McMaster University | Family Medicine
David Braley Primary Care Research Collaborative

Data Management Plan *(Template)*

Disclaimer: This is a general template that includes the main parts of a Data Management Plan (DMP). Depending on the study design, you may need to add additional sections to create a comprehensive DMP.

Project Details

Name of Project:

Name of Author(s) of DMP:

Name of LPI (if different from the author):

Name of Co-Investigators (respected institutions):

Name of all the team members:

Start and End Date of Project:

Funder or Funding Status:

REB Reference #:

Data Collection

What types of data will you collect, create, and/or record or acquire *(for retrospective studies)*? Provide estimates for the amount of data you aim or anticipate to collect (e.g., 20 interviews)	*Surveys, interviews, focus groups, administrative data, personal health information*
What information sources are you accessing?	*Directly from the patient, EMR, external institution*

Describe all types of file formats that your data will be in. How will you record the study data?	*Excel, word, REDCAP, SPSS, Google Survey, Audio files, video files, transcripts, physical documents. Also include the type of files data will be converted into.*
Do these formats allow for data re-use, sharing and long-term access to the data?	*Proprietary vs non-proprietary formats. Collection formats can be different then long-term storage formats (e.g., collected in Microsoft Excel, but stored as CSV files). At what point is data being transferred between different file formats? (E.g. SPSS to Excel, Word to NVivo, Hard copy to online survey)*
How will you use to structure, name and version-control your files to help you and others better understand how your data are organized? Will you use read-me files, data dictionaries, code books, any other summary documentation?	*Best practices include for naming conventions include - Version number, No spaces or special characters, Initials if reviewed, Clear title describing the file* *For instance:* *Naming Conventions 2016-07-18 Draft V01 CI.doc* *Meeting Agenda 2016-05-31 Draft V02 CI LC.doc*
What documentation will be needed for the data to be read and interpreted correctly in the future? *Data dictionary, specific software etc.*	
How will you make sure that documentation is created or captured consistently throughout your project?	*Will regular data checks be performed and by who?*
If you are using a metadata standard and/or tools to document and describe your data, please list here. *Will you use a data management tracking sheet to track version controls and users of the documents? Where is this stored?*	*Code books, read me files, data dictionaries, protocols or protocol publications*

Data Storage *Data containing identifiable information must be securely stored at all times. Indicate the steps to protect all files containing identifiable information*	
What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.)? McMaster RDM Storage Finder Tool	*Consider ethical requirements through HiREB for length of time storing.*
How and where will your data be stored and backed up during your research project? e.g: REDCap, Sharepoint, Physical data files etc.	*For Physical Data: Describe Physical location of records.* *For electronic data: Computers/digital recorders must be housed in a locked location, provide details.*
Please indicate the computer security methods that will be used to prevent unauthorized access to identifiable data. *For HiREB Language please see last section of this document.*	
If you plan to de-identify the data, when will this be done? What will be the process of de-identification?	
How will the research team and other collaborators access, modify, and contribute data throughout the project?	*Describe how data and information will be shared among collaborators/research team members. Who will be the data steward, who will have access to what? E.g., Sharepoint, dropbox, email, secure data transfer services (e.g., macdrive), fax*
For how long will the data be stored?	*General guidance (i.e., HiREB) is to destroy identifiable data as soon as possible, but no limit on retain unidentifiable data. Consider data security of storing data and grant/journal/etc requirements.*

Will identifiable data be sent to an external institution/organization than where it was collected and/or will you be receiving data from other sites (e.g. in the case of a multi-site study where you are the coordinating site receiving data)?	
How will the security of the identifiable information be protected during transmission? (if applicable)	
Will you destroy any of the research data at the end of the research or storage period? If yes, how and when will this happen?	
Data Preservation and Sharing	
How will you ensure your data is preservation read? (If applicable)	*Preservation-friendly file formats* *Anonymization and de-identification, inclusion of supporting documentation (eg. codebooks, readme files that explain your project, how the data is set up, any special information)*
Some funders/ journals may require data sharing. What data will you share. What format will it be in?	*Raw, processed, analyzed, final, or multiple formats*
Where will you deposit your data to facilitate preservation and access at the end of your research project? (if applicable)	*Recommended data sharing sites for FM: https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data*

Have you considered what type of end-user license to include with your data?	*Creative commons, open data commons licenses*
How will you help the research community know that your data exists? -	*Knowledge translation strategies, Publications, DOI for dataset included in publications*
Responsibilities and Resources	
Who will be responsible for managing this project's data during and after the project? What data management tasks will they be responsible for?	*Principal investigator, co-investigators, research coordinators* *Also see:* https://dataoneorg.github.io/Education/bestpractices/define-roles-and
What other resources will you need to implement your data management plan?	*See RDM cost estimator:* https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management
Who will be responsible for managing data activities if substantive changes happen in the personnel overseeing the project's data?	
If your project includes sensitive data, how will you make sure it is securely managed and accessible to approved members of the project?	
Ethical and Legal Compliance	
How will you manage legal, ethical, and intellectual property issues?	*Note: Data collected within Indigenous communities has specific guidelines. Please See CARE Principles and First Nations OCAP Principles.*

If applicable, what strategies will you undertake to address secondary uses of sensitive data?	*Consider data licence* *Consider what is needed for ethics applications and consent processes to allow for secondary data use* *Also see: Sensitive Data Toolkit: Sensitive Data Expert Group. (2020) for language for informed consent*
HiREB Language for Secure Storage	
Description of REDCap	A secure data collection and management software hosted by the Department of Family Medicine at McMaster University. REDCap servers are regularly backed up.
Description and security features of SharePoint	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. We use BitLocker. The BitLocker is a full volume encryption feature included with Microsoft Windows. We use University enterprise level firewall as well, SonicWall NSA series appliance to protect our networks.