October 5, 2023

Women's Advanced Risk-Assessment in Manitoba (WARM) Hearts

A Data Management Plan created with DMP Assistant

Data Management Planning Expert Group

Information

Creators: James Ducas, Jacqueline L. Hay and Todd Duhamel

Affiliation: University of Manitoba

Funder: St. Boniface Hospital Research Foundation

Template: Portage Template

J L Hay ORCID iD: 0000-0002-4284-7141

T Duhamel ORCID iD: 0000-0001-8467-4800

Project abstract: WARM Hearts is a prospective, observational cohort study of 1000 Manitoban women over the age of 55 without a prior history of cardiovascular disease (CVD) and interested in participating in a CVD screening program. Screening data will be collected from the women in person, and then health outcomes data will be collected 5 years after screening through a health data repository. The objectives of this prospective observational cohort study are to: (1) determine if individual markers of CVD risk collected in the WARM Hearts cardiovascular screening protocol identify women 55 years of age and older who are at an elevated risk for adverse cardiovascular outcomes within 5 years after screening; (2) create a novel CVD risk score for women incorporating the best markers of CVD risk (including measures of gender) collected in the WARM Hearts study; (3) compare the predictive ability of the WARM Hearts novel risk score and the Framingham Risk Score for predicting future adverse cardiovascular outcomes in women; and (4) investigate the correlation between reproductive history and CVD risk. The study will also seek to identify novel biomarkers that may indicate a woman is at increased risk for CVD.

Identifier: WARM Hearts

Start date: 01-10-2019

End date: 20-12-2022

Last modified: 13-12-2022

URL: ClinicalTrials.gov Identifier: NCT03938155

Acronyms used in this document				
CRFs	Case Report Forms		**PWA**	Pulse Wave Analysis
.CSV	Comma Separated Value file format		**SC**	Study Coordinator
CVD	cardiovascular disease		**SOPs**	Standard Operating Procedures
HRV	Heart Rate Variability		**RAs**	Research Assistants
MCHP	Manitoba Centre for Health Policy		**REDCap**	Research Electronic Data Capture
PHIA	Personal Health Information Act		**WARM**	Women's Advanced Risk-assessment in Manitoba
PHIN	Personal Health Information Number		**.XLS**	Microsoft Excel spreadsheets
PI	Principal Investigator			

Data Collection

What types of data will you collect, create, link to, acquire and/or record?

Quantitative data

1. Participant metadata
    - Content from the consent form, including content from the consent form, such as consent to be contacted for future studies, Personal Health Information Numbers (PHINs), and contact information
2. Demographic data, such as education and gender identity
3. Data from self-completed questionnaires, including detailed indicators of reproductive history
4. Data from in-person health assessments, including:
    - Pulse wave analysis (PWA) (i.e., measures of arterial stiffness)
    - Accelerometry data (i.e., sleep, physical activity, and other related metrics)
    - Resting and exercise heart rate variability (HRV) data
5. Data extracted from blood samples, including clinical biochemistry
6. Data extracted from stool samples, including measures of gut microbiota
7. Data obtained through linkage to administrative records, including demographic and socioeconomic information from the administrative record and frequency/type of health care utilization over 5 years

A complete itemized list of the variables collected/generated for this study is tracked in a study codebook.

Biobank

- Blood and stool samples remaining after planned analyses will be preserved and retained in a frozen (-80oC) repository (biobank) for 10 years after data collection ends. Potential analyses include metabolomics, myokine/cytokines, extracellular vesicles, and DNA methylation[1].
- The results of future biobank analyses will be stored on St. Boniface Research Centre or University of Manitoba data servers and linked to the Research Electronic Data Capture (REDCap) database. REDCap is a web-based program designed to support electronic data capture for research data management. It meets the Personal Health Information

[1] This DMP includes provisions for both research data and research materials (in this case, biological samples from study participants). While the two concepts are linked, there are distinctions between them. For more detail, see 1c in the FAQ https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html

Act (PHIA)[2] database requirements for audit trails of data access. The REDCap server is housed within the Secure Research Environment of the Rady Faculty of Health Sciences, University of Manitoba. However, the REDCap database can be accessed from anywhere through the internet because the Secure Research Environment is a virtualized system maintained by the Medical Information Technology team of the Max Rady College of Medicine. A security token is required to access the REDCap database.

What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?

Open and proprietary data formats will be collected. Additionally, some data is gathered in paper, whereas others are collected via devices or samples. All data collected via paper will be transcribed to REDCap, and any proprietary format will be saved as open formats in addition. Further description, including all file formats matched to data types described earlier, is included in internal documentation available to the research team.

What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized?

Folder naming conventions

- Folders are named according to data type.
    - Examples: Blood Results, 6 Minute Walk and HRV Data, Accelerometer Data

File naming conventions

- REDCap exports are saved using the following naming convention: WARMHeartsStudy_DATA, year-month-day, time, file extension.
    - Example: WARMHeartsStudy_DATA_2020-04-28_1844.CSV
- Accelerometry, HRV, .GT3X, .AGD, and .CSV format file names follow this convention:
    - Participant ID_data capture timeframe_bodily location(where needed)_date file extension_or epoch (.AGD files) and .CSV heart rate data
        - Example: W0235_HRV_Wrist (2020-03-11).GT3X
        - Example: W0235_24Hr(2020-03-11)60sec.agd
    - Note that .AGD files are the proprietary format created by taking the .GT3X files and processing them using proprietary filters. This file type is used within ActiLife for wear time validation, data, and sleep scoring. However, it is important

[2] PHIA (https://www.gov.mb.ca/health/phia/) is provincial legislation that protects an individual's right to access information about their own health, and to expect that the privacy of their health information will be protected by those who have access to it. Each province has a similar mechanism in place to protect health information.

to note that .GT3X files can now be read and analyzed utilizing non-proprietary software utilizing the GGIR package.

- Non-clinical biochemistry data extracted from blood and stool samples have not been processed as of this date, and as such, a naming convention does not yet exist for these data.
- Due to the COVID-19-induced pauses in the research, some participants completed certain parts of the WARM Hearts study twice (i.e., If the participant had completed the Day 1 assessment, but had not yet completed the Day2 assessment prior to the research shutdown, they may have re-done some parts of the protocol). Since we want to use data from the appointments completed within the intended timeline (Day 1, followed by Day 2 approximately 7–14 days later), we must organize and name the files approximately. Data from participants that partially completed the study in 2020 or 2021 will be retained but will be labeled as the following:
- W####_COVIDRSCH_DONTUSE_(MM_DD_YEAR)

The new files containing data from the participants affected by COVID delays will be labeled using this standard. This approach will enable all of the data to be available if there is a reason to access it.

Documentation and Metadata

What documentation will be needed for the data to be read and interpreted correctly in the future?

The Call Log and Participant Log are needed to review contacts with study staff, and participant details such as appointment dates, contact information and study participation details.

- The content and functionality of the WARM Hearts REDCap database are documented in the data dictionary and study codebook. The majority of data items, including data from in-person health assessments, self-completed questionnaires, and clinical biochemistry, are referenced and stored using the indexing and naming conventions described by the project's REDCap data dictionary. The REDCap data dictionary is a .CSV file that represents the detailed structure of the database and is used to program its functionality. The REDCap study codebook is a human-readable data dictionary version that allows quick reference to the variables and functionality detailed by the data dictionary.
- Standards for data structures, metadata, and documentation for biobank analyses will need to be developed with future research protocols using these samples.
- Notes capturing miscellaneous documentation of special cases in the data are recorded on CRFs to participant files and comments, which are subsequently transcribed into REDCap.
- Biobank samples will be maintained in accordance with the Standard Operating Procedures (SOPs) developed by the Duhamel lab.

How will you make sure that documentation is created or captured consistently throughout your project?

Methodologies for data collection / generation

Data collected during in-person visits are conducted by appropriately trained and delegated research assistants (RAs).

Quality assurance

Data quality assurances are conducted on a rolling basis, with 10% of the entered participant data checked against the original data collection source for accuracy.

SOPs are developed for data collection and processing steps that are to be performed by more than one individual. These SOPs are kept on an online repository that RAs may access and review at any time.

As part of routine practice, the SC will audit a random selection of consent forms and CRFs for accuracy and completeness. Additionally, all RAs conducting data entry into REDCap will report any irregularities they find to the SC during the completion of their data entry duties. The SC will

issue query reports to the RAs, who will work to resolve identified irregularities, and then return the reports to the SC, who will retain them for documentation purposes.

If you are using a metadata standard and/or tools to document and describe your data, please list here.

Documentation and data description processes follow in-house standards, as outlined and maintained in the REDCap data dictionary and study codebook. These documents would be shared with any data made available for subsequent use to ensure readability and operability.

PWA data will use the metadata standard of Hypertension Management Software Client Server, version 5.2., and accelerometer data will use the metadata standard of ActiLife 6, version 6.13.4. Biobank analyses will use in-house standards. These data will be incorporated into existing documentation at the point of data sharing or archiving.

Dataverse metadata schemas will be applied once study data is archived. Exact formatting plans have not been established at this stage.

Storage and Backup

What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.) and the length of time you will be storing it?

It is expected that the study will generate, at most, a terabyte of digital data for one complete copy of the dataset. This digital data volume is well within the existing storage capacities of this research group. Similarly, the paper documents generated (consent forms, CRFs, PHINs) are expected to fill 3 large, locked filing cabinets, which the PI already has. However, if this is underestimated, all digital and paper data storage locations can be scaled to at least double their current capacity without difficulty.

Paper documents will be retained at the PI's storage facilities at the University of Manitoba Fort Garry campus for a maximum of 10 years after data collection ceases.

Digital data will be kept on the PI's allocation of the University of Manitoba and St. Boniface Hospital Research Centre data servers, and the backup devices for the study, for a maximum of 10 years after data collection ceases.

Approximately 210 boxes, or less than 13 cubic feet, will be required to store the biological samples in the freezer (-80oC).

How and where will your data be stored and backed up during your research project?

Data storage

The PI has locked laboratories at the University of Manitoba Fort Garry campus and St. Boniface Research Centre that house paper documents and provide access points to digital servers administered by the respective parent organization.

Backup

We are following the 3-2-1 backup rule; we have at least 3 copies of our data; the copies are stored on two different media, and one backup copy is kept offsite.

UM Standard: ISS-004 Backup and Recovery Strategy mandates the backup of the data housed on the server regularly. Backups will also be held on two additional devices, with copies held in different secured locations. The primary backup device is an encrypted 4-terabyte external hard drive housed in the PI's locked lab at the University of Manitoba Fort Garry campus. The secondary backup is an encrypted 1-terabyte external hard drive held in the PI's locked lab at the St. Boniface Research Centre. In order to minimize the possibility of theft or mechanical failure, these external drives are kept in locked filing cabinets and are only connected to laboratory computers when in use. The total data generated by the study is not expected to exceed the capacity of the external hard drives, however these devices can be scaled up as needed.

How will the research team and other collaborators access, modify, and contribute data throughout the project?

Study processes and protocols are in place for accessing each type of data collected, in conjunction with guidelines for which project roles have responsibility for data modification and contribution.

Individual team members and collaborators will not store data on their own devices, in physical spaces, or outside of the project REDCap space.

Preservation

Where will you deposit your data for long-term preservation and access at the end of your research project?

Paper documents will be retained at the PI's storage facilities at the University of Manitoba Fort Garry campus for a maximum of 10 years after data collection ceases. At that point, it will be confidentially shredded.

Digital data will be kept on the PI's allocation of the University of Manitoba and St. Boniface Hospital Research Centre data servers, and the backup devices for the study, for a maximum of 10 years after data collection ceases. At this point, it will be deleted, and the storage devices deleted and reformatted.

We will use the University of Manitoba Dataverse to make our research data available to collaborators at the end of our study. Dataverse is a university-administered and controlled access repository that facilitates research data creation, management, and dissemination. The University of Manitoba version of Dataverse is locally hosted.

Indicate how you will ensure your data is preservation ready. Consider preservation-friendly file formats, ensuring file integrity, anonymization and de-identification, inclusion of supporting documentation.

Where possible, all data will be stored in open or industry standard file types (i.e., .CSV or .XLS). The original data collected will be retained in its original pre-processed format (i.e., accelerometer .GT3X files) utilizing the previously described naming conventions for anonymization. Any processed files will contain information denoting processing procedures, decisions (if any), and the date of the analysis or processing.

- Data kept in the REDCap databases are exported as .CSV text files, which is an open format that is readable by Microsoft Excel.
- Accelerometry activity/sleep and HRV data are collected using ActiGraph wGT3X+ and wGT3X-BT accelerometers, respectively. The data are collected at a 90 Hz sampling rate. Accelerometer data are downloaded using software proprietary to ActiGraph (ActiLife 6, version 6.13.4) and saved as proprietary file formats (.AGD and .GT3X). Data can then be exported into non-proprietary .CSV formats, i.e., either in raw accelerations (g) or epochs (range selected can be 1 –60 second). We chose this flexibility as there is no gold standard approach.
- Heart rate data are exported from ActiLife 6 and saved in the .CSV file format and then HRV data is then cleaned, assessed, and processed using open Kubios software.
- Pulse wave velocity data are exported from the Hypertension Management Software Client Server, version 5.2, and saved on Microsoft Excel (.XLS) files.

- Data generated from biobank blood and stool samples will be kept on Microsoft Excel (.XLS) files

Sharing and Reuse

What data will you be sharing and in what form? (e.g. raw, processed, analyzed, final).

Raw, processed and analyzed datasets may be shared, in compliance with ethics documentation. Analyzed data will all be shared in UM Dataverse. Data sharing will comply with end-user licenses developed for the study; see the following answer.

Have you considered what type of end-user license to include with your data?

Before research data or biological samples are shared, the PI must consult with the University of Manitoba Office of Research Services to request their involvement in creating a Data Sharing Transfer Agreement or a Material Transfer Agreement, respectively. The collaborating parties and their respective institutions must sign and adhere to these agreements. The University of Manitoba Libraries Research Services will also be consulted during the drafting of agreements to ensure that the wording is consistent with the policies that govern the usage of Dataverse.

What steps will be taken to help the research community know that your data exists?

Avenues of discovery will be made available to potential collaborators through public postings, conferences, and presentations. Publications will describe the goals, methods, and planned analyses of the study and include the contact information of the PI. A posting for the study on the clinicaltrials.gov database already exists (ClinicalTrials.gov Identifier: NCT03938155), which informs researchers that they can request access to our data. A protocol paper describing the WARM Hearts trial published in BMJ Open (DOI: 10.1136/bmjopen-2020-044227) also informs researchers that they can request access to our data. Once analyzed data has been deposited in Dataverse, DOIs will allow for the data to be searched as well as linked to publications.

Responsibilities and Resources

Identify who will be responsible for managing this project's data during and after the project and the major data management tasks for which they will be responsible.

The PI assumes ultimate legal and ethical responsibility for all aspects of data management. The PI's responsibilities include developing and implementing data collection and management procedures that ensure the data are credible and that the rights and safety of participants are preserved.

The PI may delegate staff to manage any aspect of the study's data management on his behalf, except for the assumption of ultimate legal and ethical responsibility. Delegated staff must document training in their areas of responsibility and formally demonstrate competency for responsibilities that involve participant interaction or laboratory procedures. Delegation and training logs are kept, facilitating this documentation.

How will responsibilities for managing data activities be handled if substantive changes happen in the personnel overseeing the project's data, including a change of Principal Investigator?

The study PI will identify a colleague or collaborator of equivalent qualification to take stewardship of the project in the event of his unforeseen incapacity. This individual is Dr. Rakesh Arora. In the event that the PI is unable to fulfill his responsibilities, the university will inform the University of Manitoba Health Research Ethics Board of a vacuum in study leadership, request guidance from the Board appropriate to the specific circumstances of the situation, and formally request transition of responsibilities to the appointed steward.

What resources will you require to implement your data management plan? What do you estimate the overall cost for data management to be?

The data management plan will require support from the PI and research team, University of Manitoba Office of Research services, University of Manitoba Libraries Research services and the continued support and provision of the Dataverse, MCHP data centre, REDCap, and Duhamel labs biobanks. The resources needed to support the DMP are in place and the responsibility to maintain them rests with the PI or the University of Manitoba. No service interruptions are anticipated.

Ethics and Legal Compliance

If your research project includes sensitive data, how will you ensure that it is securely managed and accessible only to approved members of the project?

The main data security risk of this study is the potential for participants' confidential personal health information to be unintentionally disclosed to third parties. This risk is mitigated by storing physical data in secured buildings, locked rooms, and locked filing cabinets; and by storing digital data on password-protected computer systems and password-protected encrypted storage media. The REDCap database servers are administered by the University of Manitoba and The George & Fay Yee Centre for Healthcare Innovation. The database servers are password-protected and rely on secured connections between servers and clients. The risk of unintentionally disclosing confidential personal health information is considered adequately mitigated. Nevertheless, during the consent process, participants are made aware that, despite these precautions, a possibility of unintentional disclosure exists. In light of this, participants are welcome to decline or withdraw consent.

Administrative analyses provided by MCHP are covered by their security and confidentiality policy/processes.

All study staff will protect participant confidentiality as a condition of their employment.

Study staff must complete and adhere to the PHIA and Tri-Council Policy Statement. All research staff must complete PHIA and Course on Research Ethics (CORE) training provided through the University of Manitoba and the Government of Canada, respectively.

Additional relevant policies include University of Manitoba network security policies, St. Boniface Research Centre network security policy (available upon request to the system administrator), The George & Fay Yee Centre for Healthcare Innovation REDCap security policy (available upon request to the system administrator), University of Manitoba Office of Research Services, University of Manitoba Health Research Ethics Board, and MHCP Privacy, Ethics & Data Provider Approvals / Permissions policy. Laboratory and study SOPs, and other relevant documentation, will be made available upon request to the study PI.

If applicable, what strategies will you undertake to address secondary uses of sensitive data?

Study participants were provided with an opportunity to consent that their data and biological samples could be shared with collaborators. For those participants who provided this consent, their data and biological samples will be made available to collaborators following the creation of a Data Sharing Transfer Agreement or a Material Transfer Agreement, respectively.

Data obtained through linkage to administrative records are kept at the MCHP. This data is only accessible at the MCHP. Individual-level data will not be shared with anyone outside of that organization.

How will you manage legal, ethical, and intellectual property issues?

The collaborating parties and their respective institutions must sign and adhere to Data Sharing Transfer Agreements or Material Transfer Agreements, respectively. Templates for these agreements can be found on the University of Manitoba Office of Research Services website (https://umanitoba.ca/research/ors/research_forms_online.html).

If these template agreements do not meet the needs of the PI or potential collaborators, a non-template contract may be drafted by the University of Manitoba Office of Research Services and its counterpart at the collaborator's institution. All contracts must be approved by the Office of Research Services and reviewed by the University of Manitoba Office of Fair Practice and Legal Affairs.