The MI onset Study was a multicenter epidemiologic study of the determinants of myocardial infarction onset and prognosis. In this study, patients with acute myocardial infarction were recruited from 64 medical centers across the US and interviewed about lifestyle and health characteristics within an average of 3 days of hospitalization. In this analysis, you will evaluate the relationship between reporting ever using marijuana on the baseline interview and the incidence rate of death from cardiovascular cause.

|  |  |
| --- | --- |
| **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.

**TUESDAY**

1. What is the prevalence of having a history of reporting ever using marijuana at the time of the baseline interview.

Graphical user interface, text, application

Description automatically generated

2a. Compute the incidence rate of death from cardiovascular causes among those with and without a history of reporting ever using marijuana at baseline and describe the relationship on the multiplicative scale.

9.705 per 1,000 person-years

22.902 per 1,000 person-years

Text

Description automatically generated

2b. Calculate and interpret the incidence rate ratio and its 95% confidence interval.

2c. Calculate and interpret the incidence rate difference and its 95% confidence interval for the association between reporting a history of ever reporting marijuana on the baseline interview and the incidence rate of death from cardiovascular causes.

3a. Evaluate whether the crude association you computed above is confounded by age Based on your subject matter knowledge, illustrate in a directed acyclic graph (DAG) the relationship between age, history of ever using marijuana and the incidence rate of death from cardiovascular causes.

3b. Does age have the characteristics of a confounder of the marijuana --> cardiovascular mortality relationship? Evaluate the relevant associations. Note that the variable age\_cat is defined as age\_cat=1 if age<50; age\_cat=2 if age >=50 to <65; age\_cat=3 if age>=65.

3c. Is there a violation of positivity in these data? If so, drop the relevant strata from the dataset and calculate the relevant relationships.

3d. Based on the associations you computed above, do you think the estimate accounting for confounding by age will be larger or smaller (on the absolute scale) than the crude estimate?

**THURSDAY**

4a. Compute the incidence rate ratio for the association between reporting ever using marijuana at baseline and the incidence rate of death from cardiovascular causes after stratifying by age (dichotomize age as less than vs. greater than or equal to 50). Examine the point estimates and confidence intervals. Is there evidence for effect measure modification on the multiplicative scale?

Graphical user interface, application

Description automatically generated

4b. Now conduct a formal hypothesis test:

* + 1. State the null and alternative hypothesis for the test for effect measure modification on the multiplicative scale.
    2. How many degrees of freedom does this test have?
    3. What is the value of the H-statistic?
    4. Interpret the p-value itself, and then state the result of the Neyman-Pearson hypothesis test.

Text

Description automatically generated