## HRA Member Data-Sharing Survey Results

**Purpose**: In January 2021, HRA members were surveyed about their data-sharing policies and practices. a survey. Of the 29 organizations who responded to this survey, 18 of them reported have a data-sharing policy-practice in place. This document is intended to report the raw responses to the survey. It is for HRA membership only and not to be distributed.

***Thank you to each organization who responded to this survey. You are helping all of us make progress towards our data-sharing goals!***

# Contents

| Organization | Data-Sharing Policy?* | | | Applicable to.. | | When do Researchers Share Data-Sharing Plan? As part of… | | | Influence Scoring | Costs Associated with Data-Sharing Allowed in Budget | | Compliance |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Y | N | Thinking about it | Entire Portfolio | Specific Grants | Application | Terms | Progress Report | Y | Costs Allowed | Provide Additional $ | Y |
| AFAR | | X | | | | | | | | | | |
| Alex's Lemonade Stand Foundation | R | | | X | | X | X | X | X | X | | |
| American Brain Tumor Association | E | | | | X | X | | | | X | | |
| American Heart Association | R | | | | X | X | | | | X | | X |
| American Society of Hematology | | X | | | | | | | | | | |
| Barth Syndrome Foundation | | | X | | X | X | | | | | | |
| Cancer Research Institute | | | X (E) | | | | | | | | | |
| Children's Tumor Foundation | R | | | X | | X | | | X | X | | X |
| Circle of Service Foundation | E | | | | X | | X | | | | | |
| CMT Research Foundation | | | X | | | | | | | X | | |
| Crohn's & Colitis Foundation | R | | | | X | | | | X | | | X |
| Damon Runyon Cancer Research Foundation | E | | | | X | | | X | | X | X | |
| Doris Duke Charitable Foundation | E | | | | X | X | | | X | | X | X |
| Health Resources in Action | E | | | X | | | | | | X | | |
| Hydrocephalus Association | | E | | | | | | | | | | |
| Kenneth Rainin Foundation | | X | | | | | | | | | | |
| Lupus Research Alliance | R | | | X | | | X | X | | X | | X |
| Lymphoma Research Foundation | E | | | X | | | | | | X | | |
| Melanoma Research Alliance | R | | | X | | X | | | | X | | |
| Misophonia Research Fund of The REAM Foundation | R | | | X | | X | X | X | X | X | | |
| NYSCF | | X | | | | | | | | | | |
| Pershing Square Sohn Cancer Research Alliance | R | | | X | | | | | | X | | |
| Simons Foundation | R | | | X | | X | X | X | | X | X | |
| St. Baldrick's Foundation | E | | | X | | X | | X | | X | | X |
| Susan G. Komen | R | | | X | | X | X | X | X | X | | X |
| The Gerber Foundation | | X | | | | | | | | | | |
| The Kavli Foundation | | X | | | | | | | | | | |
| The V Foundation for Cancer Research | E | | | | X | X | X | X | | X | | |
| The Vallee Foundation | | X | | | | | | | | | | |

*Please note that respondents have consented to sharing this information with HRA Membership. Not for distribution.*



*R = Require / E = Encourage

## AFAR

**Primary Survey Contact**: Odette van der Willik, Deputy Executive Director
**Email**: odette@afar.org

*Notes: Questions 5-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. Enforcement and staff capacity
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments
11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## Alex's Lemonade Stand Foundation

**Primary Survey Contact:** Anna Greene, Director of Science
**Email**: a.greene@alexslemonade.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    - We currently require a resource sharing plan as part of grant applications, but are rolling out a more formalized policy later this year.

2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    - N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    - See attachments after question 22.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    - N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    - We currently require a resource sharing plan as part of grant applications, but are rolling out a more formalized policy later this year.
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    - Educating applicants/grant recipients on what is meant by resource and data sharing, educating grant reviewers on rating resource sharing plans, conducting an internal review of plans before funding, and ensuring compliance with sharing plans/policies.
9. Do you specifically require or encourage your awardees to share data?
    - Require
10. Comments
    - N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    - We encourage the use of established repositories to the extent possible.
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    - Part of the application, terms and conditions and the progress report.
13. Comments
    - The plan is a requirement of the application. We have verbiage requiring sharing in our Grant Conditions document, and we ask them for evidence of sharing on their progress and final reports.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    - No later than the time of publication or within 12 months after the conclusion of the grant funding, whichever comes first.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    - Yes - it is part of the peer review
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    - Yes
17. Comments about training material
    - We provide a written rubric, though I don't think it's often used. We need to strengthen how we train for this; we plan to implement a reviewer training call this year for our grant categories, including resource sharing in this training session.
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    - Yes
19. Comments

- I think this is important, but I rarely see PIs take advantage of this unfortunately. This is likely an opportunity for improved education in this space.

20. Do you provide any supplemental financial support to facilitate data sharing?
    - No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    - No
22. What are the consequences for noncompliance with your policy?
    - Right now we say that future grant funding is contingent upon compliance, but we are uncertain how best to monitor compliance across a large grant portfolio. I would love to hear about other's experience in this space.

*Notes: We've been using this for ~2 years now. I'm also attaching a copy of our newly developed Resource Sharing Policy. We plan to roll this out later this year along with our Resource Portal (https://resources.alexslemonade.org/), where we will require grantees to list any resources developed under our funding to make them discoverable and requestable by others in the field. We're really excited about it! We haven't finalized the sharing policy, and the portal isn't fully developed yet, so please keep in mind they may change before final release.*

# ALEX'S LEMONADE STAND FOUNDATION

# RESOURCE SHARING POLICY

Resource sharing is an expected outcome for grants funded by Alex's Lemonade Stand Foundation (ALSF). Sharing research outputs accelerates scientific advances, improves the quality and reproducibility of the work, and reduces redundancies in the research process. The goal of the ALSF Resource Sharing Policy is to enable faster translation of research discoveries into cures for children with cancer. Resources to be shared encompass all unique research outputs developed, including but not limited to: model organisms, cell lines, plasmids, protocols, software, and data. We expect that, where available, resources will be deposited and archived in standard public repositories. Repositories ensure that resources are discoverable and archived, while reducing researcher burden for maintenance and distribution of a resource over time. We strongly discourage sharing upon request. Furthermore, we expect that sharing will be timely and unbiased, and outputs shared will be of the highest quality possible.

This policy defines sharing expectations which should guide applicants' Resource Sharing Plans that are submitted as part of new applications for ALSF grant funding. The Resource Sharing Plans are reviewed as part of the peer-review process as a scored criterion. Past sharing behavior plus the planned sharing approach are reviewed. Grant recipients are asked to report on adherence to their sharing plan in progress and final reports and are required to post unique research resources to the ALSF Childhood Cancer Research Resources Portal to make them discoverable and requestable. Future ALSF grant funding is contingent upon faithfully adhering to the approved Resource Sharing Plan and the resource sharing policies described herein.

## OPEN ACCESS POLICY

ALSF is committed to sharing research information to ensure research transparency and enable unrestricted access to research results. Grant recipients must submit all publications, excluding non-research articles such as review articles, that were in part or fully funded by ALSF as a preprint to bioRxiv, medRxiv or a similar preprint sharing service prior to or at the time of initial journal submission. Preprints should have a CC BY or CC BY-NC license applied. This allows researchers to immediately begin building upon these results to accelerate the pace of scientific discovery.

## ALSF CHILDHOOD CANCER RESEARCH RESOURCES PORTAL

The ALSF Childhood Cancer Research Resources Portal has an overarching goal to bring ultimate transparency to childhood cancer research resources available in the community, allowing users to post and request resources within the portal. All unique research outputs developed with ALSF grant funding within the grant period are required to be posted to the ALSF Childhood Cancer Research Resources Portal prior to or at the conclusion of ALSF grant funding. Resources developed with ALSF funding that are finalized after the grant end date are also required to be listed on the portal. To make posting resources straightforward, the portal allows researchers to import resources that are already listed with public repositories. Resources that are not stored in a repository can be listed manually to make the community aware that it exists and is available for request. Grant recipients are encouraged to continue using the portal, posting and requesting resources as desired, after the grant period ends.

## Material Resources

We require that unique resources (e.g. cell lines, plasmids/clones, antibodies, transgenic organisms, and other reagents) generated with ALSF funding be shared openly with the research community no later than the date of publication or within 12 months after the end of grant funding, whichever comes first. We expect that, where available, resources will be deposited and archived in public, widely-used repositories, such as Addgene for plasmids/DNA reagents/viruses and Jackson Labs for model systems lines, rather than distributing materials "on request".

Distributing materials "on request" should be avoided if possible. If distribution will occur upon request, applicants, in their Resource Sharing Plan, should specify the expected response time for requests for the resource, how the response time will be measured, who will be responsible for maintaining and sharing the resource, and how sharing will happen after the grant ends. If the lab will be maintaining the material or reagent, methods that will be used to authenticate the reagent should be specified, as well as when the authentication will occur and who is responsible (distributor, recipient).

## Data

Data generated through ALSF-funded research projects are to be made publicly available with as few restrictions as possible and easily accessible online through an appropriate license, such as CC0 or CC BY. We support FAIR data standards stating that data should be Findable, Accessible, Interoperable, and Reproducible. Data types include, but are not limited to: all "omics" data, imaging data, screening data, and any underlying data needed to accurately and independently reproduce experiments and findings. Data should not be made available "upon request". Data must be shared to a public repository no later than the time of publication or within 12 months after the conclusion of ALSF grant funding, whichever comes first. Furthermore, a data accessibility statement that states where the data are available and how they can be accessed must be included in all publications referencing data collected through ALSF funding.

We understand that some data cannot be shared publicly and must instead be shared in a controlled access manner. There are also limited circumstances in which data cannot be shared, when it would violate human subjects privacy regulations, superseding regulations (laws or institutional policies), intellectual property grounds or financial grounds, if sharing would cause undue financial burden. Any limits to data sharing must be explained in the Resource Sharing Plan upon grant application.

### Metadata

For data to be maximally useful, there must be metadata that describe the data (i.e. the methodology for collecting the data, definitions of variables, units of measurement, etc). We recommend using standard terminology to describe and structure data.

### Data Repository requirement

We expect that raw data will be deposited in a public, accessible data repository that assigns a permanent study identifier such as an accession number or DOI (digital object identifier). A best practice is to use public repositories already established in a research field (e.g., GEO or ArrayExpress for gene expression data, SRA for RNA-Seq data, etc.). When there is no public, widely-used repository available, a general purpose archival repository such as Figshare, Dryad or Zenodo should be used. See here for NIH's guide for the selection of data repositories.

### Patient Consent and Data Transfer Agreement Criteria

Patient consent information must be provided with the application and any limitations with regards to data use/sharing as per the consent should be highlighted in the ALSF Resource Sharing Plan. Any consent forms that will be used as part of a funded study must include provisions regarding data sharing with "promoting research initiatives at other institutions" as one of the stated intended uses of the data referenced in the Consent form. With regard to Data Transfer Agreements, it is critical that data sharing is not unnecessarily hindered by the language of the Agreement. To further data sharing, researchers must commit to suggesting language that is minimally restrictive, such that any limitations on data sharing are narrowly tailored to actual risk to privacy or re-identification.

### Considerations for making data publicly available:

Take into consideration the following to determine if data should be public or controlled access.

1) Publicly Available: the key here is to allow public access to all information that is de-identified, for example:
    a. clinical information that cannot be used to identify the patient, maximum amount of data allowable under HIPAA guidelines (e.g. while date of birth is not allowable, age in month and year is permissible unless this information is somehow uniquely identifiable)
    b. Omic profiling data for tumors, such as gene expression data, somatic mutations, tumor specific copy number and structural alterations, comparative genomic hybridization data, epigenetic and proteomic profiling data and related types of genomic data that are output oriented, and de-identified
    c. Specific molecular diagnosis, tissue pathology data, small molecule screenings and profiling
2) Controlled Access: the key here is data that poses a significant risk of re-identifying the patient
    a. patient/tumor information and unverified or raw molecular data (e.g. certain array-based and sequencing files) that pose a significant risk of patient re-identification
    b. some specific patient information such as demographic and phenotypic information (depending on disease rarity)
    c. Germline variants
3) Not at all:
    a. Any data that could be identifiable, for which a patient has not provided consent

## Source Code

Source code developed with ALSF funding must be stored in a version control system and made available through a version control service (e.g. GitHub, Bitbucket, or similar). A permissive license for the source code (such as MIT, BSD 2-Clause Plus Patent License, or Apache v2.0) must be specified because otherwise copyright is retained by the researcher and thus no one else is able to reproduce, distribute, or create derivatives from the work. All pre-existing and derivative code should be licensed under the most permissive license possible, given the licensing terms of the pre-existing code. Source code must be archived to an archival service (e.g. Zenodo) at the time of submission of manuscripts that rely on the source code and the conclusion of the grant.

## Protocols

Experimental protocols generated by ALSF-funded work should be publicly shared through a protocol sharing service, such as protocols.io, prior to or by the date of publication or within 12 months after the end of grant funding, whichever comes first. These services allow protocols to become living records of the core experiments underlying research results

and make reproducing experiments easier and more transparent for the labs that develop the methods as well as for others looking to build upon the work.

## Clinical Trial Reporting

ALSF-funded clinical trials, funded in part or in full, must be registered with ClinicalTrials.gov prior to the initiation of the study and must publicly report summary results there no later than 12 months after trial completion. Furthermore, in addition to positive trial results, negative and inconclusive results must be published in a timely fashion.

Clinical trial data should be made available at the time of publication or no later than 12 months after the completion of the trial, defined as submission of the final trial report to regulatory authorities. Any restrictions to this must be outlined in the Resource Sharing Plan at application submission.

## Resource Sharing Costs

We recognize that comprehensive resource sharing has an associated cost (e.g. curation fees, data repository fees, shipping costs, etc.). Therefore, researchers are encouraged to and should provide a budget in their application that reflects a reasonable assessment of these costs.

# ALSF Resource Sharing Form and Guidelines



Questions regarding the Resource Sharing application section or review criteria can be directed to Grants@AlexsLemonade.org.

# Application Resource Sharing Section: Track Record and Plan

Application reviewers will be asked to consider the manner in which resources will be shared and the extent to which that plan, as well as the investigator's track record* of sharing useful outputs, will increase or decrease the impact of the proposed project. This will depend on the extent to which sharing enhances or diminishes the perceived value of the work.

- Complete relevant categories for unique research outputs expected from this grant.
- Delete unused categories; use "Other" for additional categories.
- You should delete the instruction text in italics when completing this form.
- Copy and insert the completed form into the Resource Sharing section of the application outline.

**Early Career investigators applying for Young Investigator, 'A' Award or Psychosocial Launch grants are encouraged to describe past experience; however, it is understood this may be limited. The review will focus on how you would share outputs from this project.*

## FORM (1-page maximum)

**Data Sharing:**

- *Highlight how you have shared data publicly – i.e., not upon request – and how those data have been reused. Illustrate with reuse metrics such as citation counts, downloads, or other such data if available.*
- *Discuss how you plan to share the outputs from this proposal and how the data will be archived (via the recognized repository for the type of data or, for data without such a repository, via Zenodo, FigShare, or similar archival services). How will data be licensed (i.e., CC0 or another license). You must discuss how and when data that you generate during the course of this project will be shared. If access will be controlled via a data access committee or other such structure, describe the conditions under which data will be shared and specify how relevant metrics (number of requests made, number of requests approved, time to respond to requests) will be stored and reported to us and the scientific community.*

**Protocol Sharing:**

- *Highlight how you have shared protocols openly – i.e., not upon request – and how those protocols have been used by others. For example, you may have posted them to protocols.io or a similar service.*
- *Discuss how and when you plan to share the outputs from this proposal. Not all projects will result in protocols. If yours does not, this section can be deleted.*

**Material and Reagent Sharing:**

- *Highlight how you have shared materials and reagents and how those reagents have been reused.*
- *Discuss how and when you plan to share the reagents and materials developed in your group as part of the proposal (e.g. deposit plasmids in Addgene, deposit cell lines in the appropriate cell bank). Not all projects will produce new materials and reagents. If yours does not, this section can be deleted.*

**Source Code Sharing:**

- *Highlight how you have shared source code, software, and computational workflows openly – i.e., not upon request – and how the source code has been used by others. For example, you may have uploaded them to GitHub or a similar service.*
- *Discuss how and when you plan to share the outputs from this proposal. How will software be licensed (i.e., MIT or another license)? Are there plans to produce a polished software package? If so, how will that be distributed? Not all projects will result in source code. If yours does not, this section can be deleted.*

**Clinical Trials Reporting:** *If you propose a clinical trial discuss how you will maintain up to date records on the relevant repositories (e.g., clinicaltrials.gov). Discuss the trials you have run in the past and the extent to which those records have been maintained. Please provide links.*

**Other Outputs:**

- *Highlight how you have shared other outputs and how those outputs have been used by others.*
- *Discuss how and when you will share other expected outputs from this work.*

# Resource Sharing Example:

**Data Sharing:** In previous projects, we performed gene expression analysis of treated and untreated cell lines. We uploaded our data to NCBI's GEO repository at the time the data were collected, and we made these data openly available with the publication of our manuscript [1]. In GEO these have been assigned the identifiers GSE1245, GSE1246, and GSE1247. We annotated these data with treatment date, processing batch, cell line, and treatment type. These data were downloaded and reanalyzed by Doe et al. [2] and Smith et al. [3] to identify additional targets. These data were integrated into a larger analysis of multiple datasets by Patel et al. [4]. In this project we will perform RNA-seq analysis of XYZ cell lines. We will upload sequencing data to SRA and link the raw data to summary information in NCBI's GEO repository. We will annotate experimental metadata using terms from the Experiment Factor Ontology (EFO) where relevant terms are available. We will make these data publicly available to the community at the time of publication.

# Rubric for Reviewers:

Please use the full range of scores (1-9) for this criterion. We expect that very few applications will receive a perfect score in this area.

## General Track Record:

- Do the applicants have a track record of sharing resources that are remarkable for their richness, granularity, or quality such that those resources are particularly inviting to people who wish to use them.
- Do the applicants have a track record of sharing resources in a manner that is as easy as possible for people to re-use within ethical and legal constraints.
- Have the applicants shared resources that have *already been reused* by other investigators to answer a new question?
- Early Career Grants: Young Investigator, 'A' or Psychosocial Launch. Applicants are encouraged to describe past experience; however, it is understood they may not have a track record. The reviewer should focus on the Sharing Plan.

## General Resource Sharing Plan:

- Do the authors use an established repository for the resource? (See AHA guidelines on repositories for questions https://goo.gl/2UCZ43). A lab website is not acceptable.)
- Is the resource distributed in a way that maximally facilitates reuse?
- Will the resource as described have sufficient metadata available to promote reuse?
- For resources that must be maintained, is there a plan in place to maintain the resource?

## Data Sharing:

- Public, widely-used repositories should be used if possible (e.g., GEO or ArrayExpress for gene expression data, SRA for RNA-Seq data, etc.).
- If no public, widely-used repository is available for the data type in question, a general purpose archival repository (e.g., FigShare, Zenodo) should be used.
- For more detailed discussion, the guidelines provided by F1000 research for authors are an excellent resource: https://f1000research.com/for-authors/data-guidelines
- If authors or reviewers have questions, please feel free to contact the Childhood Cancer Data Lab (ccdl@alexslemonade.org) which is happy to seek available options.

## Materials and Reagents:

- Public, widely-used repositories (Addgene, cell banks, etc.) should be used if possible.
- On request should be avoided if possible. If distribution will occur upon request, specify the expected response time for the resource, how the response time will be measured, how it will be discussed in progress reports, and how sharing will happen after the grant.
- If the lab will be maintaining the material or reagent, methods that can be used to authenticate the reagent should be specified. When the authentication will occur and who is responsible (distributor, recipient) should also be specified.

## Protocols:

- How protocols will be distributed should be specified.
- How protocols will be maintained and clarified should be specified.
- If there exists an appropriate service (e.g., protocols.io) it should be used. Lab websites should generally not be used to distribute protocols.

## Source Code:

- Source code should be stored in a version control system and made available through a version control service (e.g., GitHub, Bitbucket, or similar).
- Source code should be archived to an archival service (e.g., Zenodo) at the time of submission and the conclusion of the grant.
- A license should be specified.

## American Brain Tumor Association

Primary Survey Contact: Heather Calderone, Director, Research and Grants
Email: hcalderone@abta.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. Requirements of Applicants and Grantees: a. All applicants to ABTA Research Collaboration Grants, Discovery Grants, and Basic Research Fellowships are required to submit a Data Management and Sharing Plan as part of the Full Application. b. Lead mentors on ABTA Jack & Fay Netchin Medical Student Summer Fellowships are required to include in their letter of support, submitted with the Full Application, a description of the data management and sharing practices within their laboratory and comment on how the data generated in the project will be shared. Grantees generating omics data (including, but not limited to, genomic, transcriptomic, proteomic, epigenetic, etc.), nanomaterials, or imaging (MR, PET, etc.) data are encouraged to deposit all data needed for independent verification of published results into an ABTA approved repository within six (6) months of publication. All data should be shared in accordance with the Grantee's approved Data Management and Sharing Plan. Grantees are encouraged to deposit all data, not just the data included in publications, into an approved data repository. The list of ABTA-approved repositories can be found on the ABTA website in the Resources for Applicants and Grantees https://www.abta.org/research/for-researchers/. Applicants and Grantees may recommend other repositories for approval by contacting grants@abta.org. Changes to the Data Management and Sharing Plan: Grantees may make changes to their Data Management and Sharing Plan at any time during the grant term by submitting a Change of Data Management and Sharing Plan Request Form in proposalCENTRAL. Changes to the Data Management and Sharing Plan must be approved by the ABTA. If the requested change to the plan will require a change in budget, grantees will also be required to submit a Budget Change Request Form in proposalCENTRAL. 4. Reporting: Grantees will be asked to report on the progress made on their Data Management and Sharing Plan in their annual and/or final progress reports. Reports should include digital object identifiers (DOIs) for all preprints, open access publications, and shared data. For reporting on outputs that fall after the end of the grant term, grantees are encouraged to send to the ABTA the digital object identifiers as soon as they are available. The ABTA may request updates on the required research outputs for up to five years after the end of the grant term.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. We hope to start requiring grantees to share data in the next couple of year, but have to develop procedures and determine how much staff time it will take to implement the requirement.
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?

    a. Specific grants
7. If just some, to which grants/programs do they apply and why?
    a. The policy applies to all of our grants except for our Medical Student Summer Fellowship. The reason being that they are small grants ($3,000) with a very short term (three months); however, we encourage the students to submit a data sharing plan and ask the mentors to describe, in their letter of support for the application, the data sharing practices in the lab.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. None so far, but our policy is new and does not (yet) require that data be shared.
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. Yes – ArrayExpress, BRAIN Commons, cBioPortal, ClinicalTrials.gov, COSMIC, dbGAP, Figshare, Flow Repository, GDC, GenBank, GENIE, GitHub, Metabolomics Workbench Nanomaterial Registry, NCBI BioProject, NCBI BioSample, NCBI GEO Datasets, NCBI Protein NCBI Reference Sequence, Open Science Framework, PRIDE, PubChem, Sequence Read Archive, The Cancer Imaging Archive, Zenodo, Approved Protocol Repositories: Bio-protocol Protocol Exchange, Protocols.io, Scientific Protocols
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. Grantees are encouraged to deposit all data needed for independent verification of published results into an ABTA approved repository within six (6) months of publication
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. Our reviewers are encouraged to look at and provide feedback to the applicants on the data sharing plan, but at this time, we ask that they not consider the merits of the data sharing plan when determining the score of the application. The main reasons for not scoring the data sharing plans are that we first want to assess how effective our form is for the data sharing plan and also because we are not yet prepared to train reviewers on reviewing data sharing plans.
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?

    a. No
22. What are the consequences for noncompliance with your policy?
    a. Not yet. Our policy is only two years old and so very few of our grantees have shared data. We intend to establish practices for following up with grantees so that we may be more effective in implementing policies once we require data sharing.

## American Heart Association

**Primary Survey Contact**: Rachel McEnany, Research Information Manager
**Email**: rachel.mcenany@heart.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. See AHA's Data-Sharing Policy and Open Science FAQs after question 22 below.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Specific grants
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Require
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. Yes, See list of AHA's Approved Repositories after the AHA Data-Sharing Policy Below
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. N/A
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?

    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes
22. What are the consequences for noncompliance with your policy?
    a. N/A

## Open Science Policy Statements for AHA Funded Research

**Public Access**

**"Outbound" Public Access**

**The American Heart Association (AHA) requires that all journal articles resulting from AHA funding be made freely available in PubMed Central (PMC) and linked to an AHA award within 12 months of publication.** It is the responsibility of the awardee to ensure journal articles are deposited into PMC.

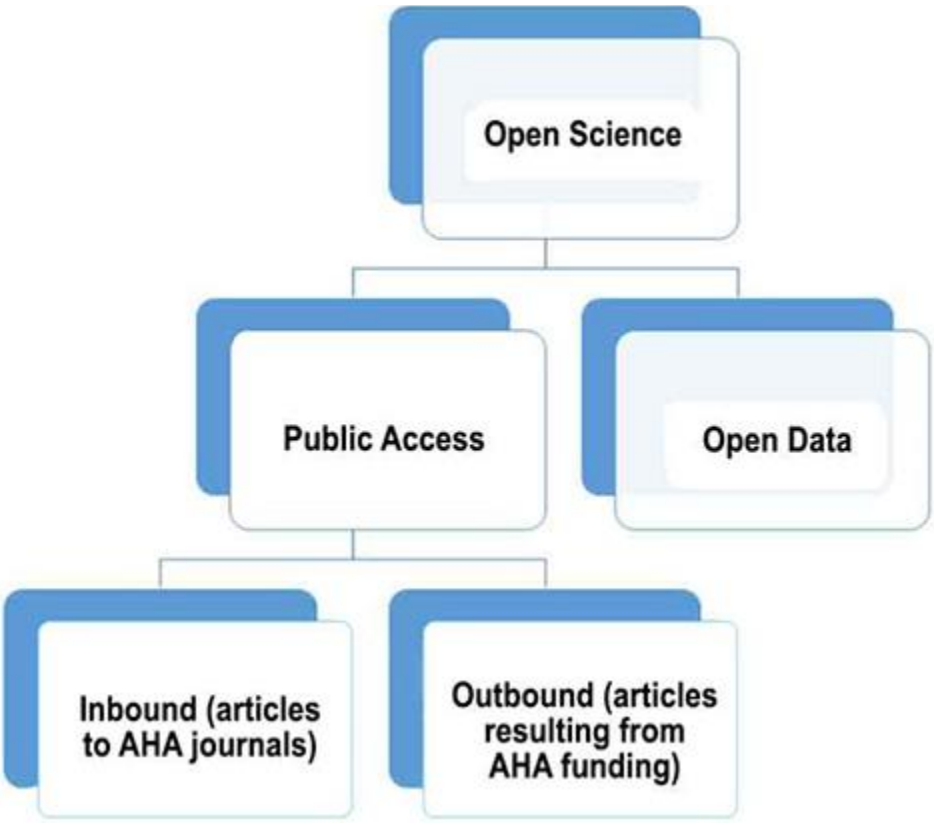
**Mechanism**

The AHA has adopted the procedures established by the Health Research Alliance (HRA) which enable awardees to deposit publications directly into PMC.

**"Inbound" Public Access**

All original research articles in the 11 subscription-model AHA journals ("inbound" research) are made freely available on each respective journal website 6 months after publication. All non-original research articles are made freely available on each respective journal web site 12 months after publication. Scientific statements and clinical practice guidelines are made freely available immediately on publication.

The Journal of the American Heart Association (JAHA) is the AHA open access journal. Because it utilizes an author pays model, the author pays for publication costs and retains copyright. The AHA is granted a nonexclusive license of all rights of copyright in and to the article. JAHA articles are deposited in PMC on publication.




**Open Data**

**The AHA requires certain applicants to include a data sharing plan with the application. Any factual data that is needed for independent verification of research results must be made freely and publicly available in an AHA-approved repository within 12 months of the end of the funding period (and any no-cost extension)**. Recipients of the following early-career awards are exempt from this policy: AHA Predoctoral Fellowships, AHA Postdoctoral Fellowships, and Institutional Undergraduate Student Fellowship Program.

**Compliance Threshold**

Applicants will be prompted to answer each of the following questions when completing a data plan in the application:

1. What data outputs will the research generate?
2. When will the data be shared?
3. Where will the data be made available?
4. Are any limits to data sharing required?

Any costs associated with preparing data for sharing may be covered by award funds and must be tracked in the project budget and financial reports. In most circumstances, additional grant funds will not be required to comply with the Open Data Policy.

**Opt-Out Conditions**

Certain applicants may seek exemption from the Open Data policy. These applicants must submit an opt-out request with the application to explain why the Open Data policy should be waived. Although the applicant may provide other rationale, most opt-out requests fall into one of the following four categories:

- Human Subject Grounds: As the National Science Foundation explains, "[H]uman subject's protection requires removing identifiers, which may be prohibitively expensive or render the data meaningless in research that relies heavily on extensive in-depth interviews." Data sharing may not violate privacy regulations stipulated by HIPAA or fail in any way to safeguard the rights of research participants. It is the responsibility of the applicant to make a case for why the use of the HIPAA Safe Harbor de-identification method would not be feasible for their data.
- Superseding Regulations Grounds: Governing laws or institutional policies may limit the release of certain data.
- Intellectual Property (IP) Grounds: Although data sharing may not protect IP, opt-out requests citing protection for potential or anticipated IP will not be approved until after IP rights are established.
- Financial Grounds: Data sharing should not cause an undue financial burden for the awardee.

**Data Types**

Because the AHA supports a wide range of research, the nature of the data collected varies greatly. Any factual data that is needed for independent verification of research results must be included in the data sharing plan.

All data must be properly documented. This documentation, sometimes referred to as metadata, is necessary for others to properly use and interpret the data. Consistent with NIH guidelines, the metadata must provide "information about the methodology and procedures used to collect the data, details about codes, definitions of variables, variable field locations, frequencies, and the like. The precise content of documentation will vary by scientific area, study design, the type of data collected, and characteristics of the dataset."

**Embargo Period**

Awardees should make their data publicly available as soon as possible. All pertinent data must be made freely and publicly available within 12 months of the end of funding period (and any no-cost extension).

**Acceptable Repositories**

Because of the wide range of projects funded by the AHA, no single repository is universally preferable. Therefore, the AHA grants awardees flexibility to select the repository most compatible with their data. Any repository approved by the AHA must meet the following criteria:

- *Re-Use:* The repository must guarantee to any interested party free access to the data without restriction on research reuse.

- *Security*: The repository must describe how datasets are stored and confidential information is protected.
- *Stability:* The repository must assure that the data will be available for the indefinite future, regardless of whether the repository is dismantled.
- *Subject Focus:* The repository should be compatible with the subject matter. If a repository emerges as the standard resource in a field (e.g., GenBank for DNA sequences), the awardee is encouraged to use that repository to better disseminate the research results to like-minded investigators interested in building upon the research.
- *Metadata:* The repository must require the awardee to provide sufficient metadata to explain the data to others. These metadata must be searchable so that repository visitors can easily locate desired datasets.
- *File Formats:* The repository should accommodate all file types generated by the awardee.
- *Machine Extraction*: Preferably, the repository will feature machine-readable and machine-interpretable functionality to enable third-party users to more easily locate the data.
- *Reception to AHA Data:* The repository must be willing to accept data submitted by AHA-funded researchers.

*If a desired repository is not currently approved by the AHA, the applicant may request in the application that the AHA consider the repository for approval. If the repository is approved, it will be added to the list of AHA Acceptable Data Repositories.*

Recipients of non-exempt awards beginning January 1, 2015, and after must comply with both the outbound Public Access policy and the Open Data policy.

## Principles and Background Information

**What is open science?** Open science is the idea that scientific knowledge of all kinds should be freely and openly shared as early as is practical in the discovery process. Open science is characterized by the notion that some or all aspects of the research and discovery lifecycle should be made available with little or no restrictions to the scientific community.

**What is AHA's approach to open science?** The AHA's policies focus on two critical aspects of open science – increasing the availability of both scholarly research outputs and the raw data from which these outputs are derived. The former, typically referred to as "open access" or "public access", focuses on the rapid online availability of peer-reviewed research results, permitting any users to read, download, copy, distribute, print, search or link to the full text of these articles, crawl them for indexing, pass them as data to software or use them for any other lawful purpose. The latter, typically referred to as "open data", promotes the accessibility and reuse of the raw data generated during the scientific discovery phase. Open data typically applies to a range of non-textual materials, including datasets, statistics, transcripts, survey results, and the metadata associated with these objects. The data is, in essence, the factual information that is necessary to replicate and verify research results. Open data policies also usually encompass the notion that machine extraction, manipulation, and meta-analysis of data should be permissible. This extends beyond the human abstraction of facts.

**What is the AHA's public access policy as it relates to the research it funds?** The AHA requires that all journal articles resulting from AHA funding should be made freely available in PubMed Central within 12 months of publication. It is the responsibility of the author to ensure this occurs.

To be clear, the AHA's policy advocates the rapid, but not the immediate, free dissemination of journal articles. The plan is therefore more accurately characterized as a public access policy rather than an open access policy. Public access, unlike open access, does not demand the immediate availability of peer-reviewed articles. Rather, it recognizes that the realities of the current publishing system allow for publishers to embargo access for some period of time.

**What is AHA's public access policy as it relates to the research it publishes?** All research articles in AHA journals are made freely available on that AHA journals website after six months. All other articles in AHA journals are freely available on the AHA journal web site one year after publication. Statements and clinical practice guidelines are always freely available upon publication.

Unique among AHA publications, the Journal of the American Heart Association (JAHA) is an open access journal. Because it is open access using an author pays model, the author's paper is accepted the author pays for publication costs. The author also keeps copyright and the AHA has a license to publish the article. Articles are deposited in PubMed Central upon publication.

**What is the AHA's open data policy specifically**? The AHA requires grant applicants to include a data sharing plan as part of the application process. Any research data that is needed for independent verification of research results must be made freely and publicly available within 12 months of the end of the funding period (and any no-cost extension). Specific early career awards are currently exempt from this requirement.

**Why is the AHA implementing these policies?** Quite simply, public access and open data align with the American Heart Association's strategic goals. The AHA invests heavily in research in order to accelerate the pace of scientific discovery, encourage innovation, enrich education, and to improve the public good. We recognize that scientific investigation advances only through sharing of methods and results, and the value of an investment in research is only maximized through wide use of this information. At present, the building blocks of science research outputs and the raw data supporting these outputs are not made available rapidly to the broadest community of potential users. Internet technologies provide an increasingly cost-effective opportunity to bring these components to a wider audience, and to use these materials in new, innovative ways. Open science policies can facilitate increased discoverability and reusability. This reduces the gaps in the research cycle and makes it easier for interested parties to pursue promising investigative directions. It lessens the likelihood that multiple laboratories will be pursuing duplicative research in siloed environments. It decreases the potential for data miscalculation, misinterpretation, manipulation, and fraud by opening raw results up to the broader community. It also encourages the broadest possible audience to access and build upon research results.

From a practical standpoint, open science policies demonstrate a tangible return on investment. The AHA relies on private contributions to support our activities. Disseminating research outputs and data in a highly visible manner that promotes sharing, discussion, and follow-up science is a clear way to demonstrate the effective use of donations. It tangibly exhibits the American Heart Association's commitment to the good stewardship of donors' investments.

**Open Data**

**Are there examples of acceptable data plans?** Yes. View the sample data plans.

**Does the AHA's open data policy apply to all grant recipients?** Specified early-career awards are exempt from this policy. The programs that are currently exempt are AHA Predoctoral Fellowships, AHA Postdoctoral Fellowships, and Institutional Undergraduate Student Fellowship Programs. For other programs, if a data sharing plan is not included as part of the application process, the applicant should provide a rationale for why it is unnecessary or inappropriate. "Opt-out" requests will be evaluated according to established guidelines outlined in the policy document.

**What types of research data must be shared in accordance with the AHA's policy?** The AHA supports a wide range of research. As such, the nature of the data collected in conjunction with these projects varies greatly. Consistent with NIH guidelines, the AHA defines research data as recorded factual material commonly accepted in the scientific community as necessary to validate research findings. This does not include laboratory notebooks, partial datasets, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as gels or laboratory specimens. Like NIH, we leave it to the individual's discretion to determine the factual information that is necessary to replicate and verify research results.

Regardless of the type of research data to be shared, it must be accompanied by proper documentation. This documentation, sometimes referred to as metadata, is necessary to allow others to use the data properly and without confusion. Again consistent with NIH guidelines, the metadata must provide "information about the methodology and procedures used to collect the data, details about codes, definitions of variables, variable field locations, frequencies, and the like. The precise content of documentation will vary by scientific area, study design, the type of data collected, and characteristics of the dataset."

**My data will have limited or no value to others. Why must I share it?** The entire concept of open data is grounded in the notion that the market for the building blocks of science research outputs should not be artificially restricted. Who knows where the next innovation will come from, or what combination of datasets will produce a scientific breakthrough/ open data policies maximize the information the research community has at its disposal to pursue new leads, build upon the scientific record, and accelerate discovery.

**Does the AHA policy apply only to data associated with published articles?** No. Your data sharing plan must encompass all research data, consistent with the NIH definition offered immediately above, from funded research that can be shared without compromising individual subjects' rights, regardless of whether the data have been used in a publication. From a practical standpoint, if the AHA open data policy covered only data associated with published articles; this would slow down the availability of AHA data dramatically. The peer-review process, which often involves submission and rejection to multiple journals and long periods of review and revision, would unnecessarily and artificially delay the sharing of the underlying data. More fundamentally, the failure of a researcher to translate research into a published paper may not be due to the data's inaccuracy or insignificance. The AHA policy aims to surface all the research data it funds so that the broadest possible scientific community can access, interpret, and build off of it.

**My research involves human subjects. What steps must I take to safeguard their privacy?** Your data sharing plan must safeguard identifying information related to research subjects in order to comply with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy Rule. The HIPAA Privacy Rule establishes the conditions under which "protected health information" may be used or disclosed. In

order to ensure demonstrable compliance, please follow the guidelines set forth in the AHA's HIPAA Compliance Guidelines (PDF).

**What is the process for submitting my data sharing plan?** Grant applicants should complete the Data Sharing Plan section of the online application.

**Where should I deposit my research data?** Given the wide range of projects funded by the AHA, no single deposit location is universally applicable. Instead, the AHA provides its grant recipients with a degree of latitude in selecting the most appropriate repository to house their research data. In order for a repository to be approved by the AHA, it must be deemed appropriate across each of the following dimensions:

- Re-Use. The repository must allow any interested party to freely access the data without restriction on research reuse. This should be codified in the repository's terms of use.
- Security. The repository must describe how datasets are stored, as well as how any confidential information is protected.
- Stability. The repository must have a clearly articulated funding mechanism or business plan to provide reasonable assurances that the data will be available for the indefinite future. It should also have a continuity plan addressing what will happen to the data in the event the repository is discontinued.
- Fee Structure. What is the cost, if any, to deposit data in the repository? Is the payment one-time or recurring? Does the size of the dataset impact the cost? The repository must define its rates and explain how these fees ensure financial stability.
- Subject Focus. There are hundreds of topic-specific repositories in operation at this writing. The grant recipient should endeavor to deposit his/her data in a repository that is appropriate for the subject matter in question. Further, if a repository has emerged within a specific research community as the default resource in that field (e.g., GenBank for DNA sequences), grant recipients should, as a general rule, utilize that repository. This optimizes the ability of others to discover and build upon the data.
- Metadata. The repository must require a depositor to provide sufficient metadata provided to enable the dataset to be used by others. These metadata should be searchable so that repository visitors can easily discovery appropriate datasets.
- File Formats. The repository should be able to accommodate all aspects of the grant recipients' dataset, regardless of file type.
- Machine Extraction. The data stored in the repository should ideally be available in a machine-readable and machine-interpretable format.
- Willingness to Accept AHA Data. Finally, the repository must be willing to accept data submitted by AHA-funded researchers.

**What are some examples of acceptable data repositories?** As previously mentioned, the AHA funds a wide range of projects with a broad array of data outputs. Therefore, it is difficult to identify a comprehensive list of "pre-approved" repositories that can accommodate all the possible datasets covered by this policy. There are, however, a number of examples that can be cited as meeting the criteria outlined in the section immediately prior.

We currently have a list of vetted, free repositories that are recommended by NIH and the Wellcome Trust. If you would like to recommend a specific repository for AHA evaluation, please contact us here.

View the list of *AHA Acceptable Data Repositories*.

**Can I opt out of data sharing?** There may be certain instances in which grant applicants seek to be exempted from the data sharing policy. Grant applicants seeking waivers should complete the out-out-request portion within their application to indicate the grounds on which they are lodging these requests. The opt-out request information will be provided with the application in lieu of the data plan fields. Broadly speaking, waiver requests will be expected to fall into one of four predetermined categories, although the applicant can provide additional rationales as appropriate:

- Human Subject Grounds. As the NSF spells out in its exemption guidelines, "[H]uman subject's protection requires removing identifiers, which may be prohibitively expensive or render the data meaningless in research that relies heavily on extensive in-depth interviews." Data sharing cannot violate privacy regulations (e.g., HIPAA) or in any way fail to safeguard the rights of research participants.
- Superseding Regulations Grounds. Governing laws or institutional policies may limit the release of certain data elements.
- Intellectual Property Grounds. Under certain circumstances, data sharing may violate existing IP rights. Opt-out requests would not be approved for potential or anticipated IP rights. Once IP rights are established, awardees can still seek opt-out approval up until the deadline for depositing their data.
- Financial Grounds. Data sharing should not cause an undue financial burden for the grant recipient.

If the opt-out request is not approved and the application receives funding, the awardee will be required to submit a data plan.

**When must I make my research data available?** Any data that is needed for independent verification of research results must be made freely and publicly available within 12 months of the end of the funding period (and any no-cost extension). The AHA's open data policy is designed to provide you with prolonged – but not indefinite - first use of the data.

**What if there is an issue with making all the relevant research data available within 12 months of the end of funding period?** The AHA open data policy requires any data needed for independent verification of research results to be made freely and publicly available within 12 months of the end of the funding period (and any no-cost extension). It is the AHA's intention to require compliance in all but the most exceptional circumstances. If you believe you have a justifiable reason for delaying the release of your data, please contact us here.

**What if I am not able to publish a paper using my research data within 12 months of my grant's expiration. Why can't I hold my data privately until I get my article published?** The AHA recognizes the unique interest grant recipients have in extracting value from their original data. However, this must be balanced against the interest that the AHA, the research community, and the general public share in accelerating the pace of scientific discovery. Some research projects never generate publishable articles. Others take many years to do so. A policy that gave grant recipients an unlimited exclusive window to produce published articles off of their data would, on the whole, create unnecessary delays in the sharing of data. Grant recipients who feel they have a compelling reason to delay their data's release should contact us here.

**Why can't I make the research data available as a supplementary file on a publisher's website?** Publishers, be they not-for-profit or commercial, are unlikely to satisfy all of the conditions enumerated in the AHA's acceptable repository description. Of particular concern are issues of re-use restrictions, long-term preservation, and machine extraction. It is critical for your data to be made widely available, in the near term and the long term, to any interested party. Publishers' websites are unlikely to meet these criteria.

**Why doesn't making my paper freely available via PubMed Central satisfy AHA's open data requirement?** PubMed Central is an archive of publicly-accessible journal articles. While it is an extremely valuable resource, it is not a data repository. In addition, the AHA's policy extends beyond data associated with published articles.

**Why can't interested parties just contact me directly to get a copy of the research data?** As above, it is essential for your data to be made widely available, in the near term and the long term, to any interested party. Third parties repositories that have an established track record of fulfilling this role are a more suitable vehicle for making your data available.

**Data derived from this research may be eligible for intellectual property protection. Should that be indicated in my data plan as a limit to data sharing? Is this justification for seeking opt-out approval?** Potential or anticipated IP protection is not justification for opting out of our open data policy requirements or even limiting your data sharing. If IP rights are obtained for some or all of the data before the deadline for depositing it in the approved repository, the data plan can then be modified or an opt-out request submitted.

**What if no appropriate repository exists to house my research data?** No one will be expected to develop a de novo database. The list of AHA-approved data repositories is growing and we anticipate that it will grow to cover a wide spectrum of research areas. If you do not see a repository on this list that is an appropriate home for your data, we recommend the following next steps:

- Go to the Registry of Research Data Repositories and look for a subject-appropriate match. NOTE: Not all of the repositories listed in this database will meet AHA standards. If you identify a subject-appropriate repository, please check its policies against the AHA's criteria for acceptable data repositories. If you feel that the identified repository fits the AHA's guidelines, you can submit the repository as "other" in the data plan portion of your application and AHA will review it.
- If you cannot identify an appropriate subject-focused repository, consider depositing your data in your institutional repository (IR). Many universities have developed IRs to house works created by their affiliated researchers. The OpenDOAR Directory provides a searchable database of institutional repositories – you can query it to see if your university has an IR. As with the Registry of Research Data Repositories, not all of the repositories listed in this database will meet AHA standards. Please check your IR's policies against the AHA's criteria for acceptable data repositories. If you feel that the IR fits the AHA's guidelines, you can submit the repository as "other" in the data plan portion of your application and AHA will review it.
- Currently there are a few AHA-approved data repositories that are considered "general" vs. "subject-focused". These include Zenodo, figshare, and Dataverse. All of these will accept image sets. "Subject-focused repositories, when available, are preferred over general repositories."

Our open data policy is new and we are not able to make recommendations to applicants for repositories other than the AHA-approved list above. We will vet additional repositories that are recommended by our applicants and add new ones to the list as they are approved.

**How much will it cost to deposit my research data? Who pays for it?** All of the data repositories in the vetted list above are free of charge to both deposit and access data. Although we anticipate that most deposits will occur in free repositories, the AHA recognizes that there is an effort associated with data preparation. The Open Science Committee would need to approve any request to deposit data in a repository that is not on our approved list (including any that charge for deposit). Any costs associated with preparing the data for sharing should be covered in your award and tracked for reporting. Given the strong likelihood that researchers are already performing fundamental tasks to organize their data for internal consumption and the plethora of free data archiving repositories, it is expected that, in most circumstances, no additional grant funds shall be required to fulfill the data sharing plan.

**Once I share my research data, who owns it?** It is generally held that factual data cannot be copyrighted. Anyone can access the data from the repository in which you deposit it, use the data, and build upon it, provided they provide appropriate attribution.

**What about the possibility of scientific misuse or unsophisticated analysis of the research data?** Data sharing can actually decrease the potential for data miscalculation, misinterpretation, manipulation, and fraud by opening raw results up to the broader community. Any conclusions that a scientist draws from the data can be checked, questioned, and countered by the broader community.

**Will source code be required to be included in the data sharing plan?** All unprocessed data that is needed for independent verification of research results must be included in the data sharing plan. This may in some instances include source code if special programs have been developed to analyze or manipulate data associated with the research. GitHub is currently the default location for open source code and is on our list of pre-approved repositories.

**Some of the AHA pre-approved data repositories restrict the use of the data and are not always "freely and publicly available" as required by the AHA open data policy. Will this result in my data plan being rejected?** Some of the repositories such as dbGaP provide two levels of access - open and controlled - in order to allow broad release of non-sensitive data, while providing oversight and investigator accountability for sensitive data sets involving personal health information. The assumption is that AHA funded data would fall under the open category unless there are exceptional circumstances.

**What if my research builds on restricted data and/or leverages cohorts that include consented human data?** It is sometimes required that the research data generated be redeposited in a repository that is not publicly available, such as BioLINCC or the controlled section of dbGaP. These data would still be available for use in other approved studies, but would not be publicly available as required by our open data policy. In these situations we expect the PI to address in their data plan whether ANY data COULD BE deposited in one of our approved open repositories. If ALL the resulting data is restricted, then an opt-out request should be submitted.

**I have selected an approved repository for my data, but what if over the course of the award I want to modify the data plan. Is it possible to change repositories or add repositories to my data plan?** It is possible to make changes to your data plan through Grants@Heart as needed going forward, so it would

not be an issue to change the selected open repository if you find a better fit or want to add additional repository selections sometime before the data needs to be deposited. New or modified data plans or even opt-out requests can be submitted prior to the deadline for depositing the data.

**What happens if I fail to comply with the AHA's open data policy?** In situations where the AHA or members of the research community feel that AHA-funded researchers are not sharing data in a manner consistent with our policy, you may be asked by the AHA to demonstrate your compliance. If you cannot do so, this may affect future funding.

**Public Access**

**To whom does the AHA's public access policy apply?** The AHA requires that all journal articles resulting from AHA funding must adhere to the public access policy. This means that any researcher receiving AHA funds must comply with the policy.

**My paper is based on research only partially funded by AHA. Is the paper still bound by the AHA's public access policy?** Yes, the public access policy applies to any journal article manuscript that arises from any amount of direct funding from the AHA.

**What is PubMed Central?** PubMed Central is a free digital archive of full-text biomedical and life science literature in biomedical and life sciences. PubMed Central was developed and is maintained by the U.S. National Library of Medicine (NLM), a division of the National Institutes of Health.

**What are the benefits of posting peer-reviewed papers to PubMed Central?** PubMed Central is a highly visible database that has emerged as a research destination for scientists, clinicians, patients, educators, and students interested in accessing full-text articles on a wide range of biomedical subjects. The inclusion of AHA papers in PubMed Central provides a prominent discovery path for the research we support.

**How do I deposit my papers in PubMed Central?** The AHA will be adopting the deposit procedures put in place by the Health Research Alliance (HRA). We anticipate that this will be available for AHA researchers beginning in 2015. The HRA has arranged for its member organizations (including the AHA) to establish a mechanism by which grant recipients can deposit papers directly into PubMed Central. View more information about depositing papers in PubMed Central.

**Do I need to do anything to retain rights with the publisher so that I can make the article available in Pub Med Central within 12 months of publication?** Grant recipients should endeavor to make publishers aware of the access conditions associated with their funded research at the earliest possible juncture. A template letter may be adapted and included during the manuscript submission process for these purposes. This notification should not influence acceptance of your article by any reputable journal. After your manuscript has been accepted by a journal, you should ensure that the publication agreement you sign is consistent with the AHA's public access policy. The terms of such agreements vary widely from publisher to publisher. If the agreement does not provide for your right to make the article publicly available in PubMed Central no later than 12 months after the official date of publication, you will need to adapt it. Simply add the following language to the agreement before returning it to the publisher:

*As an author whose research is funded in part or in whole by the American Heart Association, I am obligated as a condition of my grant to reserve certain rights. The Journal therefore acknowledges that Author retains the right to provide a copy of the authors' final manuscript, including all modifications from the publishing and peer review process, to PubMed Central at the NIH upon acceptance for Journal publication for public archiving as soon as possible but no later than 12 months after publication by Journal. _____ [Publisher Name] accepts these terms and agrees that the terms of this agreement are paramount and supersede any provisions to the contrary any publication agreement for this article, already signed or to be signed at a later date that may conflict.*

In the unlikely event the publisher balks at accepting a modification of their standard agreement, you should reiterate the conditions of your funding and remind them that the requirement is similar to that of the National Institutes of Health, Howard Hughes Medical Institute, and other funders. There is no evidence of financial or other material harm to publishers as a result of embargoed free public access to research articles. If the publisher still balks, please contact your funder for assistance.

**Does it matter what version of my article I make available in PubMed Central?** The AHA encourages authors to deposit the final, published version of an article as it appears on the journal's web site into PubMed Central. However, some publishers require deposits to PubMed Central to be made in the form of an author's final, peer reviewed manuscript of an article. Check your publishing agreement with the publishers for details. We prefer the version of record to be deposited, but we will accept the penultimate version as a Plan B if that is all a publisher allows.

**Does the AHA public access policy apply to new grants only or to existing grants as well?** All new grants awarded after January 1, 2015, must comply with the outbound public access policy. However, active grant holders are strongly encouraged (though not required) to adhere to the policy. Doing so will demonstrate a commitment to the acceleration of the pace of scientific discovery

## AHA Approved Data Repositories

ArrayExpress - The ArrayExpress Archive is a database of functional genomics experiments including gene expression where you can query and download data collected to MIAME and MINSEQE standards. Gene Expression Atlas contains a subset of curated and re-annotated Archive data which can be queried for individual gene expression under different biological conditions across experiments.

BioModels - BioModels Database is a repository of computational models of biological processes. Models described from literature are manually curated and enriched with cross-references. All models are provided in the Public Domain.

CellML - The purpose of CellML is to store and exchange computer-based mathematical models. CellML allows scientists to share models even if they are using different modeling tools. It also enables them to reuse components from one model in another, thus accelerating model development.

ClinicalTrials.gov - ClinicalTrials.gov is a Web-based resource that provides patients, their family members, health care professionals, researchers, and the public with easy access to information on publicly and privately supported clinical studies on a wide range of diseases and conditions.

ClinVar - ClinVar is a freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting evidence. ClinVar facilitates access to and communication about the relationships asserted between human variation and observed health status, and the history of that interpretation.

COSMIC - COSMIC is designed to store and display somatic mutation information and related details and contains information relating to human cancers.

Dataverse (general) - The Dataverse Network is an open source application to publish, share, reference, extract and analyze research data.

database of Genotypes and Phenotypes (dbGaP) (open section) - The database of Genotypes and Phenotypes (dbGaP) was developed to archive and distribute the results of studies that investigate the interaction of genotype and phenotype. Such studies include genome-wide association studies, medical sequencing, molecular diagnostic assays, as well as association between genotype and non-clinical traits. dbGaP provides two levels of access - open and controlled - in order to allow broad release of non-sensitive data, while providing oversight and investigator accountability for sensitive data sets involving personal health information. The assumption is that AHA-funded data would fall under the open category unless there are exceptional circumstances.

database of Single Nucleotide Polymorphisms (dpSNP) - In collaboration with the National Human Genome Research Institute, the National Center for Biotechnology Information has established the dbSNP database to serve as a central repository for both single base nucleotide substitutions and short deletion and insertion polymorphisms.

European Genome-phenome Archive (EGA)- The European Genome-phenome Archive (EGA) is designed to be a repository for all types of genotype experiments, including case control, population, and family studies. It includes SNP and CNV genotypes from array based methods and genotyping done with re-sequencing methods. This data may be either publicly available or limited access, depending on the design of the study.

EMDataBank - Electron Microscopy DataBank is a unified global portal for deposition and retrieval of 3DEM density maps, atomic models, and associated metadata, as well as a resource for news, events, software tools, data standards, validation methods for the 3DEM community.

European Nucleotide Archive (ENA) - Europe's primary nucleotide sequence resource. The main sources for DNA and RNA sequences are direct submissions from individual researchers, genome sequencing projects and patent applications. It is one part of the European Bioinformatics Institute (EMBL-EBI), which maintains the world's most comprehensive range of freely available and up-to-date molecular data resources.

exRNA Atlas - The exRNA Atlas is the data repository of the Extracellular RNA Communication Consortium (ERCC). It is developed and maintained by the Data Management and Resource Repository (DMRR). It includes qPCR-derived exRNA profiles from human and mouse biofluids and conditions and currently stores data profiled from small RNA sequencing assays.

figshare (general) - figshare allows users to upload any file format to be made visualisable in the browser so that figures, datasets, media, papers, posters, presentations and file sets can be disseminated in a way that the current scholarly publishing model does not allow.

FlowRepository - FlowRepository is a web-based application accessible from a web browser that serves as an online database of flow cytometry experiments where users can query and download data collected and annotated according to the MIFlowCyt standard.

FlyBase - FlyBase is a database of genetic and molecular data for D. melanogaster and other Drosophila species, targeted to an audience of research professionals.

GenBank - GenBank is an annotated collection of publicly available DNA sequences through the National Center for Biotechnology Information databases. GenBank contains over 135,000,000 sequence records and is updated every two months. GenBank is part of the International Nucleotide Sequence Database Collaboration along with the DNA DataBank of Japan and the European Molecular Biology Laboratory.

GitHub (source code) - Repository for open source code.

ImmPort - ImmPort Shared Data enables searching and downloading of shared biomedical research data funded from NIAID, DAIT, DMID, other NIH agencies, and non-government sources. Additional resources include step-by-step data reuse tutorials with example R and Python analysis code, the Cell Ontology Visualizer, the Cytokine Registry, 10,000 Immunomes - a reference dataset for human immunology, and immuneXpresso - the cytokine and cell interaction literature mining tool.

IntACT - IntAct provides a freely available, open source database system and analysis tools for protein interaction data. All interactions are derived from literature curation or direct user submissions and are freely available.

International Mouse Phenotyping Consortium (IMPC) - The International Mouse Phenotyping Consortium is an international scientific endeavor to create and characterize the phenotype of 20,000 knockout mouse strains. Using a standardized phenotyping protocol, the IMPC integrates data to existing mouse and human disease resources, and provides strains and phenotype data for use by the research community.

MetaboLights - MetaboLights is a database for Metabolomics experiments and derived information. The database is cross-species, cross-technique and covers metabolite structures and their reference spectra as well as their biological roles, locations and concentrations, and experimental data from metabolic experiments.

Metabolomics Workbench - The Metabolomics Workbench serves as a national and international repository for metabolomics data and metadata and provides analysis tools and access to metabolite standards, protocols, tutorials, training, and more. The Workbench is a companion to RCMRCs and is a part of the Common Fund Initiative in metabolomics.

Nanomaterial - The Nanomaterial Registry is an authoritative, fully curated resource that archives research data on nanomaterials and their biological and environmental implications.

National Collection of Pathogenic Viruses - A wide-ranging archive of well-characterised, authenticated human pathogens which will resource the supply of viruses, and materials derived from them, to the scientific community. [The National Collection of Pathogenic Viruses comprises over 300 human pathogenic viruses available for supply which require handling at UK biosafety containment levels 2, 3 and 4.]

National Collection of Type Cultures - The National collection of Type Cultures (NCTC) is a specialized laboratory located in the Central Public Health Laboratory, Colindale. It accesses, preserves and supplies authentic cultures of bacteria and mycoplasmas that are pathogenic to man or other animals that may occur in food or water and in hospital or health related environments and which can be preserved by freeze-drying.

NCBI BioProject - The BioProject repository collects projects with biological data that relates to a single initiative that originates from a single entity or consortium. Records provide users with a single location for the links to diverse data types generated for those projects.

NCBI BioSample - The BioSample database contains descriptions of biological source materials used in experimental assays.

NCBI Conserved Domains Database - The Conserved Domains Database (CDD) contains annotations of functional units in proteins; including multiple sequence alignment models for ancient domains and full-length proteins. This collection of models includes 3D structures that display the sequence/structure/function relationships in proteins. Users can identify amino acids in protein sequences with the resources available through CDD as well as view single sequences embedded within multiple sequence alignments.

NCBI dbVar - The dbVar is a database of genomic structural variation containing data from multiple gene studies. dbVar is a structural variation database designed to store data on variant DNA ≥ 1 bp in size. It is recommended that variation data that is > 50bp be submitted to dbVar and variation data that is ≤ 50bp to dbSNP. All clinically relevant structural variation should be submitted to ClinVar or dbGaP. Users can browse data containing the number of variant cells from each study, and filter studies by organism, study type, method and genomic variant. Organisms include human, mouse, cattle, and additional animals.

NCBI Gene - Integrating information from a variety of species, records in Gene include nomenclature, Reference Sequences, maps, pathways, variations, phenotypes, and links to genome-specific, phenotype-specific, and locus-specific resources.

NCBI Genome - The Genome database contains annotations and analysis of eukaryotic and prokaryotic genomes, as well as tools that allow users to compare genomes and gene sequences from humans, microbes, plants, viruses and organelles. Users can browse by organism, and view genome maps and protein clusters.

NCBI GEO Datasets - An international public repository, GEO (Gene Expression Omnibus) DataSets archives and distributes microarray, next-generation sequencing, and other forms of high-throughput functional genomics data. The records include original submitter-supplied records (Series, Samples and Platforms) and curated DataSets. GEO aims to provide a database that efficiently store this data; offer

simple submission procedures and formats that support complete and well-annotated data deposits from the research community; and provide user-friendly mechanisms for users to find and use studies and gene expression profiles of interest. GEO DataSets provides tools to identify differences in gene expression levels and cluster heatmaps.

NCBI GEO Profiles - The Gene Expression Omnibus (GEO) database stores individual gene expression profiles from NCBI databases and is searchable by gene annotation as well as gene profile characteristics. GEO archives microarray and next-generation sequencing as well as other forms of genomic data submitted by researchers within the scientific community.

NCBI HomoloGene - The HomoloGene database provides a system for the automated detection of homologs among annotated genes of genomes across multiple species. These homologs are fully documented and organized by homology group. HomoloGene processing uses proteins from input organisms to compare and sequence homologs, mapping back to corresponding DNA sequences.

NCBI Nucleotide - The NCBI Nucleotide database collects sequences from such sources as GenBank, RefSeq, TPA, and PDB. Sequences collected relate to genome, gene, and transcript sequence data, and provide a foundation for research related to the biomedical field.

NCBI PopSet - NCBI PopSet collects DNA sequences to analyze the ways that populations are related by evolution. Such sequences indicate if populations originate from different members of the same species or from organisms of different species entirely.

NCBI Probe - Probe database provides a public registry of nucleic acid reagents as well as information on reagent distributors, sequence similarities and probe effectiveness. Database users have access to applications of gene expression, gene silencing and mapping, as well as reagent variation analysis and projects based on probe-generated data. The Probe database is constantly updated, with over 11,000,000 probes available.

NCBI Protein - The Protein database collects protein sequences related to biological structure and function. The sequences in NCBI Protein come from the translations from annotated coding regions in GenBank, RefSeq, and TPA, and records from SwissProt, PIR, PRF, and PDB.

NCBI Protein Clusters - The Entrez Protein Clusters database contains annotation information, publications, structures and analysis tools for related protein sequences encoded by complete genomes. The data available in the Protein Clusters Database is generated from prokaryotic genomic studies and is intended to assist researchers studying micro-organism evolution as well as other biological sciences. Available genomes include plants and viruses as well as organelles and microbial genomes.

NCBI Reference Sequence - The Reference Sequence database provides explicitly linked nucleotide and protein sequences, as well as comprehensive and annotated sequence sets with genomic DNA, proteins and transcripts. Users have access to a wealth of resources for gene identification, comparative analysis and genome research. Reference Sequences are available for naturally occurring DNA, RNA and protein sequences in organic species worldwide.

NCBI Structure - The Structure database provides three-dimensional structures of macromolecules for a variety of research purposes and allows the user to retrieve structures for specific molecule types as

well as structures for genes and proteins of interest. Three main databases comprise Structure-The Molecular Modeling Database; Conserved Domains and Protein Classification; and the BioSystems Database. Structure also links to the PubChem databases to connect biological activity data to the macromolecular structures. Users can locate structural templates for proteins and interactively view structures and sequence data to closely examine sequence-structure relationships.

NCBI Taxonomy - Currently covering about 10 percent of the described species on the planet and more than 175,000 taxa, Taxonomy is a curated classification and nomenclature for all organisms in the public sequence databases. Taxonomy gives species names and higher-level classifications of the organisms represented in the Entrez sequence databases. It maintains a phylogenetic classification (containing only monophyletic groups if possible). Most species are represented only by a small piece of sequence data that's insufficient to construct a full phylogeny, but some species contain complete genomes.

NCBI Trace Archive - The Trace Assembly Archive stores pairwise alignment and multiple alignment of sequencing reads, linking basic trace data with finished genomic sequence.

NITRC - NITRC facilitates finding and comparing neuroimaging resources for functional and structural neuroimaging analyses. NITRC and its components—the Resources Registry (NITRC-R), Image Repository (NITRC-IR), and Computational Environment (NITRC-CE) offer researchers PET/SPECT, CT, EEG/MEG, optical imaging, clinical neuroinformatics, computational neuroscience, and imaging genomics software tools, data, and computational resources.

Online Mendelian Inheritance in Animals (OMIA) - Online Mendelian Inheritance in Animals contains textual information, references, links, and relevant records related to genes, traits, and inherited disorders in animals.

Online Mendelian Inheritance in Man (OMIM) - OMIM contains authoritative medical data on all known mendelian disorders as well as full-text and referenced overviews on the relationship between phenotype and genotype. Users can search the OMIM database by chromosome as well as narrow their search results by known gene sequences, phenotypes and gene map locus; as well as searching using only clinical synopses containing any combination of 22 specified criteria. The information contained in OMIM is available to download for personal, educational and research uses.

Open Science Framework (general) - Open Science Framework is a repository hosted by Center for Open Science, which is a non-profit technology company providing free and open services to increase inclusivity and transparency of research.

Protein Data Bank in Europe - The EBI Protein Structure Database in Europe is a project for the collection, management and distribution of data about macromolecular structures, derived from the Protein Data Bank (PDB). It is one of the founding members of Worldwide Protein Data Bank (wwPDB).

ProteomeXchange - The ProteomeXchange Consortium was established to provide globally coordinated standard data submission and dissemination pipelines involving the main proteomics repositories, and to encourage open data policies in the field. ProteomeXchange fully supports both MS/MS proteomics and SRM data submission. Submissions of other types of proteomics data is also possible using the Partial Submission mechanism.

PRoteomics IDEntifications database (PRIDE) - The PRoteomics IDEntifications (PRIDE) database at EMBL-EBI is a centralised, standards compliant, public data repository for proteomics data. It has been developed to provide the proteomics community with a public repository for protein and peptide identifications together with the evidence supporting these identifications. PRIDE is also able to capture details of post-translational modifications. It is a core member in the ProteomeXchange (PX) consortium.

PubChem - PubChem is an open chemistry database at the NIH. It accepts and stores information on chemical structures, identifiers, chemical and physical properties, biological activities, patents, health, safety, toxicity data, and many others.

Rat Genome Database - The Rat Genome Database houses genomic, genetic, functional, physiological, pathway and disease data for the laboratory rat as well as comparative genomics between rat, human and mouse.

RCSB Protein Data Bank (PDB) - Protein Data Bank (PDB) archive is the single worldwide repository of information about the 3D structures of large biological molecules, including proteins and nucleic acids.

Sequence Read Archive (SRA) - The Sequence Read Archive stores the raw sequencing data from such sequencing platforms as the Roche 454 GS System, the Illumina Genome Analyzer, the Applied Biosystems SOLiD System, the Helicos Heliscope, and the Complete Genomics. It archives the sequencing data associated with RNA-Seq, ChIP-Seq, Genomic and Transcriptomic assemblies, and 16S ribosomal RNA data.

The Cancer Imaging Archive (TCIA) - TCIA is a service which de-identifies and hosts a large archive of medical images of cancer accessible for public download. The data are organized as “Collections”, typically patients related by a common disease (e.g. lung cancer), image modality (MRI, CT, etc) or research focus. DICOM is the primary file format used by TCIA for image storage. Supporting data related to the images such as patient outcomes, treatment details, genomics, pathology, and expert analyses are also provided when available.

TriTrypDB - TriTrypDB is an integrated genomic and functional genomic database for pathogens of the family Trypanosomatidae, including organisms in both Leishmania and Trypanosoma genera.

UK Data Archive - The UK Data Archive (UKDA) is a centre of expertise in data acquisition, preservation, dissemination and promotion and is curator of the largest collection of digital data in the social sciences and humanities in the UK.

WikiPathways - WikiPathways is an open, collaborative platform dedicated to the curation of biological pathways. Building on the same MediaWiki software that powers Wikipedia, the platform has a custom graphical pathway editing tool and integrated databases covering major gene, protein, and small-molecule systems.

Wormbase - WormBase is an online biological database about the biology and genome of the nematode model organism Caenorhabditis elegans and contains information about other related nematodes.

Zenodo (general) - ZENODO builds and operate a simple and innovative service that enables researchers, scientists, EU projects and institutions to share and showcase multidisciplinary research results (data

and publications) that are not part of the existing institutional or subject-based repositories of the research communities.

ZFIN - ZFIN serves as the zebrafish model organism database, on-line database of information for zebrafish researchers.

Subject-focused repositories, when available, are preferred over general repositories.

SOURCES:

- https://wellcome.ac.uk/funding/guidance/developing-outputs-management-plan
- http://www.ncbi.nlm.nih.gov

## American Society of Hematology

**Primary Survey Contact**: Cole Thompson, Manager Program Evaluation and Data Analysis
**Email**: jthompson@hematology.org

*Notes: Questions 5-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. Not Planning.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments
11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material

18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## Barth Syndrome Foundation

**Primary Survey Contact**: Erik Lontok, Director of Research
**Email**: erik.lontok@barthsyndrome.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Thinking about it
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. Nothing formal, however, we internally connect researchers with potentially complementary datasets together to foster discussion.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. None at the moment. We have taken a wait and see approach with specific grantees/researchers developing their own protocol language, and seeing if their IRB accept them.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. On paper and in concept, it is a noble and ideal approach, however, our attempts to move forward have encountered considerable challenges in the form of all the different inertias. 1) Institutional - the concern that any data sharing, research-generated samples will be slowed and hampered by institutional roadblocks and red tape, and often manifests as, "I don't know if my IRB would approve that." 2) Publish or perish - the concern that providing access to one's existing dataset would lead to getting scooped. 3) Time - the concern that such a policy adds research burdens that are not supported by additional funding to make it happen, basically an unfunded mandate. 4) Language - the concern that developing Foundation language will result in extended negotiations with funded grantees/individual institutions, resulting in additional cost to implement. 5) Pandora's box - initial attempts at developing a policy resulted in affected individual/community concerns regarding its extension into a Return of Research Results policy.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. Yesterday. Don't mean to be sassy, but I consider each day without one a missed opportunity and just backloaded work for the future when it (will) thornily comes up.
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Specific grants.
7. If just some, to which grants/programs do they apply and why?
    a. We are primarily concerned with research studies that involve the participation of affected individuals. As a foundation that serves a rare disease community, and one that advocates for participation in research, we feel that it is equally our responsibility to ensure that use of

said data be equally maximized. It is only fair that both researchers and participants make the best use of such critical data.

8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. #5 is the closest answer I have to this.
9. Do you specifically require or encourage your awardees to share data?
    a. Neither
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application
13. Comments
    a. This is the first step we are considering, so as to soften the ground amongst our research community.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. N/A
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. No
19. Comments
    a. But we are open to this discussion.
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No
22. What are the consequences for noncompliance with your policy?
    a. N/A

## Cancer Research Institute

**Primary Survey Contact**: Ryan Godfrey, Grants Administrator
**Email**: rgodfrey@cancerresearch.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Thinking about it.
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A

3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. Within the year.
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. N/A
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. Outside of data submitted with their publications, should we be identifying and requiring specific repositories for our grantees to deposit their data? What additional costs will this impose on the grantee?
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. N/A
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. N/A
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. N/A
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. No
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No
22. What are the consequences for noncompliance with your policy?
    a. N/A

## Children's Tumor Foundation

**Primary Survey Contact**: Salvatore La Rosa, CSO
**Email**: slarosa@ctf.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. Our grant guidelines explain that every grantee will automatically be part of the NF-Open Science Initiative, an alliance that was formed with the goal of fostering a paradigm of open science in the neurofibromatosis and schwannomatosis research community.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. XIII. Data and Resource Sharing  Data: CTF believes in making data from all its funded projects freely accessible irrespective of whether the findings were positive or negative. Normally CTF allows for 12 months embargo on the data from the end of the award. During the embargo period, only the awardees have access to the data after which the data will be opened to the community. Towards this, in 2018 CTF launched the NF Open Science Initiative (NF-OSI, https://www.synapse.org/#!Wiki:syn17083165/ENTITY/587186), an alliance created with the goal of fostering a paradigm of open science in the neurofibromatosis and schwannomatosis research community. It is the culmination of a collaboration effort initiated by CTF and the Neurofibromatosis Therapeutic Acceleration Program (NTAP). All CTF-funded awards require participation in the NF-OSI. To participate, awardees need to acquire a Synapse account (https://www.synapse.org/#!Synapse:syn17083165/wiki/591070), register their project's basic information as a new study, and include a Data Sharing Plan (DSP) for their proposal. Resources: CTF incentivizes the sharing of resources (reagents, cell lines, etc), especially for those developed during the grant. It is an optional opportunity for the awardee to make this resource available to the community and list the resources on our website and the data portal.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire Portfolio
7. If just some, to which grants/programs do they apply and why?
    a. We started with selected grants, now we extended it to the entire portfolio.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. Educate our grantees on how to do data-sharing. For every new grant cycle, we are now delivering on-boarding sessions with our partners at Sage Bionetworks to explain to new grantees how to share (use of nf data portal) and what to share.
9. Do you specifically require or encourage your awardees to share data?
    a. Require

10. Comments
    a. We require participation in the NF-OSI and a data-sharing plan at the submission.
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. Yes, our NF data portal.
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. Part of the application
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. We ask the grantee to share as soon as the data is ready, but it can be embargoed up to 12 months after the end of the award.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. Yes - it is part of the administrative review
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. We rely on our expert at Sage for the evaluation of the data sharing plan.
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. Through our partnership with Sage, the team can provide some support to the grantees.
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes - At the conclusion of every award cycle, we will coordinate with our team at Sage to check what data has been uploaded to the portal. If there is no data, then we reach out and ask for the task to be completed.
22. What are the consequences for noncompliance with your policy?
    a. We haven't had any issues with non-compliance. The most common issue is that we cannot track the completeness of data generated versus shared.

## Circle of Service Foundation

**Primary Survey Contact**: Josh McGowan, Program Officer
**Email**: josh@cosfoundation.org

*Note: COSF does not support researchers directly, but rather organizations that fund researchers. The language included in our grant terms relate to organizations have a data-sharing policy/plan.*

Language Included in Grant Terms: No Year One payment shall be due with respect to the Grant unless and until you submit an initial "Organizational Data-sharing Plan". The Organizational Data-sharing Plan should detail how you plan to encourage and/or require funded researchers to (1) craft data-sharing plans and (2) share applicable data that are generated from their funded projects. No Year Two payment

shall be due with respect to the Grant unless and until you submit an update on the progress towards finalizing and implementing the Organizational Data-sharing Plan.

## CMT Research Foundation

**Primary Survey Contact**: Keith Fargo, Chief Scientific Officer
**Email**: keith@cmtrf.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Thinking about it
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. We informally encourage it, but do not yet have a formal policy. We just started working on one with our SAB.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. The main challenge we are having right now is how to write a policy that encourages or insists upon data sharing without discouraging people from applying to us for funding. We also need to think through a process for exceptions when data sharing would run afoul of confidentiality and non-disclosure agreements.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. Not for at least a year to work through all of these issues. We don't want to rush this and make a misstep that could haunt us later.
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. N/A
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Neither
10. Comments
    a. We encourage informally, but not with anything in writing.
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. N/A
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. N/A
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. We have not set specific timelines. That is part of what we need to think through in developing our policy.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. N/A

16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. All our grants are reviewed by our SAB, and we have just started this conversation with them (literally earlier this week).
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. We would allow this, but this is not something we generally see in the budgets we are receiving. If/when we create a policy requiring data sharing, we would ask that it be included in the budget, and we would cover the costs. Same for open access publishing.
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. N/A
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Not now, but we will if/when we implement a policy.
22. What are the consequences for noncompliance with your policy?
    a. N/A

## Crohn's & Colitis Foundation

**Primary Survey Contact**: Angela Dobes, Vice President IBD Plexus
**Email**: adobes@crohnscolitisfoundation.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. We have policies for our IBD Plexus program. Raw data generated for a project, such as data derived from biosamples, must come back into IBD Plexus once data exclusivity terms expire, which all IBD Plexus across industry and academia can then access Exclusivity Terms Non-targeted data Data that is not directed at a particular hypothesis but rather is a preliminary exploration of the data 6 months exclusivity Targeted data Data that has been generated to answer a focused hypothesis Data that has been generated using a non-commercially available assay 18 months exclusivity and opportunity for additional 12-month extension Also accommodates need to abide by journals and funders data sharing requirements If required, to post data to a 3rd party database Only limited, required clinical data should be posted Data can only be used for validation or reproduction of the research results Data is subject to a Creative Commons license with all of the attribution non-commercial and share alike features or their equivalent.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?

    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Specific grants
7. If just some, to which grants/programs do they apply and why?
    a. I selected grants because there were no other options but the data sharing rules only apply to IBD Plexus projects.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Require
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. Yes, IBD Plexus
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. N/A
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. See exclusivity terms above.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. Yes - it is part of the administrative review
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. No
19. Comments
    a. We do this for free
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes
22. What are the consequences for noncompliance with your policy?
    a. Can no longer access IBD Plexus resources

## Damon Runyon Cancer Research Foundation

**Primary Survey Contact:** Anja Armache, Scientific Director
**Email**: anja.armache@damonrunyon.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. I am not sure I understand this correctly but we do not require any data sharing practices as part of our application, we ask for information in their progress reports (not for all award programs) We also encourage publishing in open access journals by offering reimbursement of costs (annual fund of $25,000 that is available on a first-come, first-served basis to current awardees to pay fees incurred by publishing in open access journals) and we always try to encourage collaborations during our meetings.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. Do you have any plans to share your data, either before or after its publication in a peer reviewed journal? If so: a. What type of data is it (genome/exome sequence, images, structures, statistical software or algorithms, etc.)? b. Where will it be shared (such as Figshare, Dataverse or Open Science Framework (OSF))? c. Who will be in charge of preparing the data to share? d. What are the expected costs of preparing and sh
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Specific grants
7. If just some, to which grants/programs do they apply and why?
    a. We started with a smaller number to understand how to best collect these data and what we can learn. We picked different award programs to reflect the variety of research we fund.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. As this is a question and not a requirement we did not encounter any barriers.
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the progress report
13. Comments
    a. I would not call it a requirement for us (see wording above) but rather a question.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. N/A
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?

    a. No
17. Comments about training material
    a. This is not part of our application
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. Again data sharing is not a requirement but it can be part of a budget. We also have an Open Access Fee Reimbursment Policy that awardees can use (annual fund of $25,000 that is available on a first-come, first-served basis to current awardees to pay fees incurred by publishing in open access journals.)
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. Yes
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No
22. What are the consequences for noncompliance with your policy?
    a. N/A

## Doris Duke Charitable Foundation

**Primary Survey Contact**: Sindy Escobar Alvarez, Senior Program Officer for Medical Research
**Email**: sescobar@ddcf.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. The Medical Research Program of the Doris Duke Charitable Foundation does not require sharing of data derived from its grants. However, we communicate that we value data sharing by: 1) making a data sharing plan a required section of research proposals, 2) indicating that data sharing plans are evaluated when considering the impact of the proposal, and 3) offering additional funding to cover data sharing costs for recipients of our major research grant programs.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Specific grants
7. If just some, to which grants/programs do they apply and why?
    a. We offer data sharing grants only to recipients of our major research grants as a pilot program. These grants could be offered to recipient of training grants in the future.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?

    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application
13. Comments
    a. We only require a data sharing plan
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. We do not have requirements about timing. Timing depends on each project.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. Yes, it is part of peer review
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. No
19. Comments
    a. No, cost of data sharing are considered separately.
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. Yes
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes
22. What are the consequences for noncompliance with your policy?
    a. N/A

## Health Resources in Action

**Primary Survey Contact**: Lara Bethke, Chief Scientific Officer
**Email**: lbethke@hria.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. We encourage researchers to share data and other research outputs but we do not link this to grant payments.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. Figuring out how to enforce the policy in a meaningful and cost effective way, and making a convincing business case to our clients to pay us to do this.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire Portfolio
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. Lack of bandwidth to enforce
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. N/A
13. Comments
    a. We are adding a question directly aimed at this to our progress reports.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. Would love to know what other people are doing for this, and how it is working
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No, we plan to start asking directly in progress reports
22. What are the consequences for noncompliance with your policy?
    a. None

## Hydrocephalus Association

**Primary Survey Contact**: Jenna Koschnitzky, National Director of Research

**Email**: research@hydroassoc.org

Note: *Questions 13-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. Most of the research we fund is basic science/pilot studies, while most of the data-sharing plans seem to be geared to genomic and other clinical data. We also do not have funding to help researchers implement data sharing plans.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. N/A
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## Kenneth Rainin Foundation

**Primary Survey Contact**: Jackie Hausman, Program Officer
**Email**: jackie.hausman@krfoundation.org

Note: *Questions 2-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments
11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## Lupus Research Alliance

**Primary Survey Contact**: Diomaris Gonzalez, Director of Grant Programs
**Email**: dgonzalez@lupusresearch.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A

3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. "This is part of our of our Terms and Conditions: Data Sharing. LRA is committed to sharing the results and accomplishments of the projects that it funds with the general public and the research community. The wide dissemination of data results in the expedited translation of research results into knowledge, products and procedures that hold the promise of improving human health. LRA endorses the sharing of final research data to serve these and other important scientific goals and requires the timely release and sharing of final research data from LRA supported studies for use by other investigators. "Timely release and sharing" is defined as no later than the acceptance for publication of the main findings from the final data set.

       All recipients of LRA awards must agree to this principle and must take concrete steps to facilitate availability of data and materials. Consequently, a Data Sharing Plan must be completed and approved by the LRA prior to grant activation. This data sharing plan should be submitted using the current NIH-style Data Sharing Plan.

       LRA recognizes that data sharing may be complicated or limited, in some cases, by organizational policies, policies pertaining to the protection of intellectual property, local IRB rules, and local, State and Federal laws and regulations, including privacy and confidentiality policies.

       The rights and privacy of individuals who participate in LRA sponsored research must be protected at all times and Grant recipient and Sponsoring Institution agree to comply with all relevant HIPAA provisions as they relate to patient data. Thus, data intended for broader use should be free of identifiers that would permit linkages to individual research participants and variables that could lead to deductive disclosure of the identity of individual subjects.

       A variety of institutional based and government organized databases (e.g. dbGaP, ImmPort) are now being made available to expedite the sharing of data. Where possible, subject to HIPAA regulations, data from LRA funded studies should be submitted to these publicly available databases."
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire portfolio
7. If just some, to which grants/programs do they apply and why?
    a. Entire portfolio, however, we hope to be able to strengthen how we review and follow up with grantees when it comes to data sharing.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Require

10. Comments
    a. All recipients of LRA awards must agree to this principle and must take concrete steps to facilitate availability of data and materials. Consequently, a Data Sharing Plan must be completed and approved by the LRA prior to grant activation. This data sharing plan should be submitted using the current NIH-style Data Sharing Plan.
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. We encourage the use of established repositories to the extent possible. A variety of institutional based and government organized databases (e.g. dbGaP, ImmPort) are now being made available to expedite the sharing of data. Where possible, subject to HIPAA regulations, data from LRA funded studies should be submitted to these publicly available databases.
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the Terms and Conditions and the progress report.
13. Comments
    a. All recipients of LRA awards must agree to this principle and must take concrete steps to facilitate availability of data and materials. Consequently, a Data Sharing Plan must be completed and approved by the LRA prior to grant activation. This data sharing plan should be submitted using the current NIH-style Data Sharing Plan.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. "Timely release and sharing" is defined as no later than the acceptance for publication of the main findings from the final data set.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. Would be very interested in how and what other organizations are doing about data sharing guidelines, review and how it is implemented.
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes - We ask PIs to share as part of progress reports.
22. What are the consequences for noncompliance with your policy?
    a. None

## Lymphoma Research Foundation

**Primary Survey Contact**: Whitney Steen, Associate Director, Research and Scientific Programs
**Email**: wsteen@lymphoma.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. Chapter 15 Availability of Research Results: Publications, Intellectual Property Rights, and Sharing Research Resources It is LRF policy that the results and accomplishments of the activities that it funds should be made available to the public. The LRF Grantee, Mentor, Sponsoring Institution acknowledge the foregoing and, as applicable, agree to share all information, discoveries, or ideas arising out of research funded in whole or in part by LRF with LRF and the medical community at large. As a means of sharing knowledge, LRF encourages grantees to arrange for publication of LRF-supported original research in primary scientific journals. For each publication that results from LRF grant-supported research, grantees must include an acknowledgement of LRF grant support. In general, rights in data resulting from a grant-supported project shall follow the Sponsoring Institution's policies to the extent such policies include applicable provisions. Any publications, data or other copyrightable works developed under an LRF grant may be copyrighted without LRF approval. One copy of each publication resulting from work performed under an LRF grant supported project must accompany the annual or final progress report submitted to LRF. 15 (a) Public Access Policy – PubMed Central In addition, LRF-funded researchers are required to submit, or have submitted for them, to the National Institutes of Health's Pub Med Central database an electronic version of the author's final manuscript including all modifications from the publishing and peer review process (the "postprint") upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication. This requirement applies to all grants awarded after May 1, 2012, whether LRF funds the research in whole or in part. All scientific progress reports must include the PMC ID number to publications supported by the Lymphoma Research Foundation starting May 1, 2012. 15 (b) Unique Research Resources Investigators conducting biomedical research frequently develop unique research resources. Categories of these resources include synthetic compounds, organisms, cell lines, viruses, cell products and cloned DNA as well as DNA sequences, and mapping information. Specific examples include specialized or genetically defined cells, including normal and diseased human cells; monoclonal antibodies; hybridoma cell lines; microbial cells and products; viruses and viral products; recombinant nucleic acid molecules; DNA probe; nucleic acid and protein sequences; certain types of animals, such as transgenic mice; and intellectual property, such as computer programs. LRF considers the sharing of such unique research resources an important means to enhance the value of LRF-sponsored research. These materials represent a valuable resource for the scientific community at large, paid for by the generous contributions of LRF's donors. The availability of these research resources directly affects the ability of the members of the scientific community to replicate the experiments of others and the pace and cost of future research. Therefore, LRF requires that when these resources developed with LRF funds and the associated research findings have been published, the LRF Grantee, Mentor, and Sponsoring Institution accept the responsibility of providing biological reagents developed during the course of LRF-sponsored research when reasonably requested to do so by other investigators. Grantees

are expected to submit unique biological information, such as DNA sequences, to the appropriate data banks so that they can be made available to the broad scientific community.

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. We are seeing more pushback from institution grant offices on even our fairly general policy and are not sure how a more robust policy requiring using of an open data repository and/or a data management plan will be received. LRF also funds grants that usually have some additional funders covering some expenses and we're trying to think through how best to resolve any conflicts between policies (for example if some trial expenses are covered by a pharmaceutical grant that has certain confidentiality expectations).
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire Portfolio
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. Staff bandwidth for tracking compliance and what resources are available. (Have managed to stay on top of publications but other data is more difficult.)
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. N/A
13. Comments
    a. We do not currently require this although we are considering adding such a requirement.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. We recommend a year from initial data publication but do not have any consequences if it takes longer than that.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. I have yet to see many of these costs in our application budgets but we would permit it under the same umbrella as publication costs.

20. Do you provide any supplemental financial support to facilitate data sharing?
    a. N/A
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No
22. What are the consequences for noncompliance with your policy?
    a. Currently none

## Melanoma Research Alliance

**Primary Survey Contact**: KristenMueller, Senior Director, Scientific Program
**Email**: kmueller@curemelanoma.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. MRA has adopted the following Data Sharing Policy:
        - MRA requests the posting of manuscripts based on or developed under an MRA Award to a pre-print server ahead of or at the time of journal submission.
        - MRA requests the posting of research outputs (data, code, software) to public data repositories at the time such research outputs are generated.
        - MRA requests that manuscripts based on or developed under an MRA Award be published in open-access journals. 2021 MRA Dermatology Fellows Award RFP Page 9 of 13
        - MRA requests that all research outputs based on or developed under an MRA Award (including publications, data, code, and software) be made available with no commercial modification rights (e.g. CC BY-NC license).
        - MRA requires that the final, accepted version of any publication based on or developed under an MRA Award be deposited in PubMed Central so that it is available 12 months after publication.
        - MRA requires that any data, code, and/or software needed for the independent verification of published research results based on or developed under an MRA Award be curated and made freely and publicly available at the time of publication. MRA will incur costs associated with policy compliance, provided these fees (e.g., article processing charges, data storage), are included in the original grant application budget.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire Portfolio

7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. We are in the process of implementing this - it appears in our RFP and applications in 2020 and will be implemented for all grants activated in 2021.
9. Do you specifically require or encourage your awardees to share data?
    a. Require
10. Comments
    a. A little of both - we require data sharing upon publication.
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. We encourage the use of established repositories to the extent possible. If asked, I would direct our awardees to a list such as this: https://www.nature.com/sdata/policies/repositories#life
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application
13. Comments
    a. We ask questions the following questions as part of our application: What data will be generated by your research? What is your plan for sharing the data? What renewable reagents will be generated by your research? What is your plan for sharing the renewable reagents for non-commercial research purposes?
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. MRA requires that any data, code, and/or software needed for the independent verification of published research results based on or developed under an MRA Award be curated and made freely and publicly available at the time of publication. MRA will incur costs associated with policy compliance, provided these fees (e.g., article processing charges, data storage), are included in the original grant application budget.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No
22. What are the consequences for noncompliance with your policy?
    a. Since we have not yet implemented this, we have formally determined this. I imagine we will not be particularly harsh in enforcement primarily because we are not staffed at a level that would allow for careful monitoring.

## Misophonia Research Fund of The REAM Foundation

Primary Survey Contact: Elizabeth Versten, Executive Director
Email: Beth@REAMfoundation.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. Convening grantees to share what they are doing
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. "Data Sharing Requirements: The Grantor expects that the Grantee Organization will share data publicly and in a timely, accessible manner. As such, the Grantee Organization agrees that it will make every effort to encourage the Principal Investigator leading the Project to share his/her data, with the understanding that some data may not be shareable for reasons related to intellectual property protection and other aspects of drug development. In that spirit, the Grantee Organization agrees to adhere as best as possible to the following requests:
        - The final research data, including the metadata and descriptors, and all analytical methodologies should be shared publicly, as specified in the approved proposal within the Grantee Organization's Data Sharing Plan, attached, with the intent to make the sharing as meaningful and usable as possible. Project results, data (de-identified, where applicable), reagents, and other research tools developed under the Grantor funding should be included in the Data Sharing Plan where applicable. The analytical methods, statistics, and tools utilized to assess the data should also be defined, de-identified where applicable, and shared.
        - Where reasonable, data should be shared as broadly as possible, such that the public has access freely, specifically to statistics, tools, reagents, biomarkers, and data (de-identified where appropriate) associated with the Project. Additionally, any failures should be included in the publicly available dataset.
        - Data generated or collected in the Project must be shared by publication or presentation by the Grantee Organization. Additionally, where reasonable, raw data, metadata, summary tables, and associated findings should be made publicly available by deposit in a publicly accessible data repository, as specified by the Grantee Organization's Data Sharing Plan.
        - Data should be shared as described above in a timely manner, and the findings and annotations should be shared within eighteen (18) months after completion or earlier termination of the Project. The Grantor reserves the right to make the Project results and data public after eighteen (18) months from the conclusion of the Grant Term.
        - The Grantor reserves the right to publicly disclose the Project results and data in press releases, announcements, and otherwise (subject to the Grantee Organization's advance approval if any such public disclosure will occur before the applicable results have been made public through publication, presentation, or otherwise). The Grantee Organization shall summarize the measures it has taken to

share data in accordance with its approved Data Sharing Plan, including the location of such data, in its progress and final reports."

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. no common repository for data, no consistent format
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. Already in progress
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire portfolio
7. If just some, to which grants/programs do they apply and why?
    a. It applies to our entire scientific research portfolio. Our other grants are not related to scientific research.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. no common repository for misophonia data, no consistent format
9. Do you specifically require or encourage your awardees to share data?
    a. Require
10. Comments
    a. The policy above is from our grant agreement.
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application, terms and conditions, and part of the progress report
13. Comments
    a. All of the above.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. As soon as possible, and no later than 18 months after project completion
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. Yes - it is part of the administrative review
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No – good idea, though
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes, absolutely
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No, we have not yet completed a round of grants yet, so have not come to that point. Good idea though.
22. What are the consequences for noncompliance with your policy?
    a. None yet, but if they don't have a plan in their proposal, we would be less likely to fund them.

## NYSCF

**Primary Survey Contact**: Kristin Smith, AVP, External Programs
**Email**: ksmith@nyscf.org

*Note: Questions 2-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments
11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## Pershing Square Sohn Cancer Research Alliance

Primary Survey Contact: Christy Barrow,Program Manager
Email: cbarrow@persq.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.

    a. In addition to our "public access" clause in our grant agreement, we ask our PIs for their ORCID number when submitting their annual progress reports. We don't yet do anything with this information, but it helps us see which grantees are part of the ORCID platform (most are) and lets them know that we are thinking about it and for data sharing. We also have our grantees convene to share their data with their peers, but of course this doesn't cover the entire scientific community, just a subset of cancer researcher in NY.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. We have a "public access" policy, which is likely not as comprehensive as many other data sharing policies, but is a start for us! Upon the recommendations from NFRI's workshops, we adopted their language for our grant agreements as of 2020: "Public Access: Awardee is required to submit, or have submitted for them, to the National Institutes of Health's Pub Med Central database an electronic version of the Awardee's final manuscript including all modifications resulting from the publishing and peer review process (the "postprint") upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication. This requirement applies to all publications related to PSSCRA-funded research whether the research was funded in whole or in part. All narrative progress reports must include the PMC ID number (PMCnnnnn) for publications, if known at the time of the report due date."
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. We likely don't have the bandwidth to ensure all of our grantees are following the policy (and/or to assist them with these efforts), and just aren't sure how much new language around this to include in our grant agreements that the institutions wouldn't push back on in any way.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire portfolio
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. None yet, but this is very new and our terms are very loose/easy to follow, as pretty much all of our grantees include their postprints on PubMed. Many also put their pre-prints up, as well, although not yet required.
9. Do you specifically require or encourage your awardees to share data?
    a. Require
10. Comments
    a. Require them to put their final manuscript (postprint) on pub med's central database, but do not require them to put a preprint or anything further (we just encourage this).
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. Yes - We specifically required Pub Med Central, but would like to encourage the use of established pre-print repositories to the extent possible.
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. N/A
13. Comments

    a. No, we do not. We would be interested in seeing templates that other foundations use, and learning at what step in the application cycle they require submission (pre or post-award).
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. No later than 12 months after the official date of publication (language recommended by NFRI)
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. This isn't specifically included in the itemized budgets we receive, but we would consider and allow it under the budget item of "publication fees".
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No. We do not specifically ask them this, but when they send us their publications from the award, I manually search for each one on PubMed and in a tracking document link the title of the article to the PubMed page.
22. What are the consequences for noncompliance with your policy?
    a. N/A (technically it would be a breach of the agreement and we would have to discuss with the PI/institution, but it is unlikely we would take such severe action and just sent an email to a grantee asking them why it's not on PubMed or a similar open access repository- this has yet to occur.)

## Simons Foundation

**Primary Survey Contact**: Richard McFarland, Director, Grants Management
**Email**: rmcfarland@simonsfoundation.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. "The Simons Foundation utilizes SFARI Base, a clearinghouse for autism and autism-related research data and biospecimens supported by the Simons Foundation Autism Research Initiative (SFARI) as well as an online portal for the submission of research recruitment requests. It contains data from the following cohorts:
        - Simons Simplex Collection (SSC)
        - Simons Searchlight
        - Simons Foundation Powering Autism Research for Knowledge (SPARK)

- Autism Inpatient Collection (AIC)

Researchers can request access to phenotypic, genetic, or imaging data; order biospecimens; and/or submit a research match application to recruit individuals or families for future research studies.

In SFARI progress reports we ask the following:

- What data were generated by your research during this funding period?
- Has there been a change in your plan for sharing all data? The plan should include a timeline and type of shared content (e.g., file types).

Our policies are as follows:

- The Simons Foundationt believes it is essential that institutions and PIs share data developed using Simons Foundation funds with other qualified investigators. PIs are required to have a renewable reagents and data-sharing plan in place prior to receiving a grant. At the foundation's discretion, the PI may be required to provide the foundation with an electronic copy of all properly de-identified research-generated data prior to the end of the grant. In all reasonable cases, the foundation will assume financial responsibility for costs associated with the data transfer.
- The Foundation encourages PIs to post preprints on recognized servers, such as arXiv https://arxiv.org/ or bioRxiv http://biorxiv.org/, in parallel with (or even before) submission to a peer-reviewed journal. The Foundation also encourages PIs to publish under Open Access licenses, which are allowable budget costs.
- Additionally, upon publication of results, the PI(s) must make every effort to deposit all research-generated data into public databases that are widely accessible, without charge, to the scientific research community. If no such databases are available that properly fit the type and content of the research-generated data, the PI(s) must make every effort to make these data available through electronic supplementary tables and figures, which are now routinely associated with publications."

3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. https://www.simonsfoundation.org/funding-opportunities/policies-and-procedures?tab=policies-and-procedures It is essential that institutions and PIs share renewable reagents and data developed using Simons Foundation funds with other qualified investigators. PIs will be required to have a renewable reagents and data-sharing plan in place prior to receiving a grant. At the foundation's discretion, the PI may be required to provide the foundation with an electronic copy of all properly de-identified research-generated data prior to the end of the grant. In all reasonable cases, the foundation will assume financial responsibility for costs associated with the data transfer. Foundation personnel may, at their discretion, release these data to other qualified investigators who agree not to publish on these data until after an embargo period expires. The length of this embargo period will be established on a case-by-case basis in consultation with the PI(s) but will generally not exceed one (1) year after the end of the grant or until publication, whichever comes first. These data may eventually be integrated with the other foundation data collections. How, when and if such data are made available to the wider research community will remain under the sole discretion of the foundation. In addition, upon publication of results, the PI(s) shall make every effort to deposit all research-generated data into public databases that are widely accessible, without charge, to the scientific

research community. If no such databases are available that properly fit the type and content of the research-generated data, the PI(s) shall make every effort to make these data available through electronic supplementary tables and figures, which are now routinely associated with publications. Genetically modified model organisms (e.g., mutant mice, rats, drosophila, C. elegans, zebrafish) as well as primary or genetically modified cells (e.g., fibroblasts, induced pluripotent stem cells, lymphoblastoid cell lines) are considered renewable reagents, and PI(s) and their institution(s) are required to share renewable reagents developed with Simons Foundation funding or used in a Simons Foundation–funded project by deposition at a third-party repository (e.g., the Jackson Labs, the Mutant Mouse Research and Resource Center, the Drosophila Genomics Resource Center, the Zebrafish International Resource Center, the Rutgers University Cell and DNA Repository) or at a repository indicated by the Simons Foundation. In instances where a legally binding restriction prevents deposition of a reagent in a third-party repository, the Simons Foundation expects the PI(s) and the PI(s) institution(s) to work with the licensed grantor, the repository and Simons Foundation staff members to identify an acceptable solution. The Simons Foundation prefers that this be done before use of the reagent. Failure to reach a mutually satisfactory solution may affect future funding decisions. To close the grant, the PI(s) must deposit the renewable reagents into the repository by the end of the grant and/or within a month of the initial publication, whichever occurs first. An embargo period for access to a model organism or a cell lines, negotiated on a case-by-case basis in consultation with the PI(s), will be maintained until the initial results are published by the PI(s), at which time it will expire. The PI(s) is responsible for transferring the animals/cells to the appropriate center. Once the transfer has occurred, the Simons Foundation will pay costs assessed by the third-party repository for maintenance of the reagent at the depository. In specific cases at the foundation's discretion, the foundation may contract with the repository to pay the additional costs necessary to maintain the model organism as a viable colony versus preservation in a frozen repository or to backcross the model into a different strain in the case of rodent models.

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire portfolio
7. If just some, to which grants/programs do they apply and why?
    a. The only grants it would not apply to are those that do not generate data and/or renewable reagents.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. N/A
9. Do you specifically require or encourage your awardees to share data?
    a. Require
10. Comments
    a. How, when and if such data are made available to the wider research community will remain under the sole discretion of the foundation. In addition, upon publication of results, the PI(s) shall make every effort to deposit all research-generated data into public databases that are widely accessible, without charge, to the scientific research community. If no such

databases are available that properly fit the type and content of the research-generated data, the PI(s) shall make every effort to make these data available through electronic supplementary tables and figures, which are now routinely associated with publications.

11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. We encourage the use of established repositories to the extent possible. The Simons Foundation utilizes SFARI Base, a clearinghouse for autism and autism-related research data and biospecimens supported by the Simons Foundation Autism Research Initiative (SFARI) as well as an online portal for the submission of research recruitment requests. It contains data from the following cohorts: Simons Simplex Collection (SSC) Simons Searchlight Simons Foundation Powering Autism Research for Knowledge (SPARK) Autism Inpatient Collection (AIC) Researchers can request access to phenotypic, genetic, or imaging data; order biospecimens; and/or submit a research match application to recruit individuals or families for future research studies. Similar data archives at the Simons Foundation in addition to those offered by SFARI, there is the Simons Genome Diversity Project dataset. More information on that can be found at https://www.simonsfoundation.org/simons-genome-diversity-project/. The SF encourages the use of MAGMA and arXiv for our Mathematics and Physical Sciences division. More information can be found on that at https://www.simonsfoundation.org/mathematics-physical-sciences. Information on data sharing tools our Simons Collaboration on the Global Brain can be found here: https://www.simonsfoundation.org/collaborations/global-brain/data-sharing-initiatives/. Information on Data sharing efforts for the Simons Collaboration on Ocean Processes and Ecology can be found at http://scope.soest.hawaii.edu/data/. Information on the Simons Collaboration on Arithmetic Geometry Number Theory, and Computation can be found at https://simonscollab.icerm.brown.edu/data/.
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application, as part of terms and conditions, and as part of the progress report.
13. Comments
    a. In SFARI progress reports we ask the following: What data were generated by your research during this funding period? Has there been a change in your plan for sharing all data? The plan should include a timeline and type of shared content (e.g., file types).
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. It is essential that institutions and PIs share renewable reagents and data developed using Simons Foundation funds with other qualified investigators. PIs will be required to have a renewable reagents and data-sharing plan in place prior to receiving a grant. At the foundation's discretion, the PI may be required to provide the foundation with an electronic copy of all properly de-identified research-generated data prior to the end of the grant. In all reasonable cases, the foundation will assume financial responsibility for costs associated with the data transfer. Foundation personnel may, at their discretion, release these data to other qualified investigators who agree not to publish on these data until after an embargo period expires. The length of this embargo period will be established on a case-by-case basis in consultation with the PI(s) but will generally not exceed one (1) year after the end of the grant or until publication, whichever comes first. These data may eventually be integrated with the other foundation data collections. How, when and if such data are made available to the wider research community will remain under the sole discretion of the foundation. In addition, upon publication of results, the PI(s) shall make every effort to

deposit all research-generated data into public databases that are widely accessible, without charge, to the scientific research community. If no such databases are available that properly fit the type and content of the research-generated data, the PI(s) shall make every effort to make these data available through electronic supplementary tables and figures, which are now routinely associated with publications.

15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. If we or a reviewer notice that a likely-to-be-funded application has a weak data sharing plan, we will flag it in the Letter of Intent. In most cases, PI will revise and resubmit their stronger data sharing plan as part of their pre-award materials. A huge caveat of all of this is that few disciplines have standardized, accessible ways of sharing data in user-friendly ways. Genomics and neuroimaging are the exception. As such, it can require a LOT of RA effort to get the data into a format that is useful and then further effort to maintain the inevitable 'service' aspect of it (responding to other PIs requests, prepping the data etc.). This is why data sharing for most disciplines is much easier said than done.
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. The reviewers we recruit for each application are well-versed in the standard practices and norms of how that particular field shares, so they often know better than we do what a truly strong data sharing plan looks like. The exception here is probably genomics since we have such deep in-house expertise in that discipline.
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. At the foundation's discretion, the PI may be required to provide the foundation with an electronic copy of all properly de-identified research-generated data prior to the end of the grant. In all reasonable cases, the foundation will assume financial responsibility for costs associated with the data transfer. To close the grant, the PI(s) must deposit the renewable reagents into the repository by the end of the grant and/or within a month of the initial publication, whichever occurs first. An embargo period for access to a model organism or a cell lines, negotiated on a case-by-case basis in consultation with the PI(s), will be maintained until the initial results are published by the PI(s), at which time it will expire. The PI(s) is responsible for transferring the animals/cells to the appropriate center. Once the transfer has occurred, the Simons Foundation will pay costs assessed by the third-party repository for maintenance of the reagent at the depository. In specific cases at the foundation's discretion, the foundation may contract with the repository to pay the additional costs necessary to maintain the model organism as a viable colony versus preservation in a frozen repository or to backcross the model into a different strain in the case of rodent models.
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. Yes
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. In practice, this can be quite hard to police because most shareable data produced by a SFARI grant is not actually ready to share until long after the grant has ended.
22. What are the consequences for noncompliance with your policy?

a. Failure to reach a mutually satisfactory solution may affect future funding decisions. If a PI does not share a mouse model through a third-party repository SFARI may take a strong stance with a PI's non-compliance. When a PI who has not shared a mouse model through a third-party repository submits a new grant application SFARI may choose to only review the application once confirmation that the mouse model has been deposited through a third-party repository. Overall, these efforts to enforce compliance become extremely difficult outside of the field of genomics and neuroimaging because the community lacks the proper infrastructure even if the motivation is there.

   Data sharing varies enormously across different types of data sets. For some highly structured data-- such as sequencing data-- there are highly standardized ways of sharing- and we generally have no issues with PIs in this department.

   For other types of data sets-- it's far more complicated as there are fewer/no standardized ways of storing/sharing-- and we often settle for the suboptimal situation where PIs agree to share with anyone who requests. That checks a box but in the real world it's highly suboptimal and often leads to issues down the road of data access.

   I would be in favor of devoting more resources to this task-- but exactly how best to approach is not obvious.

## St. Baldrick's Foundation

**Primary Survey Contact**: Laura Chung, Director of Grants Administration
**Email**: laura@stbaldricks.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. We have added a required data sharing plan to our applications. They are also asked to report on data sharing in report of results.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. RENEWABLE REAGENTS AND DATA SHARING PLAN (Instructions for St. Baldrick's Foundation applicants)

       It is essential that St. Baldrick's Investigators share renewable reagents and data developed with St. Baldrick's funds or used in the St. Baldrick's funded project with other qualified investigators. Grantees are expected to encourage and facilitate such sharing.

       St. Baldrick's Foundation funds biomedical research to better understand the causes of pediatric cancers and to advance its prevention, treatment, and cure. The main output of this research is new knowledge. To ensure this knowledge can be accessed, read, applied, and built upon in fulfillment of our goals, St. Baldrick's Foundation encourages researchers to share data with the research community in accordance with the NIH policy on data

sharing and expects its grantees to publish their findings, including but not limited to publication in peer reviewed journals.

The St. Baldrick's Foundation is aware of the need to provide flexibility in the assessment of Renewable Reagents and Data Sharing Plans. The foundation recognizes that disciplines differ widely in their practices and expectations. The foundation is actively working with the pediatric oncology research community to understand data sharing options, obstacles and opportunities, and encourages your input; please send any feedback or suggestions to Grants@StBaldricks.org

In developing your plan, you may want to consult with university officials as many universities have explicit reagent and data sharing policies.

For additional guidance related to this topic please refer to:
https://grants.nih.gov/policy/sharing.htm

Collaborative proposals and multi project proposals should include only one combined Renewable Reagent and Data Sharing Plan, regardless of the number of non-lead collaborative proposals or projects included.

The plan must address the following questions:

- What data and reagents will be generated by your research? (examples include mouse models and other key reagents that are not commercially available, genomics, consumables, etc.).
- What is your plan for sharing the data and reagents?
- What will be your policies for access and sharing the data during the life of your award and after award closeout?
- What will be the format, mode of delivery and timetable for data distribution?
- Please include any provisions for appropriate protection of privacy, confidentiality, security or intellectual property.

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire portfolio
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. Grants that are just for salary support question why/how to fill this in.
9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. N/A
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. No

12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application and progress report.
13. Comments
    a. N/A
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. We don't have timing guidelines
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. We only ask them to review it and add any feedback but it is not scored
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. I'm not sure this has been asked - but we would likely consider it
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes - we ask them in the progress reports.
22. What are the consequences for noncompliance with your policy?
    a. None.

## Susan G. Komen

**Primary Survey Contact**: Jamie Stanford, Manager, Research Programs
**Email**: jstanford@komen.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. "DATA SHARING POLICY (pasted from our Research Policies and Procedures) To accelerate scientific discovery, research results and data should be made as widely and freely available as possible, while safeguarding the privacy of participants and protecting confidential and proprietary data. Komen's Data Sharing Policy aligns with the NIH's data sharing requirements (https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm) including policies for sharing large-scale genomics data (https://osp.od.nih.gov/scientific-sharing/policies/) and clinical trial information (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-149.html). The policy will apply to all Komen-funded research grants, regardless of awarded amount, that were awarded after February 1, 2018, and it is strongly

encouraged for all grants awarded prior to this date. Applicants are required to provide a data sharing plan as part of their grant application and may request funds necessary for data sharing and archiving in the submitted budget. Grantees will be required to report on progress toward the data sharing plan in both annual and final progress reports. Key points of the Komen Data Sharing Policy include the following:

- 9.1 What data should be shared?
    - All data from basic, translational, clinical, and other types of research studies should be considered for data sharing. This includes laboratory research and all clinical trials, regardless of study phase, type of intervention, etc.
    - Final research data, especially unique data, along with metadata and descriptors-- i.e., all material necessary to document, support, and validate research findings-- must be shared.
- 9.2 When should data be made available?
    - Data should be made available as soon as possible and for as long as possible. Data must be released no later than the time of publication of the main findings from the final dataset, although possible exceptions will be made to protect patentable and other proprietary data.
    - Clinical trials must be registered at clinicaltrials.gov no later than 21 days after enrollment of first participant and updated at least once a year. Summary results, including adverse event information, must be provided no later than one year after the trial completion date, unless regulatory approval of the product is being sought.
- 9.3 Where/With whom should data be shared?
    - Data should be shared as broadly as possible to the extent consistent with applicable laws, regulations, rules, and policies. Rights and privacy of human subjects must be protected at all times
    - Researchers may select the method(s) for data sharing.
    - Data repositories with common standards and an established infrastructure dedicated to the appropriate distribution of data are recommended.
    - Large-scale non-human genomic data must be submitted to any widely used data repository; large-scale human genomic data must be submitted to an NIH-designated data repository.
    - Clinical trials should be registered at clinicaltrials.gov. Please reference Appendix B for further details on elements to consider when preparing a Data Sharing Plan.
- APPENDIX B: KEY ELEMENTS TO CONSIDER IN PREPARING A DATA SHARING PLAN UNDER NIH EXTRAMURAL SUPPORT - Research results developed with NIH funding should be broadly available to the research community for furthering research. This resource document is intended to assist applicants by outlining certain key elements that should be addressed in any data sharing plan. While the precise content of a data sharing plan may vary depending on the data being generated and collected, addressing the basic questions of What, Who, Where, When, and How can assist researchers and research administrators in formulating a meaningful data sharing plan that communicates essential information about:
    - What data will be shared?
    - Who will have access to the data?

- Where will the data to be shared be located?
- When will the data be shared?
- How will researchers locate and access the data? What data will be shared? To optimize the benefits of data sharing, final research data along with metadata and descriptors should be shared to make sharing meaningful and usable by other researchers.

- In describing what data will be shared, a data sharing plan should indicate:
  - What types of data are to be collected in the study and shared (e.g., genetic, physiological, clinical, medical history, etc.)?
  - Will the study include unique data that cannot be readily duplicated (e.g., large surveys that are too expensive to replicate; studies of unique populations, such as centenarians; studies conducted at unique times, such as a natural disaster; studies of rare phenomena, such as rare metabolic diseases; etc.)?
  - Will individual-level data or raw data also be shared, and if so, will the whole data set be shared? Will aggregate data (e.g., summary statistics or tables) also be shared? Will the analytical methods used (tools and parameters) be defined?
  - What data quality control measures will be implemented?
  - What data documentation will be shared (e.g., metadata, descriptors, schema) so that others can understand and use the dataset and to prevent misuse, misinterpretation, or confusion?
  - What commonly accepted data standards or standardized vocabularies will be used to enable others to interpret the data and improve interoperability with other data systems?
  - What format will be used to encode the data? Will this format be consistent with extant, commonly used standards?
  - In addition to final research data, what other data will be available? Who will have access to the data? To maximize the benefits of data sharing, data should be shared as broadly as possible to the extent consistent with applicable laws, regulations, rules, and policies.
- In describing who will have access to data, a data sharing plan should indicate:
  - Will the general public have access to some or all of the data?
  - Will access to certain data or certain components of the data be restricted to qualified researchers, e.g., to address specific rules, laws, regulations or policies (e.g., IRBs, human subjects, informed consent, etc.)?
  - If data access is restricted, what are the justifications/criteria for restricting access (e.g., relevant laws (local, State, Federal, etc.), regulations, rules, institutional policies, IRB approvals, and consent documents?
  - What will researchers who seek to obtain data need to do to comply with any data access restrictions?
  - Are there any limitations on release of data that may be considered “sensitive”?
  - What data sharing agreements will be necessary to appropriately restrict the transfer of protected, sensitive, or confidential data to others and to require that data be used only for research purposes?

- Who will be operationally responsible for ensuring that no personally identifiable information is made available (e.g., principal investigator, independent curator)? Where will the data to be shared be located? To minimize additional administrative workloads for sharing of data, data repositories with common standards and an established infrastructure dedicated to the appropriate distribution of data would generally be ideal for data sharing.

- In determining where data to be shared will be located, a data sharing plan should indicate:
  - Will an existing database, data repository, data enclave, or archive be used to store and disseminate the data (e.g., dbGaP, National Database for Autism Research (NDAR)), and if so, how the policies and procedures in place for others to access the data are consistent with applicable NIH policies?
  - Will a new repository need to be developed, and if so, who/what will maintain the repository?
  - Will the data be distributed directly by an investigator to those who request it (e.g., through an electronic file)? When will the data be shared? To optimize the timely and broadest usage of data, data should be made available as soon as possible and for as long as possible.
- In determining the timeframes for data sharing, a data sharing plan should indicate:
  - The schedule for release of data:
    - What data, if any, will be shared prior to publication?
    - What data will be shared upon acceptance for publication?
    - If using a repository, when will data be submitted to the repository?
  - Will data from ongoing longitudinal studies be released in increments as data become available?
  - Will the timing of data sharing be specifically linked to other relevant policies concerning the timing of release of data (e.g., NIH GWAS policy, ClinicalTrials.gov, specific requirements in the funding opportunity announcement (FOA))?
  - How will data maintenance and access be ensured after the award ends?
    - Will there be support for continued sharing of data (e.g., through grant applications, administrative supplements, or other sources) or planned migration of data to another database, data repository, etc.? How will researchers locate and access the data? To optimize usage of the data, researchers need to be able to easily identify locations of relevant data and to be able to easily access the data.
- In describing how researchers will learn about, locate, and access the data, a data sharing plan should indicate:
  - What steps will be taken to help researchers know that the data sets exist?
    - Will registries, repositories, indexes, word-of-mouth, publications, and/or other approaches be used to publicize the availability and accessibility of the data?
    - Will these be linked and cross-referenced so other researchers can readily find them?
  - How will the data be accessed (web service, ftp, etc.)?"

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. N/A
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Entire portfolio
7. If just some, to which grants/programs do they apply and why?
    a. N/A
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. Grantees inform us if they are following our policy, but we don't have the bandwidth to confirm that the data is being shared as it should.
9. Do you specifically require or encourage your awardees to share data?
    a. Neither
10. Comments
    a. Both - Our policy states this: The policy will apply to all Komen-funded research grants, regardless of awarded amount, that were awarded after February 1, 2018, and it is strongly encouraged for all grants awarded prior to this date.
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. We encourage the use of established repositories to the extent possible. "Where/With whom should data be shared?
        - Data should be shared as broadly as possible to the extent consistent with applicable laws, regulations, rules, and policies. Rights and privacy of human subjects must be protected at all times.
        - Researchers may select the method(s) for data sharing.
        - Data repositories with common standards and an established infrastructure dedicated to the appropriate distribution of data are recommended.
        - Large-scale non-human genomic data must be submitted to any widely used data repository; large-scale human genomic data must be submitted to an NIH-designated data repository.
        - Clinical trials should be registered at clinicaltrials.gov."
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application, terms and conditions and part of the progress report.
13. Comments
    a. We require a data sharing plan as part of the application and request an update on how it is being following in the annual progress report. It is also part of our Research Policies and Procedures.
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. "When should data be made available?
        - Data should be made available as soon as possible and for as long as possible. Data must be released no later than the time of publication of the main findings from the final dataset, although possible exceptions will be made to protect patentable and other proprietary data.
        - Clinical trials must be registered at clinicaltrials.gov no later than 21 days after enrollment of first participant and updated at least once a year. Summary results, including adverse event information, must be provided no later than one year after

the trial completion date, unless regulatory approval of the product is being sought."

15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. Yes - it is part of the administrative review
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. N/A
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. Applicants may request funds necessary for data sharing and archiving in the submitted budget.
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. Yes, Grantees will be required to report on progress toward the data sharing plan in both annual and final progress reports.
22. What are the consequences for noncompliance with your policy?
    a. Progress reports are only approved if completed, which includes reporting any progress toward the data sharing plan. And, approved reports are needed before we send payment. However, outside of holding payments, there are no other consequences for noncompliance with our policy. As stated earlier, one of the challenges we have is that the grantee may report they are meeting our policy, but we do not follow up in detail (investigate/confirm what, when and where data is being shared) to ensure the grantee is meeting our policy.

## The Gerber Foundation

**Primary Survey Contact**: Catherine Obits, Program Manager
**Email**: cobits@gerberfoundation.org

*Note: Questions 5-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. N/A
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. N/A
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. Internal expertise to guide this activity
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?

6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments
11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## The Kavli Foundation

**Primary Survey Contact**: Chris Martin, Director, Physical Sciences & Assoc VP Operations
**Email**: cmartin@kavlifoundation.org

*Note: Questions 5-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. We support and welcome our grantees' efforts in data-sharing and specifically initiated and fund an effort (Neurodata without Borders www.nwb.org) to create a data-sharing standard in neuroscience. But since our other funding (outside of NWB) is unrestricted we are careful not to place restrictions or limitations on what our funding can do. If grantees want to use our funds toward data-sharing that is entirely at their discretion.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments

11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## The V Foundation for Cancer Research

**Primary Survey Contact**: Carole Wegner, Senior VP, Research and Grants Administration
**Email**: cwegner@v.org

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. Yes
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
    a. As an early step in this process, we added a survey question on both our applications and progress reports which asked if the applicant has a data sharing plan and to describe it if they did. We used these questions to understand whether individuals were tending toward more data sharing even if it was not required. As NIH implements more requirements, researchers are implementing those requirements across their research activities so NIH appears to be the biggest driver here.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
    a. " V Foundation for Cancer Research- Data Sharing Policy Final version, eff. July 2020
    Rationale for a Data Sharing Policy

    The primary driver for funding biomedical research in cancer is the translation of that research into improved cancer outcomes, higher quality of life, and reduced risk of cancer for all Americans. Disseminating the results of research through publication and the release of primary data is the best way to maximize the contribution of research, provide independent validation and ensure the reproducibility of the science supported by the V Foundation.

    1. The V Foundation will adopt a data-sharing policy to promote transparency for research grants which produce research data. Applicable de-identified research data will include:
        a. All "omics" data (for example, genomics, proteomics, metabolomics etc.), such as sequencing data with clinical annotations released through the Genomic Data

Commons, the Clinical Proteomic Tumor Analysis Consortium, the Human Tumor Atlas Network, and the Cancer Immunology Network. The formats, best practices, and protocols for data generation and submission are being developed through the NCI Cancer Research Data Commons. See https://datascience.cancer.gov/data-commons, https://datascience.cancer.gov/news-events/blog/towards-cancer-research-data-commons, https://cancerres.aacrjournals.org/content/77/21/e15

   b. Imaging datasets, such as those released through The Cancer Imaging Archive
   c. High throughput screening data, as allowable/limited by proprietary agreements.
   d. Additional research data, data about the experiments, software and algorithms, and other information needed to reproduce the experiment and results as accurately and faithfully as possible

2. As a condition of funding, we ask grantees to publicly share their de-identified data, and additional data, protocols, algorithms and code which is necessary for validation and replication underlying their articles in a FAIR-compliant data repository (see https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/), at the time of publication. To achieve these ends, we will:
   a. Require electronic copies of any research papers, that have been accepted for publication in a peer-reviewed journal, and are supported in whole or in part by The V Foundation for Cancer Research, to be made available through PubMed Central. If feasible, the primary datasets used for the paper must be released through an appropriate data sharing site, for instance the publisher's website, concurrent with publication. Requests for a period of embargo beyond 12 months may be granted on a case by case basis with prior permission from The V Foundation.
   b. Encourage authors and publishers to license research publications using the Creative Commons Attribution license (CC-BY) so they may be freely copied and re-used (for example, for text- and data-mining purposes or creating a translation);
   c. Encourage authors to use a preprint server like BioArxiv to disseminate the work of the author(s) and to make prepublication versions of the paper available to the research community as described in this NIH policy: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-050.html;
   d. Affirm the principle that it is the intrinsic merit of the work, and not the title of the journal or the publisher with which an author's work is published, that should be considered in making funding decisions.
   e. Require grantees to list their patient clinical trials in available registries. (e.g. ClinicalTrials.gov for human trials and AVMA Animal Health Studies Database for veterinary clinical trials).
3. The V Foundation will require grantees to explain how they plan to collect, annotate and share their primary dataset. Specifically, we:
   a. Expect our researchers to maximize the availability of research data, code, software and materials, and protocols. The data underpinning research papers should be made freely and publicly available in an established, open repository at the time of publication as well as any original code and software that is required to view datasets or to replicate analyses;

b. Require anyone applying for V Foundation research funding to include their approach to data sharing in the data management plan. All research applications submitted to the V Foundation will include a plan for how applicants will prepare and share data. The V Foundation will include a template of required questions to be addressed including:
    i. What data will be produced in the course of research?
    ii. Where will it be shared, and who will be in charge of preparing the data to share and ensuring confidentiality of subjects?
    iii. Which metadata will be released (i.e. information about the variables and also key information about the study)?
    iv. What data standards will be used, and why were they chosen?
    v. What are the expected costs of preparing and sharing data?
c. Consider whether researchers have managed and shared their research outputs in line with our requirements, as a critical part of the end of grant reporting process.

4. The V Foundation will commit to financially supporting data preparation activities. All V Foundation research grant applications will be expected to include, within the grant budget, specific details regarding how data will be collected, curated and stored and made available to the research community for at least five years as well as costs associated with open publication, if any.
5. The V Foundation will help to promote ways to reward grantees professionally for sharing the results of their research. The V Foundation recognizes the importance of sharing research resources such as plasmids, animals, cells, and data. Reviewers will be instructed to consider the grantees resource sharing activities as part of the favorable criteria for recommending funding, along with more traditional outputs such as published manuscripts.
6. Exceptions to this policy: Mission grants which are not research projects and are not expected to generate data are exempted from this policy."

4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
    a. We have a policy but there is a significant barrier to rolling it out to the entire grants portfolio in that the scientific advisers for the foundation are not convinced this is the best approach and may hurt the professional development of the up and coming leaders in the field.
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
    a. N/A
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
    a. Specific grants
7. If just some, to which grants/programs do they apply and why?
    a. The policy has been required for only one program which supports research grants to the Comparative Oncology Research Consortium which is a limited group of multiple institutions that engage in collaborative team research between institutions. Due to the collaborative nature of this group, a shared data base seemed possible and was perceived as an opportunity to pilot the policy.
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
    a. The established academic thought leaders that advise the foundation are not convinced it is useful or would result in academic/career harm to the up and coming thought leaders.

9. Do you specifically require or encourage your awardees to share data?
    a. Encourage
10. Comments
    a. Most of the policy uses the words encourage. We do require that they list clinical trials on clinical trials.org but I believe that is already an NIH requirement. We also require that published papers be made available through Pub Med Central. Generally, we didn't require anything new that wasn't already required, except for the requirement that they provide a data sharing plan (enforced only for applications received from CORC).
11. Do you recommend or require specific data repositories? If yes, which one(s)
    a. We encourage the use of established repositories to the extent possible. "2. As a condition of funding, we ask grantees to publicly share their de-identified data, and additional data, protocols, algorithms and code which is necessary for validation and replication underlying their articles in a FAIR-compliant data repository (see https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792175/), at the time of publication."
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
    a. As part of the application, part of the terms and conditions and part of the progress report.
13. Comments
    a. Data Sharing plan requested as a survey question in applications and progress reports throughout the research grant portfolio. IF data sharing is required as a condition of the award, it is attached to the grant contract as an addendum (to date only applies to one program).
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
    a. Usually data sharing required at time of publication or within a year after, with permission.
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
    a. No
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
    a. No
17. Comments about training material
    a. Re: Q 15- We did ask the review committee to consider the strength of the data sharing plan as part of the review process for CORC applications only. It is not part of the routine evaluation and we have not provided training materials for reviewers, except access to the written Data Sharing Policy.
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
    a. Yes
19. Comments
    a. N/A
20. Do you provide any supplemental financial support to facilitate data sharing?
    a. No
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
    a. No
22. What are the consequences for noncompliance with your policy?
    a. Currently, there are none, since most of our policy statements encourage but do not require compliance with the policy statements.

## The Vallee Foundation

**Primary Survey Contact**: Alexa Mason, Executive Director
**Email**: amason@thevalleefoundation.org

*Note: Questions 2-22 are No or N/A.*

1. Do you have a data sharing policy (formal or informal) for your awardees?
    a. No
2. If you have informal practices that your organization implements outside of a formal policy but in the spirit of data-sharing describe them here.
3. If you have a policy or guidelines, please paste in the relevant text, or include a link here. Include any relevant literature including FAQ's, T&C's, etc.
4. If you don't have a policy yet, or are still thinking about it, what are the challenges to developing and implementing one?
5. If you are planning to implement a data sharing policy when do you anticipate implementing it?
6. Does your data sharing policy apply to your entire grant portfolio or just specific grant types?
7. If just some, to which grants/programs do they apply and why?
8. What barriers have you encountered in implementing data sharing in your grant portfolio?
9. Do you specifically require or encourage your awardees to share data?
10. Comments
11. Do you recommend or require specific data repositories? If yes, which one(s)
12. As part of the application? As part of the Terms and Conditions of the award? As part of the progress report?
13. Comments
14. If you require or recommend data sharing, what are the timing guidelines you set for your grantees?
15. Do you factor the strength of a researcher's data sharing plan in grant scoring/your award decision?
16. Does your grant review committee receive training on or guidance around what a strong data-sharing plan looks like?
17. Comments about training material
18. Are costs associated with data management and data sharing allowable under the budget for the proposed project?
19. Comments
20. Do you provide any supplemental financial support to facilitate data sharing?
21. Do you follow up with awardees to ensure they are sharing data? If yes, how?
22. What are the consequences for noncompliance with your policy?

## Recommendations from HRA Members

The last question in the survey asked respondents for recommendations and suggestions for how HRA can assist member institutions in developing or implementing data sharing policies. The bulleted list below is the collective set of responses. *Not all respondents answered this question.*

- Providing templates, best practices, enforcement
- Collecting and compiling data/resource sharing policies from members to share with all HRA members, and from this, consider developing a basic template with the commonalities between

policies so that organizations new to the development of a sharing policy have a place to start. A similar document could be developed for how funders are implementing and monitoring compliance to policies.

- Learning from orgs that required data sharing - what exactly their policies are, how they implement, what exceptions they allow (if any), and how much staff time it takes to implement the policies, would be really helpful as we consider requiring data sharing.
- Hold regular webinars on how to create/enforce policies, have members with developed policies share the documentation they used to create their policy.
- Template language would be helpful. Ways to navigate the aforementioned inertias would be helpful. Understanding how major funders (non-profits & governmental) approach this would also bolster our case with recalcitrant researchers.
- How are other HRA members ensuring compliance with this policy? How are other members judging the quality of the data that is submitted under their policies.
- By helping to facilitate information sharing from those institutions that have already started implementing the process
- If we have samples or a template, that would be great.
- We have DUA and MTDA policies that we can share.
- Continue doing what you are you are doing, offer information, stimulate discussions via webinars, sub-groups, sessions in member meetings
- Keeping a bank of data sharing policies is helpful to see what other organizations have implemented and to know who to reach to for questions and lessons learned.
- Share policies and materials such as the relevant section of application forms, report forms, and terms of award, as well as successful strategies to enforce policies without overburdening staff or researchers
- It will be great to see how others implement and have structured their policies.
- As we move towards a more structured policy would love to hear about member experiences with different open data platforms and advice on whether choosing one or allowing multiple platforms is the best option.
- I found this to be really helpful and used it as a guide to shape our policies: Open Research Funders Group How Open Is It? https://www.orfg.org/policy-development-guide
- Yes, more information about how members have put these policies into practice, lessons learned from those that have been doing it a long time, success stories from data that was actually shared, and maybe a panel of members who have successfully implemented these policies.
- This might exist in the resources and I just haven't noticed, but a concise, streamlined paragraph or two to include in award agreements (that institutions have also approved, to limit the back-and-forth) to ensure comprehensive data sharing practices. Is it just putting up data on a public access, post-print server liked PubMed, or is it also requiring them to put it on a pre-print server, or is it more than that? What considerations around IP and privacy do we as funders need to be mindful of?
- Continue to share best practices.
- It would be useful for members considering this approach to have information from other members about barriers they encountered and how they overcame barriers, other barriers that

persist in implementation and any templates related to the policy that could be used to create their own policies. It would be also good to have feedback from those who have active policies about any benefits to the Foundation or to their grantees that they gained from adopting data-sharing policies.