Trends and Disparities in Cirrhosis Mortality and Alcohol Use Prevalence in the United States [2002 - 2012]

INFO 201 - Final Project Presentation

Group 2AA Members: Solvvy Liu, Vincent Ma, Sophie Whikehart MPHc, Lenny Zhao and Jiayi Zhu

[8/7/24]

Abstract

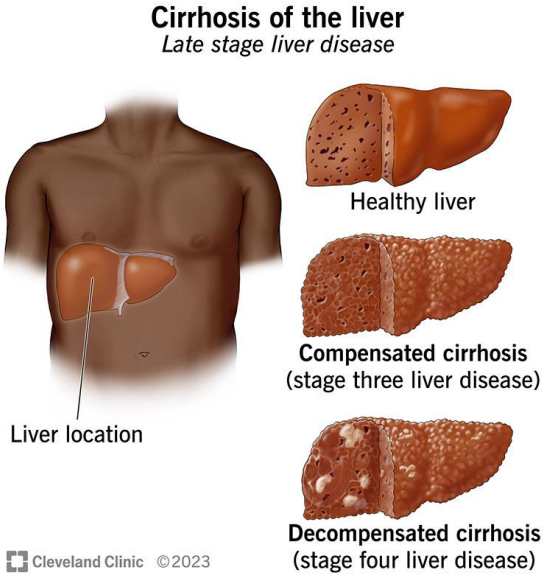
- Our project analyzed **trends** and **disparities** in **cirrhosis mortality** rates and **alcohol use prevalence** in the United States from 2002 to 2012.
- **Cirrhosis mortality** rate estimates were generated using population and death data from the National Center of Health Statistics.
- **Alcohol use prevalence** estimates was produced by applying small area models to data from Behavioral Risk Factor Surveillance System from 2002 - 2012.
- Our analysis used data from the **Institute of Health Metrics and Evaluation [IHME] Global Health Data Exchange [GHDx]**.
- Our analysis revealed significant differences in mortality rates by race/ethnicity, alcohol use and geographic location.

IHME

Centers for Disease Control and Prevention
National Center for Health Statistics

Introduction

- Cirrhosis is a serious condition where scar tissue replaces healthy liver tissue often caused by chronic alcohol abuse, viral hepatitis [i.e - hepatitis B and C] or other liver diseases (*Cirrhosis - Symptoms and Causes*, n.d.).
- Cirrhosis leads to liver failure and is a significant contributor to mortality globally.
- Studies have shown that In the US, cirrhosis mortality rates vary by different demographic factors such as race/ethnicity, age and geographic location (Nassereldine et al., 2024).
- Understanding the disparities in cirrhosis mortality rates is crucial for effective public health interventions.
- Cirrhosis is a major public health issue in the United States with increasing mortality rates over the past two decades.



Research Questions

- The project addresses the following research questions:

  1. How how cirrhosis mortality rates changed from 2002 to 2012 in the United States overall?

  2. What are the trends in cirrhosis rates by race and ethnicity and how do these trends differ across states?

  3. Are there notable differences in cirrhosis mortality rates between rural and urban counties?

  4. How does age affect cirrhosis mortality rates and has there been a shift in age distribution among cirrhosis related deaths over time?

  5. What regional patterns emerge in cirrhosis mortality rates and what factors might contribute to these variations?

Data Explanation

- Dataset 1: **IHME United States Cirrhosis Mortality by County, Race and Ethnicity 2000-2019**.
  - Cirrhosis mortality rate estimates were produced at the county level in the US, by racial and ethnic populations for each year between 2000-2019.
    - Estimates were generated using population and deaths data from the National Center for Health Statistics and extracted by Institute of Health Metrics and Evaluation.
  - https://ghdx.healthdata.org/record/ihme-data/us-cirrhosis-county-race-ethnicity-2000-2019
  - Data can be accessed by the public on IHME website using their Global Health Data Exchange database [GHDx].
  - Dataset includes:
    - CSV file of county, state and national level estimates of cirrhosis mortality rates for each age group, sex, year and racial and ethnic populations.
- Dataset 2: **IHME United States Alcohol Use Prevalence by County 2002-2012**
  - IHME research produced alcohol use prevalence estimates by county, year and sex for 2002-2012.
  - Estimates were produced by applying small area models to data from the Behavioral RIsk Factor Surveillance System [BRFSS]
- https://ghdx.healthdata.org/record/ihme-data/united-states-alcohol-use-prevalence-county-2002-2012
  - Dataset includes
    - CSV file of estimates for all states and counties for any drinking, heavy drinking and binge drinking
- We had some issues merging the two datasets because the years and location and location_name was not consistent between the two datasets

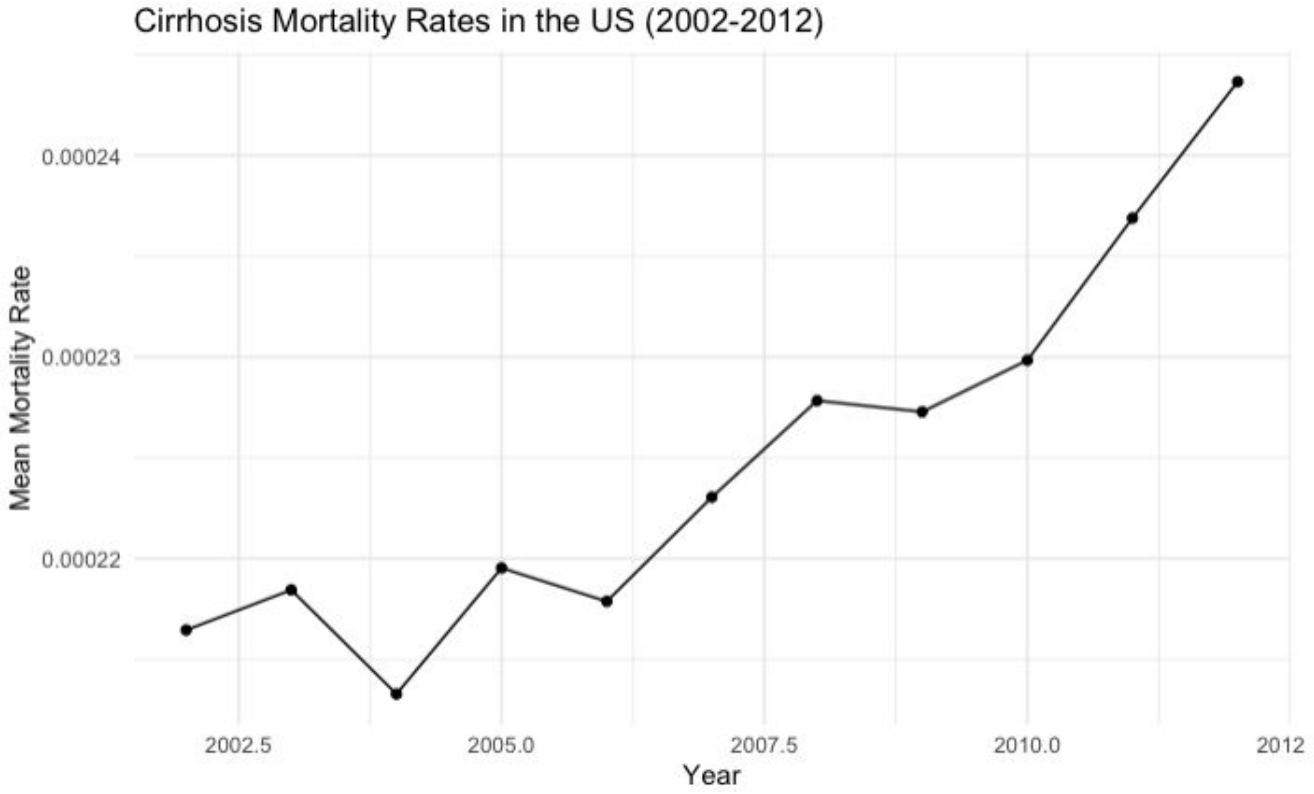
Data Description

- Cirrhosis mortality cleaned dataset included
  - Rows: 2280432 Columns: 9
- Alcohol prevalence cleaned dataset included
  - Rows: 9537 Columns: 42
- Combined dataset includes:
  - Rows: 8841 Columns: 31
- Each row in our dataset represents either cirrhosis mortality data for specific location and year or alcohol use prevalence data for a specific location for a year from 2002 - 2012
- Relevant variables are:
  - State [State]
  - Location [Location]
  - Location [location_name]
  - Year [year]
  - Cirrhosis mortality rate [val]
  - Alcohol prevalence from 2002 - 2012 Both sexes
- There were 10,798,896 missing values that we removed during data cleaning from the full datasets.

Methods

- Data cleaning & preprocessing
  - Our first dataset was in a folder with multiple csv files for each year so we needed to loop through the folder and merge all the files from 2002:2012
  - Our second dataset was one file but we created a new column called “location_name” which we would later use for merging
  - We cleaned and removed NA values
- Categorical and numerical variable creation
  - We created a new categorical variable for alcohol use prevalence
  - We also created a continuous numerical variable for change in alcohol use from 2002 - 2012 by State
  - We then created a summarization data frame with these numbers and visualization
- Complex statistical methods
  - For question 3 we wanted to analyze if there was a correlation between cirrhosis mortality rates and alcohol use prevalence in 2012 so we used a linear regression model
    - Average mortality by state was created by filtering for the year 2012 and grouping by state then summarizing the average cirrhosis mortality rate

1. How have cirrhosis mortality rates changed from 2002 to 2012 in the United States overall?



2. What is the average alcohol use in the US in 2002 & 2012 and what is the change of alcohol use from 2002 - 2012 for both sexes by State?

2002

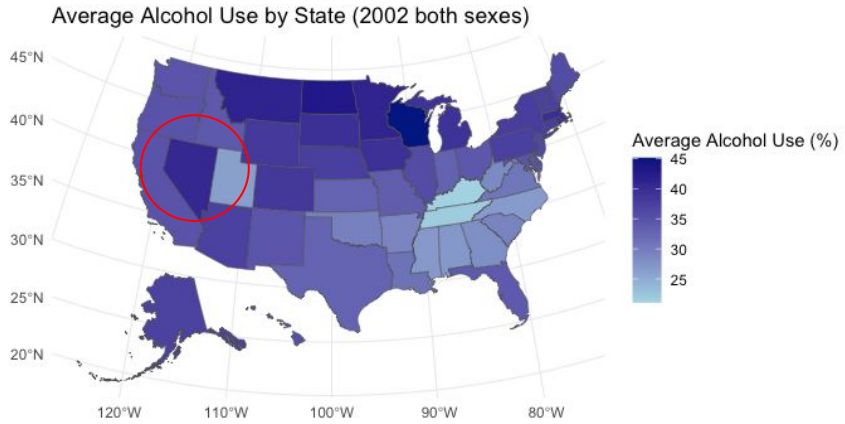
State <chr>	average_alcohol_use <dbl>
Wisconsin	45.10137
North Dakota	42.80988
Montana	41.35497
Minnesota	41.18902
Nevada	40.65000
Massachusetts	39.81556
Rhode Island	39.67778
South Dakota	39.63881
Iowa	39.55500
Michigan	39.21032

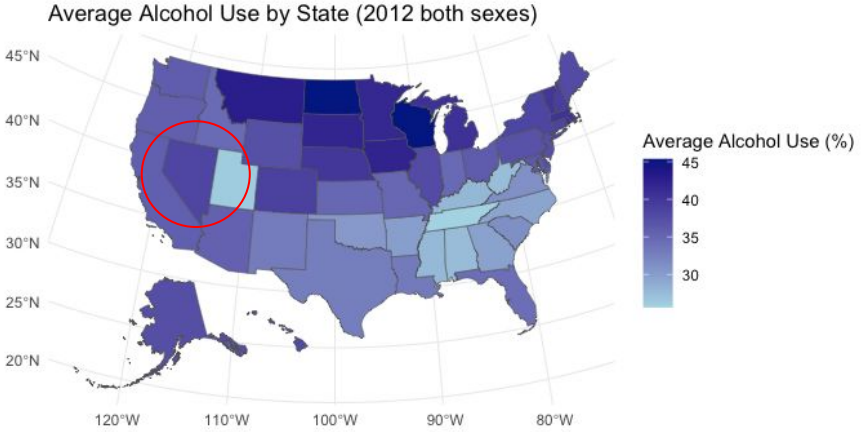
2012

State <chr>	average_alcohol_use <dbl>
Wisconsin	45.36027
North Dakota	45.33951
Montana	42.82164
Iowa	42.08667
South Dakota	41.98706
Minnesota	41.75644
Michigan	40.58889
Rhode Island	40.54444
Massachusetts	40.35333
District Of Columbia	40.30000

2002 - 2012

State <chr>	average_alcohol_change <dbl>
Wisconsin	0.25890411
North Dakota	2.52962963
Montana	1.46666667
Iowa	2.53166667
South Dakota	2.34825871
Minnesota	0.56742424
Michigan	1.37857143
Rhode Island	0.86666667
Massachusetts	0.53777778
District Of Columbia	3.26666667


Average Alcohol Use by State (2002 both sexes)
45°N
40°N
35°N
30°N
25°N
20°N
120°W
110°W
100°W
90°W
80°W
Average Alcohol Use (%)
45
40
35
30
25

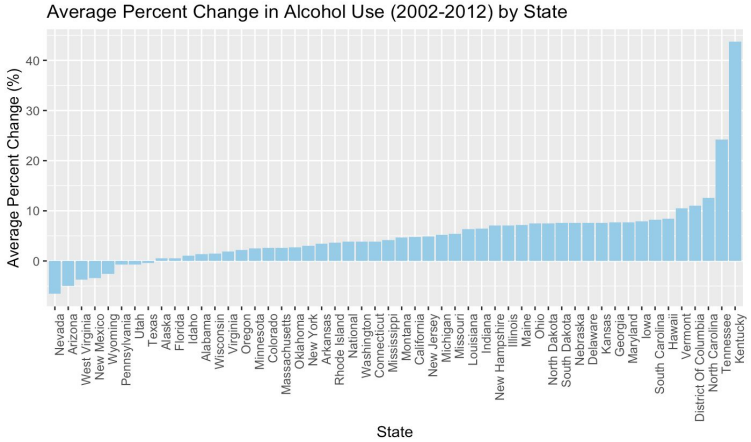

Average Alcohol Use by State (2012 both sexes)
45°N
40°N
35°N
30°N
25°N
20°N
120°W
110°W
100°W
90°W
80°W
Average Alcohol Use (%)
45
40
35
30

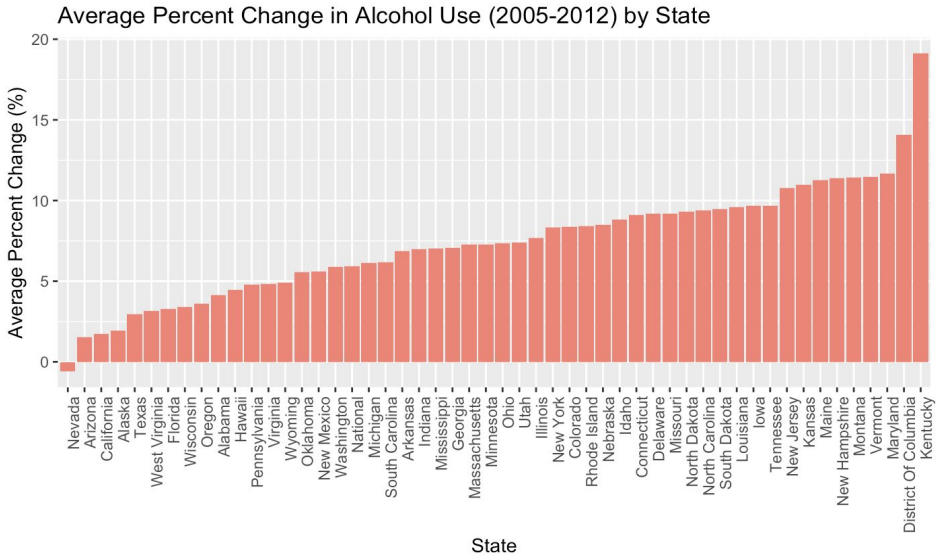
3. What is the average percent change in alcohol use from 2002 to 2012 and from 2005 to 2012 for both sexes, females and males by State?

A tibble: 52 × 3

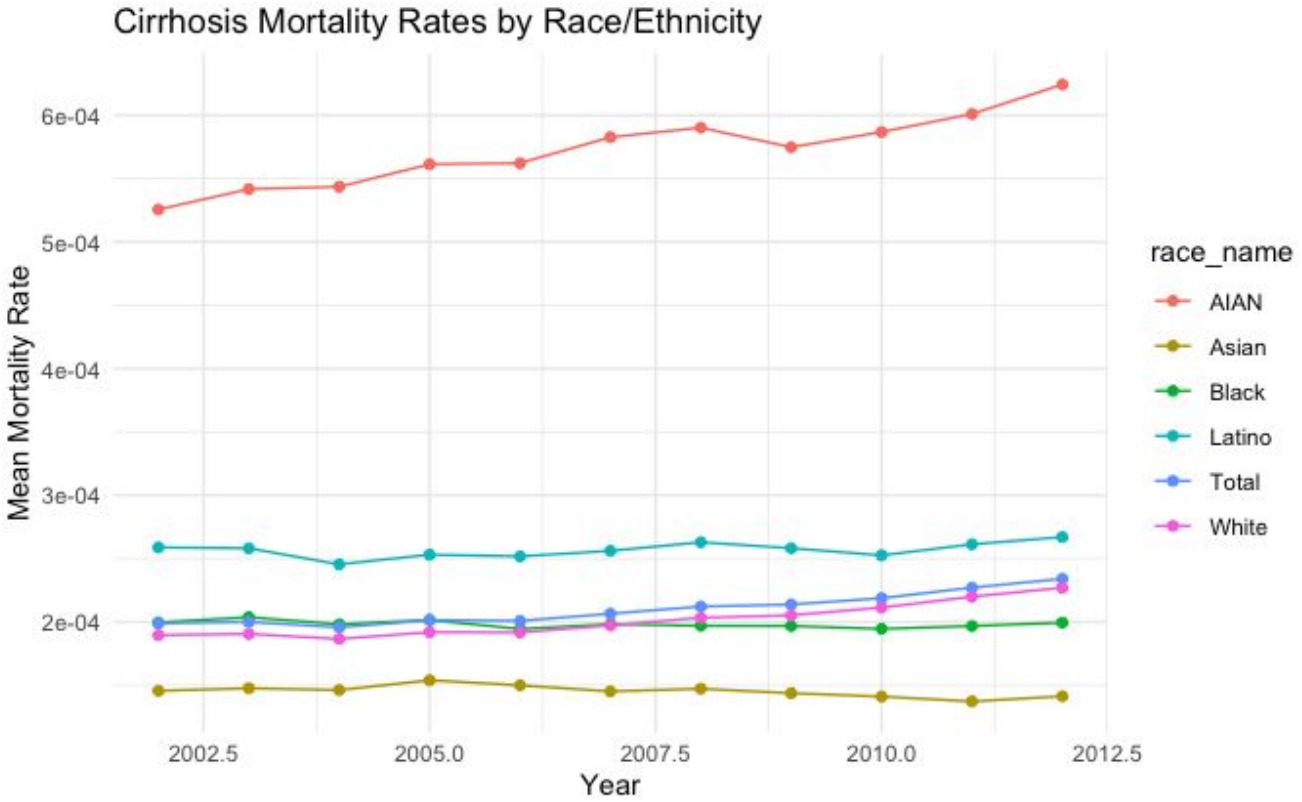
State <chr>	avg_percent_change_2002_2012 <dbl>	avg_percent_change_2005_2012 <dbl>
Kentucky	43.7245179	19.1236915
Tennessee	24.2354167	9.6916667
North Carolina	12.5825083	9.3970297
District Of Columbia	11.0666667	14.0666667
Vermont	10.4888889	11.4577778
Hawaii	8.3933333	4.4533333
South Carolina	8.1716312	6.1836879
Iowa	7.9073333	9.6873333
Maryland	7.7146667	11.6573333
Georgia	7.7112500	7.0660417

1–10 of 52 rows





4. What are the trends in cirrhosis rates by race and ethnicity?



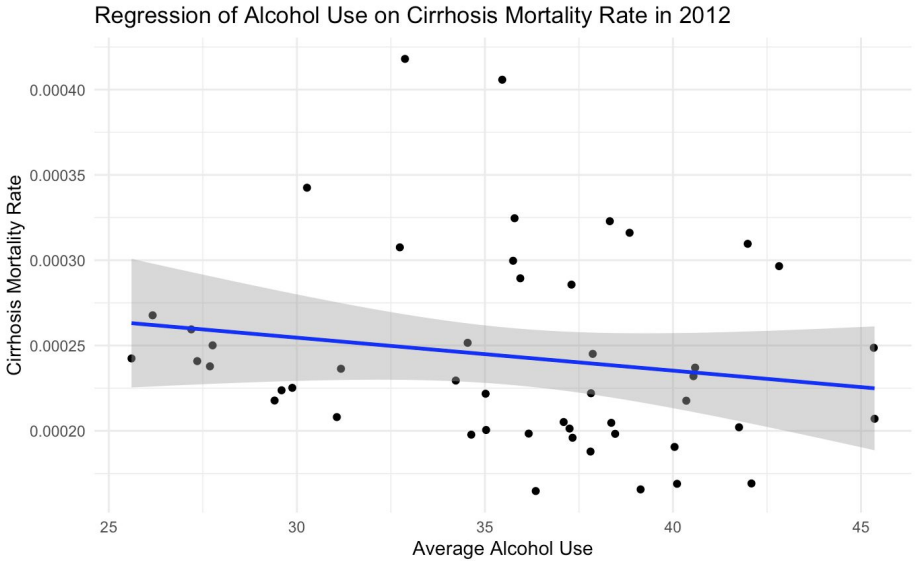
5. What is the correlation between alcohol prevalence and cirrhosis mortality rates in 2012?

```
Call:
lm(formula = average_mortality ~ average_alcohol_use, data = merged_data)

Residuals:
       Min         1Q     Median         3Q        Max
-7.763e-05 -3.946e-05 -1.767e-05  2.909e-05  1.690e-04

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          3.126e-04  5.988e-05   5.220 4.18e-06 ***
average_alcohol_use -1.933e-06  1.661e-06  -1.164    0.251
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.741e-05 on 46 degrees of freedom
Multiple R-squared:  0.0286,    Adjusted R-squared:  0.007485
F-statistic: 1.354 on 1 and 46 DF,  p-value: 0.2505
```



Findings

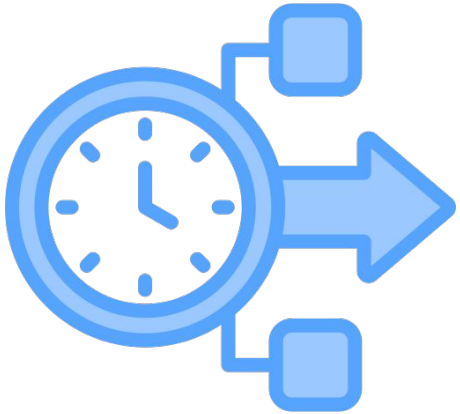
- **As we worked with the data our research questions changed a bit**
    1. How have cirrhosis mortality rates changed from 2002 to 2012 in the United States overall?
    2. What is the average alcohol use in the US in 2012 and what is the change of alcohol use from 2002 - 2012 for both sexes by State?
    3. What is the average percent change in alcohol use from 2002 to 2012 and from 2005 to 2012 for both sexes, females and males?
    4. What are the trends in cirrhosis rates by race and ethnicity?
    5. What is the correlation between alcohol prevalence and cirrhosis mortality rates in 2012?
- **We have addressed these questions as we went through our graphs.**

Limitations?

- Regional Differences
- Demographic Factors
- Quality of Data
- And more !

Future work?

- Additional Risk Factors
- Demographic Subgroup Analysis
- Machine Learning Models
- And more!



Summary

- Cirrhosis mortality is not solely driven by alcohol use, but is also influenced by broader social determinants of health, which must be considered in any effort to reduce these rates.

- We observed significant increase in cirrhosis mortality rates by ethnicity.

- Geographically we saw certain states have higher alcohol use but did not always have the highest cirrhosis mortality rates.

Thank You | Q & A

Sources

Cirrhosis—Symptoms and causes. (n.d.). Mayo Clinic. Retrieved August 7, 2024, from https://www.mayoclinic.org/diseases-conditions/cirrhosis/symptoms-causes/syc-20351487

Nassereldine, H., Compton, K., Li, Z., Baumann, M. M., Kelly, Y. O., Motte-Kerr, W. L., Daoud, F., Rodriquez, E. J., Mensah, G. A., Nápoles, A. M., Pérez-Stable, E. J., Murray, C. J. L., Mokdad, A. H., & Dwyer-Lindgren, L. (2024). The burden of cirrhosis mortality by county, race, and ethnicity in the USA, 2000–19: A systematic analysis of health disparities. *The Lancet Public Health*, *9*(8), e551–e563. https://doi.org/10.1016/S2468-2667(24)00131-2

United States Alcohol Use Prevalence by County 2002-2012 | GHDx. (n.d.). Retrieved August 7, 2024, from https://ghdx.healthdata.org/record/ihme-data/united-states-alcohol-use-prevalence-county-2002-2012

United States Cirrhosis Mortality by County, Race, and Ethnicity 2000-2019 | GHDx. (n.d.). Retrieved August 7, 2024, from https://ghdx.healthdata.org/record/ihme-data/us-cirrhosis-county-race-ethnicity-2000-2019