Opening the data drawers: Request for Data in Neurofibromatosis
2020 GUIDELINES

A critical challenge in neurofibromatosis (NF) research has been the scarcity of comprehensive, reusable data. The NF Open Science Initiative, powered by the Neurofibromatosis Therapeutic Acceleration Program (NTAP), Children's Tumor Foundation (CTF), and other organizations, has made great strides in fostering collaborative and open scientific research practices (https://go.nature.com/3g6Wb8D). A selection of recent examples of freely-available raw data resources that have been made available through this effort include genomic, gene expression, methylation, SNP array, drug screening, kinomics, and several other types of data:

- Pollard et al, Nat. Sci. Data, 2020 (data: https://bit.ly/NF1Biobank)
- Carrio et al, Stem Cell Rep, 2019 (data: https://bit.ly/pNFiPSCs)
- Kraniak et al, Expt. Neurology, 2017 (data: https://bit.ly/3DpNF)
- The Synodos for NF2 Consortium et al, PLOS One, 2018 (data: https://bit.ly/SynodosNF2)
- Fuse et al, Neuro-oncology, 2018 (data: https://bit.ly/NF2Drugs).

However, there is a body of untapped scientific data that, for a variety of reasons, has not been published. These reasons include null results regarding the specific questions posed in a study, data that replicates prior research, studies that were never finalized to the point of generating conclusive results, and historic expectations regarding raw data sharing, among other reasons. Consequently, these data are "lost to time" and are not available for the broader research community to explore and mine for novel insights. NTAP and CTF are seeking to identify and publicly share these "lost to time" datasets. Therefore, we are accepting applications from investigators who are interested in sharing previously unpublished but quality assured datasets to spur new secondary research. These data will be curated and support will be given for potentially publishing in a peer reviewed journal as a scientific resource as well as shared publicly with the broader research community to add to the knowledge base available via the NF Open Science Initiative. The shared datasets will be attributed to the contributor so that they can be cited when any of the data is used for secondary research.

The datasets selected through this RFA will be curated to make the data Findable, Accessible, Interoperable, and Reusable (FAIR) so that they can become a valuable resource to the contributing investigator as well as to the research community for the purpose of accelerating NF research.

Award details

- Selected contributors will receive up to **$10,000** with justification
- Award will be fulfilled on successful upload and annotation of data

Eligibility for application:

- Investigators with quality data directly relevant to in Neurofibromatosis 1, Neurofibromatosis 2, and Schwannomatosis

Opportunities to use the released data

Depending on the nature of data, the released data will be used to power one or more of the following:

1. Publishing the data in a peer-reviewed data release journal (optional, depending on the investigator's interest and discretion)

2. Sharing publicly via the NF Data Portal with proper attribution to the data contributor (DOI generated for the data can be easily cited by other publications)

3. A hackathon or Challenge to engage the broader scientific community for exploring the data to answer specific questions. (If selected for a DREAM Challenge, the outcomes can be published with the scientific lead)

Eligible data types

Only raw data files will be accepted for release.

Data from Cutaneous Neurofibroma, Plexiform Neurofibroma, Neurofibromatosis type 2, and Schwannomatosis are of particular interest but other manifestations of Neurofibromatosis type 1 will also be considered.

The RFA seeks data that fall under one or more of the following categories:

Genomics:
Whole-genome sequencing, RNA sequencing, Whole-exome sequencing, methylation data from tumor samples or cell line models

Imaging:
Whole-body MRI, fMRI, anatomical MRI, FDG-PET, automated microscopy datasets from tumor samples or cell line models

Drug Screening:
High-throughput in vitro dose-response data from cell line models

Other datasets:
Please submit a presubmission inquiry to nfosi-datadrawer@sagebionetworks.org with a statement describing the potential value for reuse.

Note: If an investigator has tissue samples that they would like to use for future data generation, please submit a presubmission inquiry to nfosi-datadrawer@sagebionetworks.org with a statement describing the samples to explore longer-term funding opportunities.

Selection criteria

The awardees will be selected based on the following criteria:

- The value proposition of data (eg. data fills in gaps in existing data, data shows null results for a certain hypothesis)
- Evaluation of reason for not publishing the data
- Evaluation of the data sharing plan
- Evaluation of quality control/quality assurance statements

Data contribution: What's involved?

NF-OSI Data Upload and Curation: What's involved?

Application process

The application process has been simplified through the use of a straightforward google form. The application can be accessed here.

TERMS OF AWARD

I. Applicant Notification

Applicants will be notified by email as soon as possible as to the outcome of the review. The target timeframe being about 2-4 weeks after the end of the submission period. All applicants, both funded and not funded, will be provided with a Statement Summary of feedback highlighting the key comments from the review.

II. Award Activation and Payment

The award will be paid in one single payment as soon as the data is uploaded and all the requirements (metadata, note, etc) are fulfilled. Data will be viewable on nfdataportal.org. It is anticipated that the data upload will be completed within 6 months from award assignment.

III. Award Cancellation or Early Termination

Failure to upload, annotate and publicly share data will result in the cancellation of the award.

IV. Commitment to data sharing

As a condition of their award, grantees (data contributors) are expected to upload the proposed dataset into their designated Synapse NF-OSI project space (Synapse Project) as soon as possible (within 6 months of award notification). For this award, there will be no embargo and data will be immediately available to the public. The data will be made publicly available to the broader research community through the Synapse platform and findable on the NF Data Portal. If the investigator is legally restricted from openly sharing the data (e.g. a patient informed consent form does not allow to openly share the data, but instead requires authorization), the investigator must specify the conditions to make them available to the research community in writing to the funder upon application. For this specific RFA, if conditions will not allow data to be openly shared, the funder (CTF/NTAP) will reserve the right to cancel the award.

V. Data Sharing Plan

Although a Data Sharing Plan is NOT required at the application submission stage, applicants are encouraged to describe the nature of the data that wish to contribute (species, tissue type, tumor type, number of samples), and the reasons for not publishing it elsewhere, a statement of quality assurance (summarize QA/QC test results of the data) when completing the submission form.

Data contributors will provide a data sharing plan (thorough description) using the NF-OSI Data Sharing Plan Template when the grant is awarded. Data contributors should upload the raw data from their experiments to enable the data to be re-analyzed by a third party. This includes any information necessary to interpret the data, such as a description of the cohort or the model system, experimental protocols, descriptions of study instruments and equipment, and code used to generate processed data. *It is permitted, but not necessary to share aggregated/finalized results (e.g. summary statistics, figures).*

VI. Data licensing

Upon public release, non-human subjects data will be made available under CC-BY 4.0 license by default. This permits others to use the data for commercial or non-commercial use provided that users attribute the data contributor. Read more about CC-BY 4.0 here. The grantee can select a different open license, such as CC0, which is equivalent to depositing the data into the public domain.

VII. Guidelines for human data (if applicable)

Prior to submitting data, the grantee will confirm that:

- The data submission and subsequent data sharing for general research purposes are consistent with applicable national, tribal and state laws and regulations as well as institutional policies.
- An Institutional Review Board/Privacy Board or equivalent body, as applicable, has assured that submission and subsequent sharing of human data for general research purposes are consistent with the informed consent of study participants from whom the data was obtained.
- The identity of research participants will not be disclosed to NF-OSI members and/or Sage Bionetworks:
    - All submitted human data has been de-identified according to HHS 45 CFR 46.102(f) regulation and the HIPAA privacy rules so that individuals cannot be ascertained by NF-OSI members or through secondary data use.
    - The de-identified data has been assigned random, unique identifier codes that can be made visible to anyone on the web. (i.e. the identifier codes do not include any information that can be used to re-identify the study participant).
    - Sage Bionetworks will be informed by the grantee if any metadata contains potentially identifiable information.
    - The submitter will review data after submission, before it is released, to verify that no identifying information has been submitted accidentally.
    - Filenames will not contain identifiable information and can be made visible to anyone on the web.
- Sage Bionetworks will be notified if there are quality concerns about the data that has been submitted (not limited to identifiable information).
- Data quality corrections will be submitted to Sage Bionetworks as soon as possible.

VIII. ORCID ID

As a new requirement, and following a recent announcement by the NIH that will implement this requirement starting from 2020 (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-19-109.html), all applicants are asked to acquire a personal ORCID ID (https://orcid.org/) and connect it to CTF in ProposalCentral. The connection will allow an easy

transfer of information between the Foundation and the applicant record information. ORCID is a global, community-led organization that provides digital identifiers (ORCID iDs) for researchers, so that they, as well as the organizations, can benefit from being uniquely identified and reliably connected with their professional information (affiliations, awards, publications, and more). All applicants are strongly encouraged to keep their ORCID ID record up to date with all information available, especially in the Education, Funding, and Works (publications) sections. CTF is a member of the ORCID US Community, a national community of practice for ORCID adoption and integration.

IX. Extended leaves of absence

Should the awardee need to take a leave of absence for over a month (i.e. maternity/paternity leave, illness, etc.), the Foundation must be informed of the date of departure and expected date of return.

X. Publications or Exhibits of an Awardee

CTF or NTAP should be notified at the time of public disclosures by the awardee based on their work supported by a Foundation. This includes when a paper is published in a scientific or medical journal, or when a presentation (e.g. poster, slide presentation) is made before a professional scientific or medical organization. All such publications or presentations based on work supported by the funders must credit in acknowledgements support from Children's Tumor Foundation or from Neurofibromatosis Acceleration Therapy Program. The funder requests that the awardee submits to the funder an electronic (PDF or Word) copy of the paper, abstract with slides presentation, or copy of the poster materials. This material should be forwarded to the funder - if possible once accepted for publication or presentation, and certainly, immediately following publication or presentation, together with the name of the publication or the organization accepting it, and the time and place of the meeting. Information should be sent to grants@ctf.org or info@n-tap.org. This information shall be considered confidential by the funder until publicly presented or published by the awardee.

XI. Public Access Policy

The Children's Tumor Foundation ("CTF") and the Neurofibromatosis Therapeutic Acceleration Program ("NTAP") fund biomedical research in order to better understand the causes of neurofibromatosis, how to resolve its symptoms, diagnosis and treatment. 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, CTF and NTAP expect its researchers to disseminate their findings, including publishing in peer-reviewed journals. Data will be publicly shared through to the nfdataportal.org website. A Digital Object Identifier (DOI) will be assigned to the project that can be used to uniquely cite the data.