Current Health Trends and Risk Behavior Analysis in American Youth: Using a Large Sample

Deanna Schreiber-Gregory, North Dakota State University, Fargo, ND

Abstract

The current study looks at recent health trends and behavior analyses of youth in America. Data used in this analysis was provided by the Center for Disease Control and Prevention and gathered using the Youth Risk Behavior Surveillance System (YRBSS). This study outlines demographic differences in risk behaviors, health issues, and reported mental states. Interactions between risk behaviors and reported mental states were also analyzed. Results included reporting differences between the years of 1991 and 2011. All results are discussed in relation to current youth health trend issues. Data was analyzed using SAS 9.2. This paper is intended for any level of SAS user, introductory statistics will be used. This paper is also written to an audience with a background in psychology and/or statistics.

Introduction

The Youth Risk Behavior Surveillance System (YRBSS) was developed as a tool to help monitor priority risk behaviors that contribute substantially to death, disability, and social issues among American youth and young adults today. The YRBSS has been conducted biennially since 1991 and contains survey data from national, state, and local levels. The national Youth Risk Behavior Survey (YRBS) provides the public with data representative of the United States high school students. On the other hand, the state and local surveys provide data representative of high school students in states and school districts who also receive funding from the CDC through specified cooperative agreements. The YRBSS serves a number of different purposes. The system was originally designed to measure the prevalence of health-risk behaviors among high school students. It was also designed to assess whether these behaviors would increase, decrease, or stay the same over time. An additional purpose for the YRBSS is to have it examine the co-occurrence of different health-risk behaviors. This particular study exams the co-occurrence of suicidal ideation as an indicator of psychological unrest with other health-risk behaviors. The purpose of this study is to serve as an exercise in correlating two different variables across multiple years with large data sets.

Methods

YRBSS provided data sets free to the public online and instructions on how to download the data sets, as well as how to apply the formatting. In order to apply the formatting, the researcher needed only to specify libraries for the data sets and formats:

```
libname mydata 'I:\RRSC_MWSUG_Analytics_2012\MWSUG_2012';     /* Tells SAS where the data is */
libname library 'I:\RRSC_MWSUG_Analytics_2012\MWSUG_2012';     /* Tells SAS where the formats are */
```

This enabled SAS to read all the formatting as well as output the variable names, questions, and answers in a very clean manner

Concatenating Data Sets

In order for data from all of the years to be used in this analysis, concatenating the 11 data sets was necessary. The researcher chose which questions would be used in the analysis based on the whether or not the questions in all of the national surveys. Only questions consistently asked starting in the 1991 survey and included through the 2011 survey were included in this analysis in order to reduce the number of questions explored in each model. This is a very simplistic way of reducing variables and the researcher acknowledges that it can create some bias. However, for the purposes of this practice, this method was used. The questions used were then given new names in order for the appropriate questions to be concatenated together. This was necessary because even though the questions used were present in all the surveys, the order in which the questions appeared in the survey differed between each year. The coding to rename each of the variables for the 1991 survey is given below:

```
data YRBS1991;
    set mydata.YRBS1991;
    alcohol1=q33;     alcohol2=q34;     alcohol3=q35;     demo1=q1;
    demo2=q2;         demo3=q3;         demo4=q4;         demo5=' ';
    demo6=' ';        drugs1=q37;       drugs2=q38;       drugs3=q40;
    drugs4=q41;       drugs5=q43;       drugs12=q44;      drugs13=q45;
    exercise2=q68;    exercise3=q71;    exercise6=q74;    exercise7=q75;
```

```
        mood2=q19;        mood3=q20;        mood4=q21;        mood5=q22;
        sexuality2=q48;   sexuality3=q49;   sexuality4=q50;   sexuality5=q51;
        sexuality6=q52;   sexuality7=q55;   tobacco1=q23;     tobacco2=q28;
        tobacco3=q29;     tobacco4=q30;     tobacco5=q26;     tobacco6=q31;
        vehicle1=q10;     vehicle2=q6;      vehicle4=q11;     vehicle5=q12;
        vehicle7=q8;      violence1=q14;    violence8=q16;    violence9=q18;
        weight1=q57;      weight2=q58;
run;
```

The coding to concatenate the years together is given below:

```
data YRBS_Total;
    set   YRBS1991    YRBS1993    YRBS1995    YRBS1997    YRBS1999
          YRBS2001    YRBS2003    YRBS2005    YRBS2007    YRBS2009
          YRBS2011;
run;
```

Proc Surveyfreq

To begin the analysis, the researcher used proc surveyfreq to find the frequency of occurrence for each demographic variable in the data set. Frequencies of occurrence for the different answers to selected questions were also gathered though not used for the purpose of this study. The code used is provided below:

```
proc surveyfreq data=YRBS_Total;
        cluster psu;
        strata stratum;
        tables   demo1-demo6;
        weight weight;
    run;
```

Weighted values were used for each demographic variable in the data set in order to create a nationally representative analysis.

Proc Surveylogistic

The researcher then wanted to test to see if an interaction could be seen between mental illness and certain health-risk behaviors across the years. A logistic analysis was conducted for this purpose. The logistic analysis was written in a manner so that a multiple regression analysis could be performed, given that the particular variables used were categorical. Also, given that the variables used are in a complex survey format, surveylogistic was a necessary procedure to employ for this analysis as it accounts for complex survey designs. The options order=data and link=probit was used in order to control for the fact that the variable answers were given in an ordinal format defined within the data. The following SAS code was applied to the previously mentioned target question:

```
proc surveylogistic data=YRBS_Total;
        class mood2 mood3 mood4 mood5;
        cluster psu;
        strata stratum;
        model mood2 (order=data) = mood3 | mood4 | mood5 / link=probit;
        weight weight;
    run;
```

The variable mood2 represents the question "During the past 12 months, did you make a plan about how you would attempt suicide?" This question was chosen for this study as a representation of poor mental health. The above model was used to test correlations between suicidal thoughts and independent effects as well as interactions between the following variables: making a plan for suicide, number of times attempted suicide, and if severity of suicidal event lead to treatment by a medical professional. Each of these variables contain the values of each answer for their particular question. The results were outputted as Wald Chi-Square values.

Results

The results of the above SAS coding procedures are given below. The main interest of these results are the presence of significant correlations between the suicidal ideation variables and different health-risk behaviors. When viewing the Wald Chi-Square statistics and subsequent p-values, the researcher would like to note that the model parameters were set to 0 since each variable was a linear combination of other variables in the model.

Suicidal Ideation Correlations

For the model correlating suicidal ideation and alcohol use, the following questions were found to be significant: "During your life, on how many days have you had at least one drink of alcohol?" (Wald Chi Square = 102.89, p-value <.0001), and "During the past 30 days, on how many days did you have 5 or more drinks of alcohol in a row, that is, within a couple of hours?" (Wald Chi-Square = 41.64, p-value=<.0001).

In exploring the correlations between suicidal ideation and drug use, the following questions were found to be significant. "During your life, how many times have you used marijuana? (Wald Chi-Square= 21.28, p-value = .0003), "During your life, how many times have you used any form of cocaine, including powder, crack, or freebase?" (Wald Chi-Square = 33.55, p-value = <.0001), and "During the past 30 days, how many times have you sued any form of cocaine, including powder, crack, or freebase?" (Wald Chi-Square = 7.0075, p-value = .0081).

To explore the incident of suicidal thoughts leading to suicidal actions, the researcher explored correlations between suicidal ideation and self injurious behavior. Significant variables are given as follows: "During the past 12 months, did you make a plan about how you would attempt suicide?" (Wald Chi-Square: 25.05, p-value = <.0001), and "During the past 12 months, how many times did you actually attempt suicide?" (Wald Chi-Square = 20.87, p-value = <.0003).

For exploring correlations between suicidal ideation and sexual behaviors the following variables were found to be significant: "Have you ever had sexual intercourse?" (Wald Chi-Square = 19895.31, p-value = <.0001), "During the past 3 months, with how many people did you have sexual intercourse?" (Wald Chi-Square = 111.081, p-value = <.0001), and "Did you drink alcohol or use drugs before you had sexual intercourse the last time?" (Wald Chi-Square = 114.829, p-value = <.0001).

One of the greatest correlations between suicidal ideation and health-risk behaviors was found in tobacco use. The following significant variables in the tobacco use model are: "Have you ever tried cigarette smoking, even one or two puffs?" (Wald Chi-Square = 45.12, p-value = <.0001), "During the past 30 days, on how many days did you smoke cigarettes?" (Wald Chi-Square = 18.6671, p-value=.0093), "Have you ever tried to quit smoking cigarettes?" (Wald Chi-Square = 49.002, p-value = <.0001), "Have you ever smoked cigarettes daily (regularly), that is, at least one cigarette every day for 30 days?" (Wald Chi-Square = 34.1067, p-value = <.0001), and "During the past 30 days, did you use chewing tobacco, snuff, or dip?" (Wald Chi-Square = 79.4438, p-value = <.0001).

For exploring the correlations between suicidal ideation and vehicle safety, the following variables were significant: "During the past 30 days, how many times did you ride in a car or other vehicle driven by someone who had been drinking alcohol?" (Wald Chi-Square = 44.5917, p-value = <.0001), and "During the past 30 days, how many times did you drive a car or other vehicle when you had been drinking alcohol?" (Wald Chi-Square = 17.67, p-value = .0034).

Lastly, for exploring correlations between violence or injurious behavior and suicidal ideation, the following variables were found to be most significant: "During the past 30 days, on how many days did you carry a weapon such as a gun, knife, or club?" (Wald Chi-Square = 316.8309, p-value = <.0001), "During the past 12 months, how many times were you in a physical fight?" (Wald Chi-Square = 27.3977, p-value = .0006).

Conclusion

Given the above results from the first models correlating suicidal ideation and health-risk behaviors, we can see that this topic is worth looking further into it. A more complete model combining these significant variables along with demographic variables would logically be the next step in this analysis. Backward analysis could then be applied to slowly weed out less significant variables so that a final model representing the most significant variables contributing to suicidal ideation across the YRBSS analysis years could be obtained. More in depth analysis could then be conducted on variables not given throughout the years but that could possibly prove as significant correlates with suicidal ideation. Demographic variability across these variables could also be explored.

Acknowledgments

References

Introduction to SAS. UCLA: Academic Technolgy Services, Statistical Consulting Group. From
http://www.ats.ucla.edu/stat/sas/notes2/ (accessed August, 2012).

Sanchez, M. E. "Analyzing the effects of adolescent risky behaviors on suicidal ideation." (2006). *Mathematics Theses*. Paper 21. http://digitalarchive.gsu.edu/math_theses/21

YRBSS (2011). Retrieved April, 2012, from http://www.cdc.gove/yrbs

SAS Institute Inc. 2008. SAS/STAT® 9.2 *User's Guide*. Cary, NC: SAS Institute Inc.

Contact Information (Header 1)

Your comments, questions, and suggestions are valued and encouraged. Contact the author at:

Deanna Schreiber-Gregory
North Dakota State University
E-mail: annies94ranger@hotmail.com