1. Full Title: Epidemiology of Liver-Related Hospitalizations in a community-based population

b. Abbreviated Title (Length 26 characters): risk of Liver-related hospitalizations

2. Writing Group:
   Writing group members: Mariana Lazo; Jeanne M. Clark; Lynne Wagenknecht; Elizabeth Selvin; others welcome.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _ML_ [please confirm with your initials electronically or in writing]

First author: Mariana Lazo, MD, PhD
Address: Department of Medicine and Epidemiology
          Welch Center for Prevention, Epidemiology & Clinical Research
          Johns Hopkins School of Public Health
          2024 E. Monument St. Suite 2-600
          Baltimore, MD 21287
          Phone: 410-614-4096. Fax: 410-955-0476
          E-mail: mlazo@jhsph.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: Elizabeth Selvin, PhD, MPH
Address: Department of Epidemiology
          Johns Hopkins Bloomberg School of Public Health
          Welch Center for Prevention, Epidemiology, and Clinical Research
          2024 E Monument Street, 2-600
          Phone: 410-614-3752 Fax: 410-955-0476
          E-mail: lselvin@jhsph.edu
3. **Timeline:** We expect to complete the analyses and submit the manuscript to the ARIC publications committee within 6 months after the approval of the manuscript proposals.

4. **Rationale:**

Over the last two decades liver-related mortality ranked among the top 12 causes of death in the U.S., with approximately 34,000 deaths per year, and recent mortality statistics suggest a 3.2% increase from 2010-2011(1). The three main risk factors for chronic liver disease are: viral hepatitis, alcohol consumption and obesity. While the burden of viral-related liver disease has decreased over time, liver disease associated to obesity and diabetes, commonly referred to as nonalcoholic fatty liver disease (NAFLD) has been increasing. NAFLD is now considered the most common chronic liver condition in the U.S. and other western countries, with prevalence estimates ranging between 15-30% in U.S. adults. Indeed, NAFLD is now the third most common indication for liver transplantation in the US, accounting for up to 10% of cases(2-4).

Long-term prospective studies assessing the burden of and risk factors for liver disease and liver-disease related hospitalizations in the community are scarce. Factors that have been previously associated with liver disease include age, alcohol consumption, obesity, insulin resistance, type 2 diabetes, and weight gain (5)(6-15). And given that the development of end-stage liver disease (liver disease requiring liver transplantation) may take decades to occur, a more thorough characterization of the liver-related morbidity in the general population is relevant for the design of interventions to alter the incidence and progression of liver diseases.

Strictly speaking, liver biopsy is the gold standard method to diagnose and stage liver diseases (16;17). However, for practical reasons, liver enzymes levels (alanine aminotransferase, aspartate aminotransferase and gamma-glutamyl tranferase) together with clinical data (e.g. absence of alcohol consumption) are frequently used in clinical practice and research as surrogate markers of the disease.

Using data on liver enzymes and hospital discharges from the Atherosclerosis Risk in Communities (ARIC) Study, we aim: 1) to quantify the absolute and relative risk of liver-related hospitalizations in a biracial community sample of the US, by sex, race, and age , 2) to determine the association between modifiable factors (alcohol consumption, obesity, diabetes, hypertension) and the risk of liver-related hospitalization 3) To characterize the association between levels of liver enzymes and the risk of hospitalization.

We foresee presenting the results of aims 1 and 2 in one paper, and for aim 3 we will write a separate manuscript, the rationale for doing this is twofold: 1) in ARIC the levels of liver enzymes are currently available for visit 4 only, thus aim 3 will only include data collected at visit 4 and after (~224 liver-related hospitalizations of 442 incident liver-
related hospitalization through 2011). 2) The first 2 aims can be best addressed by having a larger sample and longer follow up.

5. **Main Hypothesis/Study Questions:**

Aim 1) Descriptive: Risk for liver-related hospitalization in a bi-racial community based sample.

- We hypothesize that in this bi-racial community sample, there are racial, sex and age differences in the liver-related hospitalizations rates.
- We further hypothesize that alcohol consumption and obesity will be associated with increased risk of liver related hospitalization in a non-linear fashion.

Aim 2) To determine the predictors of incident and recurrent liver-related hospitalization in the community.

- We hypothesize that the following factors will be associated with increased risk of liver-related hospitalization: alcohol consumption (elevated or increase) obesity, diabetes, high blood-pressure.

Aim 3) To characterize the association between liver enzymes levels and the risk of liver-related hospitalization

- We hypothesize that there will be an increase in the risk of liver-related hospitalization even within what is considered normal range of liver enzymes

6. **Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).**

**Aims 1 and 2**

**Study design:** Prospective cohort study

**Baseline:** Visit 1

**Exposures:**
- Sex, race, at baseline
- Age, BMI, alcohol consumption (current, former, never; and usual alcohol intake [grams] per week), at baseline and updated at visits 2, 3 and 4.
- Changes in BMI and alcohol consumption

**Outcomes:**
Using the most recent cohort surveillance files we will identify hospitalization with a liver-related ICD-9 code mentioned in the discharge summary.

We will include the following ICD-9 codes based on prior reports(18): "155.0 Primary liver cancer", "456 Esophageal varices", "570 Fulminant liver disease", "571 Chronic liver disease and cirrhosis", "572.2 Hepatic coma", "572.3 Portal hypertension", "572.4 Hepatorenal syndrome", "572.8 Other sequelae of chronic liver disease", "573.3 Hepatitis, unspecified", "573.8 Other specified disorder of the liver", "573.9 Unspecified disorder of liver".

We will use the date of discharge to calculate the person-time.

We will identify the first occurrence and also recurrent events.

**Exclusions:** participants with missing data on sociodemographic characteristics, BMI and alcohol consumption; participants with race other than black or white.

**Covariates**
Other variables of interest include education, center, smoking status, waist circumference, triglycerides, HDL- and LDL- cholesterol, complete blood count, albumin.

**Main analyses:**
Descriptive statistics such as incidence rates and 95% confidence intervals will be calculated using Poisson regression.

We will use Cox proportional hazards regression models, with and without time-varying exposure to estimate adjusted hazard ratios for liver-related hospitalization associated with the level of alcohol consumption, obesity, diabetes and hypertension. In exploratory analyses, we will examine the presence of interactions by sex and race.

**Aim 3**

**Study design:** Prospective cohort study

**Baseline:** Visit 4

**Exposures:**
-Liver enzymes: ALT, AST and GGT
Liver enzymes will be analyzed as continuous variables, using clinical cut-points given by the laboratory performing the assays, and general cut-offs recommended by other groups.

**Outcomes:**
Hospitalization with a liver-related ICD-9 code mentioned in the discharge summary.
We will identify the first occurrence and also recurrent events.
**Exclusions:** participants with missing data on sociodemographic characteristics, BMI and alcohol consumption; participants with race other than black or white.

**Covariates**
Other variables of interest include baseline sex, race, education, center, Visit 4: smoking status, BMI, waist circumference, triglycerides, HDL- cholesterol, albumin, c-reactive protein.

**Main analyses:**

We will use Cox proportional hazards regression models, to estimate adjusted hazard ratios for liver-related hospitalization associated with the levels of liver enzymes.

Given that diabetes and hypertension may be in the causal pathway between liver-disease and liver-related hospitalization we will not adjust for these variable in the base model.

We will build models with progressive degrees of adjustment:
-Model 1: Sociodemographics
-Model 2: Variables in Model 1+ BMI
-Model 3: Variables in Model 2 + Diabetes
-Model 4: Variables in Model 2 + Hypertension

We will present the results stratified by alcohol consumption, sex and race.

We will explore non linearity of the association between liver enzymes and risk of hospitalization by using restricted cubic spline models with knots placed at clinically – relevant cut-offs.

**Limitations:**

-Relatively small number of liver-related hospitalizations in the ARIC cohort (preliminary analyses suggest there are ~ 450 incident liver-related hospitalizations). This will limit our ability to explore subgroups defined by sex or race.

-Availability of liver enzymes at only one visit and their use to define liver disease.

-Lack of data on viral hepatitis.

-Use of self-reported alcohol consumption.
7.a. Will the data be used for non-CVD analysis in this manuscript?  __X__ Yes  ____ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = “CVD Research” for non-DNA analysis, and for DNA analysis RES_DNA = “CVD Research” would be used?  __X__ Yes  ____ No
(This file ICTDER03 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript?  ____ Yes  __X__ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = “No use/storage DNA”?  ____ Yes  ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.csc.unc.edu/ARIC/search.php  __X__ Yes  _______ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  __X__ Yes  ____ No

11.b. If yes, is the proposal  __X__ A. primarily the result of an ancillary study (list number #2006.15 )  ____ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __________  __________  __________)

*ancillary studies are listed by number at http://www.csc.unc.edu/aric/forms/

12. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.


