Number of Sexual Partners and Depression

Abby Mutic

[abby.mutic@emory.edu](mailto:abby.mutic@emory.edu)

**Introduction/Background**

As a PhD student actively involved in the Center for Children’s Health, the Environment, the Microbiome and Metabolomics and a Certified Nurse Midwife practicing in the Grady Teen Clinic, I am interested in studies involving children or adolescents, their behaviors, and how their behaviors may influence health outcomes. I am particularly interested in mental health outcomes and better understanding specific areas to target for effective clinical intervention.

Research Question

For this project, I aim to answer the research question: Is the number of sexual partners experienced in adolescence (≤ 12th grade) associated with depression and/or anxiety in adulthood? I hypothesize that the more sexual partners an individual has at an early age, the greater amount of depressive or anxiety cases they will experience in their adulthood. As a part of this project I would like to learn how to handle longitudinal data with multiple time points and use complex survey weights during analysis.

**Methodology/Approach**

Dataset

The data used in this project is pulled from The National Longitudinal Study of Adolescent to Adult Health (Add Health) public database. It is a longitudinal study of United States adolescents in grades 7-12 during the 1994-95 school year. The Add Health cohort has been followed at various time points with the most recent being wave IV in 2008. Wave IV includes only those who were retained in the study from the original data collection period in 1994-1995. Since the Add Health database is open to the public, the file was easily downloaded and added to my personal repository on GitHub. The repository was then pulled into a new project in R studio in preparation for data wrangling and analysis. The Add Health database can be found at https://dataverse.unc.edu/dataset.xhtml?persistentId=doi:10.15139/S3/11900.

Analysis Plan

The statistical analysis will be performed using R. Implementation of basic statistical procedures in this platform was central to my training. Before visualizing the data, I organized the data by creating a data.frame where the rows represent each individual case and the columns represent the variables of interest. Next, I renamed the given variable names and choose names that will make coding and interpretation easier. This will be done by running the attributes of the data.frame and then recoding the names as desired. The variable I will use for age will need to be recoded to display an age number rather than the year of birth as it is currently given in the dataset. The dichotomous dependent variables- depression and anxiety- will need an additional variable that combines depression and anxiety showing a person has both mental health disorders. This can be accomplished by creating a vector and manually adding together those who report yes for depression and yes for anxiety. I do not anticipate wrangling the independent variable needed for analysis.

After my variables of interest have been renamed for analysis, I plan to look at a subset of the data and print this subset by using knitr::kable(). Next, I can look at a summary of the numeric statistics and frequencies of the factor variables. To view summary statistics of factor variables, I will convert the variables to numeric form as needed.

The analysis plan will start by summarizing and describing the sample using descriptive statistics in order to visualize the data in a meaningful way. The information will be placed in a table similar to the shell labeled Table 1. Next, inferential statistics will be conducted to test the hypotheses. The continuous independent variable- number of sexual partners- will be compared to the dependent dichotomous variables- depression and anxiety- by using multinomial logistic regression. Multinomial logistic regression will help predict the probability of depression, anxiety, or both (3 outcomes) occurring as a result of the number of sexual partners in adolescence. Odds ratio(s) will be reported and used for interpretation. An odds ratio >1 indicates as the number of sexual partners increases, the odds of depression, anxiety, or both also increases. An odds ratio <1 indicates as the number of sexual partners increase, the odds of depression, anxiety, or both decreases. Results of the logistic regression will be displayed in Table 2. Careful consideration of confounding variables will be included in model building. Some confounders identified a priori include personal and/or family history of mental illness, personal and parental socioeconomic status, social support, parental support, education, and sexual abuse. Missing data

**Data Summary/Results**

**Table 1.** Characteristics of The National Longitudinal Study of Adolescent to Adult Health cohort

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Wave I**  **1994-95**  **(n = 6504)** | | **Wave IV**  **2008**  **(n =5114)** | |
| **Characteristics** | **n** | **(%)** | **n** | **(%)** |
| Age (mean, range) (m=3) | (16, 13-21) | | (29, 25-34) | |
| Gender (m=1)  Male | 3147 | (48.4) | NA | NA |
| Female | 3356 | (51.6) | NA | NA |
| Race (check all that apply) | | | | |
| Hispanic or Latino | 743 | (11.4) | NA | NA |
| White | 4294 | (66.0) | NA | NA |
| Black or African American | 1619 | (24.9) | NA | NA |
| American Indian or Native American | 236 | (3.6) | NA | NA |
| Asian or Pacific Islander | 270 | (4.2) | NA | NA |
| Other | 425 | (6.5) | NA | NA |
| Felt depressed in past week (m=20)  Most/all the time  A lot of the time  Sometimes  Never/rarely | 193  444  1853  3994 | (3.0)  (6.8)  (28.5)  (61.4) | 108  211  1178  3616 | (2.1)  (4.1)  (23.0)  (70.7) |
| Depression diagnosis (m=1)  Ever  Never | NA  NA | NA  NA | 827  4286 | (16.2)  (83.8) |
| Anxiety diagnosis  Ever  Never | NA  NA | NA  NA | 639  4474 | (12.5)  (87.5) |
| Ever had sexual intercourse (m=86)  Yes  No | 2565  3853 | 39.4  59.2 | NA  NA | NA  NA |
| Sexual partners (mean, range) (m=4874) | (6.4, 1-500) | | NA | |

**Table 2:** Characteristics associated with Depression Diagnosis

|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristic** | **Odds Ratio** | **CI** | **p-value** |
| Female Gender | 1.33 | (0.97, 1.70) | 1.27e-12 |
| Feeling depressed in past week  Most or all the time  A lot of the time  Sometimes | 2.37  1.97  1.25 | (1.55, 3.21)  (1.32, 2.63)  (0.88, 1.61) | 1.66e-08  2.87e-09  3.05e-11 |

Graphs:

1. A bar graph could be used to display the relationship between number of sexual partners (x axis) and depression, anxiety, or both categories (y axis). Further, each bar could show the differences between male and female if interesting.

* A bar graph was produced but the distribution shown was not what I hoped it would look like.

1. A histogram may be used to display the distribution of number of sexual partners variable.

* This histogram was not very helpful in looking at the distribution of number of sexual partners.

**Interpretation, Limitations, and Conclusions**

* Females are 1.3x as likely to be diagnosed with depression than males.
* For every unit change in being depressed in the past week, the log odds of depression increases by 2.4 (most or all the time), 2.0 (a lot of time), 1.2 (sometimes)
* Number of sexual partners doesn't seem to be associated in model 1

From this point, I can:

* Make more models to compare fit.
* Use the train and test datasets I made to make predictions using the model of best fit.

**Code, Associated Data and Files, and Github Repository**

My repository can be found at the following link.

<https://github.com/amutic/AddHealth>

My rmd file can be found at https://github.com/amutic/AddHealth/Milestone2\_Mutic/Addhealthproject03\_28.Rmd

Code book or data dictionary

Record of data from raw to tidy