 yrs, 3: 65+ yrs)
female	Female (1: female, 0: male)
married	Married (1: yes, 0:no)
educ	Educational Attainment (1: <HS, 2: HS, 3: >HS)
dm	Diabetes (1: yes, 0:no)
htn	Hypertension (1: yes, 0:no)
phys_activity	Frequency of Physical Activity (0: <1/wk, 1: 1-3/wk, 2: 4+/wk)
evermarj	Ever use marijuana (1: yes, 0:no)
follow_up	Duration of follow-up (years)
dead	Death within 10-years (1: died, 0: survived)
cvdeath	Death from cardiovascular causes (1: CVD death, 0: did not die of CVD)

The dataset name is MI_Onset_10 and is available for download from the course website in several file formats including CSV, R, SAS and Stata. You may use SAS, STATA, R, the EPI202 calculator or any other statistical analysis software package of your choosing.

1. Calculate the crude cumulative incidence ratio for the association between reporting a sedentary lifestyle (physical activity < 1 time per week versus ≥ 1 time per week) at baseline and the cumulative incidence of death by the end of the 10-year follow-up period. (HINT: you will find it helpful to create a new variable categorizing sedentary = 1 if frequency of physical activity is less than once per week; and sedentary = 0 if the participant reported physical activity at least once per week).

$\widehat{CIR} = (\frac{1021}{1021+2178})/(\frac{72}{72+441}) = 2.2740$ over the 10-year follow-up period

NAME: Diego Liang

2. In this question, you will empirically evaluate if age possesses the properties of confounding and the direction of that confounding. **In this question, you will use age as a categorical variable with three levels (as coded in the data).**
    a. You will first evaluate the relationship between age and sedentary lifestyle at baseline (physical activity < 1 time per week versus ≥ 1 time per week).
        i. Calculate the prevalence of reporting sedentary lifestyle, **separately for each of the three age groups**.
        ii. Calculate the prevalence ratios **using the youngest age group as the referent.**
        iii. Using the results from i and ii: is age positively or negatively associated with a sedentary lifestyle in the study base?

NOTE: There is no need to compute confidence intervals or construct hypothesis tests.

Among those ages < 50 years, $Prevelance_1 = 558/(558 + 179) = 0.757$ cases at baseline

Among those ages 50-64 years, $Prevelance_2 = 1171/(1171 + 221) = 0.841$ cases at baseline

Among those ages ≥ 65 years, $Prevelance_3 = 1470/(1470 + 113) = 0.929$ cases at baseline

Using the youngest age group as referent,

the prevalence ratio between those ages 50-64 years and the referent is $PR_{2\ vs\ 1} = Prevelance_2/Prevelance_1 = 1.111$

the prevalence ratio between ages ≥ 65 years and the referent is $PR_{3\ vs\ 1} = Prevelance_3/Prevelance_1 = 1.227$

Age is positively associated with a sedentary lifestyle in the study base because the prevalence of reporting a sedentary lifestyle gets higher with the increment of age and $PR_{3\ vs\ 1} > PR_{2\ vs\ 1}$.

    b. Next, you will assess the relationship between age and the 10-year cumulative incidence of death.
        i. Calculate the 10-year cumulative incidence of death **separately for each of the three age groups** and,
        ii. Calculate the **cumulative incidence ratios using the youngest group as the referent.**
        iii. Using the results from i and ii: is age positively or negatively associated with the 10-year cumulative incidence of death among the non-exposed (those who reported physical activity at least once per week).

NOTE: There is no need to compute confidence intervals or construct hypothesis tests.

Among the non-exposed,

for those ages < 50 years, $CI_1 = 9/(9 + 170) = 0.050$ cases over the 10-year follow-up period for those ages 50-64 years, $CI_2 = 28/(28 + 193) = 0.127$ cases over the 10-year follow-up period

for those ages $\geq$ 65 years, $CI_3 = 35/(35 + 78) = 0.310$ cases over the 10-year follow-up period

Among the non-exposed, using the youngest age group as referent,

the 10-year cumulative incidence ratio between those ages 50-64 years and the referent is

$CIR_{2\ vs\ 1} = CI_2/CI_1 = 2.525$

the 10-year cumulative incidence ratio between ages $\geq$ 65 years and the referent is

$CIR_{3\ vs\ 1} = CI_3/CI_1 = 6.163$

Age is positively associated with the 10-year cumulative incidence of death among the non-exposed, because the 10-year cumulative incidence of death gets higher with the increment of age and $CIR_{3\ vs\ 1} > CIR_{2\ vs\ 1}$.

c. Given the results from 2.a. and 2.b. is it reasonable to be concerned that the crude association you computed in question 1 was confounded by age? Explain your answer using the three properties of confounders.

According to the results of question 2.a and question 2.b,

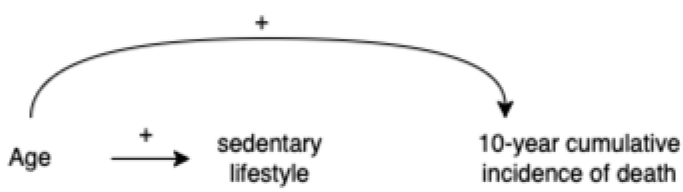
1. Age is(positively) associated with a sedentary lifestyle in the study base
2. Age is (positively) associated with the 10-year cumulative incidence of death among the non-exposed

Also,

3. Age is not the downstream consequence of a sedentary lifestyle or the 10-year incidence rate of death.

Hence, age is a confounder in the crude association computed in question 1.

d. Predict the direction that the association between reporting a sedentary lifestyle at baseline and the 10-year cumulative incidence of death will change after adjusting for age. Explain your answer in words and by drawing a DAG with signs (+) or (-) in the arrows that emerge from age.



The association between reporting a sedentary lifestyle at baseline and the 10-year cumulative incidence of death will go down after adjusting for age because the bias is upward based on this DAG.

3. Based on our results from 2), we are planning to stratify on age the association between sedentary lifestyle at baseline and 10-years all-cause mortality (with age as a categorical variable with three levels, as coded in the data). Is there any evidence that the *positivity assumption* is violated in at least one stratum? Explain your answer by checking whether the positivity assumption holds or not *in the data*.

	< 50 years			50-64 years			≥ 65 years		
	Has Sedentary lifestyle	No Sedentary lifestyle	Total	Has Sedentary lifestyle	No Sedentary lifestyle	Total	Has Sedentary lifestyle	No Sedentary lifestyle	Total
Cases	68	9	77	279	28	307	674	35	709
Non-cases	490	170	660	892	193	1085	796	78	874
Total	558	179	737	1171	221	1392	1470	113	1583

Based on the data above, the positivity assumption holds because every stratum has at least one case or non-case.

4. Next, analyze the data after adjusting for age in 3 categories.
   a. Calculate the Mantel-Haenszel summary cumulative incidence ratio for the association between sedentary lifestyle at baseline and all-cause mortality after adjusting for age.
   b. Calculate the 95% confidence interval for the Mantel-Haenszel summary cumulative incidence ratio after adjusting for age.
   c. Calculate the summary cumulative incidence difference after adjusting for age.
   d. Calculate the 95% confidence interval for the summary cumulative incidence difference after adjusting for age.

   Interpret all numerical results (a to d) in words.

$$\widehat{CIR}_{MH} = \frac{\sum_{i=1}^{I}(\frac{a_i N_{0i}}{T_i})}{\sum_{i=1}^{I}(\frac{b_i N_{1i}}{T_i})} = \frac{(\frac{68*558}{737})+(\frac{279*1171}{1392})+(\frac{674*1470}{1583})}{(\frac{9*179}{737})+(\frac{28*221}{1392})+(\frac{35*113}{1583})} = 1.7325 \text{ over the 10-year follow-up period}$$

After adjusting for age, the cumulative incidence of death over the 10-year follow-up period among those who reported a sedentary lifestyle at baseline is 1.7325 times the cumulative incidence of death over the 10-year follow-up period among those who did not have a sedentary lifestyle at baseline, assuming no EMM, residual confounding, confounding by other variables, selection bias, or information bias.

The 95% confidence interval for the $\widehat{CIR}_{MH}$:

$$X \pm 1.96\sqrt{\widehat{Var}(X)} = ln(\widehat{CIR}_{MH}) \pm 1.96\sqrt{\widehat{Var}(ln(\widehat{CIR}_{MH}))}$$

NAME: Diego Liang

$$= ln\left(\frac{\sum_{i=1}^{I}\left(\frac{a_i N_{0i}}{T_i}\right)}{\sum_{i=1}^{I}\left(\frac{b_i N_{1i}}{T_i}\right)}\right) \pm 1.96\sqrt{\frac{\sum_{i=1}^{I}(M_{1i}N_{1i}N_{0i})/T_i^2}{\left(\sum_{i=1}^{I}\frac{a_i N_{0i}}{T_i}\right)\left(\sum_{i=1}^{I}\frac{b_i N_{1i}}{T_i}\right)}} = (0.3366,\ 0.7625)$$

$$e^{(0.3366,\,0.7625)} = (1.4002,\ 2.1436)$$

After adjusting for age, these data are not consistent with cumulative incidence ratios ranging from 1.4002 to 2.1436 with 95% confidence for the association between sedentary lifestyle and all-cause mortality, assuming no EMM, residual confounding, confounding by other variables, selection bias, or information bias.

5. Present the point estimates and 95% confidence intervals for the cumulative incidence ratios and cumulative incidence difference for the association between sedentary lifestyle at baseline and 10 year all-cause mortality for each age category (keep using age as a three level variable).

	Cumulative Incidence Ratios for sedentary lifestyle and all-cause mortality (**95% CI**)	Cumulative Incidence Differences for sedentary lifestyle and all-cause mortality (**95% CI**)
<50 years	2.4237 (1.2347, 4.7580)	0.0716 (0.0296, 0.1136)
50-64 years	1.8805 (1.3107, 2.6980)	0.1116 (0.0614, 0.1617)
65 or older	1.4803 (1.1179, 1.9602)	0.1488 (0.0598, 0.2377)

6. For the stratified data in the table above, perform **the test of homogeneity** for the cumulative incidence ratio.
    a. State the null and alternative hypotheses.
    b. What is the value of the H statistic, and how many degrees of freedom does it have?
    c. Find the p-value for the test statistic.

H_0: There is no effect measure modification of the association between sedentary lifestyle at baseline and 10-year all-cause mortality across the strata of age on the multiplicative scale. $(CIR_{<50yrs} = CIR_{50-64yrs} = CIR_{65+yrs})$

H_A: There is the effect measure modification of the association between sedentary lifestyle at baseline and 10-year all-cause mortality across the strata of age on the multiplicative scale. $(CIR_{<50yrs} \neq CIR_{50-64yrs},\ or\ CIR_{<50yrs} \neq CIR_{65+yrs},\ or\ CIR_{50-64yrs} \neq CIR_{65+yrs})$

Test statistic:

$$H = \sum_{i=1}^{I}\frac{(\widehat{X}_i - \widehat{X}_{summary})^2}{\widehat{Var}\ (\widehat{X}_i)} = \sum_{i=1}^{I}\frac{(ln(\widehat{CIR}_i) - ln(\widehat{CIR}_{MH}))^2}{\widehat{Var}\ (ln(\widehat{CIR}_i))} = \sum_{i=1}^{3}\frac{\left(ln\left(\frac{a_i}{N_{0i}}/\frac{b_i}{N_{1i}}\right) - ln\left(\frac{\sum_{i=1}^{I}\left(\frac{a_i N_{0i}}{T_i}\right)}{\sum_{i=1}^{I}\left(\frac{b_i N_{1i}}{T_i}\right)}\right)\right)^2}{\frac{c_i}{a_i N_{1i}} + \frac{d_i}{b_i N_{0i}}} \sim \chi^2_{I-1} \text{ under the } H_0$$

Degrees of freedom = # of strata -1 = 3-1=2 degree of freedom

$$\widehat{CIR}_{MH} = \frac{\sum_{i=1}^{I}\left(\frac{a_i N_{0i}}{T_i}\right)}{\sum_{i=1}^{I}\left(\frac{b_i N_{1i}}{T_i}\right)} = \frac{\left(\frac{68*558}{737}\right)+\left(\frac{279*1171}{1392}\right)+\left(\frac{674*1470}{1583}\right)}{\left(\frac{9*179}{737}\right)+\left(\frac{28*221}{1392}\right)+\left(\frac{35*113}{1583}\right)} = 1.7325 \text{ over the 10-year follow-up period}$$

$$\widehat{Var}_i(ln(\widehat{CIR}_i)) = \frac{c_i}{a_i N_{1i}} + \frac{d_i}{b_i N_{0i}}$$

$$\widehat{Var}\ (ln(\widehat{CIR}_{<50yrs})) = \frac{170}{9*558} + \frac{490}{68*179} = 0.0741$$

$$\widehat{Var}\ (ln(\widehat{CIR}_{50-64yrs})) = \frac{193}{28*1171} + \frac{892}{279*221} = 0.0204$$

$$\widehat{Var}\ (ln(\widehat{CIR}_{\geq 65yrs})) = \frac{78}{35*1470} + \frac{796}{674*113} = 0.0120$$

$$\text{So, } H = \frac{(ln(0.4126)-ln(0.5772))^2}{0.0741} + \frac{(ln(0.5318)-ln(0.5772))^2}{0.0204} + \frac{(ln(0.6755)-ln(0.5772))^2}{0.0120} = 2.3557 \sim \chi^2_2$$

under the null

$p = 0.3079 > 0.05$

	Exposed	Unexposed	Total	Exposed	Unexposed	Total	Exposed	Unexposed	Total
Cases	68	9	77	279	28	307	674	35	709
Non-cases	490	170	660	892	193	1085	796	78	874
Total	558	179	737	1171	221	1392	1470	113	1583
	CIR=	**2.423735564**		**CIR=**	**1.880535562**		**CIR=**	**1.480310982**	
	CID=	**0.07158447**		**CID=**	**0.111561067**		**CID=**	**0.148768888**	
Cases									
Non-cases									
Total									
Cases									
Non-cases									
Total									

	Crude Data		
	Exposed	Unexp.	
Cases	1021	72	1093
Non-cases	2178	441	2619
Total	3199	513	3712
Crude CIR	**2.2740**		
90% CI =	1.8905	2.7353	
95% CI =	**1.8248**	**2.8338**	
99% CI =	1.7031	3.0364	
Crude CID	**1.788E-01**		
90% CI =	1.502E-01	2.075E-01	
95% CI =	**1.447E-01**	**2.129E-01**	
99% CI =	1.340E-01	2.236E-01	
Hypothesis Test:			
Z-square =	68.04		
dof =	1		
p-value =	**<0.0001**		

Summary Risk Ratio using Mantel-Haenszel weights		
Summary CIR	**1.7325**	
90% CI =	1.4490	2.0715
95% CI =	**1.4002**	**2.1436**
99% CI =	1.3098	2.2917
P-value for homogeneity:		**0.3079**

Summary Risk Difference using Mantel-Haenszel-style weights		
Summary CID	**1.080E-01**	
90% CI =	7.964E-02	1.364E-01
95% CI =	**7.421E-02**	**1.418E-01**
99% CI =	6.360E-02	1.524E-01
P-value for homogeneity:		**0.1558**

Hypothesis Test	
Z-square =	30.05
dof =	1
p-value =	**0.0000**

Tests of Homogeneity			
Risk Ratio		**Risk Difference**	
H =	2.36	H =	3.72
dof =	2	dof =	2
p-value =	**0.3079**	**p-value =**	**0.1558**

NAME: _________Diego Liang__________

7. For the stratified data in the table above, perform **the test of homogeneity** for the cumulative incidence difference.
   a. State the null and alternative hypotheses.
   b. What is the value of the H statistic, and how many degrees of freedom does it have?
   ~~c.~~ Find the p-value for the test statistic.

H_0: There is no effect measure modification of the association between sedentary lifestyle at baseline and 10-year all-cause mortality across the strata of age on the additive scale. $(CID_{<50yrs} = CID_{50-64yrs} = CID_{65+yrs})$

H_A: There is the effect measure modification of the association between sedentary lifestyle at baseline and 10-year all-cause mortality across the strata of age on the additive scale. $(CID_{<50yrs} \neq CID_{50-64yrs},\ or\ CID_{<50yrs} \neq CID_{65+yrs},\ or\ CID_{50-64yrs} \neq CID_{65+yrs})$

Test statistic:

$$H = \sum_{i=1}^{I} \frac{(\widehat{X}_i - \widehat{X}_{summary})^2}{\widehat{Var}\ (\widehat{X}_i)} = \sum_{i=1}^{I} \frac{(\widehat{CID}_i - \widehat{CID}_{MH})^2}{\widehat{Var}\ (\widehat{CID}_i)} = \sum_{i=1}^{3} \frac{\left(\left(\frac{a_i}{N_{0i}} - \frac{b_i}{N_{1i}}\right) - \left(\frac{\sum_{i=1}^{I}\left(\frac{a_iN_{0i}-b_iN_{1i}}{T_i}\right)}{\sum_{i=1}^{I}\left(\frac{N_{1i}N_{0i}}{T_i}\right)}\right)\right)^2}{\frac{a_ic_i}{N_{1i}^3} + \frac{b_id_i}{N_{0i}^3}} \sim \chi^2_{I-1} \text{ under the } H_0$$

Degrees of freedom = # of strata -1 = 3-1=2 degree of freedom

$$\widehat{CID}_{MH} = \frac{\sum_{i=1}^{I}\left(\frac{a_iN_{0i}-b_iN_{1i}}{T_i}\right)}{\sum_{i=1}^{I}\left(\frac{N_{1i}N_{0i}}{T_i}\right)} = \frac{\frac{68*558-9*179}{737} + \frac{279*1171-28*221}{1392} + \frac{674*1470-35*113}{1583}}{\frac{179*558}{737} + \frac{221*1171}{1392} + \frac{11381470}{1583}} = 0.1080 \text{ over the 10-year}$$

follow-up period

$$\widehat{Var}\ (\widehat{CID}_i) = \frac{a_ic_i}{N_{1i}^3} + \frac{b_id_i}{N_{0i}^3}$$

$$\widehat{Var}\ (\widehat{CID}_{<50yrs}) = \frac{9*170}{558^3} + \frac{68*490}{179^3} = 0.0058$$

$$\widehat{Var}\ (\widehat{CID}_{50-64yrs}) = \frac{28*193}{1171^3} + \frac{279*892}{221^3} = 0.0231$$

$$\widehat{Var}\ (\widehat{CID}_{\geq 65yrs}) = \frac{35*78}{1470^3} + \frac{674*796}{113^3} = 0.3718$$

$$H = \frac{(0.0716-0.1080)^2}{0.0058} + \frac{(0.1116-0.1080)^2}{0.0231} + \frac{(0.1488-0.1080)^2}{0.3718} = 3.7190 \sim \chi^2_2 \text{ under the null}$$

$$p = 0.1558 > 0.05$$

NAME: ________Diego Liang__________

	Exposed	Unexposed	Total	Exposed	Unexposed	Total	Exposed	Unexposed	Total
Cases	68	9	77	279	28	307	674	35	709
Non-cases	490	170	660	892	193	1085	796	78	874
Total	558	179	737	1171	221	1392	1470	113	1583
	CIR=	2.423735564		CIR=	1.880535562		CIR=	1.480310982	
	CID=	0.07158447		CID=	0.111561067		CID=	0.148768888	
Cases									
Non-cases									
Total									
Cases									
Non-cases									
Total									

Crude Data			
	Exposed	Unexp.	
Cases	1021	72	1093
Non-cases	2178	441	2619
Total	3199	513	3712
Crude CIR	**2.2740**		
90% CI =	1.8905	2.7353	
95% CI =	**1.8248**	**2.8338**	
99% CI =	1.7031	3.0364	
Crude CID	**1.788E-01**		
90% CI =	1.502E-01	2.075E-01	
95% CI =	**1.447E-01**	**2.129E-01**	
99% CI =	1.340E-01	2.236E-01	
Hypothesis Test:			
Z-square =	68.04		
dof =	1		
p-value =	**<0.0001**		

Summary Risk Ratio using Mantel-Haenszel weights		
Summary CIR	**1.7325**	
90% CI =	1.4490	2.0715
95% CI =	**1.4002**	**2.1436**
99% CI =	1.3098	2.2917
P-value for homogeneity:		**0.3079**

Summary Risk Difference using Mantel-Haenszel-style weights		
Summary CID	**1.080E-01**	
90% CI =	7.964E-02	1.364E-01
95% CI =	**7.421E-02**	**1.418E-01**
99% CI =	6.360E-02	1.524E-01
P-value for homogeneity:		**0.1558**

Hypothesis Test	
Z-square =	30.05
dof =	1
p-value =	**0.0000**

Tests of Homogeneity			
Risk Ratio		**Risk Difference**	
H =	2.36	H =	3.72
dof =	2	dof =	2
p-value =	**0.3079**	**p-value =**	**0.1558**

8. Is there statistical evidence of effect measure modification by age on either the additive or multiplicative scale? Explain briefly.

Based on the results of questions 6 and 7, we fail to reject both null hypotheses at the $\alpha = 0.05$ level. We do not have sufficient evidence to conclude that there is effect measure modification by age on the additive or multiplicative scale for the association between sedentary lifestyle at baseline and 10-year all-cause mortality, assuming no confounding, selection bias, or information bias.