

P8157 Analysis of Longitudinal Data, Fall 2019
Take Home Data Analysis Exam
Due Friday, December 6 Before 5pm

Alcohol has been a part of human culture since the beginning of recorded history. Almost all societies that consume alcohol show related health and social problems. The industrialization of production and globalization of marketing and promotion of alcohol have increased both the amount of worldwide consumption and the harms associated with it. These developments have led to an interest among public health professionals in outlining the public health problems caused by alcohol and possible strategies to reduce the harmful use of alcohol. The net effect of alcohol consumption on health is detrimental, with an estimated 3-8% of all global deaths and 4-6% of global disability-adjusted life-years attributable to alcohol.

Alcohol is viewed a risk factor for the global burden of disease and injury, especially for the so-called alcohol-use disorders—i.e., alcohol dependence and harmful use of alcohol as outlined in the International Statistical Classification of Disease tenth revision (ICD-10). Alcohol-use disorders, especially for men, are among the most disabling disease categories for the global burden of disease. However, this disease category is not the only one linked to alcohol: more than 30 ICD-10 classified diseases include alcohol in their name or definition, indicating that alcohol consumption is a necessary cause. Furthermore, more than 200 ICD-10 classified diseases exist in which alcohol is part of a component cause.

Alcohol consumption is a major risk factor for burden of disease. The average volume of alcohol consumption and patterns of drinking, especially heavy drinking occasions, contribute to this disease burden. Alcohol is linked to many disease categories, but alcohol-use disorders, cancer, cardiovascular disease, liver cirrhosis, and injury are the most important disease categories causally affected by alcohol. The net effect of alcohol on cardiovascular disease might be beneficial in regions in which alcohol is regularly consumed lightly to moderately without heavy drinking occasions, but this benefit is restricted to older people only. In other regions, where this is not the case, no net protective effect on ischaemic heart disease is expected, and the overall effect of alcohol on cardiovascular disease will be detrimental because of its harmful effect on haemorrhagic stroke and hypertensive disorders. Even in regions in which the net effect on cardiovascular disease is beneficial, the overall effect of alcohol on the burden of disease is detrimental. Globally, the effect of alcohol on burden of disease is about the same size as that of smoking, but it is greatest in developing countries. Further global consumption of alcohol is increasing, especially in the most populous countries of India and China.

In light of the above facts, a group of researchers conducted a study to compare three types of interventions and to see if these interventions were helpful in reducing alcohol dependence. A randomized trial was conducted as follows. 314 subjects known to suffer from alcohol dependence were recruited. The total number of drinks consumed in the 30 days immediately prior to randomization

was recorded. After the assignment of treatment, the total number of drinks consumed in the past 30 days were recored at 30 days and 60 days since the beginning of treatment (i.e. each observation is the total number of drinks consumed in the last 30 days preceding the day of reporting). The subjects were randomly assigned into one of three treatment groups as follows:

1. Treatment 1 : Subjects were given DVD, pamphlets and other instructional materials related alcohol dependency. The subjects kept track of their daily alcohol consumption and reported the total number of drinks consumed in the last 30 days on the 30th and 60th day.
2. Treatment 2 : Subjects were given DVD, pamphlets and other instructional materials related alcohol dependency but were also required to report the number of drinks consumed at the end of every week.
3. Treatment 3 : Subjects were given DVD, pamphlets and other instructional materials related alcohol dependency but were also required to report the number of drinks consumed daily using a toll-free phone number.

The subjects were also followed up 6 months after the end of treatment. Based on an a priori set criteria, the researchers classified the subjects as having relapsed into alcohol dependence or not.

The data are available in the file ALCDEP on *Courseworks* as an EXCEL and as a tab-delimited text file. The data are presented in the form of one data record per subject. The columns of the data set are as follows:

1. sid - Subject ID
2. Treatment - Treatment type (1,2,3)
3. Gender - (0=Male, 1=Female)
4. ND0 - Total Number of drinks consumed in the 30 days prior to randomization
5. ND30 - Total Number of drinks consumed in the 30 days following randomization (beginning of treatment)
6. ND60 - Total Number of drinks consumed in the 30 days between the 30th and 60th day follow-up.
7. Relapse - Relapse after 6 months (1=Yes, 0=No)

The researchers had several questions, stated as follows:

- Is there evidence to suggest that the treatments differ in their effects on alcohol dependence, as reflected by the number of drinks consumed in a given 30 day period?
- Is there a difference in the pattern of change in the number of drinks consumed between the various treatment groups over the duration of the study?
- Alcohol-use disorders are among the most disabling disease categories for the global burden of disease especially for men. Is there evidence to suggest that males tend to have a higher alcohol dependence than females?
- Do men and women respond differently to treatment ?

- Is there any evidence to suggest that the treatments differ in their effects on subjects with regard to relapsing into alcohol dependence ?
- Even in the case that the treatments might differ in their pattern of change or on how subjects relapse into alcohol dependence, is there any evidence to suggest that any of the treatments might be beneficial once the treatment has stopped.

As the statistical consultant, you have been called in to carry out an appropriate analysis of the data from the study in order to address these questions. Your job is to conduct a thorough analysis and write a report for the researchers describing what you did, why you did it, and what inferences may be drawn regarding the questions of interest. Although the researchers know about some basic statistical methods such as t-tests and linear regression, they have limited knowledge of any other methods. Explain your results in such a way that the researchers can understand with minimal effort on their part to post-process any results.

This is a “**closed**” take-home exam. Thus, you **may not** collaborate with or discuss the assignment with any other person (whether in the class or not), **nor** may you consult with anyone. You **may** use the class notes, books, or any other reading material you like. **Please review the Project Guidelines.**