



Preventive Medicine Grant Writing Seminar Series

# Data analysis strategies and how they can enhance the proposal



**Fridtjof Thomas, PhD**

Professor

Division of Biostatistics, Department of Preventive Medicine  
College of Medicine

The University of Tennessee Health Science Center

April 11, 2025

# About the presenter

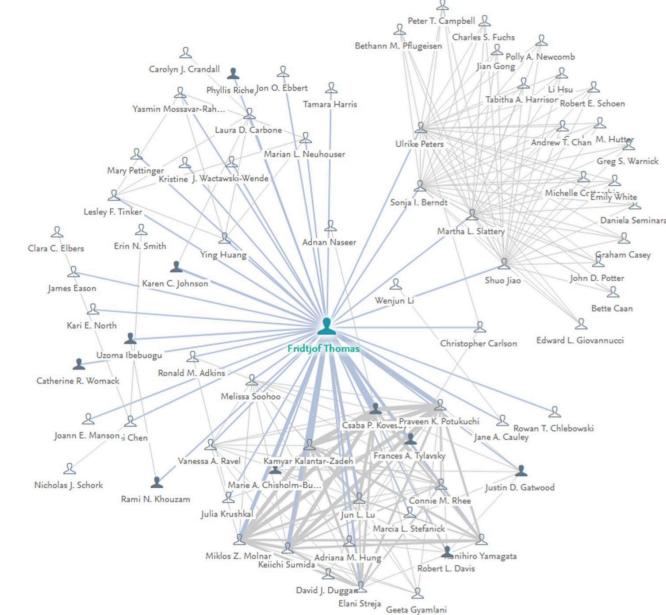
- Professor in the Division of Biostatistics, Dept. of Preventive Medicine
- At UTHSC since 2007
- NIH HEAL Data Stewardship Group that specifically works with assuring that the nationwide hundreds of HEAL-projects comply to the NIH Policy for Data Management and Sharing that has taken effect Jan. 25, 2023. (HEAL = Helping to End Addiction Long-term Initiative; [heal.nih.gov](https://heal.nih.gov)).
- Design and Analysis Committee of the EARLY trials (2010-2016 – “Early Adult Reduction of weight through LifestYle intervention,” a collection of seven randomized clinical trials funded by the National Heart, Lung, and Blood Institute (NHLBI) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health (NIH))
- Member of the Biostatistics Collaborative Core at the Southeast Regional Center of the NIH-NHLBI-funded Women’s Health Initiative (WHI) study that has recruited over 160,000 women in over 40 clinical centers nationwide. (2010-2017)
- Grant review experience since 2012 from
  - Department of Defense’s Congressionally Directed Medical Research Program (DoD CDMRP)
  - NIH Epidemiology of Chronic and Infectious Disease Study Section
  - NIH Neurological, Aging, and Musculoskeletal Epidemiology (NAME) Study Section
- 15+ years Associate Editor of *The Journal of Statistical Computation and Simulation* (JSCS; a Taylor & Francis print journal since 1972)

# Network of collaboration (Elsevier Pure “Fingerprint” 10/09/2019)



## Research output network – organizational units

Department of Preventive Medicine	77
Department of Medicine, Division of Nephrology	23
Department of Pediatrics	6
Department of Pharmaceutical Sciences	5
Department of Clinical Pharmacy – Memphis	4
Department of Clinical Pharmacy – Nashville	4
Neuroscience Institute	4
Department of Ophthalmology	4
Department of Medicine, Division of Cardiovascular Diseases	4
Department of Medicine, Division of General Internal Medicine	3
Department of Acute and Tertiary Care	2
Department of Health Promotion and Disease Prevention	2



# Outline

- What makes or breaks a proposal?
- Outcomes, Missing Data, Surrogate endpoints, Adverse events
- Intend-to-treat analysis (ITT)
- Clustering/grouping of observations
- Pre-planned vs. ad hoc subgroup analyses
- Heterogeneity of treatment effects (HTE) with respect to sex and race
- Data Management and Sharing plan (DMS)

# What makes or breaks a proposal?

- Most importantly:
  - Are you using **sound science**?
  - What is your **innovation**?
- The analytical plan cannot save your proposal, but it will sink it in a hurry if not put together thoughtfully
- Reviewers need to find “objective reasons” why they don’t like a proposal: the fewer targets the analytical plan offers for that, the better
- Make sure that
  - reviewers easily find what they are looking for
  - the analytical plan is connected to your research question
  - measures (including time points) are consistent throughout
- It is difficult to get funded!
- It is not all that difficult to have a better proposal than most I have seen when reviewing grants – but that takes time and effort!

# How strong will your derived evidence be?

## Levels of evidence

- 1a Systematic review of high quality RCTs with similar results and effect sizes for many different RCTs.
- 1b Individual high quality RCT with high precision (narrow confidence interval)
- 1c All or none

---

- 2a Systematic review of cohort studies with similar results and effect sizes.
- 2b Individual cohort study or low quality RCT (e.g., <80% follow-up)
- 2c “Outcomes Research” and ecological studies (based on average exposures etc. of populations of geographical or temporal units)

---

- 3a Systematic review of case-control studies
- 3b Individual case-control study

---

- 4 Case-series and poor quality cohort and case-control studies

---

- 5 Expert opinion (unless critically appraised or based on “first principles”)

*Source: Oxford Centre for Evidence-based Medicine*

*<https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>*

# All or none: Example “Bubble Boy” disease

- Babies born without functional immune system.
- SCID-X1: 1 in 50,000-100,000 affected; caused by a mutation in a gene (IL2RG)
- Most die within first year of life. (Only about 20% have access to a suitable sibling for a bone-marrow transplant as the existing cure.)

St. Jude announced April 18, 2019: Gene therapy cure for babies with X-linked severe combined immunodeficiency

“The gene therapy, produced in the Children’s GMP, LLC, manufacturing facility on the St. Jude campus, involved use of a virus to transport and insert a correct copy of a gene into the genome of patients’ blood stem cells. Following the treatment, the children began producing functioning immune cells for the first time, according to St. Jude, and most patients were discharged from the hospital within one month.” [All 8 babies started to produce complete sets of immune cells.]

<https://www.stjude.org/inspire/news/bubble-boy-scid-x1-cure.html>

<https://www.stjude.org/research/news-publications/research-highlights/2019-research-highlights/st-jude-gene-therapy-holds-promise-for-treating-several-diseases.html>

# Outcomes

- Outcomes: Be ambitious but realistic
  - pilot studies go a long way!
  - An R21 is not an underfunded R01
- Missing Data:
  - Missing covariate information vs. missing endpoints
  - 20% attrition in behavioral intervention studies might be acceptable, 10% is better – requires work to achieve!
  - Differential loss to follow-up needs to be addressed in the analytical plan
- Surrogate endpoints: e.g., progression free survival
- Adverse events:
  - Always occur and need to be reported/summarized (may or may not be related to the intervention)
  - How do adverse events impact on your outcome/collection of measures?

# Missing Data

- Keep it to a minimum
- Mention the word “multiple imputation”, know what it means, and have a realistic approach to it
- Complete case analysis or “last observation carried forward” doesn’t cut it in most cases
- Include sensitivity analyses

# Bias: What do we mean by bias?

- Statistical theory (not further covered here)
  - Sample variation
  - Estimation: if expected value = true value, estimator is unbiased
  - Asymptotically unbiased if value converges to true value if  $n \rightarrow \infty$
  - Bias-variance trade-off in prediction
- Publication bias (not further covered here)
- Bias due to “distorting” true relationships: <https://catalogofbias.org/> (CEBM/University of Oxford)
  - E.g., confounding incl. “confounding by indication”
  - E.g., selection bias (incl. “immortal time bias”)
  - E.g., measurement bias/assessment bias
  - (...)
- Bias analysis (E-values! - VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017 Aug 15;167(4):268-274. doi: 10.7326/M16-2607. Epub 2017 Jul 11. PMID: 28693043.)

# “Distorting” true relationships

What is a “true relationship”?

Example: What is the true odds ratio (OR) for a specific exposure and a specific event?

Operational definition: The true OR is the odds ratio that would be observed in a perfectly executed randomized clinical trial (RCT) that is large enough to make sample variation practically unimportant.

# “Distorting” true relationships (cont.)

The following aspects have to be addressed:

- Selection bias
- Measurement bias (misclassification – outcome or covariate; exposure misclassification; systematically missing activities/episodes, e.g., in activity data; recall bias, telescoping bias, etc.)
- Blinding/masking of evaluators
- Cases and non-cases/controls need identical ways to determine covariates and outcomes!
- “Immortal time bias” in time-to-event analyses
- Differential loss-to-follow-up
- “Artefacts” due to utilizing existing data for a different purpose (e.g., billing data; medical prescription data to determine adherence to medication)

# “Distorting” true relationships (cont.)

Related to data analysis:

- Collection/extraction of data
- Should the analyst be blinded to treatment group? (“Group A” vs “Group B” instead of explicit “Treatment” and “Control”)
- Unintentional programming errors
- Undetected problems with the convergence of computational algorithms
- Validation of the data incl. approaches to unusual observations
- “fishing expeditions” (and multiplicity in testing in general)
- Any form of “P-value maximization approaches”, such as
  - thresholds for continuous variables to achieve “maximal significance”
  - picking definitions for events, exposure measures, etc. based on resulting “significance”
  - etc., etc.,...

Make sure your statistical plan does not raise these concerns!

# Intend-to-treat analysis (ITT)

- Evaluation of participants as randomized
- Protects against bias due to, e.g.,
  - differential adherence to the treatment
  - differential loss to follow-up
- Breaks down if missing data is present!
- Often combined with a “per-protocol” or “as-treated” analysis (see Hulley et all ch. 11)

# Clustering/grouping of observations

- Is there any grouping structure you should account for?
- Group/cluster randomization? (Contamination of groups; not practical to individually randomize)
- Adjustment in other grouping structures (e.g., physician office even when individually randomized) that can/should be addressed in the analysis? (Often in form of mixed effects models with grouping structure being a random effect)

# Pre-planned vs. ad hoc subgroup analyses

- Ad hoc analyses will be made and should be acknowledged (e.g., when safety issues arise during the trial)
- Pre-plan what is important to you and adjust for the multiplicity in testing accordingly
- Think about details, e.g.:
  - ANOVA with post-hoc analysis using Tukey's honest significant difference
  - Dunnett's test to adjust for a single comparison/control group but several active intervention groups

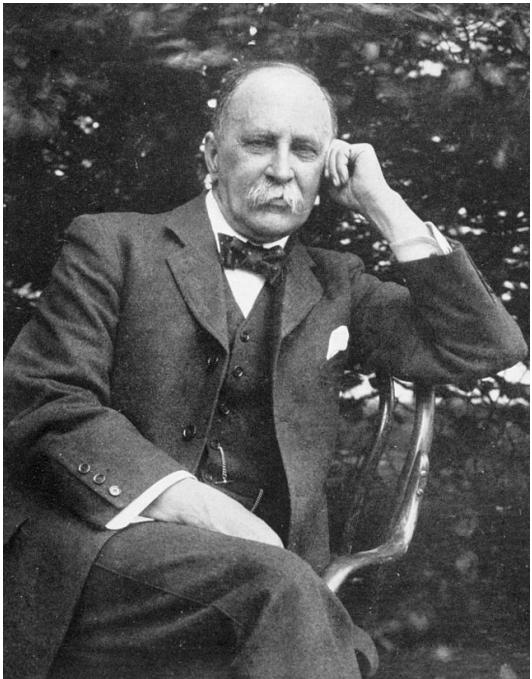
# You can easily do better!

Follow this principle to avoid the worst pitfalls:

***“Draw your assumptions before your conclusions.”***

Miguel Hernán, Prof. of Biostatistics and Epidemiology, Harvard CAUSALab

# Heterogeneity of Treatment Effects (HTE)



"William Osler c1912" by Unknown - [1]. Licensed under CC BY 4.0 via Wikimedia Commons - [http://commons.wikimedia.org/wiki/File:William\\_Osler\\_c1912.jpg#mediaviewer/File:William\\_Osler\\_c1912.jpg](http://commons.wikimedia.org/wiki/File:William_Osler_c1912.jpg#mediaviewer/File:William_Osler_c1912.jpg)

*"Variability is the law of life, and as no two faces are the same, so no two bodies are alike, and no two individuals react alike and behave alike under the abnormal conditions which we know as disease"*

**Sir William Osler**

**1849 – 1919; Canadian physician and “Father of modern Medicine”**

# Heterogeneity of Treatment Effects (HTE)

*Contains Nonbinding Recommendations*

## Evaluation of Sex-Specific Data in Medical Device Clinical Studies

### Guidance for Industry and Food and Drug Administration Staff

Document issued on March 31, 2025.

Document originally issued on August 22, 2014

For questions about this document regarding CDRH-regulated devices, contact CDRH Health of Women at [CDRHHealthofWomen@fda.hhs.gov](mailto:CDRHHealthofWomen@fda.hhs.gov). For questions about this document regarding CBER-regulated devices, contact the Office of Communication, Outreach, and Development (OCOD) at 1-800-835-4709 or 240-402-8010, or by email at [ocod@fda.hhs.gov](mailto:ocod@fda.hhs.gov).



U.S. Department of Health and Human Services  
Food and Drug Administration  
Center for Devices and Radiological Health  
Center for Biologics Evaluation and Research

#### Source:

<https://www.fda.gov/files/medical%20devices/published/Evaluation-of-Sex-Specific-Data-in-Medical-Device-Clinical-Studies---Guidance-for-Industry-and-Food-and-Drug-Administration-Staff.pdf> ; accessed 04/09/2025.

### A. Why consider sex differences?

Certain medical products elicit different responses in women compared to men. Differences may be attributable to intrinsic factors (e.g., genetics, hormones, body size, sex-specific physiology), extrinsic factors (e.g., diet, sociocultural issues, environment) or interactions between these factors. For example, there may be medical conditions that are unique to a certain sex, which should be considered in study recruitment and in reporting of results.

# Heterogeneity of Treatment Effects (HTE)

- Race and sex! (Unless you have an acceptable reason for not considering these variables).
- NIH: NIH Policy and Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research
  - <https://grants.nih.gov/policy-and-compliance/policy-topics/inclusion/women-and-minorities>
  - <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-18-014.html>
- FDA: Evaluation of Sex-Specific Data in Medical Device Clinical Studies  
<https://www.fda.gov/files/medical%20devices/published/Evaluation-of-Sex-Specific-Data-in-Medical-Device-Clinical-Studies---Guidance-for-Industry-and-Food-and-Drug-Administration-Staff.pdf>
- Patient-Centered Outcomes Research, especially PCORI
- PCORI Methodology Standards
- HTE:
  - How are HTEs assessed?
  - Should there be adjustments for multiplicity in testing?
  - Which type of subgrouping variables?
  - Should an HTE-analysis be reported even when the overall treatment effect is not statistically significant?
  - When reported, what should be reported and why?

# Heterogeneity of Treatment Effects (HTE)

**FDA: Evaluation of Sex-Specific Data in Medical Device Clinical Studies**  
<https://www.fda.gov/files/medical%20devices/published/Evaluation-of-Sex-Specific-Data-in-Medical-Device-Clinical-Studies---Guidance-for-Industry-and-Food-and-Drug-Administration-Staff.pdf>

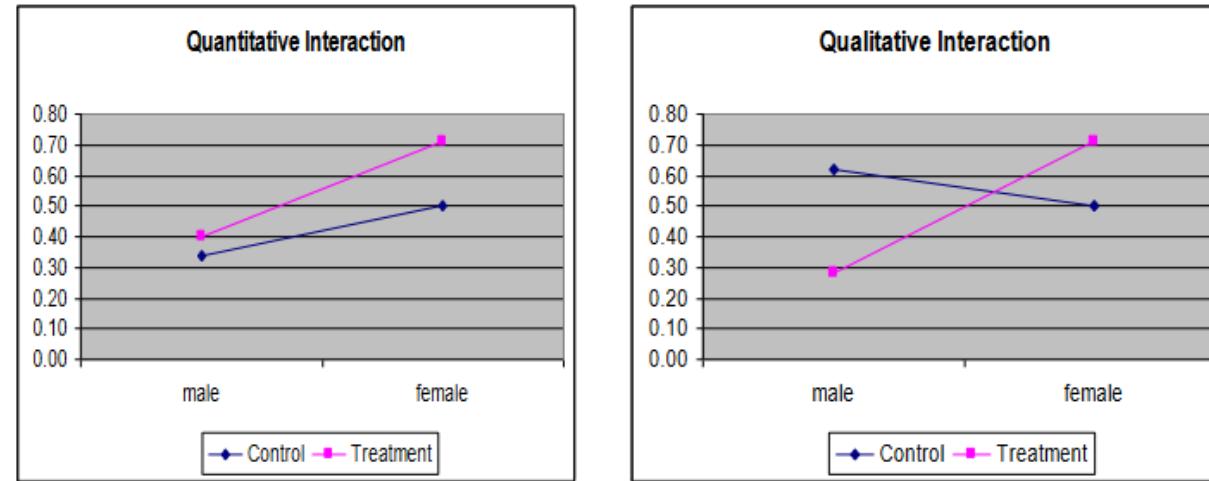


Figure 1. Illustrations of quantitative (left graph) and qualitative (right graph) interactions.

Statistical hypothesis tests of treatment by sex interaction have been widely utilized to detect treatment effect heterogeneity across sex. Most of the tests of interaction in common use have as their null hypotheses the absence of treatment by sex interaction. As statistical tests, their significance levels should be pre-specified in the investigational plan. Note, however, that the power of such tests may be unspecified. Therefore, lack of statistical significance for a test of treatment by sex interaction may not convincingly evidence the absence of clinically relevant interaction. By the same token, moderate statistical significance may not convincingly evidence the presence of clinically relevant interaction. While statistically significant interactions will be investigated for their clinical meaningfulness, interactions without associated statistical significance may also be examined for clinical reasons specific to the design and endpoint.

# PCORI's HTE guidelines

Varadhan R, Stuart EA, Louis TA, Segal JB, Weiss CO. **Review of Guidance Documents for Selected Methods in Patient Centered Outcomes Research: Standards in Addressing Heterogeneity of Treatment Effectiveness in Observational and Experimental Patient Centered Outcomes Research.**

Source: <http://www.pcori.org/assets/Standards-in-Addressing-Heterogeneity-of-Treatment-Effectiveness-in-Observational-and-Experimental-Patient-Centered-Outcomes-Research.pdf>: PCORI; 2012. Accessed 01/23/2015.

# Data Management and Sharing Plan (DMS)

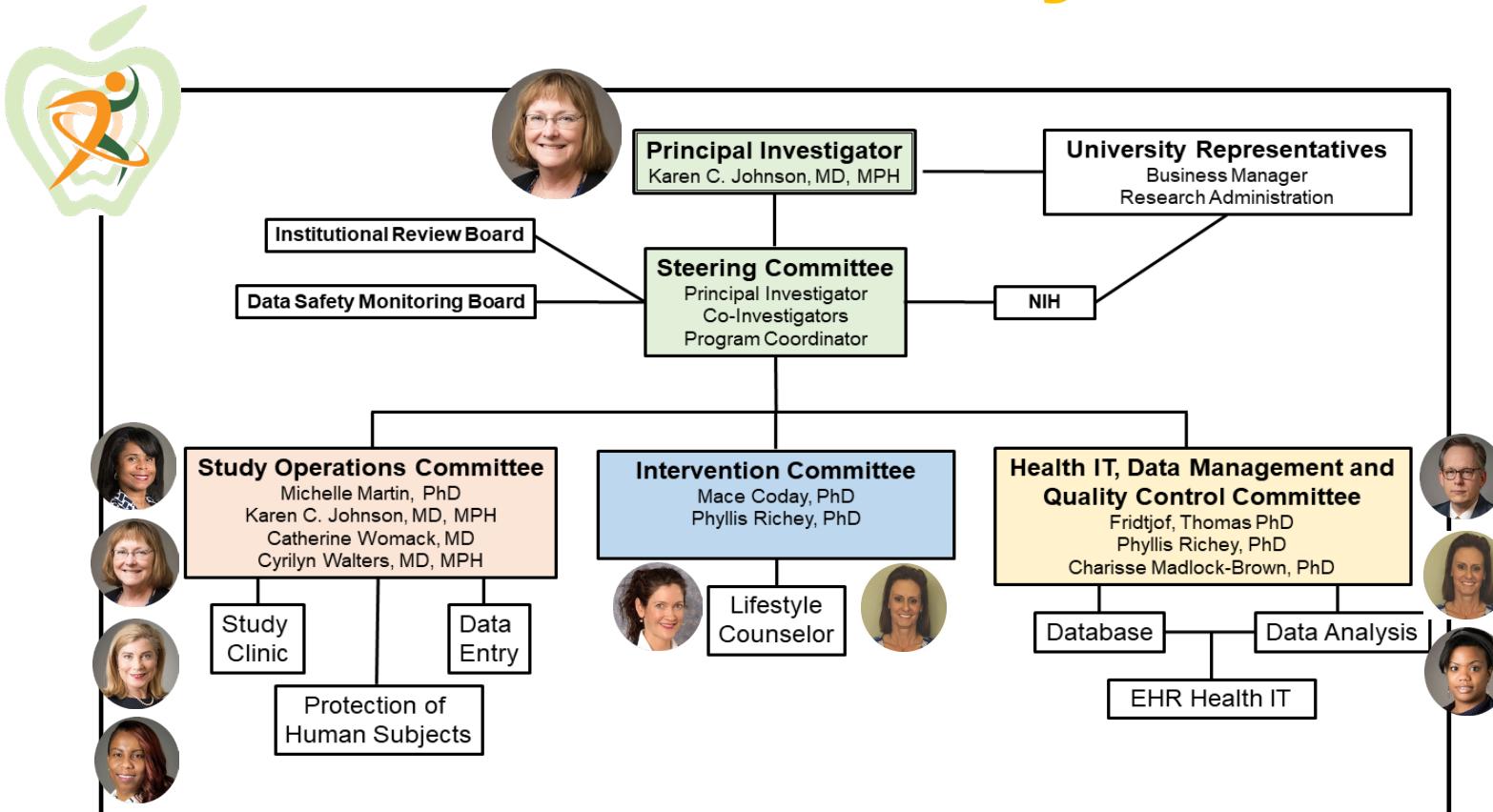


The screenshot shows the homepage of the NIH Scientific Data Sharing website. The URL in the browser is <https://sharing.nih.gov>. The page features a dark blue header with the NIH logo and the text "SCIENTIFIC DATA SHARING". Below the header is a navigation bar with links for "DATA MANAGEMENT AND SHARING POLICY", "GENOMIC DATA SHARING POLICY", "OTHER SHARING POLICIES", "ACCESSING DATA", and "ABOUT". The main content area has a blue background with the text "Expediting the Translation of Research Results to Improve Human Health." and a network of nodes graphic. To the right, there is a grid of six cards: "Scientific Data" (laptop and microscope), "Genomic Data" (DNA helix), "Research Tools" (hands using a tool), "Model Organisms" (microscopic view of cells), "Clinical Trials" (nurse and patient), and "Research Publications" (laptop screen). A yellow sidebar on the right contains the text "Explore the areas in which NIH has sharing policies." and a "Find which policies apply to you" button.

<https://sharing.nih.gov/>

# Example from CHAMPS study

## Structure of Study Team



# Example from CHAMPS study

## MOP

Chapter 13 CHAMPS Data Analysis, Sharing, and Monitoring Plan

### CHAMPS Data Analysis, Sharing and Monitoring Plan

1. Power and Sample Size Estimate
2. Data Management Plans
  - 2.1. Data Confidentiality
3. Analysis Plans
  - 3.1. Primary Analysis
  - 3.2. Missing and Incomplete Data
  - 3.3. Censoring
4. Data Sharing (NOT-OD-21-013)
  - 4.1. Data Type
  - 4.2. Related Tools, Software and/or Code
  - 4.3. Standards
  - 4.4. Data Preservation, Access, and Associated Timelines
  - 4.5. Access, Distribution, or Reuse Considerations
  - 4.6. Oversight of Data Management and Sharing

2-page DMP merged  
into the MOP



# CHAMPS

CHOOSING HEALTHY ACTIVITIES  
AND LIFESTYLE MANAGEMENT  
THROUGH PORTAL SUPPORT



#### 4. DATA SHARING (NOT-OD-21-013)

The CHAMPS study will prepare and distribute an electronic dataset with scientific data, incl. electronic versions of any paper forms that might have been used in data collection. Confidentiality of individual participants will be maintained with all releases of data. The final CHAMPS study analytical database will be processed according to HIPAA definitions for public data sharing. During this process, the participant data will be de-identified by using processes which include but are not limited to: removal of identifiers, translation of dates to delta time values, and assignment of random study identifiers. Out of this process will be a series of de-identified data files representing the final analytical data set. These data files will be provided in a standard format that is readable across a variety of applications and operating system platforms. Documentation that will be provided along with the data sharing file that may include but is not limited to: data dictionary, data code book, valid variable ranges, the protocol, MOP, and intervention manual or programs. The data release documentation will provide detailed, organized documentation of study variables and clear instructions on how to install and access the data. CHAMPS intents to make all data available as outlined in the Final NIH Policy for Data Management (NOT-OD-21-013). **However, CHAMPS has been funded before Jan 25, 2023, and is therefore technically not covered by the new NIH data management and sharing policy.**



## 4.1 Data Type

The scientific data to be generated is from 250 randomized participants and originates from questionnaires (socio-demographics, medical history, medications, diet/DHQ III, physical activities/GPAQ and sedentary behavior, alcohol and smoking, neighborhood environment, quality of life/PROMIS, social determinants of health, technology use), clinic measures (vital signs, weight, height, waist and hip circumference), and others (adverse events, intervention participation/video viewing through webpage analytics).

Data that will be preserved and shared consists of questionnaire and clinic visit data on the individual level, as well as processed data summarizing intervention exposure/web-viewing on the individual level.

Metadata is publicly available at the U.S. National Library of Medicine in the ClinicalTrials.gov database (Identifier: NCT05410353; <https://www.clinicaltrials.gov/ct2/show/NCT05410353>) and will be updated throughout the conduct of the study. In addition, structured metadata will be created as required by the respective repositories that will be used.

DHQ III = Diet History Questionnaire III

GPAQ = Global Physical Activity Questionnaire

PROMIS = Patient-Reported Outcomes Measurement Information System (NIH)



## 4.2 Related Tools, Software and/or Code

All data will be shared as comma-separated values (CSV) files with UTF-8 encoding that do not require specific software for reading/encoding. Data sets will be of rectangular arrangement with rows corresponding to participants.



### **4.3 Standards**

Mostly standardized questionnaires are used such as the DHQ III, GPAQ, or PROMIS. A detailed account is available in the Manual of Procedures (MOP), Chapter 11: Data Collection. Unique participant identifiers will link entries in the various data sets. Data dictionaries will be provided for the available data sets.

DHQ III = Diet History Questionnaire III

GPAQ = Global Physical Activity Questionnaire

PROMIS = Patient-Reported Outcomes Measurement Information System (NIH)

## List of all nutrients and food groups available from DHQ III

[Expand All](#)[Collapse All](#)[Carbohydrate Constituents](#)[Carotenoids and Tocopherols](#)[Dietary Constituents from Supplements \(reported separately in output files\)](#)[Fats, Fatty Acids, & Cholesterol](#)[Food Pyramid Equivalents](#)[Macronutrients & Energy](#)[Minerals](#)[Other](#)[Protein Constituents](#)[Vitamins](#)

### — Minerals

- Calcium
- Phosphorus
- Magnesium
- Iron
- Zinc
- Copper
- Selenium
- Sodium
- Potassium
- Manganese

### — Fats, Fatty Acids, & Cholesterol

- Cholesterol
- Total saturated fatty acids
- Total monounsaturated fatty acids
- Total polyunsaturated fatty acids
- Solid Fat
- SFA 4:0 (Butanoic)
- SFA 6:0 (Hexanoic)
- SFA 8:0 (Octanoic)
- SFA 10:0 (Decanoic)
- SFA 12:0 (Dodecanoic)
- SFA 14:0 (Tetradecanoic)
- SFA 16:0 (Hexadecanoic)
- SFA 18:0 (Octadecanoic)
- MFA 16:1 (Hexadecenoic)
- MFA 18:1 (Octadecenoic)
- MFA 20:1 (Eicosenoic)
- MFA 22:1 (Docosenoic)
- PFA 18:2 (Octadecadienoic)
- PFA 18:3 (Octadecatrienoic)
- PFA 18:3N3 (Alphalinolenic)
- PFA 18:4 (Octadecatetraenoic)
- PFA 20:4 (Ficosatetraenoic)





## 4.4 Data Preservation, Access, and Associated Timelines

Metadata available at ClinicalTrials.gov (Identifier: NCT05410353) will be archived by the U.S. National Library of Medicine. The collected human subject data is identifiable protected health information (PHI) under The Health Insurance Portability and Accountability Act of 1996 (HIPAA).

To the extent possible, data that we will make available will be de-identified of PHI using the Safe Harbor de-identification method as defined in §164.514(a) of the HIPAA Privacy Rule. We foresee to be able to furnish such limited data sets at the point in time of publications and these will contain the information needed to reproduce the main analyses in the respective publication. Because limited data sets remain protected health information under HIPAA, access to the data will need to be controlled by data use agreements (DUAs) and disclosure will only be for the purpose of research, public health or health care operations. These data sets are intended to be posted in a repository "close" to the publication (e.g., in PubMed) or in a repository explicitly named in the publication.

A comprehensive scientific data set will be posted in a repository by the end of the project performance period. That comprehensive scientific data will include electronic versions of any paper forms that might have been used as well as scientific data not (yet) used in the study publications: e.g., the DHQ-III data output with comprehensive information about carbohydrates, vitamins, minerals, protein constituents, etc. will be released as available in this study. To protect the privacy of our participants, we will remove "facial" identifiers such as names, medical record numbers, etc., but will possibly not create a data set formally classified as limited data under HIPAA, and access will have to be controlled.

We anticipate that the produced data is of value for the larger research community, institutions, and/or the broader public, and therefore prefer repository-solutions with anticipated "indefinite" storage such as ClinicalTrial.gov, PubMed, or other NIH-recommended repositories. The scientific data will be findable and identifiable by the references in the respective publications as well as through the metadata provided to the repositories (incl. ClinicalTrials.gov; Identifier: NCT05410353; <https://www.clinicaltrials.gov/ct2/show/NCT05410353> )

Supporting material that will also be made available includes the Study Protocol/Manual of Operations (MOP), Statistical Analysis Plan (SAP), Informed Consent Form (ICF), and the analytic code for main analyses.



#### 4.5 Access, Distribution, or Reuse Considerations

The IRB approved informed consent allows for usage of the collected scientific data for other medical research (secondary research). Individual details obtained from participants can be provided in publications or presentations, but they cannot be discussed in a way that would allow to identify the participant.

All our released data will remain protected health information under HIPAA, and access to the data will need to be controlled by a data use agreement (DUA; for limited data sets), an IRB approval (for data not de-identified by the Safe Harbor method as defined in §164.514(a) of the HIPAA Privacy Rule), and disclosure will only be for the purpose of research, public health or health care operations.

**FAIR** = Findable, Accessible, Interoperable, and Reusable  
<https://fair-research.org/>

Is your research reproducible?  
Is your data processing and are your statistical analyses  
reproducible?



#### **4.6 Oversight of Data Management and Sharing**

Compliance with this DMS plan is monitored by the PI of the study, Dr. Karen Johnson, during the project performance period. The HIPAA Privacy Rule requires appropriate safeguards to protect the privacy of protected health information and sets limits and conditions on the uses and disclosures that may be made of such information without an individual's authorization. This liability is transferred to the respective repository when data sets are surrendered to these repositories.

# Health Sciences Library Research Support

<https://www.uthsc.edu/library/research.php>

<https://libguides.uthsc.edu/data/uthsc-data>

**General Support Services, email [library@uthsc.edu](mailto:library@uthsc.edu)**

Literature searches, systematic reviews, citation managers

**Data Support Services, email Sarah Newell ([snewell@uthsc.edu](mailto:snewell@uthsc.edu))**

Assessing & improving research impact (h-index, citation metrics, etc)

Writing Data Management Plans (DMPs)

Selecting appropriate data repositories for storage and sharing

Basic data wrangling in Excel and OpenRefine

Enhancing data visualizations

Locating data for re-use

Identifying & evaluating journals for publication

Open Access publishing

Copyright and licensing



# Biostatistics, Epidemiology, and Research Design (BERD) Unit

The Biostatistics, Epidemiology, and Research Design (BERD) unit of the Department of Preventive Medicine provides easily-accessible consultation and data analysis to help investigators produce scientifically rigorous, reproducible, and publishable research. The unit combines experts in epidemiology, biostatistics, biomedical informatics, community engagement and molecular informatics. We aim to efficiently and effectively connect investigators to practical analytical & technical resources. We encourage offer a range of user-friendly educational opportunities for individuals of all skill levels. The BERD unit also promotes innovations in tools and methods that address barriers to clinical and translational research.

## [BERD Clinic](#)

- Sign-up today for free one-on-one advice
- Available in-person and online

## [BERD Consulting](#)

- Available as consultants and collaborators
- Please [see here for details on individual members](#)

<https://berd.uthsc.edu/>

## BERD Clinic – Easily-Accessible Consultations

With advance registration, UTHSC faculty and residents can meet with specialists for biostatistics, epidemiology, and research design consultations. These consultations are limited in scope and intended to help researchers understand the underlying statistical aspects of their research design so they can adequately plan projects (e.g. study design, analytic approaches, simple power analyses, see below). For investigators looking for a collaborator, or consulting services, an initial meeting through the BERD Clinic is required to connect you to those services.

[See an example BERD Clinic project workflow and review common questions for first time meetings.](#)

[Sign up for a Virtual BERD Clinic](#)

Clinics with a Biostatistician or Epidemiologist

[Virtual BERD](#)

Clinics with Other Specialists

[Biomedical Informatics](#)  
[Molecular Bioinformatics](#)  
[Omics Epigenetics and Metabolic Phenotypes](#)  
[Health Services Research](#)  
[Quality Improvement](#)

Personalized Advice On

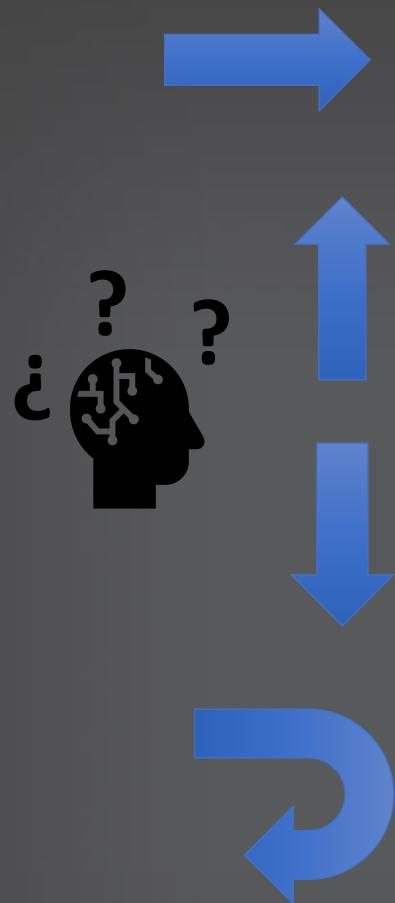
Statistical Analysis Design

- Formulating a statistical analysis plan
- Appropriate selection of statistical tests
- Interpretation of statistical results
- Advice on display of data and results
- Responding to reviewers

Study Design

- Choice of study population
- Sample size and power calculations
- Randomization planning

- [berdclinic@uthsc.edu](mailto:berdclinic@uthsc.edu)
- BERD Consulting Manager – Tristan Hayes [thayes@uthsc.edu](mailto:thayes@uthsc.edu)



Thank you!

Questions?