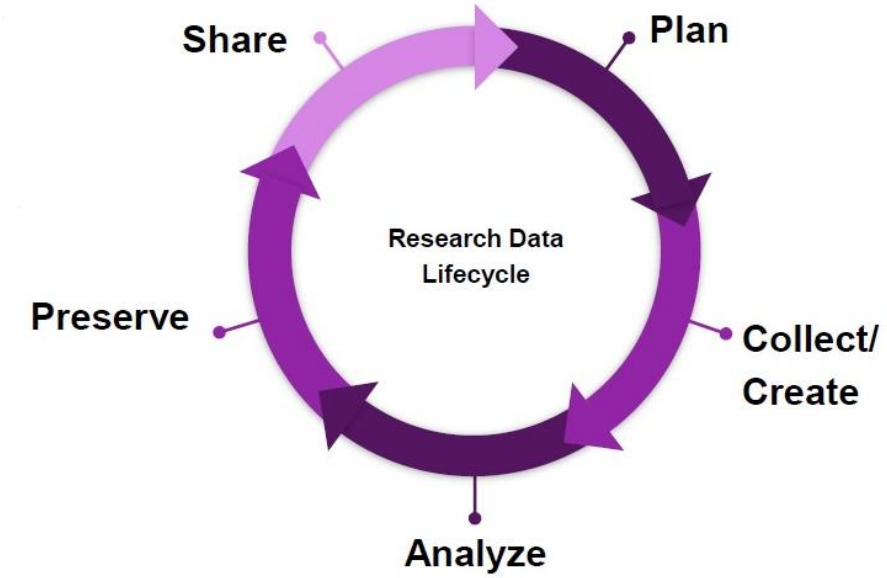
McMaster University | Family Medicine
David Braley Primary Care Research Collaborative

Title	Research Data Management
SOP Code	SOP012
Effective Date (dd/mm/yyyy)	01/09/2016
Last Updated	September 7, 2021

1.0 Purpose

This Standard Operating Procedure (SOP) is intended to guide researchers in determining necessary procedures for Research Data Management (RDM), in compliance with the Hamilton Integrated Research Ethics Board, Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans and the Personal Health Information Act. RDM refers to the security, storage, access, and preservation of data produced from a given investigation. Data management practices cover the entire lifecycle of the data, from planning and conducting the data collection and backing up data as it is created to long term security and storage of data even after the research investigation has concluded.



From: Best Practices for Managing Data Best Practices for Managing Data in your Research-McMaster RDM (https://scds.github.io/intro-rdm/watch)

2.0 Scope

This SOP is applicable to all studies undertaken within the Department of Family Medicine, McMaster University, and to the Principal Investigator/Research Associates/Research Coordinator/Research Assistant/Students/Volunteers responsible for managing study data.

3.0 Regulation of Research Data

The Tri-Agency Statement of Principles on Digital Data Management promotes excellence in data management practices, outlining expectations regarding RDM, and the responsibilities of those involved. These principles emphasize the importance for data to be stored in secure formats, and for data to be stored in a manner that enables preservation of and access to the data after the research project is completed.

In Ontario, the Personal Health Information Protection Act (PHIPA) 2004 regulates the collection, use, and disclosure of personal health information, stressing the importance for appropriate measures of data storage and security relative to the level of data sensitivity. All researchers must comply with this legislation.

4.0 Responsibilities

The Principal Investigator (PI) is responsible for ensuring that their RDM meets all requirements for data security listed in 3.0 above. Any or all parts of this procedure may be delegated to appropriately trained study team members but remains the ultimate responsibility of the PI. The Department of Family Medicine performs regular audits to verify that the proper procedures are being followed by research teams.

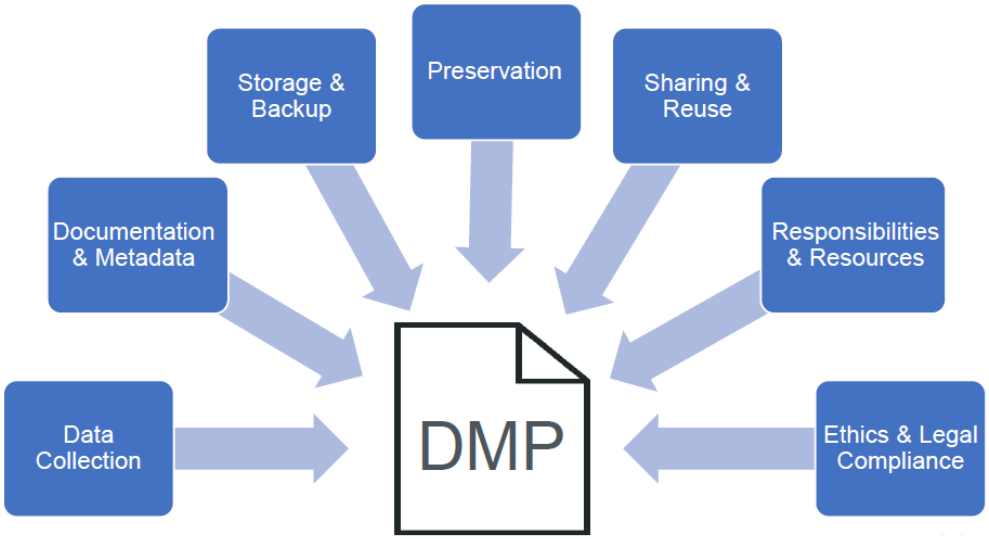
5.0 Considerations for Research Data Management

RDM is the active organization and maintenance of data throughout its lifecycle--from collection to interpretation, dissemination, long-term preservation (storage), and re-use. The application of RDM best practices improves research efficiency, enables verification of research results, and fosters innovative and interdisciplinary research through data reuse.

Why is RDM relevant? Researchers often collect personal, confidential, or sensitive data; it is important to ensure that appropriate measures are established to prevent unauthorized access to identifiable information to protect participants, researchers, and the University.

5.1 Data Management Plans

It is highly recommended that each research team completes a data management plan (DMP) for your research data in the planning phase of the research project. DMPs aid in outlining strategies and tools you will implement to effectively manage your data during the active phase of your research, and the mechanisms you will use to preserve and share your data at the end of the project. A DMP is a “living” document that can be modified throughout your project to reflect any changes that have occurred.



From: Best Practices for Managing Data Best Practices for Managing Data in your Research-McMaster RDM (https://scds.github.io/intro-rdm/watch)

The McMaster University Library works in conjunction with the Portage Network of Canada in providing a DMP Assistant Tool that researchers have free access to. More information can be found here: DMP Assistant Tool; Library Services RDM

6.0 Best Practices

Best practices for RDM are outlined below. These practices should be used in conjunction with your DMP. IT is available to meet with new project teams to advise on the best set up for projects.

For new research projects, it is best practice to create a Sharepoint site (described below) on the DFM instance of Microsoft 365 for your research data. Sharepoint is recommended over other storage options when using sensitive data. Small projects (i.e., pilot projects) have the option of using the established DFM-Research Sharepoint site. **Note: All requests to utilize the DFM instance of Microsoft 365 (SharePoint/OneDrive/Teams) must be requested through DFM IT: contact Bongi Dube dube3@mcmaster.ca.**

6.1 Digital Data Security

i. *Computer storage*

- Ensure that all computer systems storing confidential data are password protected.
- All computers should be locked when an employee steps away from their computer. To lock your computer, hold down the windows key and “L” on the keyboard at the same time. Another method is to hold Ctrl, Alt & Delete at same time and select “Lock”. You will need to re-enter your password to unlock the computer.
- Ensure computer software is up to date, including anti-virus software. Windows 10 has built-in anti-virus and anti-malware software. Ensure that Windows 10 is up to date by running Windows Update regularly. You can also set Windows Update to automatically update on a set schedule.
- Individual files containing confidential study data should be password protected, encrypted, or stored on an encrypted drive. Implement password protection and

controlled access to data files. For example, 'no access', 'read-only', 'read and write' or 'administrator-only' permission. In this manner, control access to files, folders, or entire hard drives.

- **Participant identifiable information should not be stored on the local computer's hard drive.** All sensitive information should be stored on a secure server, such as DFM's Microsoft 365 instance.
- **For files that are accessed and updated over time, it is best practice to save updated versions of the data, with previous versions kept in an archived folder on the secure server.**

ii. Cloud-based storage

- McMaster hosts cloud-based storage options which include **Microsoft 365** (SharePoint, OneDrive, and Teams) and **MacDrive**. Any data that is stored in these systems that has participant identifiable information or research data must be password protected or encrypted. Note: When utilizing cloud-based storage options, on your HiREB application it should state that the data is being stored using McMaster Microsoft 365 or MacDrive.
- **Microsoft 365**: McMaster University provides off-premises cloud-based storage via Microsoft 365. DFM has implemented a Microsoft 365 instance for research teams which includes **SharePoint**, **OneDrive**, and **Microsoft Teams**. SharePoint, OneDrive, and Teams are applications within Microsoft 365 that can be used individually but are designed to work together to form a toolkit to give your research team members integrated and flexible ways to work for their projects and tasks. The data stored using Microsoft 365 is stored on Microsoft servers located in Canada. **Note: All requests to utilize the DFM instance of Microsoft 365 (SharePoint/OneDrive/Teams) must be requested through DFM IT: contact Bongi Dube dube3@mcmaster.ca**
    a. **SharePoint** acts as a secure website for your research project that can be accessed from any device. On SharePoint, your research data is securely stored, organized, shared, and accessed by all members of your research team. **We recommend creating a SharePoint site for your research project instead of other storage options, particularly when handling sensitive information.**
    b. The system McMaster University – DFM SharePoint document repository is fully protected with McMaster / Microsoft security policies i.e. only users with Mac Id's can access the file repository SharePoint platform. Microsoft SharePoint in part of McMaster University enterprise license with Microsoft 365. Users communication with the document repository across the Internet uses SSL/TTL connections. All SSL connections are established using 2048-bit keys encryption.
    c. Similar to SharePoint, **OneDrive** allows you to store information in a secure cloud and access the information from any device. However, OneDrive is more individual-focused, with limited collaborative/sharing capabilities compared to SharePoint when used on its own. **We do not recommend** using only OneDrive for your research data, but rather SharePoint for your project storage.
    d. **Microsoft Teams** lets you quickly pull together people inside your research team, chat with team members, and securely share documents. Data can be stored on Teams, however **we recommend against using Teams for data storage**, but rather use SharePoint for your research data.
- **MacDrive**: McMaster University provides a centrally supported, cloud-based service called MacDrive. MacDrive has the advantage of being tied into the University's central authentication system. All staff and faculty can access MacDrive with their MacID. MacDrive is an individual storage space, where files can be shared with others (both internal and external to McMaster) via links. **We do not recommend MacDrive for research project file sharing**, but rather SharePoint for your research project data. MacDrive Libraries can be password protected; however, IT cannot retrieve passwords or file information, leaving project data vulnerable. MacDrive should only be used for temporary storage and should not be used for long-term storage of project data.

- It is best practice to avoid the use of commercial storage systems, like Dropbox and Google Drive when working with research data. This is particularly important when sensitive data is being used.

iii. *Network-based storage*
- DFM has network-based storage which was the main storage solution prior to implementation of Microsoft 365. While this storage solution is still available, it is being phased-out in preference of SharePoint. To access data that is stored on the DFM network, a department-provided laptop and VPN are required, and employees need to be mapped to specific research project drives.

iv. *Working Remotely with External Collaborators*
- It is advised that external collaborators, including researchers from other McMaster departments, personnel in external agencies, and unpaid volunteers not be given access to a project's Office 365 data. Files can be shared within McMaster using MacDrive (described above). Research data that is to be shared with external collaborators can also be done using MacDrive or using a Secure FTP. Note: If the information being shared is research data, ensure that a data sharing agreement (or non-disclosure) is in place.
- **Secure File Transfer Protocol (SFTP)** can be used to securely share documents remotely. Access to SFTP sites can be done using software such as WinSCP (Windows) or Fugu (Apple). This method requires both sender and receiver to have login access (username, password) to a specified folder on a McMaster-hosted server. Access to an SFTP site is controlled by Research Team Leads and those who require access need the appropriate credentials. On the DFM-Research SharePoint: go to>IT, Equipment and Software>Secure File Transfer Protocol sFTP>How to Use DFM Research Secure File Transfer Protocol.doc

6.2 Physical Data Security

i. Paper documentation of participant identifiable information (e.g., signed consent forms) will be stored in a locked filing cabinet, in a locked office. The location of the locked office at DBHSC, is 5th Floor, Room 5001/N. Filing cabinets will not be labelled with a study name and will be locked at the end of each day. Access to buildings, rooms, cabinets where data, computers, media, or hardcopy materials are held is controlled. *It is recommended that research staff log every instance of removal or access to media or hard copy material in storage rooms*.

ii. Documents with participant identifiable information will not be taken off-site unless required by the study procedures and security measures are in place to protect the participant's identity (see below).

iii. **HIREB security measures for transporting paper-based participant identifiable information**: Place information in a sealed envelope, preferably one envelope if able. Transport the documentation to the secure location, without stopping along the way, and place in locked filing cabinet, in locked office. In situations where data is collected off-site and immediate transport to office is not possible, information is to be kept in a locked briefcase or similar bag. Documents with identifiable information should not be kept in the same bag as study data documents (e.g., surveys, questionnaires).

iv. **Digitalizing physical documents**: it is best practice to digitize physical data if feasible. Transferring physical documents to a digital form requires such data to be secured under the guidelines for 6.2 Digital/Technical Security (below)

6.3 Data transfer

If electronic participant identifiable or research data needs to be physically transported between sites, the data must be password protected or encrypted, and stored on an encrypted device (USB Flash

Drive; tablet; laptop; digital recorder; cellphone; video recorder; camera). After transfer, the data must be deleted immediately.

Alternatively, the data can be sent via a secure online method:

i. **SFTP**: Participant identifiable information or research data in digital format can also be transferred via the DFM SFTP (described above). The researcher is responsible for ensuring that all uploaded files are password protected and are deleted within 24 hours so that the file cannot be accessed externally afterwards.
ii. **REDCap**: REDCap's "Send-It" feature can be used to send research data files up to 32 MB securely to team members via email. Send-It allows multiple recipients to download a file in a secure manner. Each recipient will receive an email containing a unique download URL, along with a second follow-up email with the password (for greater security) for downloading the file. The file will be stored securely and then later removed from the server after the specified expiration date. Senders require a REDCap account; however, recipients do not.

6.4 Time Requirements for Data Storage

HiREB guidance on Data Storage: It is up to the PI to determine how long the study data should be retained. Any clinical trial with Health Canada oversight must retain the study information for 25 years. Apart from that legal requirement, it is a decision of the research team.

Consider retaining the data for 2-3 years post-publication. It is worth considering not destroying data, given trends for combining and re-analyzing data sets. **Note: Identifiable data should be destroyed as soon as possible. De-identified data may be retained indefinitely**.

7.0 Administrative Procedures

A DMP is useful in outlining required administrative steps for RDM. These steps include:

- Development of organizational rules about who has access to research data.
- Imposing Non-disclosure Agreements for managers or users of confidential data. In many (but not all) cases, a Confidentiality Agreement is sufficient.
- Outline best practices (6.0 above) to the research team to ensure proper RDM procedures are being followed.

7.1 Subcontractors

Subcontractors are individuals and agencies that provide services to our research projects, such as transcription or translation.

- Subcontractors must be set up as an approved supplier with DFM. A contract will need to be in place which will include confidentiality and data transfer agreements, as applicable.
- Data transfer will adhere to the procedures outlined in Item 6 above.
- Subcontracted Employees must submit any new or altered research data to the PI on or before the last day of their contract and then destroy any research data held locally

8.0 Violation of SOP

Violation of this policy by investigators, research assistants, coordinators, associates, data managers, subcontractors and other associated research staff, students, volunteers, and affiliates constitute grounds for corrective action up to and including termination of employment, participation in research, or loss of data access privileges. Unauthorized release of confidential information may also have personal, civil and/or criminal liabilities and legal penalties attached.

9.0 Resources

For answers to Frequently Asked Questions read the MREB Data Storage and Security Tools document developed in 2018. Much information in this SOP was adopted from that document. A summary of resources is listed below:

i. MREB Data Storage and Security Tools
ii. Tri-Agency Statement of Principles on Digital Data Management
iii. Ontario Personal Health Information Protection Act, 2004
iv. Library Services RDM
v. DMP Assistant Tool
vi. McMaster Microsoft 365: OneDrive/SharePoint/Teams
vii. MacDrive
viii. Secure FTP: dfmresearch>IT, Equipment and Software>Secure File Transfer Protocol sFTP>How to Use DFM Research Secure File Transfer Protocol.doc
ix. Portage Training Guides

10.0 Revision History

SOP Code	Effective Date (dd/mm/yyyy)	Pages	Summary of Changes
SOP002-01	(15/01/2017)	2 - 3	Inclusion of procedure for research data transfer with research team members (non-employees) Item 6.0
SOP002-01	(27/02/2017)	2	Inclusion of procedure for research data transfer via newly developed sFTP (Items 4.4 and 5.0
SOP002-01	(04/12/2020)	1	Inclusion of remote data storage and access
SOP002-01	(16/01/2021)	overall	Addition of updated McMaster procedures
SOP002-01	(25/02/2021)	2-4	Revised aspects of Item 5.0
SOP002-01	(30/04/2021)	All	Updated name of SOP (formerly Data Security) Added section 5.0 on RDM Updated Sections 1-5
SOP002-01	(06/05/2021)	All	Updated Sections 5-7 Added 6.4 Added Section 9.0 Resources
SOP002-01	(20/05/2021)	All	Revised 1.0 (added in the data life cycle figure) Revised 5.1 (added in the DMP figure) Revised text of 6.2, 6.4 and 7.1 Added additional resource to 9.0
SOP002-01	(07/09/2021)	3-6	Revised section 6.0 Switched sections 6.1 and 6.2.