
DHVI - Duke BIA Core Facility

A Data Management Plan created using DMPTool

Creators: Brian Watts, First Name Surname

Affiliation: Duke University

Funder: Various Sponsors/Cost Recovery Cores

Template: Duke University Data Management Plan Template and Guidance

Last modified: 04-06-2020

DHVI - Duke BIA Core Facility - Roles and Responsibilities

Roles and Responsibilities

Who will be implementing data management procedures and what are their individual responsibilities?

All students, postdoctoral fellows, research staff, research trainees, visiting scholars, faculty, and principal investigator(s) should follow the guidelines within this data management plan as contributing members of their assigned role in the Biomolecular Interaction Analysis (BIA) Core Facility in the Duke Human Vaccine Institute.

- **Data Generators** - Trained Independent Users (students, postdoctoral fellows, research staff, research trainees, project managers, and/or PIs), BIA Core Facility Operators, and BIA Core Facility Manager.
- **Data Managers** - BIA Core Facility Personnel (Operators, Manager, and/or PI).

The BIA Core Facility PI and Manager have the authority to establish this procedure and are responsible for ensuring that all Data Generators are trained on the best practices described in this document. All Data Managers are responsible for reading and understanding this document and are responsible for implementation of the procedures described. It is important to note that some projects, grants, contracts, etc. may have specific requirements for data management that are equal to or better than those described here. In these situations, the data management policies of the sponsor, regulatory agency, etc. will supersede those described here.

Prior to initiating a project, BIA Core Facility Personnel will work with the project originator to establish the objectives of the project and outline appropriate methodologies for experimental design, sample analysis, and data procurement. All BIA Core Facility Personnel are responsible for maintaining a culture that prioritizes data and research integrity, which includes voicing opinions regarding the strengths/weaknesses of research questions, the methods utilized, and the accurate interpretation and presentation of results. Additionally, all BIA Core Facility Personnel are expected to propose methods for minimizing bias in research activities and voice concerns regarding the proper collection and storage of unbiased research data. Any concerns regarding data and research integrity should be brought immediately to the attention of the BIA Core Facility PI and/or Manager.

How will new people be trained and ongoing communication be encouraged amongst all team members with respect to implementing best practices in data management?

Upon gaining access to the BIA Core Facility, whether as newly trained Independent Users or BIA Core Facility Personnel, the PI and/or manager will meet individually with any Data Generators and/or Managers to review the expectations of data and research integrity in the facility including the data management best practices that are applicable to the role of the individual in the BIA Core Facility. Data Managers will monitor data deposition as outlined in this document and will periodically perform detailed audits to ensure compliance. The Data Managers and PI will co-sign that the proper training has taken place. Documentation of training and audits will be available in the laboratory and are open to public scrutiny.

New Data Generators will be asked to review the [Duke University Library's online data management guidelines](#) and consider attending an in-person data management Responsible Conduct of Research (RCR) forum through the Library or the ASIST program.

Before departing, Data Managers should organize and index the data and research records for which they have been responsible, and discuss with the BIA Core Facility PI and/or Manager whether it is permissible to retain any copies. In addition, there should be at least one designated person that reviews these files with the soon-to-be-departing person, to ensure there is a good understanding of the organization and location of the remaining data and research records.

DHVI - Duke BIA Core Facility - Research Methods and Data Description

Research Methods and Data Description

What type of research does this plan cover? (i.e., briefly summarize your project scope or research unit's focus area that will be covered under this DMP)?

The Duke Biomolecular Interaction Analysis (BIA) Core Facility provides specialized applications and support in real-time, label-free analyses of biomolecular interactions through the Surface Plasmon Resonance (SPR) and Biolayer Interferometry (BLI) platforms. The BIA Core Facility offers both Independent Use, in which individuals are trained in the independent operation of the Biacore T200 SPR shared resource, and Core Services, in which submitted samples are completed by BIA Core Facility personnel.

The scope of this Data Management Plan applies to all data generated in the BIA Core Facility, including data directly generated from instruments, collected from visual observation from the scientist collecting the data, and all data analyses, tables, and figures resulting from interpretation of the underlying raw data.

How will the research, summarized above, be carried out? (i.e., provide a general, "big picture" overview for the methodologies in practice)

Independent Use Projects - The Independent User training process will include instruction on data management best practices based on the guidelines described here. The BIA Core Facility Manager will be available to consult with the project originator to suggest appropriate methodologies for experimental design, sample analysis, and data procurement. Outside of specific collaborations and consultations, Independent Users will be expected to independently design and execute all data collection and analysis.

Core Service Projects – Prior to initiating a project, BIA Core Facility Personnel will work with the project originator to establish the objectives of the project and outline appropriate methodologies for experimental design, sample analysis, and data procurement. Whenever possible and especially when observer bias may influence results, strategies such as replication by an alternate operator, sample ID blinding, and randomization should be utilized to facilitate the collection of unbiased results. BIA Core Facility Personnel will complete data collection and analysis according to the guidelines described here.

What types of data will your research team be generating/using including the format, size, and source?

Raw Data - Raw data refers to the direct readouts from an experiment prior to any processing, interpretation, and/or manipulation. Raw data file types can include: .blr (SPR), .frd (BLI), .dsc (DSC), .xls (Excel), .jpg (Image), .txt (Text), and/or .doc (Word). Raw data files are typically < 50 MB.

Processed/Analyzed Data - Processed/analyzed data refers to the indirect readouts from an experiment after processing, editing, analysis, and/or export. Processed/analyzed data file types can include: .bme and .ble (SPR), .fsd (BLI), .csc (DSC), .xls (Excel), .ppt (PowerPoint), .txt (Text), and/or .doc (Word). Processed/analyzed data files are typically < 100 MB.

DHVI - Duke BIA Core Facility - Storage and Organization Workflow

Storage and Organization Workflow

Where and how will the data, documentation, and resources be stored while you are actively using the data?

The primary source document for all experiments is the individual lab notebook. All scientists conducting experiments and collecting data are required to document experiments in a notebook in real time.

- **Physical Laboratory Notebook** - The physical lab notebook should be utilized for every experimental procedure each day. The notebook should be bound, with numbered and dated pages, filled out in continuous fashion without inter-current blank pages. Corrections should be lined through and not erased, and dated when the correction is made. Back dating should not occur. For those working under clinical trial or manufacturing best practices, the notebooks should be kept according to the network/project specific guidelines. Attempts should be made to support continuity by cross-referencing experiments by lab notebook and page numbers.
- **Electronic Notebook** - Electronic record keeping systems can also be utilized for wet lab data conform to Data Provenance principles and meet Best Practices. An example of this is the implementation of electronic notebook systems (a 21 CFR Part 11 compliant, fully auditable electronic notebook system that maintains raw data integrity and secure data sharing) coupled with a secure data management system (ECM, secure repository). These systems provide automated file indexing and version control, secure and centralized permanent archive, and a fully traceable audit trail for all files generated from data acquisition from an instrument to reporting data for manuscripts and clinical trials.
- **Electronic Data Storage** - All raw and analyzed data should be maintained on DHVI network servers (DHVI N:\ Drive). All files should be arranged in a format that allows for an independent third party to logically find data related to a given experiment. Data is backed up and managed by Duke Office of Information Technology (OIT).

All primary source data including laboratory notebooks, assay record sheets, electronic spreadsheets etc. are the property of the University, not the investigator or issuer of a given notebook, and as such are not to be removed from Duke property. If it is necessary to work from written primary source documents off site a photocopy of the original document must be made. All electronic data can only be accessed via secure connection to the Duke network and should not be stored on any personal devices. All laboratory books and electronic files are open to members of the originating laboratory, PI and relevant BIA Core Facility/DHVI leadership, unless restricted by regulatory bodies.

- **Independent Use Projects** - Each individual is responsible for their current working lab notebook which should be stored in their lab or at their desk/office. Independent Users are responsible for management, storage, and archiving of their data in compliance with their lab-specific data management guidelines set forth by their individual PIs or Unit/Lab managers. The BIA Core Facility will periodically back-up any original raw data, analyzed data, metadata, and documentation, that is stored in user/lab-specific folders to the DHVI network servers (DHVI N:\ Drive). However, continued storage and archiving will not be guaranteed indefinitely.
- **Core Service Projects** - Each individual is responsible for their current working lab notebook which should be stored in the lab or at their desk/office and be made available upon request. All completed lab notebooks should be indexed and stored in the lab or with the BIA Core Facility PI or Manager. Written records and lab notebooks will be kept for at least 10 years. All electronic copies of data are stored on DHVI network servers and located on lab-specific network drives, or in a defined archive location, for at least 10 years after it has been published. Data is backed up and managed by Duke Office of Information Technology (OIT).

How will the data, documentation, and resources be organized within and across file systems?

The Duke BIA Core Facility has a comprehensive folder structure on the DHVI N:\ Drive mentioned above. Folder access and privileges are set by the DHVI IT team at the written request and approval of the team/lead unit manager and PI. Instrument and project specific directories are established such that personnel will know how and where to store new data sets. All files will be named and stored in such a way that an independent third party would be able to logically find data related to a given experiment.

Raw Data - Raw data refers to the direct readouts from an experiment prior to any processing, interpretation, and/or manipulation. Raw data should be preserved for all experiments and stored on the DHVI N:\ drive. Raw data file types can include: .blr (SPR), .frd (BLI), .dsc (DSC), .xls (Excel), .jpg (Image), .txt (Text), and/or .doc (Word).

- **Independent Use Projects** - Raw data files should be named by the user with conventions such that the experiment date, operator, project, and/or lab notebook reference are indicated. All Biacore SPR data will be organized in user/lab-specific folders within a common folder (BIA Users), stored locally on the T200 control computer, and periodically backed up to the DHVI N:\ Drive by the BIA Core Facility Manager.
- **Core Service Projects** - Raw data files should be named such that the experiment date, operator, project, and/or lab notebook references are indicated. All Biacore SPR and BLI data must be stored on the N:\ drive in instrument-specific backup folders
 - Biacore 3000: N:\Investigators\Alam\Laboratory\B-BiaCore BackUp Data
 - Biacore S200: N:\Investigators\Alam\Laboratory\B-S200
 - Biacore T200: N:\Investigators\Alam\Laboratory\B-T200
 - Biacore T200 Core: N:\Investigators\Alam\Laboratory\B-T200 Core Backup
 - ForteBio Octet-Red: N:\Investigators\Alam\Laboratory\B-Octet-Red BackUp Data

Processed/Analyzed Data - Processed/analyzed data refers to the indirect readouts from an experiment after processing, editing, analysis, and/or

export. Processed/analyzed data should be preserved for all experiments and stored on the DHVI N:\ drive. Processed/analyzed data file types can include: .bme and .ble (SPR), .fsd (BLI), .csc (DSC), .xls (Excel), .ppt (PowerPoint), .txt (Text), and/or .doc (Word).

- **Independent Use Projects** - All data generated by processing/analysis of raw data should be stored and backed up using a similar methodology as described above. Filenames of processed/analyzed data should be similar to those of the raw data files from which they were generated such that the experiment date, operator, project, and/or lab notebook reference are indicated.
- **Core Service Projects** - All data generated by processing/analysis of raw data should be stored and backed up using a similar methodology as described above. Filenames of processed/analyzed data should be similar to those of the raw data files from which they were generated such that the experiment date, operator, project, and/or lab notebook references are indicated. Processed/analyzed data that has been exported to other file formats (Excel, PowerPoint, Text, etc.) should include experimental descriptions, lab notebook references, and/or descriptive filenames such that the related raw and/or processed data files can be easily identified. All data generated from any piece of equipment in the BIA Core Facility Laboratory should be saved and clearly labeled on the N:\ drive along with the processed data sets.

DHVI - Duke BIA Core Facility - Documentation and Metadata

Documentation and Metadata

What documentation or metadata will be created during the project and in what format?

Each experiment should have an initial description of the method used in sufficient detail that a third party could replicate the experiment. The initial description should take the form of a methods section entry as this level of detail will aid the subsequent construction of manuscripts. Information that is important to record includes the following: Source and catalog numbers of solutions and reagents, make and model of equipment used, software used for analysis, and observational notes of samples.

Raw data files should be named such that the experiment date, operator, project, and/or lab notebook references are indicated. Filenames of processed/analyzed data should be similar to those of the raw data files from which they were generated such that the experiment date, operator, project, and/or lab notebook references are indicated. Processed/analyzed data that has been exported to other file formats (Excel, PowerPoint, Text, etc.) should include experimental descriptions, lab notebook references, and/or descriptive filenames such that the related raw and/or processed data files can be easily identified.

When appropriate and allowed, administrative, study-level, and file-level information will be documented and stored with associated data sets, analyses, and charts/tables/figures.

- **Administrative Documentation** - Explains the provenance of the data. It includes information on the authors, grant funding, roles and responsibilities of the research team, storage locations, file organization, and naming conventions.
- **Study-Level Documentation** - Includes information on your research methods including data collection methods, processing and analysis decisions to ensure accuracy, completeness, and authenticity of research methods and results. Examples of this type of documentation include calculated variables, weighting, and other processes that have been applied to the data in the transformation from raw to analyzed data to reported results.
- **File-Level Documentation** - Defines the contents of the actual data files; e.g., labeling variables, values, fields, measurement units, etc.

DHVI - Duke BIA Core Facility - Data Sharing and Archiving

Data Sharing and Archiving

What data will you share at the conclusion of your project? Where will you make it available and under what conditions?

Independent Use Projects - All primary data, analyzed data, metadata, and documentation generated by an Independent User will be the sole responsibility of the Independent User. Users are responsible for management, storage, and archiving of their data in compliance with their lab-specific data management, guidelines set forth by their individual PIs or Unit/Lab managers. The BIA Core Facility will periodically back-up any original raw data, analyzed data, metadata, and documentation, that is stored in user/lab-specific folders to the DHVI network servers (DHVI N:\ Drive). However, continued storage and archiving will not be guaranteed indefinitely.

Core Service Projects - All primary raw data, processed/analyzed data, metadata, and documentation will be made available to the service requestor and their project PI via Duke email or other mechanism such as DukeBox. The BIA Core Facility will retain these original files/documents as described in Storage and Organization Workflow.

How will you prepare (i.e., “curate”) your data to support future access and use?

All forms of data including primary raw data, processed/analyzed data, metadata, and documentation will be backed-up on DHVI network servers (N:\ Drive) in either instrument-specific back-up folders and/or compiled into appropriate project folders. All files should be named and arranged in a format that allows for an independent third party to logically find data related to a given project.

All completed lab notebooks will be indexed and stored in the lab or with the PI or Unit/Lab manager. Written records and lab notebooks will be kept for at least 10 years. All electronic copies of data are stored on DHVI network servers and located on lab specific network drives, or in a defined archive location, for at least 10 years after it has been published. Data is backed up and managed by Duke Office of Information Technology (OIT).

When applicable, discipline-specific metadata or other forms of end-user documentation (i.e., README files, data dictionaries, instruments, etc.) will be prepared for future use. When applicable, a de-identification plan will be developed to ensure ethical data sharing practices that are consistent with relevant regulations such as the [HIPAA Privacy Rule for Protected Health Information](#).

Where will you archive your data at the conclusion of the project to ensure collaborators and Duke University research stakeholders can gain access?

If data are NOT deposited in a permanent repository or archive for broader sharing as discussed above, then data will continue to be secured and accessible during the established retention period stipulated by institutional policies and protocols or external sponsor requirements using DHVI network servers and/or cloud options approved by Duke (e.g., DukeBox, LabArchives, and OSF).

If appropriate, describe what data will be destroyed to comply with legal or policy requirements and how you will dispose of the data?

We do not anticipate that any data will be destroyed. If required to destroy data, then we will work with DHVI IT to do so consistent with Duke policies.