Dean&Francis
ISSN 2959-409X

The Relationship Between Alcohol Consumption and Liver Status, For US People in Age 18-64, Control Sex and Modified by BMI And Age

Yue Yang[1*,]

[1]Department of Global Public Health, New York University, New York, NY 10012, USA
*Corresponding author: yy5381@nyu.edu

Abstract:

Alcohol consumption and liver function have become increasingly important issues in modern society. Numerous studies have investigated their relationship; this study focuses on U.S. adults from 2017 to 2020 to examine the association between alcohol consumption and liver function, assessed by alanine aminotransferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transferase (GGT), while controlling for sex and evaluating modification by body mass index (BMI) and age. Linear regression models were applied to estimate associations. The results indicate that alcohol consumption is strongly associated with liver enzyme levels, and this relationship is confounded by sex and modified by BMI. These findings emphasize the importance of considering demographic and metabolic factors when evaluating alcohol-related liver injury and provide new evidence for targeted prevention and intervention strategies in public health. Future research should adopt longitudinal designs to clarify causal pathways and quantify the potential benefits of integrated prevention strategies.

Keywords: Alcohol Consumption; Liver Status; Linear Regression.

1. Introduction

With global economic growth, alcohol consumption has become increasingly prevalent worldwide. According to the World Health Organization (WHO), alcohol consumption is a major public health issue and contributes substantially to the global burden of liver disease. Especially the United States, some research by Nehring et al in 2025 point out that, the alcohol is the most popular consumption for the people age after 12 in US [1]. Thus, the potential health risks of alcohol use have attracted increasing global research attention. Some of the research point out that alcohol consumption has been identified as an important risk factor for illness, disability, and mortality [2]. This underscores the multifaceted impact of alcohol on

population health. Also, some research by suggest that the alcohol also relate to the burden of alcohol-attributable liver cirrhosis and liver cancer is high and entirely preventable [3]. Even some of the people suggest that the drink in moderation is good for health, but other research by Michael Roerecke and Jürgen Rehm suggest that the even a few of alcohol can also change the liver status [4]. So, it is important for us to get a scene of how the alcohol can affect the people's liver status and cause the potential harm and threaten from some liver disease.

There a lot of way to have a test for liver to shows its status, so choose some of the function and index is really important. Clinical evaluation of liver health often relies on serum biomarkers such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transferase (GGT) and by van Beek, the ALT, AST and GGT is some of the very important information data to shows the liver status of the people [5]. So, in this research, the dataset is the NHANES dataset form 2017-March 2020 to get the both alcohol consumption data and ALT, AST and GGT data from United state to explore the relationship between the alcohol consumption and the liver function maker. All the dataset is from NHNAES because the NHANES is the most authoritative and credible dataset which contain the whole United State dataset. And it is also is a nationally representative data on health behaviors, laboratory indicators, and clinical outcomes. We use the 2017-2020 dataset which is the pre-pandemic because it is the newest dataset which contain the Standard Biochemistry Profile dataset. After the Covid-19, some of the data collections are not built yet, so the newest dataset from 2021-2023 do not contain these liver index.

In this research, we use the NHANES dataset to find the relationship between alcohol use and liver index. We seek to provide insights to guide public health policy, screening strategies, and clinical counseling regarding alcohol-related liver risks. Higher alcohol consumption is significantly associated with elevated ALT, AST, and GGT levels, with sex as a confounder and BMI as a modifier, and these findings are consistent with previous studies [6–10].

2. Method

2.1 Dataset Background

2.1.1 Dataset and Population

The National Health and Nutrition Examination Survey (NHANES) aims to provide nationally representative estimates of health and nutrition status among U.S. citizens. Because we are using the NHANES dataset, our study is a cross-sectional study. Here we use the datasets combining the NHANES demographic dataset, NHANES Body Measures dataset, NHANES alcohol questionaire data and the NHANES Standard Biochemistry Profile question dataset. These two datasets are all from 2017 to March 2020 as a sample. The data are collected for the U.S. civilian non-institutionalized population. The data are collected from both personal interviews and mobile inspection centers for physical examination. The data collection is using a four-step cluster, which from countries or groups of countries, segments within countries, households within countries and the individual. The data does some of the oversampling with some kinds of people, like Hispanic, Black, Asian and others to get enough data. In the sample discussion, we can see that a total of 15560 were interviewed and some of them completed both the interview and the health examination component with a clinician-administered medical examination. It is a probability-based sampling approach because the cluster sample is a probability-based sampling. The data also uses stratification in the sample. And the data design is from the file "Plan and Operations of the National Health and Nutrition Examination Survey, August 2021–August 2023 File: Sample Design, Estimation, and Analytic Guidelines,".

2.1.2 Variables Summary

We use the Alcohol questionnaire as our independent variable. We use the answer from ALQ121 which is the how often did the responder drink. It is very hard to calculate the precise drink number for people so it is better to just use this people drink rate as our research variable. For liver index: ALT, AST, GGT, the ALT unit is U/L from 2 to 682, the AST unit is U/L from 6 to 489 and the GGT unit is IU/L from 2 to 2394. All methods were measured on the Roche Cobas 6000 (c501 module) analyzer.

We also have possible confounder and EMM. We want to control them to make more precise here. The possible confounder is sex and EMM can be BMI and age. The BMI is calculate with the weight divide by square of height. And the square of height is in the Mobile Examination Center (MEC), by trained health technicians. The bmi is from 11.9 to 92.3 in kg/m^2. The other variables like age, sex are from the using Computer-Assisted Personal Interview (CAPI) system. The sex include male and female. The age is from 0 to 80 but we only use 18 to 64.

2.2 Data Analysis and Model Fitting

Firstly, we are going to make the Descriptive (univariable) analysis which contains the means (SD) or medians (IQR or range) for continuous variables and frequencies and percentages for categorical variables. Then we can do the Bivariate analysis for both exposure and the category and

ISSN 2959-409X

numerical of outcome. The Bivariate analysis not only contain the old descriptive analysis result, but also contain new means (SD) or medians (IQR or range) for continuous variables and frequencies and percentages for categorical variables. We will use the ANOVA test for category variable vs numerical variable we use the chi-square test for 2 category variables and we use the linear regression test for two numerical variables and report their Pearson correlation coefficient. We can find the confounder suggestion and how different variables are distributed in this analysis.

Then, the Multivariable analysis will establish. For linear regression, the model will use numerical variable ALT, AST, GGT as the outcome and the exposure alcohol level, confounder sex and EMM age and BMI to be included in the model. The model will first use the variable itself to create the separate model (crude model). Then the model with the confounder is established, EMM and exposure to compare their coefficient (full model 1). If the change of coefficient is greater than 10%, we say the confounder exists. At last, the interaction variable between the exposure alcohol take and EMM sex are added in the model to determine is this interaction term is significant or not to justice EMM by check the p-value of the interaction term (full model 2). All analyses accounted for the complex survey design of NHANES, incorporating sampling weights, clustering, and stratification to ensure nationally representative estimates.

3. Result

From the EDA, the data distribution is clear. The most people have medium alcohol use 2151(40.92%), then is low1532(29.14%), never (16.32%), 715(13.6%). The majority of participants were aged 50–64 years (1,917, 36.47%), followed by 18–29 years (1,238, 23.55%), 40–49 years (1,071, 20.38%), and 30–39 years (1,030, 19.6%). More people here are female 2647(50.36%) than male 2609(49.64%). For ALT, the mean is 23.41 and AST mean is 22.09, the mean of GGT is 32.66. The BMI here is mean 30.23 (Table 1).

Table 1. Univariable analysis

Variable name	Percent (%)	Variable name	Percent (%)
Alcohol		Alcohol	
High	715(13.6%)	High	715(13.6%)
Medium	2151(40.92%)	Medium	2151(40.92%)
Low	1532(29.14%)	Low	1532(29.14%)
never	858(16.32%)	never	858(16.32%)
Age_group		Age_group	
18-29	1238(23.55%)	18-29	1238(23.55%)
30-39	1030(19.60%)	30-39	1030(19.60%)
40-49	1071(20.38%)	40-49	1071(20.38%)
50-64	1917(36.47%)	50-64	1917(36.47%)
Sex		Sex	
Male	2609(49.64%)	Male	2609(49.64%)
Female	2647(50.36%)	Female	2647(50.36%)

Table 2 shows that the 18–29 age group had the highest proportion of medium alcohol users (625, 50.48%), with a statistically significant difference compared to other age groups. The male always has medium alcohol consumption (1119, 42.89%) and this different is statistic significant. The ALT mean is greatest at high alcohol use people (27.07) and the AST (26.63) and GGT (50.93) is also greatest for high alcohol use people and this different is statistic significant. Participants with low alcohol use exhibited the highest mean BMI (31.11 kg/m²), which was also statistically significant.

Table 2. Bivariate analysis

Alcohol consumption					p-value
	High	Medium	Low	Never	
Age_group					
18-29	115(9.29%)	625(50.48%)	412(33.28%)	86(6.95%)	
30-39	152(14.76%)	427(41.46%)	329(31.94%)	122(11.84%)	<0.001
40-49	147(13.73%)	437(40.80%)	316(29.51%)	171(15.97%)	
50-64	301(15.70%)	662(34.53%)	475(24.78%)	479(24.97%)	
Sex					
Male	473(18.13%)	1119(42.89%)	604(23.15%)	413(15.83%)	<0.001
Female	242(9.14%)	1032(38.99%)	928(35.06%)	445(16.81%)	
Alt					
Mean	27.07	23.59	21.99	22.46	<0.001
median	21	18	17	18	
AST					
Mean	26.63	22.11	20.49	21.10	<0.001
median	21	19	18	18.5	
GGT					
Mean	50.93	31.03	27.17	31.31	<0.001
median	28	21	19	32	
BMI					
Mean	28.37	29.9	31.11	31.01	<0.001
Median	27.3	28.6	29.60	29.55	

In the linear regression model, we use the three linear regression model to talk about ALT AST GGT separately. For ALT, compare to the high-level alcohol use, the low use people have 5.08 less of alt (b=-5.08, std=0.9364). The sex can be treated as confounder because after add it, at least one of the coefficients of an alcohol use change more than 10%. Effect modification analysis indicated that BMI significantly influenced the relationship between alcohol use and liver enzymes (all $p < 0.05$). Evidence for interaction with age groups was limited, as only one comparison reached statistical significance ($p = 0.039$) For AST, compare to the high-level alcohol use, the low use people have 6.14 less of alt (b=-6.14, std=0.6793). The sex can be treated as confounder because after add it, at least one of the coefficients of an alcohol use change more than 10%. For EMM, it seemed that the BMI do have influence on alcohol use with all p-value <0.05. But for age group, it is hard to say that there is an interaction relationship because only three of them have low p-value (0.01,0.007,0.001). The interaction plot shows there might be a little influence with 18-29 age group, but totally is seemed same. For GGT, compare to the high-level alcohol use, the low use people have 20.53 less of alt (b=-30.53, std=3.23). The sex can be treated as confounder because after add it, at least one of the coefficients of an alcohol use change more than 10%. For EMM, it seemed that the BMI do have influence on alcohol use with all p-value <0.05. But for age group, it is hard to say that there is an interaction relationship because only three of them have low p-value (0.008,0.009,0.04). Also, the interaction plot shows totally all age group are seemed same.

4. Summary

As a conclusion, the higher alcohol use level people have a higher ALT, AST and GGT level. And this association and confound by sex, modified by BMI. The high level of alcohol use, the people in the sample are have a higher value of these data and this different is statistic significant. The result of this paper is similar with past research like Roerecke's research, Patel R 's paper, Desalegn's research, Moon SY's research and Bryazka's research. Also, the sex is confounder of data and BMI also influence this relationship. This can help people to explore more about the reason of the problem of liver function among the people in the US. However, However, the research does have some limitations. As we say, there are about half (3709) people in the data who have a NA of our different variable. So, such a great amount of missing data might cause a big bias in the data. Also, because we are using a cross-sectional study, we cannot determine that the relationship between the alcohol use and the different ALT

ISSN 2959-409X

AST and GGT. In the future, more longitudinal cohort studies should be used, the research can use more advance tools like wearable devices and mobile health applications.

References

[1] Nehring SM, Chen RJ, Freeman AM. Alcohol Use Disorder: Screening, Evaluation, and Management. [Updated 2024 Mar 16]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK436003/.

[2] Rehm J, Parry C. Alcohol consumption and infectious diseases in South Africa. Lancet. 2009;374(9707):2053. doi: 10.1016/S0140-6736(09)62150-4.

[3] Global burden of alcoholic liver diseases Rehm, Jürgen et al.Journal of Hepatology, Volume 59, Issue 1, 160 – 168.

[4] Michael Roerecke, Jürgen Rehm, Irregular Heavy Drinking Occasions and Risk of Ischemic Heart Disease: A Systematic Review and Meta-Analysis, American Journal of Epidemiology, Volume 171, Issue 6, 15 March 2010, Pages 633–644.

[5] van Beek JH, de Moor MH, de Geus EJ, et al. The genetic architecture of liver enzyme levels: GGT, ALT and AST. Behav Genet. 2013;43(4):329-339. doi:10.1007/s10519-013-9593-y.

[6] Roerecke M, Vafaei A, Hasan OSM, et al. Alcohol Consumption and Risk of Liver Cirrhosis: A Systematic Review and Meta-Analysis. *Am J Gastroenterol*. 2019;114(10):1574-1586. doi:10.14309/ajg.0000000000000340.

[7] Patel R, Mueller M. Alcohol-Associated Liver Disease. [Updated 2023 Jul 13]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-.

[8] Desalegn, H., Diaz, L. A., Rehm, J., & Arab, J. P. (2024). Impact of alcohol use on liver disease outcomes. Clinical liver disease, 23(1), e0192.

[9] Moon SY, Son M, Kang YW, Koh M, Lee JY, Baek YH. Alcohol consumption and the risk of liver disease: a nationwide, population-based study. Front Med (Lausanne). 2023 Nov 28;10:1290266. doi: 10.3389/fmed.2023.1290266. PMID: 38089863; PMCID: PMC10713769.

[10] Population-level risks of alcohol consumption by amount, geography, age, sex, and year: a systematic analysis for the Global Burden of Disease Study 2020