**OBI Data Management Plan (DMP) - Template**

Version: 1.2 - Date: July 17, 2024

|  |
| --- |
| ***Document overview:**** This document is used to clearly outline the data to be collected, documentation and metadata created, the storage and backup of the data during studies, preservation of data in the long term, data sharing and reuse plans, responsibilities and resources for data management, and the ethics and legal compliance in place
* The DMP***must*** be shared with Ontario Brain Institute (OBI) for review and to inform of the study data plan.

***Resources:**** *Brain-CODE homepage:* [*https://www.braincode.ca/*](https://www.braincode.ca/)
* *OBI Informatics Governance Policy:* [*https://braininstitute.ca/docs/Brain-CODE-Governance-Policy-version-FINAL.pdf*](https://braininstitute.ca/docs/Brain-CODE-Governance-Policy-version-FINAL.pdf)
* *FAIR principles:* [*https://www.go-fair.org/fair-principles/*](https://www.go-fair.org/fair-principles/)
 |

|  |
| --- |
| ***General instructions:*** * *This document can be used as a template to complete your study Data Management Plan (DMP).*
* *Please create one DMP per study protocol.*
* *Please complete all sections highlighted in yellow including bullet point sections.*
* *Additional instructions are provided in the grey dashed boxes such as this box. You can delete these instruction boxes once you’ve completed your DMP.*
* *Do not alter the template text unless necessary.*
* *Please share this completed DMP with OBI prior to use for the study.*
 |

**Study name:** [ENTER STUDY NAME]

**Date:** [ENTER DATE DMP LAST UPDATED]

# **Data Collection**

## **1.1 Data to be collected**

|  |
| --- |
|  *Instructions:* * *Enter the source of the data (e.g., “participants recruited from the neurology clinic”)*
* *List all Brain-CODE Common Data Elements [CDEs] (both Core and Recommended CDEs)*
* *Using the provided table template, provide a brief description of all assessments, measures and modalities that are not Brain-CODE CDEs (e.g., other questionnaires, imaging, genomics, etc.) and why they are being collected. If possible, please provide the link to any relevant validation papers as a hyperlink in the “Measures/Variables” column.*

*Note: you can make use of the Study Protocol to insert details here.* |

**Location(s) of Data Collection**

All data collected will be monitored by onsite staff at [name of hospitals, clinics, university centres].

**Common Data Elements**

|  |
| --- |
| *Instructions:** *Please refer to the Stage 1 document #3 “OBI DQF Common Data Elements Guidelines”*
 |

In addition, we will be collecting Common Data Elements including Brain-CODE CDEs defined by OBI:

**Brain-CODE Core CDEs:**

* Brain-CODE Demographics form
* Brain-CODE Medical History forms (adult and child/adolescent versions)
* WHO-QOL-BREF (for adults) OR KINDL-R (for children and adolescents)

**Brain-CODE Recommended CDEs:**

* [Enter recommended CDEs, if any]

**Data and Data Context:**

List all measures, tasks, assays and other instruments that will be collected for this study and provide contextual details in the following table:

|  |
| --- |
| *Instructions:** *Please list all measures, tasks, assays and other instruments that will be collected for this study (excluding Brain-CODE CDEs) and provide contextual details in the following table.*
* *If multiple outcomes are derived from a single instrument/assessment (e.g., multiple hormone levels from a single blood draw), please list all outcomes in a single row.*

*Column descriptions:** *Measures/Variables: name of the instrument*
* *Units of measure: description of the numeric unit or scale used to quantify the measure/variable.*
* *Data/file formats: format or file extension of data file(s).*
* *Data coding and analysis plan: brief description of what outcome measures the data will be used to generate.*
* *Vocabularies/ontologies: list of any ontologies used.*
* *Why the measure/variable is being collected.*
 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Measures / Variables | Units of measure | Data/ files format | Data coding and analysis plan | Vocabularies/ Ontologies | Why the measure/variable is being collected |
| [Example: WHOQOL-BREF] | [26 items, 5-point scale] | [REDCap eCRF] | [Secondary endpoint measuring therapeutic efficacy] | [None] | *[used to subjectively quantify therapeutic efficacy]* |
| *[Example: wrist-worn accelerometry]* | *Gravitational units* | *.cwa (converted to .edf)* | *[activity minutes using validated cut-points]* | *[None]* | *[determine relationship between activity minutes and cognitive function]* |
| *[Example: blood draw]* | *[mcg/dL]* | *[.csv]* | *[cortisol concentration]* | *[MedDRA]* | *[determine relationship between stress and cortisol concentrations]* |
| *[Example: imaging MRI T1/T2 structural]* | *TR/TE* | *DICOM* | *[QA/QC pipelines, cortical thickness measures]* | *[OntoNeuroLog]* | *[quantify cortical thickness]* |

**Special Licensing Requirements:**

|  |
| --- |
| *Instructions:** *Please list the measures and instruments that have special licensing requirements. Specify any licensing details related to instruments and potential concerns regarding data sharing or future use of data.*
 |

* […]

## **1.2 Describe when the data will be captured (please attach a Study Visit Schedule table)**

|  |
| --- |
| *Instructions:** *Please provide a Study Visit Schedule in table form for all instruments, measures, tasks, and assays (including Brain-CODE core and recommended CDEs and all other assessments), indicating the collection timepoint(s) at which each measure will be taken.*
 |

* [e.g., Clinical data will be captured at baseline visits, 2 weeks follow-up, 4 weeks follow-up and 8 weeks follow-up.]
* [Imaging data will ...]
* [...]

**Study Visit Schedule**

|  |  |  |  |
| --- | --- | --- | --- |
| Measure | Timepoint 1[description] | Timepoint 2[description] | Etc. |
| [Example: Test Questionnaire] | [x] |  | [x] |
|  |  |  |  |

## **1.3 Describe how data will be collected, tracked and versioned**

|  |
| --- |
| *Instructions:** *Please complete the list of procedures used to collect, track and version data. From the text below, please remove any data collection method that will not be used by this study. These steps may be altered if data will be transferred to Brain-CODE rather than captured on Brain-CODE systems.*

*Do not include details related to data storage (see Section 3).*  |

The following procedures will be implemented at our study sites to track data copies and versions:

* Participant recruitment and consent will be collected and entered into the Brain-CODE Subject Registry and Subject Enrollment and Informed Consent Form on REDCap.
* Clinical assessment measures will be collected on REDCap hosted on the OBI Brain-CODE platform. Data will be monitored using the REDCap toolbox. In the case where paper forms are used due to IT system outages, upgrades, or unavailability, transcription to REDCap will be conducted as soon as possible after system recovery. The version of the clinical assessment data collected will be tracked.
* Imaging data will be collected onsite by imaging technicians. These data will then be transferred to SPReD/XNAT system on the OBI Brain-CODE platform daily. The version of the imaging data collected will be tracked.
* Molecular data will be collected onsite by molecular lab technicians. The raw molecular data will be managed at the labs, but the processed data, such as genomic alignment files and variant call files, will be transferred to LabKey on the OBI Brain-CODE platform on an ongoing basis. Labs will manage the version of raw molecular data. The versions of processed molecular data will be tracked.
* Wearables data will be collected onsite by staff. These data will then be transferred to the Brain-CODE platform file system as soon as onsite data extraction is complete.
* […]

## **1.4 Describe the data processing that will be performed on the data onsite (at the clinic or in the lab) *prior* to transfer to Brain-CODE, if any**

|  |
| --- |
| *Instructions:* * *Please provide a high-level summary of the processing that will occur prior to transfer to Brain-CODE with focus on the “*what” *and not the “*how”.
 |

* [e.g., The raw wearables time series data will be synchronized with other devices' timestamp data from the same recording session at the lab prior to transfer to Brain-CODE]
* [...]

## **1.5 Describe how the data will be transferred to Brain-CODE and at what frequency**

|  |
| --- |
| *Instructions:* * *Please provide a high-level summary describing the procedure and frequency of data upload to Brain-CODE*
 |

* [e.g., Clinical measures will be captured directly on Brain-CODE REDCap.]
* [e.g., MRI scans will be uploaded to Brain-CODE XNAT on the same day as collection.]
* [e.g., Samsung wearable accelerometer data will be transferred via file upload to the Brain-CODE file system on a monthly basis.]

**1.6 Describe any processing required by your study that will be performed on the data *after* it has been collected/transferred and curated on Brain-CODE to prepare the data for study research analysis**

|  |
| --- |
| *Instructions:** *Please provide a high-level summary of categories of outcome measures or specific outcome measures that will be calculated/derived from the data for use in research analysis.*
 |

* [e.g., MRI scans collected on Brain-CODE will be processed to compute cortical thickness from the images into quantitative values.]
* [...]

## **1.7 Clearly list the direct identifiers that are being collected as part of the study for each data modality**

|  |
| --- |
| *Note:* * *Please refer to the list of direct identifiers defined in the* [*OBI Informatics Governance Policies*](https://braininstitute.ca/docs/Brain-CODE-Governance-Policy-version-FINAL.pdf) *section 1.6.4 (pages 27-28).*
 |

* [e.g. First Name and Last Name]
* [e.g. Phone number]
* [...]

## **1.8 List all research sites and primary contact at each site that are participating in the study**

|  |
| --- |
| *Instructions:** *Please list the location and primary contact for all sites (including sites involved in participant recruitment, data collection, and data processing/curation). Contact information is not required.*
 |

* [e.g. First Name and Last Name]
* [e.g. Phone number]
* [...]

## **1.9 Describe what type of informed consent document will be used**

|  |
| --- |
| *Instructions:** *Please list all informed consent documents that will be used for the study.*
 |

* [e.g., participant form, study partner form, assent form, etc.]

# **Documentation and Metadata**

## **2.1 Research methodology**

|  |
| --- |
| *Instructions:** *Please provide a 2-to-3-paragraph summary of your research methodology, the endpoints used, and assumptions made.*
 |

Our research methodology will ...

## **2.2 Metadata standards used**

|  |
| --- |
| *Instructions:** *Please add any additional metadata specifications used.*
 |

Metadata will be created for all subjects, records, and study description information will be coded in accompanying metadata files as TXT, CSV and JSON documents. These metadata will be packaged within a standard folder and file hierarchy to facilitate analysis and collaborations.

# **Storage and Backup**

## **3.1 Collection site data storage**

|  |
| --- |
| *Instructions:* * *Please describe the onsite location for data storage and plan to back-up the data, if applicable.*
* *Please provide the estimated total dataset size that will be stored on OBI’s Brain-CODE.*
 |

 Staff will be responsible for capturing or transferring collected data on OBI’s Brain-CODE platform for long-term storage, however onsite storage will also be required for the following data:

* [e.g., Storage of clinical data onsite will be located at …]
* [e.g., Based on our target of *n* participants and estimated [number] MB of data collected per participant, approximately [number] GB of total data storage will be required on Brain-CODE.]
* [...]

## **3.2 Long-term storage and preservation**

Data will be stored on the OBI Brain-CODE platform and processed to pass quality assurance and quality control. These data will then be placed into long-term storage on Brain-CODE. Brain-CODE implements best-in-class data security and privacy practices for health data storage including personal health information (PHI). Offsite backups of data will be created on a regular basis for rapid recovery in case of primary storage failures.

Ongoing access to the study data stored on Brain-CODE will be enabled for study researchers, collaborators, and third parties that have been approved for access.

# **Sharing and Reuse**

## **4.1 Data processing for sharing**

Collected data will undergo quality assurance, quality control and processing to remove errors and format data according to OBI’s Data Quality Framework. This data curation will enable data to meet [FAIR principles](https://www.go-fair.org/fair-principles/) to render them more findable, accessible, interoperable, and reusable.

Data will initially be “minimally curated” to meet high quality standards while making the data available for sharing promptly. After data are compiled on Brain-CODE for sharing, they may undergo additional post-processing to prepare them for study research analysis.

Data will be processed and made available for sharing on Brain-CODE in an ongoing way, or as often as is feasible. This approach will further promote our contribution to the field and facilitate active collaborations.

## **4.2 Data licensing**

Data will be shared in accordance with OBI’s data governance policies and made available for access according to OBI’s data use agreement terms. These are available on the Brain-CODE portal at: <https://braininstitute.ca/docs/Brain-CODE-Governance-Policy-version-FINAL.pdf>

## **4.3 Making data findable**

The study data collection team will work with OBI to ensure that data are suitably annotated and that comprehensive metadata are captured for the study. Metadata will include study description, contributors, data provenance, data dictionaries, and quality report information. The study data and metadata will be hosted on the Brain-CODE platform which implements a number of practices to ensure data are findable and reusable, such as standardized data structure, searchable metadata, search engine friendly metadata, persistent identifiers to make data citable in publications and indexable and abide by public or controlled data access models. In addition, OBI may register the data on partner platforms that may include CONP, Borealis, FAIRsharing, Zenodo, and others.

# **Responsibilities and Resources**

## **5.1 Data management team**

|  |
| --- |
| *Instructions:** *Please describe who (job title) will be responsible for carrying out the parts of your data management plan. Include details regarding team structure and training delivered. Contact information is not required.*
 |

The team responsible for managing the data at sites include [...]

OBI is providing the Brain-CODE platform and the Centre for Analytics services as an in-kind contribution to support the data collection, processing, storage, access, sharing, analysis, and informatics training for this study.

# **Ethics and Legal Compliance**

## **6.1 Duration of data storage**

Data collected on site will be stored on site for a period of [enter period]. Once data are transferred to Brain-CODE, OBI will host the data indefinitely for the purposes of data sharing and meeting the principles of FAIR data.

## **6.2 Consent and Ethics approval**

The study will obtain informed consent from all participants. The consent materials will be developed in accordance with OBI’s standard consent language and institutional agreement with OBI. OBI will review protocols and consent forms prior to submission to research ethics boards (REBs). A REB will evaluate the appropriateness of the study protocol and consent materials in accordance with the Tri-Council Policy Statement concerning the Ethical Conduct for Research Involving Humans (TCPS-2) guidelines.