Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D1PM1R

Title: DartLab Immune Monitoring and Flow Cytometry Shared Resource Data Management Plan

Creator: Daniel Mielcarz - **ORCID:** 0000-0001-8784-8386

Affiliation: Dartmouth College (dartmouth.edu)

Principal Investigator: Daniel Mielcarz

Data Manager: Daniel Mielcarz

Project Administrator: Fred Kolling

Funder: National Cancer Institute (cancer.gov)

Funding opportunity number: PAR-17-095

Grant: https://reporter.nih.gov/project-details/9855301

Template: NIH-Default DMSP

Project abstract:

The Immune Monitoring and Flow Cytometry Shared Resource (IMFCSR) or DartLab provides services ranging from providing researcher-operated flow cytometers and cell sorters to performing in vitro immune monitoring assays . While the PI funding projects is ultimately responsible for data storage and distribution, DartLab provides support for the activities as outlined in this document.

End date: 07-31-2028

Last modified: 01-22-2024

Copyright information:

DartLab Immune Monitoring and Flow Cytometry Shared Resource Data Management Plan

Data Type

Types and amount of scientific data expected to be generated in the project: *Summarize the types and estimated amount of scientific data expected to be generated in the project.*

Describe data in general terms that address the type and amount/size of scientific data expected to be collected and used in the project (e.g., 256-channel EEG data and fMRI images from ~50 research participants). Descriptions may indicate the data modality (e.g., imaging, genomic, mobile, survey), level of aggregation (e.g., individual, aggregated, summarized), and/or the degree of data processing that has occurred (i.e., how raw or processed the data will be)

DartLab performs a wide variety of immunoassays including flow cytometry, Luminex, Seahorse, ELISA, and ELISPOT. Flow cytometry assays generate individual FCS files per sample and control that can vary in size from hundreds of KB to over 1 GB depending on the complexity of the data. FCS files contain the fluorescence intensity of each marker studied in the experiment along with sample metadata in the file header. Luminex, Seashorse, ELISA, and ELISPOT assays are recorded into Microsoft Excel files or CSV files. In addition, ELISPOT assays generate .jpg files with individual images per well.

Due to DartLab's status as a shared resource, it is difficult to predict the exact number of samples that will be collected in a given period of time.

Scientific data that will be preserved and shared, and the rationale for doing so: *Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.*

Individual users of DartLab are ultimately responsible for storage of data associated with individual projects. However, DartLab archives flow cytometry data for a period of 5 years post-acquisition on external hard drives. DartLab archives other immunoassay data indefinitely on external hard drives.

Metadata, other relevant data, and associated documentation: Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

Metadata is associated directly with the relevant data files in either headers on the raw data files or included on the output Excel spreadsheets/CSV files from our instruments. Thus relevant metadata is

always associated with the data. When users are trained, they are urged to complete all available metadata fields to ease data interpretation and reusability.

Related Tools, Software and/or Code

State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

Flow cytometry FCS files require specialized software to interpret the data. There are several commercial vendors of this software (DartLab typically uses FlowJo from BD Biosciences and maintains a 50-seat site license for Dartmouth users) but free and open-source software is also available through Bioconductor.

All other data can be analyzed via Microsoft Excel or other software capable of interpreting CSV files.

Standards

State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources, and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist

All data generated in DartLab uses file formats established by instrument vendors. Flow cytometry data conforms to the FCS 3.1 file format described here:

https://onlinelibrary.wiley.com/doi/10.1002/cyto.a.20825

Other instrumentation generates data in the CSV format that can be interpreted by a myriad of data analysis tools.

Data Preservation, Access, and Associated Timelines

Repository where scientific data and metadata will be archived: Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived.

DartLab archives all flow cytometry data for at least 5 years on external hard drives. Users will be urged to upload flow cytometry data to Flow Repository when required by their grant-making organization or publisher:

https://flowrepository.org

DartLab archives all other immunoassay data indefinitely, and will urge users to upload it to an appropriate repository of their choice when necessary.

How scientific data will be findable and identifiable: Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

When shared to a repository as necessary, data will be findable and identifiable per the repositories' functionality.

When located in DartLab's archive, data will be organized by the instrument on which it was acquired, then by date, then by PI name.

When and how long the scientific data will be made available: Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

Data are available to the sponsoring PI immediately after acquisition. Flow cytometry data will be archived in a read-only archive for 5 years after acquisition. Other immunoassay data will be stored indefinitely.

Access, Distribution, or Reuse Considerations

Factors affecting subsequent access, distribution, or reuse of scientific data: NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing.

There are no anticipated factors or limitations that will affect the access, distribution or reuse of the

scientific data generated by the proposal.

Whether access to scientific data will be controlled: State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

Protections for privacy, rights, and confidentiality of human research participants:
If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

In order to ensure participant consent for data sharing, IRB paperwork and informed consent documents will include language describing plans for data management and sharing of data, describing the motivation for sharing, and explaining that personal identifying information will be removed.

To protect participant privacy and confidentiality, shared data will be completely de-identified. When samples are provided to DartLab, sponsoring PIs must either de-identify samples or agree to de-identification by DartLab, where a code number will be assigned.

Oversight of Data Management and Sharing

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

Daniel Mielcarz, PhD serves as the director of DartLab and is ultimately responsible for the management and sharing of data generated within the facility. DartLab staff perform day-to-day management tasks including data generation and migration of data to local archives.

Planned Research Outputs

Dataset - "FCS Files"

FCS files generated by our flow cytometers.

Dataset - "Microsoft Excel Files"

Data from our Bio-plex 200 Luminex instrument.

Dataset - "CSV Files"

CSV files from ELISA, ELISPOT, and Seahorse experiments.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
FCS Files	Dataset	Unspecified	Open	None specified		None specified	None specified	No	No
Microsoft Excel Files	Dataset	Unspecified	Open	None specified		None specified	None specified	No	No
CSV Files	Dataset	Unspecified	Open	None specified		None specified	None specified	No	No