Research Data Management Policy

1.0 Purpose

The purpose of this policy is to provide a principle-based framework that ensures a Data Management Plan is established for all Research Projects; all data is managed to the highest standards at all stages of the Research Project lifecycle, and that Digital Research Data are deposited into the central CAMH Repository, in accordance with the following key objectives:

- Complies with the Tri-Agency Research Data Management Policy and the NIH Policy for Data Management and Sharing;
- Supports Canadian research excellence by promoting sound Research Data Management and Data Stewardship practices;
- Facilitates transparency and accountability;
- Complies with all applicable ethical, cultural, legal and third-party constraints and obligations;
- Adopts existing relevant industry standards and best practices applicable to publicly funded healthcare institutions as appropriate, and seeks to develop internal standards of excellence in Research Data Management, while recognizing that such standards and best practices are continuously changing.

The CAMH Repository is an institutional data repository that supports the deposit of Digital Research Data by CAMH Research Personnel. The CAMH Repository is not a repository service or platform to be used by third parties as their institutional data repository, and may not be used by third parties to satisfy any data deposit requirements of the Tri-Agency or other funders.

This policy applies to all Research Projects conducted at or under the auspices of CAMH, and outlines best data management practices at the level of the Research Project and the CAMH Repository.

This is not an open data policy. Sharing/re-use of Digital Research Data from the CAMH Repository for future unspecified secondary research, including making available any Digital Research Data stored in the CAMH Repository for future unspecified secondary research, is out of scope of this policy and covered in the BrainHealth Databank procedure (forthcoming).

2.0 Persons Affected

This policy applies to all CAMH Research Personnel.

3.0 Policy

Research Data Management (RDM) is the process which ensures that research data is responsibly and securely managed and is, where ethical, cultural, legal and commercial obligations allow, available for reuse by others. CAMH supports the FAIR (findable, accessible, interoperable, and reusable) guiding principles for RDM and Data Stewardship. At the Research Project level, best practice methodologies in RDM must be described in Data Management Plans.

4.0 Definitions

CAMH Repository: The BrainHealth Databank (BHDB) is a dual purpose platform that serves as (1) CAMH's designated digital repository for Digital Research Data which is required to be deposited as per section 6.5; and (2) a platform that supports the sharing and re-use of data for future secondary research which is out of scope of this policy.

Data Custodianship: The safe custody, transport, storage, backup and retention of Digital Research Data, including the technical environment and database structure. Data custodians support technical processes to sustain data integrity and implement safeguards.

Data Management Plan (DMP): A stand-alone, living document that states how Research Project data will be collected, documented, formatted, protected and preserved; how existing datasets will be used and what new data will be created over the course of the Research Project; whether data will be shared; and where data will be deposited. *(Tri-Agency Research Data Management Policy, 2021)*

Data Stewardship: Maintaining the quality, integrity and access arrangements of Digital Research Data and Metadata in a manner that is consistent with applicable law, institutional policy and individual permissions. The practice of data stewardship includes the formalized management and oversight of data assets in concordance with knowledgeable and appropriate use, including content and context.

Digital Research Data: De-identified, electronic raw data and electronic analyzed data and results collected, observed, created or generated in connection with a Research Project, including Metadata and code, that directly support the research conclusions in journal publications and pre-prints. Examples include but are not limited to clinical research data (transcriptions only of video and audio capture of participant interviews, survey responses, REDCap data), biological research data (DNA sequencing output, medical imaging, tissue analysis), statistical analysis of large datasets, etc. Does *not* include biological samples. *(adapted from Tri-Agency Research Data Management Policy, 2021)*

Metadata: "Data about data"; data that accompanies the research data, and defines and describes the characteristics of that research data, with the intent of making the research data findable, understandable, and reusable.

Project Lead: May also be referred to as the Principal Investigator (PI). The primary individual at CAMH who is responsible for the leadership of the Research Project and is the primary individual who plans, designs and leads the team that carries out activities to complete the Research Project. At CAMH, only one (1) person can be identified as the Project Lead on any given Research Project. Should the Research Project be regulated under Health Canada, a qualified investigator (QI), based on Health Canada's definition, must also be identified if the Project Lead is not eligible to be the QI.

Research Data Management (RDM): The processes applied throughout the lifecycle of a research project to guide the collection, documentation, storage, sharing and preservation of research data.

Research Personnel: All CAMH employees (including but not limited to research coordinators, research analysts, research assistants, research associates, research methods specialists), as well as researchers/investigators, principal investigators, qualified investigators, scientific appointees, clinicians, contractors, physicians, trainees, volunteers, and affiliates who are involved in research conducted at or under the auspices of CAMH.

Research Project: Also referred to as a research study. This is the research that is proposed with an intended inquiry, examination, defined question or hypothesis. Research projects have defined start and end dates, deliverables and in most cases budgets.

5.0 Responsibilities

5.1 Krembil Centre for Neuroinformatics (KCNI), Operations Director & Scientific Director

5.1.1 Responsible for Data Custodianship and Data Stewardship for Digital Research Data deposited to and maintained within the CAMH Repository;

5.1.2 Provide or support access to repository services or other platforms that securely preserve, curate and provide appropriate access to Digital Research data;

5.1.3 Support the deposit of Digital Research Data from each Research Project as described in section 6.5;

5.1.4 Ensure that deposited Digital Research Data is maintained in an accessible format, and archived/destroyed in accordance with required retention periods;

5.1.5 Develop, deliver and/or support access to formal training and education in RDM practices;

5.1.6 Support researchers in their efforts to establish and implement Research Data Management practices in accordance with ethical, cultural, legal and commercial obligations, as well as external grant agency requirements;

5.1.7 Provide researchers with data management guidance and advice, where appropriate; and

5.1.8 Recognize that data created in the context of research by and with First Nations, Métis, and Inuit communities, collectives and organizations will be managed according to principles developed and approved by those communities, collectives and organizations, and in partnership with them.

5.2 Project Lead

5.2.1 Using the CAMH Data Management Plan (DMP) template, develop a DMP for all Research Projects sponsored/lead by CAMH;
- For research conducted by and with First Nations, Métis and Inuit communities, collectives and organizations, DMPs must be co-developed with these communities, collectives and organizations, in accordance with RDM principles or DMP formats that they accept. DMPs in the context of research by and with First Nations, Métis and Inuit communities, collectives and organizations should recognize Indigenous data sovereignty and include options for renegotiation of the DMP.

5.2.2 Comply with the NIH Policy for Data Management and Sharing for research, funded or conducted in whole or in part by NIH, that results in the generation of scientific data;

5.2.3 Responsible for Data Custodianship and Data Stewardship for Research Project data;

5.2.4 Ensure that Research Project data is collected, documented, formatted, protected and preserved as per the finalized DMP, and as per the requirements of any applicable institutional or funding agency policies and professional or disciplinary standards;

5.2.5   For multi-centre Research Projects where CAMH is the lead site or sponsor, ensure that all participating sites implement and adhere to the study-specific DMP;

5.2.6   Determine what counts as relevant Digital Research Data to be deposited into a digital repository, and which data should be preserved;

5.2.7   Deposit into the CAMH Repository Digital Research Data that directly support the research conclusions in journal publications and pre-prints that arise from CAMH Research Projects, as described in section 6.5;

5.2.8   Ensure that Research Project data is maintained in an accessible format, archived in accordance with required retention periods, and destroyed after the required retention period;

5.2.9   Ensure that any uses of Research Project data are consistent with governing funding agreements, contractual obligations, cultural, legal, ethical and regulatory requirements (including consent provisions); and

5.2.10 Ensure succession planning and appropriate handover of the above responsibilities upon ending duties on a Research Project and/or departing from CAMH.

5.3 Research Operations, Services & Support (ROSS)

5.3.1   Create an institutional RDM strategy, notify the Tri-Agency when it has been completed, and make the strategy publicly available on the CAMH website with contact information to which inquiries about the strategy can be directed. At CAMH, this policy serves as the strategy;

5.3.2   Develop and maintain a standardized institutional DMP template;

5.3.3   Engage CAMH legal services to facilitate the negotiation and execution of agreements setting out rights and responsibilities related to Research Project data, where applicable; and

5.3.4   In collaboration with KCNI, all responsibilities listed in sections 5.1.5 to 5.1.8 above.

6.0 Procedure

6.1 Digital Research Data Quality and Standards

6.1.1   All Research Project data must be collected and maintained as per research SOP *AR 1.6 – Research Record Management and Retention* (and *HSR 201 – Documentation Requirements in Human Participant Research* for human participant research);

6.1.2   Digital Research Data and Data Stewardship must be responsibly and securely managed using the [FAIR (Findable, Accessible, Interoperable, and Reusable) principles](#);

6.1.3   Where possible, Research Personnel should incorporate institutional common data elements and standardized electronic data capture across data modalities (i.e. the different types of data collected) to enable the potential for data harmonization across Research Projects;

6.1.4   Metadata and Discoverability
- To enable effective discoverability, understandability, and reusability, Metadata shall be recorded and made openly available in an internationally recognized standard whenever possible;
- Metadata standards are diverse and vary across disciplines, but should generally state who created the Digital Research Data and when, and include

information about how the data were created, their quality, accuracy and precision, as well as other features necessary to the intentions of Metadata;

- Minimum Metadata are required to populate the [Data Tags Suite 'DATS' model](#) for basic organization, findability and uniqueness;
- Published results should include information on how to access the Digital Research Data on which the results are based;
- If the Digital Research Data cannot be, or is not yet available within the CAMH Repository, the Metadata may be published in order to communicate the existence of the data;
- Recommended Metadata standards are available as options for Research Personnel to implement.

6.2 Data Custodianship and Stewardship

6.2.1 Custodianship
- CAMH is responsible for Data Custodianship of all CAMH Research Project data. CAMH delegates Data Custodianship duties as follows:
  - *Research Project data:* Data Custodianship oversight is provided by the Project Lead for Research Project data;
  - *PET and MRI images on BHIC servers:* Data Custodianship is provided by the CAMH BrainHealth Imaging Centre (BHIC);
  - *Digital Research Data deposited to the CAMH Repository:* Data Custodianship oversight is provided by KCNI. KCNI also manages the infrastructure and safeguards for Digital Research Data storage, management, back-up, and archiving, including the research storage system and secure research databases.

6.2.2 Stewardship
- *Research Project data:* Data Stewardship is the responsibility of the Project Lead.
- *Digital Research Data deposited to the CAMH Repository:* Data Stewardship is the responsibility of KCNI.

6.3 Ownership and Transfer of Research Project data

6.3.1 Ownership
- Subject to any ethical, cultural, legal or commercial obligations, all Research Project data collected at or under the auspices of CAMH is the property of CAMH. Any PHI/PI collected and used for research by the Project Lead remains the responsibility of CAMH as the Health Information Custodian (HIC), as defined by PHIPA.
  - Examples of ethical, cultural, legal or commercial obligations include, but are not limited to:
    - Considerations of Indigenous data sovereignty
    - Contractual obligations
    - Legislative requirements
    - Requirements of funders
    - Regulatory requirements (e.g. ICH E6 (R2) Good Clinical Practice (GCP))

6.3.2 Transfer (externally) of Research Project data

- Agreements should be in place stipulating the rights and responsibilities with respect to the transfer of Research Project data and/or PI/PHI *outside* of CAMH. Project Leads who wish to transfer data outside of CAMH must submit a Research Legal Services Intake Form to Legal Services;
- Project Leads must also submit a written request to their scientific/Centre director and ROSS
  - o If the scientific/Centre director supports (in writing) the transfer, the ROSS delegate will consult the Research Ethics Board, Privacy & Information Office (IPO) and Research Legal Services (as applicable) to review the request for feasibility and to specify the terms and conditions for any such transfer;
  - o Via execution of the fully signed legal agreement, ROSS shall authorize the release of a *copy* (i.e. not original documentation) of any Research Project data and/or PI/PHI only upon satisfaction that the transfer is in compliance with all applicable privacy laws (and participant consent, as applicable);
- All transfers are prohibited without a fully signed and executed legal agreement; and
- The above steps are not required for de-identified BrainHealth Databank datasets generated by KCNI for future unspecified secondary use for research (refer the BrainHealth Databank procedure (forthcoming)).

6.4 Data Management Plan (DMP)

6.4.1 Data management plans (DMPs) are intended as a method of improving Research Data Management practices. A well-written DMP has the potential to effectively organize the research process, provide consistent guidelines for handling data throughout the entire Research Project lifecycle, increase efficiency, and significantly reduce the costs of data management. DMPs indicate who is responsible for managing Research Project data and outline ethical, cultural, legal and commercial constraints that the data are subject to, and methodological considerations that support or preclude data sharing. The DMP should recognize that data may be of potential long-term value, sometimes for purposes distinct from those for which the data were created;

6.4.2 One (1) DMP must be created per Research Project;

6.4.3 The lead site/institution (non-regulated Research Projects) or the regulatory sponsor (regulated Research Projects) of the Research Project is responsible for developing the DMP, and may consult participating sites in its development, as required;

6.4.4 For Research Projects with human participants, DMPs do not require REB review/approval unless requested by the REB, in which case the Project Lead must submit the current version;

6.4.5 Where a grant agency has specified the required format, scope and contents of a DMP, or the submission timing of the DMP, the grant agency's requirements will apply

- E.g. Tri-Agency requires DMPs to be submitted at the time of application for certain funding opportunities (as outlined in the call for proposals) for consideration in the adjudication process. As such, for Tri-Agency funding opportunities, the DMP must be uploaded at the time of submission of the Research Assurances and Approvals Form (RAAF);

6.4.6   For all other Research Projects sponsored/lead by CAMH and where there is no applicable grant agency requirements, DMPs must be implemented by the time of Research Project start and filed within the investigator study binder. Research Personnel must utilize the CAMH DMP template as a starting draft;

6.4.7   DMPs are controlled documents and can be modified/amended (with appropriate date/version controlling) to accommodate changes throughout the course of a Research Project;

6.4.8   Common elements of a DMP include:
- Data collection: what data will be collected or created, and how?
- Documentation and Metadata: what standards, documentation and Metadata will accompany the data? What are the project and participant naming conventions used?
- Storage and backup: how and where will the data be stored and backed up during the research? How will access and security be managed?
- Retention, preservation and disposal: What is the long-term retention/preservation plan for the data? What is the destruction plan after the long-term retention period? What repositories will data be deposited to?
- Data sharing: If applicable, how will the data be shared? Is there a need for any data sharing restrictions?
- Responsibilities and resources: Who will be responsible for data management? What resources will be required to deliver the DMP?
- Ethics and legal compliance: how will ethical issues, copyright, and intellectual property rights issues be managed?

6.4.9   The portion of the DMP that addresses preservation and sharing must be developed with regard for what is appropriate given the nature of the Research Project data and any applicable restrictions (funding agreements, contractual obligations, cultural, legal, ethical and regulatory requirements).

6.5 Data Deposit

6.5.1   Digital Research Data that directly support Research Project conclusions in journal publications and pre-prints must be deposited into the CAMH Repository
- Any ethical, cultural, legal or commercial restrictions on the future unspecified secondary use of such Digital Research Data must be identified by the Project Lead;
- KCNI supports GitLab, a standard, internal repository for code maintenance. Code related to Research Project data will be deposited to GitLab by KCNI;
- Digital Research Data from human participants that is deposited into the CAMH Repository must be de-identified/coded (at minimum);
- CIHR-funded Research Personnel: Since January 1, 2008, recipients of CIHR funding have had to comply with the limited data deposit requirements included

in the [Tri-Agency Open Access Policy on Publications](#). Research Personnel must continue to comply with these requirements, which are specific to bioinformatics, atomic, and molecular coordinate data.

6.5.2 Determining what counts as relevant Digital Research Data to be deposited, and which data should be preserved, is often highly contextual and should be guided by disciplinary norms;

6.5.3 Should they wish to do so, Project Leads may choose to deposit Digital Research Data into additional digital repositories beyond the CAMH Repository, subject to:
- ethical, cultural, legal and commercial obligations;
- CAMH policies and procedures; and
- REB review, in the case of human participant data;

The choice of depositing into additional repositories may be guided by disciplinary expectations and the Project Lead's own judgment, but in all cases the repository must ensure safe storage, preservation and curation of the Research Project data;

6.5.4 It is recommended that Digital Research Data be deposited every 6 months, however, it must be deposited into the CAMH Repository annually (at minimum) throughout the Research Project, with the final deposit made at the time of first publication;

6.5.5 Whenever possible, Digital Research Data should be linked to the publication with a persistent digital identifier (e.g. digital object identifiers (DOIs) for journal articles; ORCID IDs for researchers, etc.);

6.5.6 Where Research Project data or biological materials from human participants are being deposited into any repository for future unspecified secondary use in research:
- participant consent must be sought;
- the Project Lead, KCNI, and future researchers share the responsibility of ensuring that the terms of participant consent are respected; and
- the repository requires REB review and is subject to continuing research ethics review, in accordance with a proportionate approach to research ethics review.

Refer to [TCPS2](#) and the BrainHealth Databank procedure (forthcoming).

6.6 Data Storage, Security and Privacy

6.6.1 Research Project data must be stored securely and protected in accordance with all CAMH policies and procedures related to information security and privacy;

6.6.2 The privacy rights of research participants must be protected at all times. Research Project data may only be collected, used and/or disclosed in accordance with all applicable privacy laws, contractual obligations, participant consent, research ethics board (REB) approvals, and CAMH policies and procedures (including consent provisions).

6.7 First Nations, Métis and Inuit communities, collectives and organizations

Research Projects that involve the collection and/or use of Indigenous data or traditional

knowledge must be conducted with respect, reciprocity, and responsibility during the entire Research Project, including:

6.7.1 Acquiring familiarity with, and adhering to the specific practices and requirements of the Indigenous community and/or organization;

6.7.2 Ensuring meaningful engagement of Indigenous communities, collective and organizations; formal, proper attribution of contributed knowledge; informed consent for the use for the use of the knowledge; and contributor control of knowledge;

6.7.3 Data Management Plans
- For Research Projects conducted by and with First Nations, Métis and Inuit communities, collectives and organizations, DMPs must be co-developed with these communities, collectives and organizations, in accordance with RDM principles or DMP formats that they accept. DMPs in the context of research by and with First Nations, Métis and Inuit communities, collectives and organizations should recognize Indigenous data sovereignty and include options for renegotiation of the DMP;

6.7.4 Data Deposit
- For research conducted by and with First Nations, Métis and Inuit communities, collectives and organizations, these communities, collectives or organizations will guide and ultimately determine how the data are collected, used and preserved, and have the right to repatriate the Digital Research Data. This could result in exceptions to the data deposit requirement set out in section 6.5.

6.8 Data Retention

6.8.1 Research Project data must be retained for a time-limited retention period required as per policy *AR 1.6 – Research Record Management and Retention*, subject to compliance with any retention periods required by specific granting and/or regulatory agencies; by a publishing journal; by any agreement or contract; to protect intellectual property rights; or by CAMH in the event of a research misconduct or conflict of interest allegation, whichever is longer;

6.8.2 Subject to legal or regulatory requirements, the record retention period for Research Project data begins at the Research Project closure date, which may be:
- After final reporting to the Research Project sponsor;
- After final financial closeout of a research award;
- After publication of research results; or
- Upon Research Project closure date with the Research Ethics Board (REB) (for human participant research).

6.9 Registration and public disclosure of clinical trial results (human participant research)

6.9.1 Clinical trials sponsored/lead by CAMH must be registered in a publicly available, free to access, searchable clinical trial registry complying with the World Health Organization's international agreed standards before the first visit of the first

participant. At CAMH, *clinicaltrials.gov* must be used as the clinical trial registry for Research Projects sponsored/lead by CAMH;

6.9.2 Public disclosure of Research Project results must be done within a mandated time frame:
- publications describing clinical trial results must be open access from the date of publication;
- summary results (aggregate data that you would expect to find in a publication; not participant-level data) must be publicly available on *clinicaltrials.gov* within 12 months from the last visit of the last participant (for collection of data on the primary outcome); and
- All publications must include the *clinicaltrials.gov* registration number/trial ID (to be specified in the article summary/abstract).

6.10 Destruction

6.10.1 Upon completion of the required retention period, the destruction of Research Project data must be carried out so that sensitive, confidential and/or personal information cannot be practically read or reconstructed. The manner and time of destruction must be documented; and

6.10.2 Digital Research Data deposited to the CAMH Repository will be preserved (i.e. not destroyed) by KCNI beyond the retention period of the Research Project data if it is intended to be made available for future unspecified secondary use in research, subject to conditions set out in the BrainHealth Databank procedure (forthcoming).

6.11 Research Data Management Costs

6.11.1 It is acknowledged that the process of meeting Research Data Management requirements is likely to incur extra costs throughout the entire data lifecycle – this includes the costs of managing data during the Research Project, and the costs of providing access, preservation and sharing once the Research Project has ended;

6.11.2 Where applicable, Research Personnel should request funding from funders to cover the costs related to Research Data Management, including but not limited to:
- Data storage;
- Data transfer and access;
- Data backup;
- Data security;
- Data sharing and consent for data sharing;
- Transcription;
- Anonymization;
- Operationalization;
- Data description, documentation and Metadata;
- Data cleaning;
- Data formatting and organization;
- Digitization; and
- Data destruction.