DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics should be deleted. There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the *format* shown below.

Element 1: Data Type

A. Types and amount of scientific data expected to be generated in the project:
Summarize the types and estimated amount of scientific data expected to be generated in the project,

We propose a randomized, double-blind, placebo-controlled, cross-over trial of copper supplementation for chronic liver disease patients with deficient copper status as well as a clinical study of untreated patients with either deficient or replete copper status. Because cirrhosis patients with low serum Cu are heterogeneous in terms of demographics, disease etiology and clinical manifestations, a cross-over design would allow control of within-person variability and enable us to detect a treatment effect with fewer subjects. Demographic, laboratory results, clinical observations, diet history, serum copper, serum ceruloplasmin and copper, zinc, iron, and selenium concentration in serum.

Serum metabolite profiles, copper enzyme activities, selected cytokines and stool microbiome profiles will be generated from patient specimens. We expect the total number of participants not to exceed 80 subjects; thus, generating a relatively small dataset even with repeated measures for specific variables.

B. Scientific data that will be preserved and shared, and the rationale for doing so:
Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.

Clinical laboratory results, clinical observations, diet history and trace element data:
Data and results from this proposal are likely to have a lasting value beyond the term of the award and will be archived and preserved, supporting access to third parties interested in replication and verification of our results and data reuse by other investigators planning future studies. We will prepare and share a complete, cleaned, de-identified copy of the final dataset used in conducting the final analyses upon which the accepted primary study publication is based.

The final dataset will include participants' treatment group assignment, demographic, laboratory results, clinical observations, diet history, serum copper, serum ceruloplasmin and copper, zinc, iron and selenium concentration in serum. De-identified, individual level data will be preserved and made available for sharing, and CSV files will be produced in the course of the project.

The final dataset will be structured to maximize future scientific value while protecting patient and health system privacy. HIPAA-specified direct identifiers will be removed or de-identified. The aim of the data sharing policy is to strive for the least restrictive plan possible while providing appropriate protection for participant privacy, health system privacy, and scientific integrity.

Metabolomics data: We will use MSU's bioinformatics and pathway modeling resources to compile the metabolomics data into metadata matrices that are easily interpretable and that will be distributed to other scientists (not directly involved in the study) who may be interested in examining or reanalyzing the metabolomics data that will be produced in this project. Additionally, the proposed metabolic pathway analyses will generate a biological system-wide interconnectedness between gene, protein, and metabolite expression profiles. The network diagrams will be made publicly available for further

hypothesis generation and rational therapeutic design. We will follow the recommendations of the NIH Common Fund Metabolomics Data Sharing Plan and will include (1) raw data i.e. NMR, MS spectrometric information, chromatographic profiles, instrument and raw data processing software; (2) analytical metadata i.e. tabulation of sample origin and metabolite preparation and extraction protocols; sample storage; detailed accounting of the analytical methods used to permit experiment reproductions by other laboratory; (3) biological data including animal management and diet information; (4) metabolite identification and matrix of known and unknown metabolites, and other relevant parameters as described in the Metabolomics Workbench recommendations.
Fecal microbiome data: Human fecal microbiome sequence data will be generated through 16S rDNA sequencing to analyze microbiome composition and presence of pathogenic species in chronic liver disease patients with normal or deficient copper status. Raw and aligned sequence data and associated experimental metadata will be deposited in the Sequence Read Archive (SRA) and made publicly available.

C. Metadata, other relevant data, and associated documentation:
Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

Clinical: The study will be registered on ClinicalTrials.gov so that the research community is aware of our study, the purpose, design, and timing of data collection. Significant updates to the protocol, enrollment, or the DMSP will be submitted to ClinicalTrials.gov at least annually by the study sponsor (Responsible Party: UW). In compliance with reporting requirements for applicable clinical trials (ACT), study results will be submitted by the Responsible Party to ClinicalTrials.gov no later than 12 months after the primary completion date. We will comply with and share data in accordance with the definitions contained in the ClinicalTrials.gov guidance document: Data Element Definitions for Interventional and Observational Studies.

Metabolomics and microbiome: analytical metadata i.e., tabulation of sample origin and metabolite or metagenome preparation and extraction protocols; sample storage; detailed accounting of the analytical methods used to permit experiment reproductions by other laboratory.

Each of these profiles will be made publicly available to the scientific community by submitting the data for deposition in publicly accessible databases and repositories as described below.
Furthermore, the NMR Metabolomics Core Facility at Montana State University has dedicated computer server space for posting complete data sets for public access. We expect to present findings at appropriate scientific meetings as the project moves forward and submit publications to peer-reviewed journals. Accepted manuscripts will be made available to PubMed Central (PMC) either through publishers or through direct author submission.

Element 2: Related Tools, Software and/or Code:
State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

Data will be made available in CSV format and/or as files adherent to FAIR principles appropriate to the data type and will not require the use of specialized tools to be accessed or manipulated.

Element 3: Standards:
State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.

Clinical

In accordance with FAIR Principles for data, we will use open file formats (CSV, TXT, PDF) for our data and metadata. A README text file will be created to describe all our data files. Our protocol will be registered on ClinicalTrials.gov and will align with standards for data elements needed by ClinicalTrials.gov for interventional studies.

Metabolomics
We will follow the recommendations of the NIH Common Fund Metabolomics Data Sharing Plan and will include (1) raw data i.e. NMR, MS spectrometric information, chromatographic profiles, instrument and raw data processing software; (2) analytical metadata i.e. tabulation of sample origin and metabolite preparation and extraction protocols; sample storage; detailed accounting of the analytical methods used to permit experiment reproductions by other laboratory; (3) biological data including animal management and diet information; (4) metabolite identification and matrix of known and unknown metabolites, and other relevant parameters as described in the Metabolomics Workbench recommendations.

Microbiome
Raw and SRA Normalized format files will be submitted to the SRA. SRA Normalized format contains base calls, full base quality scores, and alignments. This format is designed to adhere to **FAIR** (Findable, Accessible, Interoperable, Reusable) principles. Microbiome sequence data will be cleaned of human DNA sequences.

Element 4: Data Preservation, Access, and Associated Timelines

A. Repository where scientific data and metadata will be archived:
Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived; see Selecting a Data Repository).

Clinical
We will submit metadata associated with the clinical datasets to Vivli, a secure, independent data repository. The Vivli repository includes a search engine is designed to index clinical trial metadata and a secure research environment. The repository and publication journal will provide metadata, persistent identifiers, and long-term access for HIPAA compliant open access.

Metabolomics data
Metabolomics data will be deposited in the Metabolomic Workbench (including its newly created Human Metabolome Gene/Protein database (MGP).

Microbiome data
Raw and SRA Normalized format files will be submitted to the Sequence Read Archive available through the National Center for Biotechnology Information (NCBI) and multiple cloud-based servers.

B. How scientific data will be findable and identifiable:
Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

All material with potential for public use will be made publicly available to researchers, clinicians, and patients through submission to Vivli, MGP and SRA as appropriate and defined in Element 4 A. Additionally, study results and lessons learned will be disseminated to the scientific community at large through; 1) Posters and presentations at local, regional, national, and international scientific meetings, 2) PubMed Central (in accordance with the NIH Public Access Policy). Data sharing via journal will result in a digital object identifier associated with the manuscript itself; 3) Publications will also be communicated through social media.

C. When and how long the scientific data will be made available:
Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data

will be available.

We will prepare and share a complete, cleaned, de-identified copies of the final clinical, metabolome and microbiome datasets upon which the accepted primary study publication and any subsequent publications are based. Data will be made available no later than the end of the completion of the funded project period for the award or once a manuscript (or manuscripts) covering the primary study aims have been published, whichever is earlier.

Element 5: Access, Distribution, or Reuse Considerations

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See Frequently Asked Questions for examples of justifiable reasons for limiting sharing of data.

There are no anticipated factors or limitations that will affect the access, distribution or reuse of the scientific data generated by the proposal. Data will be shared as allowed by the participant's informed consents and in accord with IRB approval.

B. Whether access to scientific data will be controlled:

State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).)

As stated in Element 1B above:
The final dataset will be structured to maximize future scientific value while protecting patient and health system privacy. HIPAA-specified direct identifiers will be removed or de-identified. The aim of the data sharing policy is to strive for the least restrictive plan possible while providing appropriate protection for participant privacy, health system privacy, and scientific integrity.

C. Protections for privacy, rights, and confidentiality of human research participants:

If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

To protect the privacy and confidentiality of participants, all data to be preserved and shared will be de-identified prior to placing it in the repository. To ensure participant consent for data sharing, IRB and informed consent documents will include language describing plans for data management and sharing of data after the completion of the study, describing the motivation for sharing, and explaining that personal identifying information will be removed.

Any information sent to NIH, prepared for reports or manuscripts, or presented at scientific meetings will not contain participant names or identifying information. Electronic data will be stored on a secure server, and separate databases will be used to maintain identifying information and study data. Links between study ID and participant names will be kept only for the duration of the study and destroyed after data analysis is completed. All databases will be stored on a secure, HIPAA-compliant network and will be password-protected on a secure server.

Element 6: Oversight of Data Management and Sharing:

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

The Principal Investigators, Dr. Burkhead, Dr. Yu, and Dr. Copié, and will be responsible for oversight and compliance with all NIH policies, including data management activities and data sharing. The Data Management and Sharing Plan (DMSP) will be reviewed in detail annually when preparing progress reports. Broader issues of the DMSP compliance oversight and reporting will be handled by the PIs and general stewardship, reporting, and compliance processes.