CREATING A DATA SHARING PLAN FOR NIH

The precise content of the data-sharing plan will vary, depending on the data being collected and how the investigator is planning to share the data.

Questions to answer in a NIH Data Sharing Plan (as appropriate):

What data will be shared?

Who will have access to the data?

Where will the data be available? How will researchers locate and access the data (on your website, available via ftp download, via e-mail, or archive/repository)? Which archive/repository/central database have you identified as a place to deposit data (if appropriate)?

When will you make the data available?

Will a data-sharing agreement will be required and, if so, include a brief description of such an agreement (including the criteria for deciding who can receive the data and whether or not any conditions will be placed on their use)

Data Details as Appropriate -

Which file formats will you use for your data, and why?

What transformations (to more shareable formats) will be necessary to prepare data for preservation/data sharing?

What metadata/documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?

What other types of information should be shared regarding the data, e.g. the way it was generated, analytical and procedural, information?

Address any other specific requirements listed in the solicitation.

EXAMPLE DATA SHARING PLANS

Example 1
The proposed research will involve a small sample (less than 20 subjects) recruited from clinical facilities in the New York City area with Williams syndrome. This rare craniofacial disorder is associated with distinguishing facial features, as well as mental retardation. Even with the removal of all identifiers, we believe that it would be difficult if not impossible to protect the

identities of subjects given the physical characteristics of subjects, the type of clinical data (including imaging) that we will be collecting, and the relatively restricted area from which we are recruiting subjects. Therefore, we are not planning to share the data.

Example 2

The proposed research will include data from approximately 500 subjects being screened for three bacterial sexually transmitted diseases (STDs) at an inner city STD clinic. The final dataset will include self-reported demographic and behavioral data from interviews with the subjects and laboratory data from urine specimens provided. Because the STDs being studied are reportable diseases, we will be collecting identifying information. Even though the final dataset will be stripped of identifiers prior to release for sharing, we believe that there remains the possibility of deductive disclosure of subjects with unusual characteristics. Thus, we will make the data and associated documentation available to users only under a data-sharing agreement that provides for: (1) a commitment to using the data only for research purposes and not to identify any individual participant; (2) a commitment to securing the data using appropriate computer technology; and (3) a commitment to destroying or returning the data after analyses are completed.

Example 3

This application requests support to collect public-use data from a survey of more than 22,000 Americans over the age of 50 every 2 years. Data products from this study will be made available without cost to researchers and analysts. https://weblogin.umich.edu/?cosign-www.umich.edu&https://umich.edu/. User registration is required in order to access or download files. As part of the registration process, users must agree to the conditions of use governing access to the public release data, including restrictions against attempting to identify study participants, destruction of the data after analyses are completed, reporting responsibilities, restrictions on redistribution of the data to third parties, and proper acknowledgement of the data resource. Registered users will receive user support, as well as information related to errors in the data, future releases, workshops, and publication lists. The information provided to users will *not* be used for commercial purposes, and will *not* be redistributed to third parties.

Example 4

Sharing of data generated by this project is an essential part of our proposed activities and will be carried out in several different ways. We would wish to make our results available both to the community of scientists interested in [this disease] and the biology of [its causative agent] to avoid unintentional duplication of research. Conversely, we would welcome collaboration with others who could make use of the vaccine assessment protocols developed in [the project].

Our plan includes the following:

<u>Presentations at national scientific meetings</u>   From the projects, it is expected that approximately four presentations at national meetings would be appropriate. There is an

annual [Disease] Study Group meeting, of which the PI is secretary. This one-day meeting of interested persons presents new information on a variety of topics related to [the disease]. It is expected that the investigators from this [project] will be active participants of this focused group.

<u>Annual lectureship</u> A lectureship has brought to the University distinguished scientists and clinicians whose areas of expertise were relevant to those interested in [the disease]. Lecturers have been [list of names]. Visiting lecturers will be scheduled to interact with the investigators of the project as appropriate with their specific areas of expertise which will provide an opportunity for members to present their work to the visitor.

<u>Newsletter</u>   The [disease interest group] publishes a newsletter which currently has a circulation of [number]. The newsletter's intent is to disseminate new information regarding [the disease]. The activities and discoveries of [the project] will be allocated 20% of the newsletter's coverage.

<u>Web site of the Interest Group</u> The [interest group] currently maintains a Web site where information [about the disease] is posted. Summaries of the scientific presentation from the [quarterly project] meetings will be posted on this Web site, written primarily for a general audience. [Link to Web site]

<u>Annual [Disease] Awareness week</u> Beginning this fall during the week of [date], the [interest group] will be sponsoring a [Disease] Awareness week. As part of that program, there will be a research poster display with discussions. In future years, [the project investigators] will be active participants in this program.

<u>SAGE Library Data</u> [This project] will generate data from several SAGE libraries. It is our explicit intention that these data will be placed in a readily accessible public database. All efforts will be made to rapidly release data through publication of results as quickly as it is possible to analyze the experiments. Data used in publications will be released in a timely manner. SAGE data will be made accessible through a public site that allows querying as has been set up for a similar project. This site can be accessed at [link to Web site].