Centers for Disease Control and Prevention

Additional Requirement – 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.

Page last reviewed: June 29, 2021