SysBioCube Data Sharing and Publication Policy

Contents

Introduction

SysBioCube is an integrated data warehouse and analysis platform for diseases of military relevance developed for the US Army Medical Research and Materiel Command Systems Biology Enterprise (SBE). It brings together clinical, pathophysiological, psychological, molecular and biochemical data from animal models and human samples. SysBioCube organizes the data in a central location and provides an access portal for subsequent analysis by the SBE. It provides browsing, querying and visualization capabilities to provide better understanding of the systems biology of diseases such as PTSD, coagulopathy, inflammation and host responses to threat agents. The metadata and data are stored in an Oracle database using a modularized database schema. The web interface provides researchers with systematic information and options to interrogate the profiles of pan-omics components across different data types, experimental designs and covariates.

Data Security

SysBioCube maintains strict information technology procedures to safeguard information and data within SysBioCube and to prevent inadvertent access to it. All computers and servers are password protected so that only authorized personnel can access the datasets. Data provided to the SysBioCube repository is protected under the terms of the ABCC/FNL IT policy which is available upon request.

Login

Users need to request login information to SysBioCube by contacting the SysBioCube team at sysbiocube@mail.nih.gov. The request is processed by the SysBioCube Data Access Committee (DAC) However the final approval has to be obtained from the director of the SBE or the Principal Assistant for Research and Technology (PART).

Data Submission

All data and information will be submitted through a secure transmission/upload form available at http://SysBioCube.abcc.ncifcrf.gov.

SysBioCube accepts many different data types including:

- Raw data from imaging, biomarker (Omics), or physiologic studies.
- Normalized data from biomarker (Omics) studies.
- Analysis data from imaging, biomarker (Omics), or physiologic studies.
- Behavioral and clinical data.

Detailed documentation on each of the data types including sample data for each will be made available through the SysBioCube web interface.

Data Formats and Naming Conventions

The SysBioCube Upload and Download Standard Operating Procedure (SOP) document provides details for submitting the data files in pre-defined format with uniform and defined file naming conventions.

Data Submission Schedule

The objective of enforcing a data submission schedule is to make data available to the research community as soon as possible without compromising the ability of the research team to interpret and present their main findings.

The submission schedule for **all data** is as follows:

Data collection period	Data upload due
January 1 – April 30	May 31
May 1 – August 31	September 30
September 1 – December 31	January 30

Requirements for Data Submission

- Data submitted to SysBioCube must not include personally identifiable information (PII).
- All data collected on all human subjects and animal samples from the US Army MRMC or DoD-supported research are to be provided. These include data from control subjects.

- Individual subject-level data in addition to the summary/aggregate data are expected.
- In addition to the experimental/analysis data submissions should include:
    - The study protocol.
    - Sample information.
    - Variables measured, if any.
    - Other supporting documentation.
- All subsequent experimental and analysis data should contain the same sample identifiers as used by the original group/lab collecting the samples. This allows validation of the uploaded data files and tracking of all data related to any particular animal or human sample.

Data Quality

The quality of data within SysBioCube is crucial for ensuring its usefulness and reliability for research. The Data Provider is expected to perform the first level of quality control and certify the accuracy of the information prior to submission. Such efforts include verifying that the information received by SysBioCube is complete (i.e., not missing records intended for submission), verifying that the files contain adequate sample information and contain no personally identifiable information (PII). This will ensure that the information submitted has undergone reviews for accuracy, completeness, and availability.

Embargo Period

When SysBioCube receives Data prior to public release, SysBioCube and the data Provider will agree on a reasonable "**Embargo Period**" of up to twelve (12) months. The Embargo Period will begin on the date that Data are provided to SysBioCube and typically extends for up to twelve (12) months or to the time of acceptance of a manuscript describing the data, whichever is shorter. After the Embargo Period, SysBioCube will be permitted to post part or all of the Data in the SysBioCube.

The Embargo Period will not apply to Data already made available through publications or public disclosure.

SysBioCube Data Access

All users with login access will be able to upload data and have access to all SysBioCube functionality, including browsing, analysis and visualization, applicable to their data sets. Access to additional data in SysBioCube for research purposes will be provided through the SysBioCube Data Access Committee (DAC). Membership of the DAC will include FNL & MRMC staff with relevant expertise in areas such as the scientific disciplines, research participant protection, and privacy. The agencies anticipate that the SysBioCube DAC may be established based on programmatic areas of interest and the relevant needs for technical and ethics expertise. The SysBioCube DAC will operate according to common principles and follow similar procedures to ensure the consistency and transparency of the SysBioCube data access

process. The DAC will review the applications of each investigator requesting data and make a determination based on their affiliation with a research institution, and on the basis of the reason for the request. It is anticipated that most requests can be approved rapidly, and that only a few will require clarification. IRB compliance is the responsibility of the applicant, and requestors will be required to provide IRB numbers and expiration dates.

Access to datasets uploaded to SysBioCube will only be granted to registered users for Research and Training purposes with no restriction due to background, citizenship, affiliation or academic credentials. Investigators and institutions seeking data from the SysBioCube database will be asked to submit a written document that is signed by the investigator. Data access requests should include a short summary of the proposed research describing the intended use of the requested data and details on how the results will be shared with other SBE participants.

Investigators will agree to:

- Use the data only for the approved research;
- Protect data confidentiality;
- Follow appropriate data security protections;
- Follow all applicable laws, regulations and local institutional policies and procedures for handling SysBioCube data;
- Not attempt to identify individual participants from whom data within a dataset were obtained;
- Not sell any of the data elements from datasets obtained from the SysBioCube Database System;
- Not share with individuals other than those listed in the request any of the data elements from datasets obtained from the SysBioCube Database System;
- Agree to the list of approved research uses within the SysBioCube along with his/her name and organizational affiliation;
- Agree to report, in real time, violations of the SysBioCube policy to the DAC;
- Acknowledge the SysBioCube policy with regard to publication; and
- Provide annual progress reports on research using SysBioCube data.

The DAC will review requests to determine whether the proposed use of the dataset is scientifically and ethically appropriate. In the event that requests raise concerns related to privacy and confidentiality, risks to populations or groups, or other concerns, the DAC will consult with other experts as appropriate. It is anticipated that denials for data access will be unusual, but in circumstances where this does occur, a request to appeal the decision is allowed and will be reviewed by the director of the SBE or the Principal Assistant for Research and Technology (PART).

Publication

The DAC expects all investigators who access SysBioCube data to acknowledge the Contributing Investigator(s) who conducted the original study, the funding organization(s) that supported the work, and the SysBioCube in all resulting presentations, disclosures, or publications of the analyses.

Summary of Expectations for Data Submission and Access

The detailed expectations are enumerated in the individual sections of this policy, and summarized as follows:

Investigators submitting data are expected to:

- Obtain a login by contacting the SysBioCube team;
- Provide descriptive information about their studies;
- Provide data files which are formatted and re-named in a uniform format; and
- Provide de-identified human subject data.

Investigators requesting and receiving SysBioCube data are expected to:

- Submit a written description of the proposed research project;
- Protect data confidentiality;
- Ensure that data security measures are in place;
- Notify the Data Access Committee of policy violations; and
- Submit annual progress reports detailing significant research findings.

Inquiries

Additional information and detailed implementation guidance related to the SysBioCube can be found at https://SysBioCube.abcc.ncifcrf.gov. Specific questions about this policy should be directed to the SysBioCube Team (sysbiocube@mail.nih.gov).

Definitions

A Data "**Provider**" is the original Data producer or depositor who is permitted by his/her Institution and funder to provide the Data to SysBioCube.

A Data "**User**" is an individual who has registered with the SysBioCube and agreed to abide by all applicable SysBioCube Data Access Policies and Terms of Use.

"**SysBioCube**" is an integrated data warehouse and analysis platform for experimental data relating to diseases of military relevance developed for the US Army Medical Research and Materiel Command Systems Biology Enterprise (SBE). (https://SysBioCube.abcc.ncifcrf.gov/)

The Data Access Committee ("**DAC**") implements SysBioCube data sharing policies and handles requests for access to data. The DAC is composed of senior staff members from Army, FNLCR and NCI with relevant scientific, bioethics, and human subjects' research expertise.