# Step 1 of 9: Background/Introduction

The HHEAR Application is for investigators who wish to apply for HHEAR laboratory and data analysis services to add or broaden analyses of environmental exposures in their studies of human health.

Before submitting an Initial Application, please review the <u>policies</u> and <u>procedures</u> for accessing HHEAR services.

Contact <u>HHEARHelp@westat.com</u> with any questions about this Application.

# Instructions

- Step 1: Review background/introduction.
- Step 2: Provide information about yourself, a project contact person, and any coinvestigators.
  - If you do not have the authority to commit to the transfer of biological and/or environmental specimens and data, list contact information for those who do.
  - Upload biographical sketches for yourself and co-investigators. Use the standard NIH biographical sketch format PHS 398.
- Step 3: Complete the section on investigator eligibility. If you have questions about your eligibility, contact <u>HHEARHelp@westat.com</u>.
- Step 4: Review HHEAR policies and indicate your agreement.
- Step 5: Specify which HHEAR services you are requesting.
- Step 6: Provide information on your parent study.
  - Provide succinct responses to each item. Note word/character limits.
- Step 7: Provide information about your proposed HHEAR project.
  - Provide succinct responses to each item. Note word/character limits.
  - Upload a list of citations of key references that provide scientific premise for the proposed project including the rationale for requested services.
- Step 8: Preview application.
- Step 9: Complete application.

# **Step 2 of 9: Applicant/Investigator Information**

Project number (will be filled in automatically upon submission of this form): 2020-00500

**Principal Investigator:** Please provide contact information and biosketch for the Principal Investigator (PI). Indicate whether the PI has the authority to commit for the transfer of samples and/or data.

#### **Principal Investigator:**

Name: <u>xxxxxxxx xxxxx xxxxxxxxx</u>

Institution: xxxxxxxxx xx xxxxxxx xx xxxxxxxx

Phone: <u>xxxxxxx@xxx.edu</u>

Email: xxx-xxx-xxxx

Biosketch: [UPLOAD HERE]

#### Authority to commit to transfer

<b>Biological Samples:</b>	[X] Yes 🗌 No 🗌 N/A
Environmental Samples:	Yes No [X] N/A
Data:	[X] Yes 🗌 No

**Other Investigator:** Complete this item for the investigator with the authority to commit to transfer of biological/environmental samples and/or data, if it is not the PI. If you propose to use samples and/or data from more than one cohort, click the "Add Other Investigator(s)" button below to provide this information for all biological/environmental samples and data sources.

#### **Other Investigator:**

Name:	
Institution:	
Phone:	
Email:	
Role: Co-Investigator	
Other (specify):	

If your study will use biological samples and/or data that were collected by cohorts/institutions other than your own, identify record contact information for individuals at other institutions who have the authority to commit to transfer of samples and/or data.

#### Biosketch: [UPLOAD HERE]

### Authority to commit to transfer

<b>Biological Samples:</b>	Yes No N/A
Environmental Samples:	🗌 Yes 🗌 No 🗌 N/A
Data:	Yes No

#### [ADD OTHER INVESTIGATOR(S)]

**Project Contact Person (if different from Principal Investigator):** 

Name:	
Institution:	
Phone:	
Email:	

# Step 3 of 9: Investigator Eligibility

Please provide responses to each item below to indicate your eligibility for HHEAR services.

You have an ongoing or completed epidemiological or clinical study (parent study) with human biological and/or environmental samples linked to health outcome data.

[X]Yes	🗌 No
[A] Tes	

You want to add environmental exposure data to your parent study or need more extensive analysis of exposures to support your scientific hypothesis related to health outcomes.

[X]Yes No

Although your parent study may be ongoing, you have collected all the data and biological and environmental samples that you will provide to HHEAR for the proposed project prior to submitting your final application.

[X]Yes No

You meet at least one of the following funding criteria (mark all that apply).

 Your ongoing or completed parent study is/was funded at least in part by NIEHS extramural funds. In addition, NIEHS will consider support for studies with significant NIEHS engagement that are administered by other NIH Institutes such as the Environmental Health Disparities Centers (administered by NIMHD), the GEOHealth Centers (administered by FIC), and the ABCD study (administered by NIDA) as well as others. Applicants are encouraged to inquire about potential eligibility before submitting an application. Eligible studies supported by NIEHS may request all HHEAR services including targeted and untargeted analysis of biological and environmental samples.



 Your parent study is/was funded by the NIEHS Superfund Research Program. Studies funded by the NIEHS Superfund Research Program are eligible for targeted and untargeted analysis of biological and environmental samples.

Yes [X] No

• Your parent study is currently funded at least in part by NHLBI extramural funds. Studies funded by NHLBI extramural funds are eligible for targeted and untargeted analysis of only biological samples.

Yes [X] No

• Your parent study is currently funded by NCI extramural funds and has more than one year of funding remaining at the time you submit the HHEAR Initial Application. Studies

funded by NCI extramural funds are eligible for only targeted analysis of biological samples.

[X]Yes 🗌 No

 Your study is an ECHO-wide cohort analysis proposal that has been approved through the ECHO Publications Program, or an ECHO Opportunities and Infrastructure Fund (OIF) proposal approved through the ECHO OIF Program. Studies funded by ECHO are eligible for targeted and untargeted analysis of only biological samples.

🗌 Yes [X] No

- ECHO-wide and OIF supported projects are managed through the ECHO program and don't require an additional application for HHEAR services. ECHO cohorts may be eligible for cohort specific analyses through NIEHS, NHLBI, or NCI support through a HHEAR application.
- Email echocc-publications@dm.duke.edu for more information on the ECHO Publications Program. Email echocc-oif@duke.edu for more information on the ECHO OIF Program.

You are eligible to apply for an NIH grant at your home institution, and you have the authority to commit to documentation such as the Material Transfer Agreement, Data Submission Agreement, and Data Sharing Plan.

[X] Yes 🗌 No

You agree to share your experimental design details and supporting data, including phenotypic data at the individual level, needed to achieve the aim(s) of your proposal.

[X] Yes

No

If your study will use data or samples collected or owned by one or more institutions other than your own, obtain the informed consent form used by each institution. Confirm that the consent language is consistent with the use of the data, biological and environmental (if applicable) samples for future unspecified research; this includes the public sharing of de-identified data.

<b>Step 4 of 9: Agreement to HHEAR Policies</b> Please indicate that you have read and will comply with the <u>Policies for Access to HHEAR Services</u> by adding your eSignature in the space provided. I have read and will comply with the HHEAR policies for accessing services.	Please read <u>Policies for Access to</u> <u>HHEAR Services</u> before you submit an application and contact <u>HHEARHelp@Westat.com</u> with questions. A thorough understanding and preparation can help prevent delays in the review process.
[X] Yes No eSignature <u>*signed*</u>	/
If no, please provide an explanation:	
Are you subject to any other data sharing policies (e.g. $a$	(100 words remaining)
Are you subject to any other data sharing policies (e.g., a consortium agreement that your data must adhere to)?          Yes       [X]         If yes, please provide an explanation:	Data sharing is a critical issue for obtaining HHEAR services. HHEAR requires that you share your experimental design details and supporting data, including phenotypic data at the individual level needed to achieve the aim(s) of your proposals. If your project is subject to any other data sharing policies that conflict with this requirement, your project will not be
(100 word limit)	eligible for HHEAR services.

# **Step 5 of 9: Request HHEAR Services**

*Please indicate the HHEAR services you are requesting (select all that apply):* 

- [X] Laboratory analysis of biological samples
- Laboratory analysis of environmental samples
- [X] Statistical analysis

# **Step 6 of 9: Parent Study Information**

Please complete each item below to provide the key information that can be used by reviewers to understand the **parent study** for the proposed project. If there is more than one parent study providing data and biological and/or environmental samples, provide the information for each parent study. Do not leave any items blank.

- 1. Parent study project title: Understanding the Determinants of Racial/Ethnic Disparities in Liver Cancer and Chronic Liver Disease in Understudied and High-Risk Populations
- 2. Parent study cohort name and website link (if available): Multiethnic Cohort Study (MEC) https://www.uhcancercenter.org/mec
- 3. Parent study funding source(s), including grant number(s): NCI, R01CA228589
- 4. Parent study Principal Investigator and institution:
- 5. Parent study key publications (limit to 3; provide as PMIDs):

PMID: 31553803 PMID: 27301913 PMID: 30859154

6. Primary hypothesis of the parent study:

We hypothesize that lifestyle, genetic, social, and contextual factors explain racial/ethnic differences in liver cancer and NAFLD risks.

(18 words/20 word limit)

#### 7. Summary of main published findings for parent study:

We revealed ethnic differences in HCC incidence and NAFLD being the most common cause of liver disease and cirrhosis. We showed that diabetes is a risk factor for HCC that eliminating diabetes could potentially reduce HCC incidence in all ethnic groups with the largest potential benefit among Latinos. We found coffee consumption and better diet quality may reduce HCC incidence and that diets low in meat and cholesterol and high in fiber may reduce the risk for NAFLD and cirrhosis.

(80 words/100 word limit)

#### 8. Parent study design: (Check all that apply)

Cross-sectional	Hospital-based	Ambispective cohort
Case-control	[X] Prospective cohort	Intervention study
Population-based	Retrospective cohort	Clinical trial
Other:		

#### 9. Parent study population description:

MEC is a large prospective cohort study designed to investigate dietary, lifestyle and genetic factors in relation to cancer and other chronic conditions. The cohort includes over 215,000 men and women, aged 45 - 75 years at cohort enrollment during 1993 - 1996. It consists of participants primarily from five different racial and ethnic groups (African Americans, Japanese Americans, Latinos, Native Hawaiians and Whites) living in Hawaii and California. Blood samples were collected from about 70K participants.

(77 words/80 word limit)

a. Sample size for parent study\* (e.g., # cases and controls if case-control study, number of cohort members if cohort study): 215,000 at baseline; 70,000 with blood samples

HHEAR encourages proposals from parent studies of a wide range of sample sizes.

\*If study is longitudinal, please indicate the sample size at the first time point as well as the last time point.

- b. Age range(s) of parent study population (at study entry): 45-75 at cohort entry
- c. Has the parent study population been included in a previous CHEAR/HHEAR project or in a previous ECHO project?

Yes, specify [X] No

Specify: \_\_\_\_\_

- d. Geographic location(s) of the parent study population: Los Angeles County and Hawaii
- e. Years in which the parent study was conducted: Baseline 1993-1996 and follow up until present
- f. Method of data collection (e.g., survey, in-person visits, medical records) for the parent study: Baseline and follow up questionnaires; SEER linkage, Medicare linkage, state death files, National Death Index

- g. Number of data collection time points and interval between data collection time points for the parent study: Baseline between 1993 and 1996; follow up every 5 years until present. Blood collection in early 2000.
- **10. Main exposures (environmental and/or non-environmental) investigated for the parent study**: Demographic, lifestyle (smoking, alcohol drinking, physical activity, etc.), diet, comorbid conditions, medication use, family history of cancer, germline genetic variations.
- 11. Type of biological and/or environmental samples collected (i.e., whole blood, plasma, urine) for the parent study: blood; plasma
  - a. Years in which biological samples were collected for the parent study: Early 2000
  - b. Number of samples collected per study participant/interval between sample collection for the parent study: One time collection
  - c. Years in which environmental samples were collected for the parent study: N/A

Step 7 of 9: Proposed HHEAR Project

## **Step 7 of 9: Proposed HHEAR Project**

Please complete each item below to provide the key information that can be used by reviewers to evaluate your proposed project. High quality applications will clearly identify the significance of the proposed HHEAR project, research gap addressed, type and number of samples available, rationale for each requested analysis, and health outcomes to be evaluated.

#### Proposed Project Title: The role of PFAS exposures

in nonalcoholic fatty liver disease and hepatocellular carcinoma in the Multiethnic Cohort

**1. Abstract**: Please provide a summary (hypotheses, study design, methods and statistical analysis) of your **proposed HHEAR project** in the context of the parent study.

HCC and NAFLD rates have continued to increase over the past three decades. The health impact of the increasing incidence of HCC is compounded by its dismal survival rate. NAFLD is now recognized as a major contributor to cirrhosis and HCC development. Emerging evidence indicates that PFAS exposure disrupts lipid homeostasis in the liver and has an influence on the initiation and progression of a cascade of pathological conditions associated with NAFLD. Epidemiological evidence is scarce and there are no studies on the impact of PFAS exposure on NAFLD, cirrhosis and HCC. Our objective is to examine the associations between pre-diagnostic plasma PFAS concentrations and NAFLD, cirrhosis and HCC (n~1450) and matched controls (n~1450) in the MEC. This study is novel and cost efficient (leveraging existing data and samples from MEC), has the potential to advance our understanding of hepatotoxic effects of environmental pollutants.

(144 words/150 word limit)

#### 2. Specific aim(s) for proposed HHEAR project:

- a. **Specific aim 1**: To examine the association between plasma PFAS concentrations and risk of NAFLD, NAFLD-related cirrhosis and HCC in the MEC.
- b. Specific aim 2 (if applicable): To evaluate whether NAFLD, cirrhosis and HCC risk associated with PFAS exposures differs by genetic risk score (GRS) and diet profile

A primary outcome should be clearly defined for all studies, including those with multiple outcomes or studies for which the results will be used in the future to answer additional questions.

- c. Specific aim 3 (if applicable):
- 3. Exposures to be investigated for proposed project: PFAS, genetic risk score, diet

#### 4. Significance:

a. Describe the scientific premise for the proposed HHEAR project including the rationale for requested services (targeted, untargeted, and/or environmental analysis). Please provide citations when applicable and indicate which are "key" references for the rationale: Include the rationale for each requested lab analysis here. Successful applications will provide a clear rationale to describe how biomarkers can be used to assess the relationship between exposures and health effects.

While incidence and mortality rates have declined for most cancers, HCC rates have continued to increase over the past 35 years (1). NAFLD can progress to nonalcoholic steatohepatitis, fibrosis, cirrhosis and HCC (2). NAFLD is now a major contributor to HCC development (3). Because of the increasing incidence of NAFLD-related HCC, there is a pressing need to identify the factors responsible for HCC development in the setting of NAFLD. Emerging evidence indicates that exposure to environmental pollutants, including PFAS disrupts lipid homeostasis in the liver and has an influence on the initiation and progression of a cascade of pathological conditions associated with NAFLD. PFAS are a group of synthetic chemicals widely used in industrial applications and consumer products such as protective coatings for cookware, food packaging and furniture (4,5). PFAS are extremely resistant to degradation, they bioaccumulate in food chains and drinking water and have long half-lives in humans. Animal studies show that PFAS exposures cause liver enlargement, hepatic steatosis, and hepatocellular hypertrophy (6-13). Data in human are scarce and there are no studies on the impact of PFAS exposure on NAFLD and HCC. Crosssectional studies in adults showed that elevated serum concentrations of PFOA were associated with increased levels of alanine aminotransferase (ALT), a surrogate marker for NAFLD (14,15), and cytokeratin 18, a marker for liver apoptosis (16). These findings suggest that PFAS may be an important toxicant contributing to NAFLD and HCC.

(234 words/250 word limit)

List of citations: [Uploaded]

b. Explain how the proposed project will improve scientific knowledge of the comprehensive effects of environmental exposures on human health noting advancements over previous research on this topic or how the proposed project will address gaps in scientific knowledge. Include any information related to life stage (e.g., infants, adolescents, adults, seniors) that the project may focus on:

The goals of HHEAR are to: 1) Advance understanding of the impact of environmental exposures on human health throughout the life course, and 2) Promote characterization of the totality of the human environmental exposures called the exposome.

Successful applications will align with <u>HHEAR</u> <u>goals and research priorities</u>

With the increasing prevalence of NAFLD and NAFLD-related cirrhosis and HCC in the US, more efforts are needed to understand their etiology in order to improve prevention and early detection. Given the high prevalence of PFAS exposures in population and the link between PFAS and hepatoxicity and liver dysfunction, it is important to evaluate whether these compounds are involved in NAFLD, cirrhosis and HCC etiology. The potential association between PFAS concentrations and NAFLD, cirrhosis and HCC has not been studied in humans. The MEC provides a unique opportunity to investigate this relationship with its prospective design, long follow-up, and diverse population. This study has the potential to advance our understanding of hepatotoxic effects of environmental pollutants, and may open new avenues for NAFLD, cirrhosis and HCC prevention.

(127 words/200 word limit)

# c. Describe how the requested HHEAR analyses will enhance the findings from the parent study:

The project will provide key additional data which allow us to investigate the role of environmental pollutants in NAFLD and HCC etiology – which may help explain racial differences in risk. Furthermore, we will be able to examine genetic variants and gene-PFAS interactions in NAFLD and liver cancer risk. We hypothesize that known variants associated with HCC and NAFLD interact with PFAS exposures to further increase disease risks. This project will greatly expand the scientific scope of the parent study and will make this the first and biggest population-based study of PFAS and NAFLD and liver cancer.

(96 words/100 word limit)



#### 5. Study design of proposed HHEAR project:

The objectives and analysis plan should be clearly stated so that non-experts in the field can easily follow the project plan.

We will utilize nested case-control design in the

MEC. Three case groups: NAFLD (n=1194), NAFLD-related cirrhosis (n=154), and NAFLD-related HCC (n=101) and matched controls (n~1450). Incident HCC cases (C22.0, ICD-O-3 histology codes 8170-8175) are identified after cohort entry through the latest tumor linkage. To determine NAFLD and cirrhosis and underlying etiology of liver disease, we utilized Medicare claims data. Only cases with pre-diagnostic blood samples will be included in this study (i.e. blood collected prior to cancer diagnosis or NAFLD/cirrhosis identification). The median time between blood collection and diagnosis of NAFLD, cirrhosis or HCC is ~8 years. Eligible controls will be individually matched to each case by birth year, sex, race/ethnicity and study area (Los Angeles/Hawaii).

(116 words/250 word limit)

HHEAR encourages proposal of a wide range of sample sizes. Successful applications will justify the proposed sample size and exposure analyses to achieve the specific aims (see Item 10b).

a. Study sample size: Three case groups: NAFLD
 (n=1194), NAFLD-related cirrhosis (n=154), and NAFLD-related HCC (n=101) and matched controls (n=1449).

Include information on number of study participants expected to be included for each aim.

b. Relationship between participants (if applicable) (e.g. i.e. mother-child, siblings, family based trios):

Unrelated

(1 word/50 word limit)

c. Provide, in a narrative, a breakdown of the total number of participants with biological and/or environmental samples available for analysis by visit and/or age:

```
Three case groups: NAFLD (n=1194), NAFLD-related cirrhosis (n=154), and NAFLD-related HCC (n=101) and matched controls (n=1449).
```

(16 words/100 word limit)

6. Define your proposed HHEAR project according to the following criteria (please check all that apply):

**Hypothesis testing** (for example a new hypothesis or replication of published studies)

[X] Hypothesis generation

a. Provide an explanation to demonstrate how your HHEAR project meets the criteria for hypothesis testing, hypothesis generation, or both:

This study is novel and has the potential to advance our understanding of hepatotoxic effects of environmental pollutants. The potential association between PFAS concentrations and NAFLD, cirrhosis, and HCC has not been studied in humans. The MEC provides a unique opportunity to investigate this relationship with diverse subjects.

(48 words/100 word limit)

If you have questions about which exposure analyses are appropriate to achieve your project aims or about the suitability of your samples for proposed exposure analyses, contact the HHEAR coordinating center (<u>HHEARHelp@westat.com</u>) to schedule a pre-submission consultation with HHEAR Lab Hub scientists.

7. Proposed Project Biological Sample Characteristics and Analyses: Complete Table 1 to provide information on characteristics of participants, the associated biological samples that will be provided, and the requested laboratory analyses. Complete a separate row for each unique combination of characteristics. Provide the biological sample information in as much detail as possible. For example, if you are providing serial serum and urine samples

collected from men and women at Time 1 and Time 2, you would complete 8 rows of the Table. If you have a sample matrix that does not match a specified option, select "O-Other" and specify the matrix in the Other Comments section. If you need to add rows to the table, contact <u>HHEARHelp@westat.com</u>.

For information on which samples are suitable for each HHEAR exposure analysis and minimum volume requirements, see the <u>HHEAR</u> <u>Available Analyses by Sample Matrix</u>.

Priority Order for Analyses	Laboratory Analyses(expo sure measures) (select from	<b>Participant</b> <b>Type</b> (select from drop down)	Age/Stage at Collection (e.g., ages 0-2, first trimester)	Time in Study (T1, T2, T3)	Sample Matrix (select from drop down)	# Participants	# Samples per Participant	# Total Samples	Available Volume per Sample (with units)	Collection Method (e.g., morning void, fasting, passive drool)	Storage Temp (with units)	# of Freeze- Thaws	Sample Collection Status (All or Some)	Other Comments
1	PFAS	Adults (18+ years)	>45	T1	PE – Plasma EDTA	2898	1	2898	205 ul	fasting	-80 C	1	All collected	
	SELECT ONE:	SELECT ONE:			SELECT ONE:									
	SELECT ONE:	SELECT ONE:			SELECT ONE:									
	SELECT ONE:	SELECT ONE:			SELECT ONE:									
	SELECT ONE:	SELECT ONE:			SELECT ONE:									

#### Table 1: Proposed Project Biological Sample Characteristics and Analyses

Scroll to the right to ensure you complete all columns.

8. Provide information for each type of biological sample to be analyzed for your proposed project. Describe the type of specimen collected (*e.g., venous -whole blood, plasma, -serum, urine, saliva, etc.*), the collection method (*e.g., spot urine, passive drool saliva, etc.*) and any processing (*e.g. centrifugation, and/or aliquoting into secondary containers, etc.*). Include a description of

Successful applications will describe the types of collection containers, preservatives and/ or anticoagulants added, and how specimens were processed and stored.

any additives (*e.g. type of anticoagulant, type of preservative, etc.*) that were included with or added to the primary or secondary container during collection and/or processing. Describe the storage containers, storage temperature, length of time in storage, and number of freeze thaws. Please note that if your proposed HHEAR project progresses to the Feasibility Assessment consultation, you will be required to provide information on sample collection, processing, and storage.

We collected blood in lavender top CaEDTA tubes. Samples were collected at fasting at (one time point). Samples have been stored at -80°C since they were collected in early 2000. Samples have undergone one freeze-thaw.

(83 words/250 word limit)

Investigators requesting analysis of environmental samples are strongly encouraged to consult HHEAR Lab Hub scientists to discuss proposed analyses and sample collection, processing and storing requirements before submitting an application. Contact the HHEAR coordinating center (HHEARHelp@westat.com) to schedule a pre-submission consultation.

9. Environmental Samples and Analyses: Complete Table 2 to provide information on the environmental samples that will be provided to HHEAR and requested lab analyses. Complete a separate row for each unique combination of characteristics. Provide the environmental sample information in as much detail as possible. If you have a sample matrix that does not match a specified option, select "O-Other" and specify the matrix in the Other Comments section. If you need to add rows to the table, contact <u>HHEARHelp@westat.com</u>.

Priority Order for Analyses	Laboratory Analyses (select from drop down)	Time in Study (T1, T2, T3)	Sample Matrix (select from drop down)	# Samples	Available Total Volume/Quantit y (with units)	Storage Temp (with units)	# of Freeze- Thaws	Other Comments
	SELECT ON		SELECT ONE:					
	SELECT ON	1	SELECT ONE:					
	SELECT ON		SELECT ONE:					

#### Table 2: Proposed Project Environmental Sample Characteristics and Analyses

10. Provide information on the collection method employed, including the tools and materials used to collect the environmental sample, and any sample processing that was conducted (e.g., sieving) and the containers used to store the samples. Please also include the location where the samples were collected (e.g., if dust, where in the house was the dust collected?):

N/A

(1 word/100 word limit)

*Current and previous HHEAR studies are listed by health outcome on the <i>Data Center website*.

# **11.** In the items below, provide information about the variables that will be submitted to HHEAR for data analysis:

a. Outcome(s): (Check all that apply)

🗌 Asthma	Autism	Biomarker validation
[X] Cancer	Cardiovascular disease/risk	Diabetes
Infectious disease	[X] Liver disease	Neurologic/cognitive development
Obesity/growth	Pregnancy outcomes	Respiratory health
Other:		

b. How was/were the outcome(s) assessed? For each outcome, please indicate the clinical definitions, symptoms checklists or standardized questionnaires used to obtain outcome measures. For example: asthma diagnosis is based on spirometry lung function- FEV1/FVC; recurrent wheezing/asthma is determined by validated ISAAC questionnaire; incident cancer cases (C22.0, ICD-XX histology codes XX-XX) are identified via SEER tumor registry linkage; Type 2 Diabetes is defined as fasting serum glucose ≥126mg/dl and/or currently taking anti-diabetic medication.

Incident HCC cases (C22.0, ICD-O-3 histology codes 8170-8175) are identified via SEER tumor registry linkage. NAFLD and cirrhosis cases are identified using ICD-9 and ICD-10 codes via Medicare claims linkage. Algorithms for case ascertainment have been previously published.

(38 words/100 word limit)

c. List key covariates. Along with the covariate, please provide the name of the standardized/validated questionnaires (symptom checklists, scales, etc.), if any, used to ascertain the variable.

Age, race/ethnicity, sex, study area, body mass index, diabetes, smoking status, alcohol intake, selected dietary factors and genetic risk score

(20 words/100 word limit)

d. Provide frequency tables of key covariates and outcomes for the Proposed HHEAR study, including missing (by time point if applicable) as an attachment. This should be similar to a usual table 1 of a study population (<u>example Populations</u> <u>Characteristics Table</u>). If not currently possible, provide a description of expected missingness on all key covariates and outcomes. If you don't currently have access to this information, please explain why. *[Uploaded]* 

#### 12. Statistical analysis plan

a. Provide a summary description of the analysis strategy and statistical approaches proposed to address each aim. In your explanation, address the following points, as applicable (e.g., confounding, non-linearity, mixtures, combined effect of multiple exposures, potential interactions), and indicate how the proposed strategy will be evaluated to ensure validity, generalizability, and interpretability: Statistical methods should be well-defined and include a clear description of variables/covariates that will be used in analyses. Consider if any aspects of the data (e.g. sample size, incomplete repeated measures, missing data, highdimensionality) could present problems for the anticipated analysis (e.g. too few subjects for too many covariates).

The association between plasma PFAS concentrations and NAFLD, NAFLD-related cirrhosis and HCC in the matched case-control set will be assessed using conditional logistic regression with age, sex, race and study area as matching factors, adjusting for additional potential confounders, including BMI and other obesity related comorbidities. Adjusted odds ratios (ORs) and 95% confidence intervals (CIs) will be calculated for each case group compared to the reference group. We will test the association for the corresponding effect via a LRT. We will primarily focus on a composite PFAS variable, but will also investigate individual PFAS independently and we will use Bayesian kernel machine regression (BKMR) to further understand any non-linear relationships that may exist and to clarify which PFAS might be driving the association, if any. This later inference is possible via the posterior inclusion probabilities in the BKMR model. For aim 2, we will include selected diet variables and genetic risk score (GRS) generated with all of the known NAFLD/HCC SNPs genotyped in the parent R01. We assume that each SNP is independently associated with risk according to an additive genetic model. We will calculate GRS by summing the number of risk alleles at each locus and assume that each SNP contributed equally to the risk of NAFLD/HCC. The genetic risk score distribution among the controls with complete genotype data will be used to generate risk percentile categories.

(228 words/500 word limit)

Successful applications will demonstrate that the proposed sample size will achieve specific aims via the power analysis. Primary factors that affect power: significance level (alpha); sample size; variability in the measured response variable; magnitude of the effect of the variable.

# b. Provide power calculations (e.g., measureable effect size, sample size calculations) for each aim or explain the rationale for why the anticipated sample size is sufficient:

We assume a significance level of  $\alpha$ =0.05 for our primary composite PFAS exposure variable from a two-sided test. We also present calculation for  $\alpha$ =0.01, using a Bonferroni correction under the assumption that all PFAS exposures are independent. We have 80% power to detect OR of 1.5 with 101 cases (the smallest group, HCC) and 101 matched controls for a one standard deviation increase in an exposure variable. Assuming all PFAS are independent, we have 80% power to detect OR of 1.65. For the largest group (NAFLD: 1194 cases/1194 controls), we can detect OR of 1.15 or 1.20, respectively.

(98 words/100 word limit)

# **13**. Challenges and biases that might be encountered in conducting the proposed study analysis:

PFAS concentrations will be measured once, however, PFAS has long term half-lives in humans and are not lipophilic, thus a single measurement provides good estimate of long-term exposure.

(28 words/50 word limit)