Data Management Plan (DMP) Guidance for Award Applicants and Recipients

National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP)

Centers for Disease Control and Prevention (CDC)

Prepared by: NCCDPHP Associate Director for Science Office
Last update: September 2020

Table of Contents

List of Abbreviations

AR-25 – Additional Requirement 25
CDC – Centers for Disease Control and Prevention
DMP – Data Management Plan
FAQ – Frequently Asked Question
FY – Fiscal Year
NCCDPHP – National Center for Chronic Disease Prevention and Health Promotion
NOFO – Notice of Funding Opportunity
PII – Personally Identifying Information

Overview

Public health data collected or generated with federal funds must be made readily available for the public to access (e.g. in a non-proprietary format and with a data dictionary) and archived for a prespecified amount of time, unless there are strong reasons not to do so. These tasks are the responsibility of the awardees, not CDC. The plan for these tasks must be prepared before the data are gathered and is called a Data Management Plan (DMP). CDC will evaluate DMPs for adequacy. This document provides instructions for writing a DMP that meets CDC requirements.

Introduction

The Centers for Disease Control and Prevention – the CDC – is the United States' principal disease prevention and health promotion organization. In support of its mission, CDC collects, generates, stores, uses, and routinely provides access to public health data. CDC also works with partners and funds institutions via grants, cooperative agreements, and contracts to implement public health programs, perform research, and collect, generate, and analyze data.

Public health and scientific advancement are best served when public health data are released to, or shared with, other public health agencies, researchers, and other partners in an open, timely, and appropriate way. CDC's Policy on Public Health Research and Nonresearch Data Management and Access (CDC Data Policy) states that the results and accomplishments from federally-funded contracts, grants and cooperative agreements should be made available to the public and the scientific community. Data management and accessibility of public health data involves ensuring: 1) data quality according to established standards; 2) consideration of privacy and confidentiality; 3) appropriate security and ethical concerns are met; 4) protection of intellectual property, and 5) impartiality in the sharing of data.

Public Health Data and the Data Management Plan (DMP)

A data management plan (DMP) is a written description of the plan for the collection, protection, sharing, and long-term preservation of public health data. It is a blueprint that will assist in planning for data management and sharing in advance of the actual data generation and collection. The DMP is a *living document*, meaning it must be updated and revised as the project evolves and throughout the lifecycle of the data collected.

As of fiscal year 2017 (FY17), all CDC contracts, grants, and cooperative agreements should include a DMP if the project will collect or generate (create) public health data. Grant and cooperative agreement recipients should follow the Additional Requirement 25 (AR-25) and any additional specifications included in the Notice of Funding Opportunity (NOFO). Those awarded via contracts should follow the language in the contract.

What is public health data?

Digitally recorded factual material commonly accepted in the scientific community as a basis for public health findings, conclusions, and implementation.

What is *not* public health data?

Grantee progress reports, process monitoring data, administrative data, preliminary analyses, drafts of scientific papers, plans for future research, reports, communications with colleagues, or physical objects, such as laboratory notebooks or laboratory specimens.

In a DMP, recipients are expected to describe how they intend to manage, preserve, and make accessible data generated or collected with CDC funds. The DMP should be developed during the project planning phase prior to initiating data generation or collection activities. If the project involves the collection or generation of public health data, CDC will require the DMP with the initial funding application or soon after the award begins.

In general, all CDC-funded public health data sets are expected to be made freely available to the public in a timely manner. If the data won't be made available to the public, unrestricted, a strong justification is required.

Determining if a DMP is Necessary

Data sets that necessitate a DMP include those that contain public health data and meet at least one of the following criteria:

- Collected or generated by CDC;
- Collected or generated by other agencies or organizations funded or co-funded by CDC (e.g., through grants, cooperative agreements, contracts, similar mechanisms); and/or
- Reported to CDC by another entity (e.g., by state health departments) that become a part of a CDC data collection system (e.g., CDC surveillance systems such as the Cancer Registry).

Data sets that do not usually necessitate a DMP include those:

- Collected or generated by organizations without CDC funding and shared for informational use with CDC, with the source organization exclusively retaining all rights to the data;
- Non-public health data, e.g. administrative data, project monitoring data, et cetera;
- Secondary data obtained from an outside source, e.g., data sets purchased from, or given by, another institution. Note: DMPs are sometimes required for secondary data sets, particularly when the original data were collected in a setting that didn't require a DMP and the data will now be used as public health data, e.g., medical record data used to evaluate success of a public health program or for surveillance.

If project has no public health data, no data management plan is needed, but the application should state this with an explanation.

Most research data collections are public health data and need a DMP. For non-research projects, some collections constitute public health data and some do not.

Conducting Evaluations and Assessments

When determining whether a non-research project involves the collection or generation of public health data, the purpose(s) of the data collection activity and the intended use of the data collected must also be considered. This is especially important when determining whether the policy applies to assessments and evaluations.

In nearly all circumstances, data collection activities included in quality improvement projects; organizational performance measurement and management projects; formative or process evaluations; customer satisfaction surveys; and needs assessments will not require a DMP.

If the specific purpose of an assessment or evaluation is to improve the design or operations of a program, process, system, or service delivery mechanism, and the data are not used to determine public health impact, the data are not considered public health data and a DMP is not required.

Example evaluation activity that does not require a DMP:

In a project funded through one of CDC's cooperative agreements or grants, the following activity is described: "The recipient [Partner A] is working with other agencies to implement the State Approach to Partnership in Parenting training. The recipient plans to implement a pre- and post-assessment to all participants of the training to evaluate the training's effectiveness at reaching intended learning objectives, and plan program improvements for future training events."

If the purpose of an assessment or evaluation is to understand the public health impact of a program, policy, or intervention, and/or if the results from the data collection will be used to make policy-related decisions, then the data are considered to be public health data, and a DMP is required.

Example assessment activity that requires a DMP:

In a project funded through one of CDC's cooperative agreements or grants, the following activity is described: "Department of Health (DOH) staff will update data collection related to the HP2030 objectives. DOH will prepare, receive, and analyze data for inclusion in meetings for policy decision making and publications."

Expanding Existing Public Health Data Sets Using Federal Funding

An existing data collection is defined as an activity that has been conducted previously and is typically repeated on a regular basis (e.g., an annual survey). Expansion of an existing data collection activity may include a) adding new variables or data elements, and/or b) including new subjects, cases, populations, or sites to the existing data set.

In some circumstances, CDC funds are used by cooperative agreement and grant recipients to expand an existing data collection activity that was not previously covered by this policy because it did not have federal funding or because it began prior to fiscal year 2017.

If CDC funds are used to add new variables or data elements to an existing data collection that was not covered by CDC's data policy, and the expanded data set is consistent with the definition of public health data, a DMP focused on the new or additional variables is required. A DMP covering the complete collection is also acceptable.

Example activity that requires a DMP for new data variables/elements:

In a project funded through one of CDC's cooperative agreements or grants, the following data collection activity is described: "[State] will use federal funds to add additional questions to its existing annual survey starting in year 2021. These additional questions are designed to meet the state's specific data needs."

If CDC funds are used to add new subjects, populations, or sites to an existing data collection and the expanded data set is consistent with the definition of public health data, then a DMP covering the entire data set is required.

Example activity that requires a DMP for the entire data set:

In a project funded through one of CDC's cooperative agreements or grants, the following data collection activity is described: "[State] will expand a current data collection project to include additional cases to a data set that is currently collected and maintained through private funding. This data collection project aims to develop a typology of American health values across the state, which will inform specific marketing, policy, and program decisions in the future. The implementation of the data collection activity will be expanded to include a larger random sample of the population across the state, increasing the sample size from 250 participants to 300 participants."

Secondary Data Set Acquisition

Some projects may require the acquisition and use of secondary data. In most cases, these data sets will not require a new DMP because the organization that originally collected or generated the data was responsible for creating the DMP if one was required. However, there are circumstances in which acquisition and use of secondary data does necessitate creating a DMP. These circumstances typically involve using the existing data in such a way that it *becomes* public health data, e.g. by using the data in a way that wasn't originally intended and/or by creating a new data set from it. CDC, at its discretion, may require a DMP for secondary data acquired and used with federal funding.

If secondary data that were not collected under a public health data DMP are to be used in a manner typical of a public health data set, CDC may require a DMP for the new data set.

Example secondary data analysis activity that will likely require a DMP:

In a project funded through one of CDC's cooperative agreements or grants, the following data collection activity is described: "This data set will be derived from identifiable patient clinical care records. If possible, publicly available community indicators will be merged with the data set. Longitudinal analysis will be performed to evaluate the impact of community and clinical interventions on health status indicators (e.g., screening rates, disease self-management)."

If secondary data were acquired under a purchase, license, or data use agreement for a purpose aligned with the original intent of the data collection, a DMP for the acquired data set is not required.

Example secondary data analysis activity that does not require a DMP:

In a project funded through one of CDC's cooperative agreements or grants, the following data acquisition activity is described: "This data set will be provided by a partner academic institution under a data use agreement. The original terms of the study and its informed consent language allow for the sharing and use of the data for this purpose."

What to Submit When the Funding Application Doesn't Require a DMP

Recipients of CDC funding who collect or generate public health data are responsible for creating and submitting a DMP to CDC unless CDC states or agrees otherwise. There are some situations in which CDC states an intention to take on the responsibility of creating the DMP itself or will accept a DMP created by a third party.

If your funding application does not include a DMP, the application must state why. Listed below are phrases are offered as examples of legitimate justifications. This is not a full set of possible justifications, but they cover some common scenarios.

No public health data:
"This project will not entail collecting or generating any public health data, i.e., the data that will be collected/generated do not serve as a strong basis for public health findings, conclusions, and implementation because …"

[example] "… the only data collected and generated in this project are administrative in nature (project accounting and contact information)."

[example] "… the only data involved in this project are performance measures which CDC will independently collect and report."

Another entity is responsible for the DMP:
"This project entails collecting or generating public health data, but our organization is not responsible for creating the DMP because …"

[example] "… CDC plans to aggregate and manage the data from multiple awardees and the NOFO/contract solicitation stated that CDC will assume responsibility for one overarching DMP."

[example] "… we will perform this work as part of a consortium and another member of the consortium, [Institution Name], will create and provide the DMP on behalf of all consortium members." Note: CDC may accept this approach at its discretion.

[example] "… we are collaborating with another federal agency, [Agency Name], which is assuming the responsibility for creating the DMP."

DMP decision is pending at time of application in response to non-research NOFO:
"A DMP may be needed for this project but has not been included in the funding application because …"

[example] "… we are not yet certain whether this non-research project will entail collecting or generating public health data. We will work closely with CDC during the first 6 months of funding to finalize our work plan and evaluation plan, and we will provide a DMP, if one is needed, within that timeframe."

[example] "… although this project will probably entail collecting or generating public health data, exactly what public health data we will collect or generate cannot be determined until our work plan is completed with CDC input. During the first six months of the award, we will work with CDC to finalize a work plan and an associated DMP."

Drafting a DMP

Some parts of CDC, including NCCDPHP, have OMB-approved templates to use when creating a DMP. DMPs should include all the following information:

1. Description of the data to be collected or generated in the proposed project
2. Standards to be used for collected or generated data
3. Mechanisms for, or limitations to, providing access to the data, including a description of provisions for the protection of privacy, confidentiality, security, intellectual property, and other rights
4. Statement of the use of data standards that ensure all released data have appropriate documentation that describes the method of collection, what the data represent, and potential limitations for use
5. Plans for archiving and long-term preservation of the data, or an explanation why long-term preservation and access are not justified

Element 1: Description of the Data

The DMP should describe the topic, type of study, data elements, and anticipated time frame and frequency of collection.

The DMP should also describe where data will be maintained, and who will be responsible for the data during the project period.

Element 2: Description of Standards for Collecting Data

The DMP should describe the use of widely accepted methods and procedures for ensuring the quality of data products.

The DMP should provide a plan to ensure data quality.

Element 3: Providing Access to Data

The default expectation is that data sets will be shared in an unrestricted manner with the public, in a venue and format that is freely available (e.g., in a non-proprietary format on a public web site). However, there are circumstances in which public sharing is not appropriate; these circumstances are discussed further below.

Assuming the data set will be shared in some fashion, the DMP should describe when, where, and how the data will be available (e.g., URL for downloading public access data set, discoverability (i.e., how you will ensure that potentially interested users can find out about the data set's existence and location) and procedures for gaining access to restricted data set).

Note that CDC will typically not host awardees' data sets or handle awardees' data release unless CDC has taken on the responsibility for the DMP.

Timelines for Data Sharing

CDC policy includes timelines for sharing data, as follows:

- one-time collections / generated data sets: share within 30 months after the end of data collection or generation
- ongoing collections such as surveillance data sets: share within 12 months of the end of a collection cycle

Protecting Privacy, Confidentiality, and Data Rights

CDC and its awardees have a responsibility to protect individuals' privacy and ensure confidentiality of the data collected, as well as any rights of other organizations involved in the collection. The DMP should have appropriate discussion of data security, privacy and confidentiality (removal of PII, protection of intellectual property rights, etc.). The plan should respect promises made in informed consent forms, data use agreements, data licensing agreements, etc.

Therefore, some data sets cannot be made public but could be made available on a restricted basis; some data sets should not be shared.

What is Personally Identifiable Information (PII)?

PII is information which can be used to identity an individual person, such as name, home address, social security number, medical record number, telephone number, and email address.

PII also means multiple data elements which could be used in conjunction to identify an individual, such as a combination of medical diagnosis, occupation, sex, race, birth year, and zip code.

Access Levels

The DMP should describe what level of access by the public will be provided. The options are:

- public: e.g., downloadable on an open website;
- restricted: e.g., potential users can apply for access based on specified criteria and/or can use the data set can only in a restricted setting;
- no public access: not made available to public in any way.

It is possible for a data set to be released using more than one of these options. For example, there could be both a public-use version of the data set which lacks certain data elements that are potential identifiers of the people whose information is in the data set when combined, as well as a restricted-use data set which includes more data elements but can only be used by individuals who demonstrate a legitimate public health need for access and sign a data use agreement promising confidentiality.

If the data won't be made available to the public, unrestricted, a strong justification that is aligned with CDC's Data Policy is required. Below are some of the possible reasons for the data access level to be either *restricted* or *nonpublic* (i.e., no access):

- Potential negative consequences for data release exceed potential benefits
- Country/jurisdiction owns the data with laws or regulations which don't permit public release
- Not shareable for protection of intellectual property or trade secrets (e.g., proprietary rights); precluded by licensing or other agreement
- Removal of identifiers renders the remaining data of no value; data cannot be shared without compromising subjects' privacy
- Cost of sharing the data set outweighs the expected benefit
- Data quality is poor/inadequate
- Adequate data already available publicly
- Data no longer relevant
- Data too complex for most users

If your data set access level will be restricted or non-public (no access), you must provide an acceptable justification. CDC is responsible for determining whether justifications in DMPs for not having a public release are acceptable.

Below are some possible examples of justifications:

Restricted data release:

"We do not plan to make a fully documented, publicly available data set on a website for download because …"

[example] "…it is not possible to remove identifiers to adequately protect subjects' privacy while still retaining useful variables. Therefore, we will make the data available only to selected researchers who apply for access. The application must document a legitimate public health use or research question for the data, and the applicant must sign a pledge to keep all data private and to make no attempt to identify individual subjects. This use of the data is consistent with the language of the voluntary consent for participation agreed to by the subjects."

[example] "… our environmental scan suggests that the likely number of users is limited to our four neighboring states. However, we will contact those state departments of health to let them know that the data set is available to them upon request. The data set will similarly be provided to any other members of the public who request it. Limited technical assistance will be provided to data set recipients to explain individual variables and coding. We calculate that it will be more cost-efficient for us to provide one-on-one technical assistance to all interested users than to create full documentation and host the data set on a website."

No data release:

"The data set will not be made available because …"

[example] "…the terms of our data use agreement / license with the data source preclude us from further sharing the data."

[example] "…the rapidly-changing landscape with respect to _____ suggests that the data will be useful immediately but will have an exceedingly short 'shelf-life'. We do not anticipate demand for the data set past the close of this project."

Occasionally, plans for release (or non-release) will change over the life of the DMP. This can happen if circumstances or situational awareness change, e.g., a data set becomes of greater or lesser interest to others, data quality is found to be inadequate, etc.

Element 4: Description of Standards Accompanying Release of Data

The DMP should describe what established standards (e.g., codes such as ICD10, FIPS; data formats) will be used for data and metadata to ensure usability and interoperability of the data.

The DMP should outline what documentation will be available for analysis (e.g. data dictionary, sample software code). Appropriate documentation is required.

The DMP should outline documentation that will be available regarding data source (e.g. population studied, response rate, caveats). Appropriate documentation is required.

Element 5: Archiving and Long-Term Data Preservation

The DMP should provide a description of the planned long-term preservation; CDC does not mandate where the data should be archived, nor a timeframe for how long it should stay. If long-term preservation is not planned, provide a strong justification for why the data will not be stored perpetually or at all. As a benchmark, CDC archives most research records for 6-20 years and archives highly influential data permanently.

The DMP should describe the planned final location of the data (publicly accessible repository, institutional or governmental repository, etc.). Suitable repositories are accountable, secure, and long-lasting.

The DMP should provide a link or other contact information for the archived data with a description of when the data can be accessed, who has access, and how.

Awardees may use their own sites, privately-owned repositories, or government-based sites (e.g., NIH Repositories). CDC will not take on the responsibility for long-term preservation and archiving unless CDC has aggregated the data for awardees and has taken on the responsibility for the DMP.

Submitting a DMP to CDC

A DMP is a *living document* that must be updated, submitted, and approved for accuracy as plans solidify or change during the project's period of performance. In other words, it is acceptable to have not finalized all decisions during a project's early years. The submission and assessment of a DMP will occur:

- During the initial merit review of an application submitted in response to a contract solicitation or NOFO, and/or when a cooperative agreement evaluation plan is completed at 6-months post award **and**

- At least annually thereafter (at the time of non-competing continuation application for cooperative agreements and grants), **and**
- When the final report is submitted at the close of the funding period.

A plan addressing each element is required in every version of the DMP. A statement discussing how you intend to address each element, even if a decision has not yet been made, is adequate in the early stages of the project. These statements will indicate that the extramural awardee understands the criteria and the DMP's purpose. By the end of the project, the final DMP should contain all details and be precise.

In funding applications and reports, the DMP should be a stand-alone section with an appropriate header. When revised DMPs are submitted in response to CDC's request for a revision, they should be submitted to the CDC program officer (Project Officer, Evaluator, or Technical Monitor) as a stand-alone document.

The CDC program officer is responsible for assessing submitted DMPs and letting awardees know if a submitted DMP needs revision to bring it into alignment with CDC policy or program expectations.

Journal Publications

If the federally-funded data set is used in a peer-review journal publication, the data should be released at the time of publication unless the data set has already been released (by you or another entity) or cannot be shared for some reason. Note that publication does not create a *post-hoc* need for a DMP because the DMP is a plan for the data set itself. However, if you anticipate publishing from the outset, you should create a DMP.

Data can be released via the journal (some journals require the submission of the data set as a condition of publication and post it on their site) or by another method, such as release on your own website. The form and content of the released data can range from a machine-readable version of the tables in the paper to a full line listing of the underlying data.

Considerations When Writing a DMP

- The DMP must be consistent with any promises made by the data owner to others, as in voluntary consent language and data use agreements. Similarly, such language should be worded to be in accordance with the awardee's and CDC's expectations for the DMP.

- Recipients are fully responsible for data stewardship, data preparation and making data accessible, and any technical assistance to be provided to data users. Unless CDC has specified that it will take on any of these roles, as in cases where CDC aggregates and disseminates surveillance data from multiple award recipients, CDC will not bear responsibility for these tasks.

- Award recipients should ensure the quality of the data that they make accessible, and should follow the data quality standards set by their organization prior to making data accessible.

- Costs associated with public health data collection, access, archiving, and long-term preservation may be included as part of the total budget submitted with grant and cooperative agreement applications.

- Entities making data available to the public may use cost-recovery methods if they are justifiable, e.g., when the cost is high and the expected number of users is low.

- No funds will be available beyond the period of performance of the award. Award recipients may pre-pay vendors for hosting and repository services during the award period.

Appendix A: Frequently Asked Questions (FAQs)

Needs for DMPs for projects beginning prior to FY2017:

Q: A five-year, one-off cooperative agreement was first funded prior to FY2017 and was expected to involve collection of public health data since inception. The project is still active and collecting the originally envisioned data today. Is a DMP now needed?
A: No.

Q: A five-year, one-off cooperative agreement was first funded prior to FY2017 and wasn't expected to involve public health data when it began, so no DMP was written. The project design has now changed and will involve a collection of public health data. Is a DMP now needed?
A: Yes.

Ownership of, and responsibility for, data collected by awardees with federal funds:

Q: An awardee collects data with federal funds. Must they submit a copy of the data set to CDC?
A: Only if the NOFO or contract states that CDC must be given a copy of the data.

Q: An awardee collects data with federal funds but doesn't submit that data set to CDC. Which institution owns the data?
A: The awardee retains full rights to their data. CDC does not have rights to it.

Q: An awardee collects data with federal funds and does submit that data set to CDC. Which institution owns the data?
A: The awardee retains rights to use their data. CDC also has rights to the reported data and can use it as CDC wishes, unless the terms of the data use agreement, NOFO or contract limit what CDC can do with the data.

Q: An awardee collects public health data with federal funds and does submit that data set to CDC. Which institution is responsible for creating the DMP? Which institution is responsible for data sharing?
A: Unless otherwise stated by CDC, the responsibility for the DMP and for sharing remains with the awardee. In cases where CDC is aggregating similar data from multiple awardees, CDC may choose to be responsible for the DMP and for data sharing.

How to address inability to share data or uncertainty:

Q: What if project data cannot be made available due to security, confidentiality or privacy concerns?

A: It is understood that not all data can be made publicly available due to security, confidentiality, or privacy concerns. Should the recipient determine that the data cannot be made available for public use, a written justification is required.

Q: My project is new. What if all details of the project design are not yet certain?

A: Every element of the DMP should be addressed in each update, even if final decisions have not yet been made. An element can be described generally or stated that the final decision, e.g., for the location of the downloadable dataset, hasn't yet been made. Such entries will indicate that the extramural awardee understands the criteria and the DMP's purpose. By the end of the project, the final DMP should contain all details and be precise.

Q: If the project won't need a DMP, what should be said about a DMP in the application?

A: If the project won't collect or generate public health data, the application should state that no DMP is included because there will not be public health data or because CDC will maintain the DMP for the project.

Q: What if CDC thinks the project will require a DMP when preparing the NOFO, but the way in which a particular applicant proposes to do the project doesn't need a DMP? Or vice versa.

A: It is possible that an applicant will propose an approach CDC wasn't expecting, and this different approach could affect whether there is public health data necessitating a DMP. Whether or not a DMP is needed depends upon what the awardee will actually be doing rather than CDC's initial expectation. An applicant who proposes something other or beyond what CDC expected might need a DMP that encompasses that additional/different work.

Funding considerations:

Q: Can applicants request funds to implement data management in the proposal budget?

A: Yes. Applicants can include funds in the budget to carry out the tasks described in the DMP during the lifetime of the project; however, additional CDC funds cannot be accessed after the project closes. It is allowable for awardees to pre-pay, using award funds, for a repository for long-term access and preservation that will occur after the project closes. There is not a fixed amount or percentage allocated to data management tasks.

Q: Will CDC's assessment of the DMP affect whether an application is funded?
A: The quality of the included DMP will be assessed but this assessment will not contribute to the application's score when CDC determines which applications to fund. CDC will work with non-research awardees during the first 6 months of the award to address any deficiencies identified in the initial DMP. For research awards only, CDC may withhold some funds pending satisfactory revision of the DMP.

Appendix B: Additional Requirement - 25 (AR-25): Data Management and Access

CDC requires recipients for projects that involve the collection or generation of data with federal funds to develop, submit and comply with a Data Management Plan (DMP) for each collection or generation of public health data undertaken as part of the award and, to the extent appropriate, provide access to, and archiving/long-term preservation of, collected or generated data.

"Public health data" means digitally recorded factual material commonly accepted in the scientific community as a basis for public health findings, conclusions, and implementation. Public health data includes those from research and non-research activities.

Public health data could be quantitative, qualitative, imaging, or genomic output (for example, genome sequencing, arrays, gene expression, etc.). Public health data do not include preliminary analyses, drafts of scientific papers, plans for future research, reports, grantee progress reports, communications with colleagues, or physical objects, such as laboratory notebooks or laboratory specimens.

Data Management Plan

Consistent with the terms of and activities expected under the notice of funding opportunity (NOFO), recipients must develop and submit a DMP generally during the project planning phase, but in any event, prior to the initiation of generating or collecting public health data. Accordingly, the DMP may be evaluated during the application, study proposal, or project review process or during other times in the period of performance. For NOFOs that involve already defined projects, which include data collection or generation at the time of application, applications submitted without the required DMP may be deemed non-responsive for award. For NOFOs where CDC specifies that submission of the DMP is deferred to a later period, funding restrictions may be imposed pending submission and evaluation of the DMP. For awards where data collection or generation activities may become necessary during the period of performance, DMPs will be required to be submitted and evaluated during the period of performance of the award. These DMPs also will be required to comply with this AR. In all instances described above, the reviewing officials have to approve an acceptable DMP. Costs associated with developing and implementing a DMP, including costs of sharing, archiving and long-term preservation, may be included in the budget submissions for grants and cooperative agreements.

A DMP for each collection and/or generation of public health data funded by this award should include the following information:

- A description of the data to be collected or generated in the proposed project;
- Standards to be used for the collected or generated data;
- Mechanisms for or limitations to providing access to and sharing of the data (include a description of provisions for the protection of privacy, confidentiality, security, intellectual property, or other rights). This section should address access to identifiable and de-identified

data or justification for not making the data accessible (see below for additional information about access);
- Statement of the use of data standards that ensure all released data have appropriate documentation that describes the method of collection, what the data represent, and potential limitations for use; and
- Plans for archiving and long-term preservation of the data or explaining why long-term preservation and access are not justified. This section should address archiving and preservation of identifiable and de-identified data (see below for additional information regarding archiving).

Access to and Archiving of the Data

Recipients whose terms of award do not include submitting data to CDC are expected to plan and prepare for access to, and archiving/long-term preservation of, collected and/or generated data within the funding period, as set forth below. The final version of a collected and/or generated data set intended for release or sharing should be made available within thirty (30) months after the end of the data collection or generation, except surveillance data that should be made accessible within a year of the end of a collection cycle. In addition, recipients should ensure the quality of data they make accessible and seek to provide the data in a nonproprietary format. If data cannot be made accessible, a justification for not doing so should be provided in the final DMP. Recipients who fail to release public health data in a timely fashion may be subject to procedures normally used to address lack of compliance consistent with applicable authorities, regulations, policies or terms of their award.

For public use de-identified (removal of sensitive identifiable or potentially identifiable information) datasets, an accompanying data dictionary, codes, and other documentation relevant to use of the data set should be deposited in a sustainable repository to provide access to the data. Data that cannot be de-identified can be provided on request under a data use agreement.

Recipients will be required to inform the appropriate CDC point-of-contact identified in the award via an update to their DMP of the location of the deposited data. The DMP is a living document that should be updated throughout the life cycle of data.

For data underlying scientific publication, recipients should make the data available coincident with publication of the paper, unless the data set is already available via a release or sharing mechanism. At a minimum, release of the data set should consist of a machine-readable version of the data tables shown in the paper.

Requirements set forth in this policy are not intended to conflict with or supersede applicable grants regulations related to agency access to recipient data and records

See https://www.cdc.gov/grants/additional-requirements/ar-25.html.