

SAMPLE DATA MANAGEMENT PLAN

Project Information

This Data Management Plan (DMP) pertains to data which will be collected during a study at UNC-CH (University of North Carolina at Chapel Hill) pertaining to the experience of subjects regarding the impact of pre-existing medical conditions on group health plan benefits. Sensitive data will be collected during this study and will include personal identifiers. Potential subjects will be randomly selected from a statewide pool of potential applicants ranging from 22 to 65 years of age. Subjects who consent to participate will be anonymously surveyed using one of two methods, based upon user preference. Subjects may elect to complete the survey using either a web-based application or a pre-recorded verbal survey administered by telephone.

Data collected during this study will be archived with H.W. Odum Institute for Research located at the University of North Carolina at Chapel Hill, where it will be stored in a specific virtual archive and will be made publicly available through the Odum Institute in Dataverse, which is a public data repository. The data will be de-identified prior to submission to the archive in accordance with all requirements of the UNC (University of North Carolina) Institutional Review Board. As an established and trusted repository in the field, the Odum Institute Data Archive is a member of the Library of Congress National Digital Stewardship Alliance (NDSA), and provides a strong archival and data distribution resource to the project. Odum Institute archival staff will guide data management for this project.

General Data Management Plan information

This DMP was created on January 1, 2017 for submission to the National Science Foundation as required by NSF guidelines in the interest of securing funding for this study. The aim and purpose of this DMP is to detail and guarantee the preservation of the data collected during this study, as well as any results derived from the associated research. This DMP is intended for review by relevant NSF personnel, as well as Odum Institute staff and UNC-CH staff Odum Institute affiliated directly with this study and the collection and preservation of the associated data and research. Pending project funding, ongoing review of this plan will ensure preservation of and access to this data. This is the first iteration of this DMP associated with this data.

Policies

Comprehensive institutional and research group guidelines specified by UNC-CH were applied regarding the design of this data management plan. Policies and procedures will be followed as required by the UNC Office of Sponsored Research (<http://research.unc.edu/sponsored-research>). There are no requirements stipulated by the funding or partner organizations regarding this data. There are no additional requirements associated with the data being submitted.

Legal Guidelines and Requirements

Sensitive data collected during this study will be thoroughly de-identified. Subject confidentiality will be maintained in strict compliance with HIPAA Privacy Rules and De-Identification Standards (<https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html>), as well as UNC Office of Human Research Ethics guidelines (<https://research.unc.edu/files/2015/05/SOP-June-2-2017-bookmarked-and-TOC-links.pdf>). There are no copyright or licensing issues associated with the data being submitted.

Access, Sharing and Reuse of Data

The researchers associated with this study are not aware of any reasons that might prohibit the sharing and reuse of the data being submitted following thorough application of all policies and procedures for de-identification described previously. No additional restrictions or permissions will be required for accessing the data. Findings will be published by the researchers based on this data with an estimated publication date of July 1, 2019. An agreement is in place regarding the right of the original data collector, creator, or principal investigator for first use of the data. The specified embargo period associated with the data being submitted extends from the projected conclusion date for initial research (December 31, 2018) until six months after projected publication date for the findings (July 1, 2019). The embargo will be lifted by January 1, 2020, at which point submitted data will be made publicly available through the Odum Institute Dataverse repository. The Odum Institute Dataverse will provide tools for sharing, re-use, online analysis, discovery and download of the data.

Data Standards and Capture

Data will be stored in generally accepted formats and will be analyzed using standard statistical applications including SPSS, Stata, SAS, M-Plus, and Geo-DA.

Survey data will be reviewed by the Odum Survey Methodology group using industry standard tools including REDCap web-based survey applications and/or BLAISE CADI systems for telephone-based surveys. Extensive metadata will be generated during these processes, enabling the creation of detailed, community-standard DDI (Data Documentation Initiative) archival metadata by the Odum archive group.

Metadata

Metadata associated with the creation and development of this survey will be generated and stored with the data. Associated metadata will be manually created in XML file format, and community-standard DDI metadata will be applied.

Security, Storage, Management, and Back-Up of Data

The Odum Institute's longstanding experience with, and commitment to, secure data archiving is in keeping with established UNC Information Security Policies (<http://help.unc.edu/help/information-security-policy-summaries>). All data associated with this project will be maintained in strict compliance with these policies. During survey implementation, all associated research data will be physically stored on a password-protected secure server maintained by UNC-CH using standard SPSS file formats. No data will reside on portable or laptop devices, and no other external media/format(s) will be used for data storage.

Network-stored research data is backed up on a daily basis to a separate password-protected secure server maintained by UNC-CH. The researchers are currently responsible for storage, maintenance and back-up of the data. The long-term strategy for the maintenance, curation, and archiving of the data will be implemented when the data and associated research are migrated to the Odum Institute for Research for virtual archiving using the Odum Institute Dataverse repository.

Preservation, Review and Long-Term Management of Data

Data collected during this study will be archived with the Odum Institute at UNC-CH, where it will be stored in a specific virtual archive and made publicly available through the Odum Institute Dataverse repository. As a result of this arrangement, there are no specific financial considerations of which the researchers are currently aware which might impact the long-term management of the data. The research and archival staff of the Odum Institute will review this DMP upon accession of the data in order to ensure and demonstrate compliance. The DMP will be reviewed by Odum Institute research and archival staff prior to ingest and release into the Odum Dataverse repository.

Assessment of Data Management Plans

INSTRUCTIONS

1. Read Sample Data Management Plan
2. Fill out each section of the scoresheet: a) **include page numbers**; b) **list NA (not applicable) if appropriate**

Example:

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
1.1	Describes what types of data will be captured, created, or collected		✓ Page 1: Included transcripts, forgot survey data	

EXERCISE:

Section 1: Types of data produced

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
1.1	Describes what types of data will be captured, created, or collected			
1.2	Describes how data will be collected, captured, or created (<i>incl. methods, instruments, software</i>)			
1.3	Identifies how much data (volume) will be produced			

Section 2: Standards for data and metadata

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
2.1	Identifies metadata standards and/or metadata formats that will be used for the proposed project			
2.2	Describes data formats created or used during project			

Section 3: Policies for access and sharing

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
3.1	Provides details on when the data will be made publicly available			
3.2	Describes how the data will be made publicly available			
3.3	Describes security measures that will be in place to protect the data from unauthorized access			
3.4	Describes how long the data will be retained and made available to people outside of the project			

Section 3: Policies for access and sharing, cont'd

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
3.5	IF APPLICABLE: If the data are deemed to be of a sensitive nature, describes what protections will be put into place to protect privacy or confidentiality of research subjects			
3.6	IF APPLICABLE: Describes what intellectual property rights to the data and supporting materials will be given to the public and which will be retained by project personnel			
3.7	IF APPLICABLE: Describes factors that limit the ability to share data (e.g. commercialization or proprietary nature of the data) and those conditions. Describes to whom the data will be made available and under what conditions			

Section 4: Policies and provisions for reuse and redistribution

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
4.1	Describes the policies in place governing the use and reuse of the data			
4.2	Describes the policies for redistribution of the data			
4.3	Describes policies for building off of the data, such as through the creation of derivatives			
4.4	Describes methods for communicating policies or guidelines for reuse, redistribution, and creation of derivatives			

Section 5: Plans for data archiving and preservation of access

	Performance Criteria	Complete/detailed	Addressed issue, but incomplete	Did not address
5.1	Provides details on how the data will be archived (e.g. <i>whether or not it will be archived</i>)			
5.2	Describes how access to the archived data will be maintained			
5.3	Describes plans for archiving and preserving digital data (e.g., <i>what data; how much; where</i>)			
5.4	Identifies a timeframe for how long data will be archived			
5.5	Plan discusses the types or formats of data the investigator expects to retain in their possession			