camh | Campbell Family Mental Health Research Institute

Research Data Management Policy

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 (i.e. BrainHealth Databank), 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.

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 (DMPs).

Digital Research Data Quality and Standards

- All Research Project data must be collected and maintained as per CAMH research procedures for human participant research);
- Digital Research Data and Data Stewardship must be responsibly and securely managed using the FAIR (Findable, Accessible, Interoperable, and Reusable) principles;
- 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;
- 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.

Data Custodianship and Stewardship

- 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 BrainHealth Imaging Centre (BHIC) servers:* Data Custodianship is provided by BHIC;
    - *Digital Research Data on CAMH IT servers and/or 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.

- Stewardship
  - *Research Project data:* Data Stewardship is the responsibility of the Project Lead.
    - When working with First Nations, Inuit and Métis communities that wish to assume possession/control (in alignment with OCAP principles) of data collected within and about their community, this responsibility may be transitioned to representatives responsible for data management within the community;
  - *PET and MRI images on BrainHealth Imaging Centre (BHIC) servers:* Data Stewardship is the responsibility of BHIC;
  - *Digital Research Data on CAMH IT servers and/or deposited to the CAMH Repository:* Data Stewardship is the responsibility of KCNI.

Ownership and Transfer of Research Project data

- 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 personal health information (PHI)/personal information (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

- Transfer (externally) of Research Project data
  - Agreements must 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 to externally transfer Research Project data
    - If the scientific/Centre director supports (in writing) the transfer, the ROSS delegate will consult the Research Ethics Board, Information & Privacy 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;
    - 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 may not be required for de-identified Research Project datasets stored in the BrainHealth Databank for future unspecified secondary use for research.

Data Management Plan (DMP)

- DMPs are intended as a method of improving RDM 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;

- One (1) DMP must be created per Research Project (see CAMH Data Management Template);

- The Project Lead at 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;

- 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;

- 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 included with the funding application.

- For all other Research Projects sponsored/led 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;

- DMPs are controlled documents and can be modified/amended to accommodate changes throughout the course of a Research Project;

- 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?

- 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).

Data Deposit

- 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 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.

- 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;

- It is recommended that Digital Research Data be deposited every 6 months. It must be deposited into the CAMH Repository annually (at minimum) throughout the Research Project, and at the time of first publication (see Appendix B);

- 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.);

- 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;

- Where Research Project data from human participants are being deposited into any repository for future unspecified secondary use in research:
  - participant consent must be sought;
  - Any ethical, cultural, legal or commercial restrictions on the Digital Research Data must be identified by the Project Lead;
  - the Project Lead, KCNI (if depositing into the CAMH Repository), 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.

Data Storage, Security and Privacy

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

- 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).

camh | Campbell Family Mental Health Research Institute

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:

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

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

- For Research Projects conducted at or under the auspices of CAMH with First Nations, Métis and Inuit communities that wish to assume possession/control (in alignment with OCAP principles) over data collected within and about their community, and have the necessary infrastructure in place to do so, may assume Data Stewardship responsibilities. These responsibilities must be clearly outlined in the informed consent form and a signed agreement between CAMH and the community.

- 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;

- 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.

Data Retention

- Research Project data must be retained for a time-limited period, 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;

- 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 all human participant research).

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

- Clinical trials sponsored/led 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;

- 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).

For CAMH guidance on navigating *clinicaltrials.gov*, please refer to FAQs: Registering a Clinical Trial, and the Guidance Document: Using clinicaltrials.gov.

Public Access to Published Results of NIH-funded Research (as per NIH's Public Access Policy)

- The public must have access to published results of NIH-funded research via the NIH National Library of Medicine (NLM) PubMed Central (PMC), a free digital archive of full-text biomedical and life sciences journal literature;

- NIH-funded Project Leads must submit to PMC an electronic version of the final, peer-reviewed manuscript upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication. The final peer-reviewed manuscript is defined as the final version accepted for journal publication on or after 4/7/2008, and includes all modifications from the publishing peer review process, and all graphics and supplemental material associated with the article;

- This applies to all peer-reviewed articles resulting from research supported in whole or in part with direct costs from NIH, including research grant and career development awards, cooperative agreements, contracts, institutional and individual Ruth L. Kirschstein National Research Service Awards, SBIR/STTR awards, and NIH intramural research studies;

- Institutions and Project Leads are responsible for ensuring that any publishing or copyright agreements concerning submitted articles reserve adequate right to fully comply with this requirement;

- Applicants citing articles in NIH applications, proposals, and progress reports must include the PMCID or NIHMSID. The NIHMSID may be used to indicate compliance with NIH's Public Access

Policy in applications and progress reports for up to three months after a paper is published. After that period, a PMCID must be provided to demonstrate compliance.

Destruction

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

- 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 *GR 107 – Secondary Use of Digital Research Data and Research Samples (BHDB)*.

Research Data Management Costs

- It is acknowledged that the process of meeting RDM 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;

- Where applicable, Research Personnel should request funding from funders to cover the costs related to RDM, 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.