CSTLTS DATA MANAGEMENT PLAN: Extramural Project

This example is designed to help extramural recipients develop a data management plan for any type of CDC-funded public health data collection activity including nonresearch (public health practice) data.

Section 1: Project Information

Agency/Organization Name:

Funding Mechanism Type:

☐ Cooperative Agreement, please specify announcement name and #:

☐ Grant, please specify announcement name and #:

☐ Contract, please specify name and #:

☐ Other, please specify:

Project Contact (POC)

POC Name: POC Organization:

POC Email Address: POC Phone Number:

Section 2: Dataset Information

Date of DMP creation: *Click or tap here to enter text.*

Dataset Title – *In plain English with sufficient detail to facilitate search and discovery of the dataset, name of project data/data collection activity. This should remain the same throughout the life of the project.*

Dataset Description – *Briefly describe the data captured for this project and its purpose. This should have sufficient detail that enables readers to quickly understand whether the project or dataset is of interest to them.*

Tags/Keywords – *Terms to help users discover the project and its datasets; please include terms that would be used by technical and nontechnical users. For tags, refer to Mesh Terms, which are loaded annually from the US National Library of Medicine.*

Data Publisher/Owner

Discuss which entities own the data, who the publishing entity is, and which entities have custodial responsibility and/or gives permission to share the information collected or obtained from this project.

Owner of the Data:

Publishing Entity:

Custodial Responsibility:

Section 3: Description of the Data and Data Quality

Data Description

Data Collection Start Date – *Dates the staff will begin collection of data*.

Data Collection End Date – *Dates the staff will complete the collection of data.*

Methods – *Briefly describe the study/project design and methods, including the approach and plan to meet the objectives, such as interventions, procedures, target population or respondent recruitment, screening, and enrollment, etc.*

Data Collection Instrument(s) and Protocol – *Provide a brief description with reference to a document or website that provides detailed information. Include information such as how often the data will be collected.*

Data Management and Quality

Data Management Protocol – *Briefly describe, with reference to physical location(s) or system(s), where data will be housed (e.g., CDC shared network drive, data host system name, SQL database, etc.)*

Process for Omitting Identifying Information – *Describe what identifiers are in the database, how they will be removed, and by whom.*

Data Quality Protocol – *Describe methods for data validation and error resolution; removal or shielding of any proprietary information; removal or shielding of sensitive information; removal or shielding of any individually identifying information; including indirect identification.*

Section 4: Data Access and Protection

Proposed Public Access Level – *Briefly describe who will have access to stored information.*

Select an access level below. Select Public if the data collected/generated in this project will be released to the public in either microdata or aggregated format; select Restricted if the project data will be shared with restrictions or via CDC Research Data Center (RDC); select Nonpublic if the project data will not be released to or shared with the public

PUBLIC Release

☐ Public release – Full dataset
(Dataset can be made available without restrictions; data steward no longer controls data. ***This should be the default selection for all datasets unless justified otherwise****.)*

☐ Public release – Aggregate data
(Underlying dataset cannot be released or shared, but aggregate/summary data can be made available to public access without restriction.)

☐ Public release – Release by ad-hoc request

(Metadata will be released and the dataset is available by ad-hoc request; data requests CANNOT be denied; no data use agreement or restrictions; data steward no longer controls data.)

RESTRICTED Release

☐ Restricted use data sharing
(Dataset is available to particular parties under certain use restrictions or use agreement; data not always under CDC custody.)

☐ Restricted access data sharing
(Dataset is available only in an RDC; data need to remain under CDC custody.)

NONPUBLIC

☐ No release or data sharing

Access Rights/Restrictions

Public Access Justification – *For a Restricted Release or Nonpublic dataset, provide an appropriate justification for why the data collected/generated in this project cannot be released to/shared with the public.*

☐ Country/Jurisdiction owns the data with protections under its laws and regulations
☐ Not sharable for protection of intellectual property or trade secrets
☐ Removal of identifiers renders the remaining data of no value
☐ Other, please specify:

Data Use Type – *For a Restricted Release dataset, select the type of data use agreement that must be in place in order to release this dataset.*

☐ Data-Sharing Agreement
☐ Data Transfer Agreement
☐ Joint Statement of Understanding
☐ Memorandum of Understanding
☐ Other, please specify:

Data Use Type URL – *For a Restricted Release dataset, this is the website where the process for requesting access to the dataset can be found.*

Data Use Contact – *For a Restricted Release dataset, this is a name of a data steward/contact from whom external investigators can request access to the dataset.*

Section 5: Data Release and Documentation

Estimated Date(s) of Data Release:

Data Release Format – *Recommend to use nonproprietary format when possible, such as CSV, XML, JSON. Also specify data dictionary file format.*

External Access URL, If Known
Please provide the URL for external access to the documentation associated with this project. The information is to include protocol, data dictionary (e.g., variable names, definitions), data collection instrument and other relevant information.

Download URL(s), If Known
Provide the URL to a downloadable file of the dataset(s).
Type of Data Released – *Will the released data be the data tables associated with the publication, or will it be the raw data in either microdata or aggregated format?*

☐ Aggregated data (e.g., summary data or statistics)
☐ Microdata (e.g., information at the level of individual respondents)
☐ Tables

Data Release Documentation – *Identify the documents that will be provided to users (e.g., variable definitions, codebook, metadata file, guidance on data use). Check all that apply.*

What documents will be provided/available with the datasets?
☐ Variable definitions
☐ Codebook
☐ Data collection instrument
☐ Metadata file
☐ Guidance on data use
☐ Description of the population studied
☐ Methodology
☐ Description of dataset, such as response rates and limitations/caveats
☐ Other, describe here:

Describe Long-Term Preservation Plan: Describe the plan for archiving and long-term preservation of the data, or explain why this is not justified.