

Plan Overview

A Data Management Plan created using DMPTool

DMP ID: <https://doi.org/10.48321/D1MT0T>

Title: ProsGATE: Geriatric Assessment and Technology Evaluation in Prostate Cancer

Creator: Nabil Mir - **ORCID:** [0000-0003-0635-3985](https://orcid.org/0000-0003-0635-3985)

Affiliation: University of Chicago Medical Center (uchospitals.edu)

Project Administrator: Ryan Priest

Funder: National Institutes of Health (nih.gov)

Funding opportunity number: PAR-22-227

Grant: <https://grants.nih.gov/grants/guide/pa-files/PAR-22-227.html>

Template: NIH-Default DMSP

Project abstract:

Management of metastatic Prostate Cancer (mPCa) in older adults is complex, with frailty playing a crucial role in treatment-related morbidity. Although guidelines advocate for a Comprehensive Geriatric Assessment (GA) to evaluate frailty and adjust therapeutic dosing, less than half of these patients undergo frailty evaluation in practice. Existing GA tools are cumbersome and limited to clinic settings. The Geriatric Remote Initiative (GeRI) was established to streamline GA by integrating digital measures such as remote physical activity metrics, body composition assessments, and patient-reported outcomes into cloud-based risk-prediction models. While initial pilots have demonstrated the GeRI platform's acceptability among older adults, further studies are required to integrate vital prognostic and treatment information into GA, especially in mPCa where frailty is prevalent in 30-70% of patients. This project aims to expand upon the predictive value of wearable device-derived metrics in forewarning adverse outcomes in mPCa patients, thereby paving the way for personalized, data-driven geriatric oncology care. Through a collaboration with industry partners (EMA Wellness), we intend to explore the predictive abilities of combined activity and body composition data in older adults receiving androgen deprivation for mPCa. This project unfolds through three specific aims: 1) Determine the correlation between daily step count, lean and fat mass with clinical frailty at baseline and with treatment among older adults with mPCa; 2) Evaluate the correlation between step count and bioimpedance markers with patient-reported pain and fatigue at baseline and with treatment among this population; 3) Ascertain the predictive power of baseline step count and body composition in forecasting serious adverse events in older adults with mPCa undergoing treatment. This endeavor is significant as it seeks to establish validated digital biomarkers for

vulnerability and treatment intensity modulation, facilitating treatment customization in older adults with mPCa. The REDI K01 award will augment GeRI's predictive analytics concerning adverse outcomes in aging populations, nurturing my transition to an independent physician-scientist aiming to integrate digital biomarkers into treatment paradigms for older adults with cancer.

Start date: 02-02-2024

End date: 02-02-2026

Last modified: 01-22-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

ProsGATE: Geriatric Assessment and Technology Evaluation in Prostate Cancer

This project will produce biometric, clinical frailty, and patient-reported outcome data generated/obtained from wearable devices (Samsung Galaxy Watch 5), clinical frailty assessments (e.g., CARG score, SIOG-compliant GA Measures), and self-reported surveys (e.g., PRO-CTCAE measures). Data will be collected from 120 research participants, generating three primary datasets totaling approximately TBD in size. The following data files will be used or produced in the course of the project: raw step count and body composition data files from Samsung Galaxy Watch 5, Cancer Aging Risk Group Score data, Geriatric Assessment data, patient-reported outcomes on pain and fatigue from surveys delivered through in-person assessments and a mobile application as well as Serious Adverse Event (SAE) recorded and monitored by the investigative team. Intermediate files will include cleaned and merged datasets consolidating biometric data with clinical assessments and patient-reported outcomes. Final post-processed files will contain analyzed data segmented for statistical analysis based on the aims of the study. Raw data will be transformed by statistical analyses including logistic and ordinal logistic regression, mixed-effects models, and machine learning algorithms such as logistic regression with LASSO or random forests, and the subsequent processed dataset used for correlational and predictive statistical analysis. To protect research participant identities, aggregated and de-identified data will be made available for sharing. In Clinic data collection will be performed at clinical sites in the Chicago area(s) while smartphone/smartwatch devices will capture biometric and PRO data from a population of older adults (65 or older) diagnosed with metastatic hormone-sensitive prostate cancer scheduled for androgen deprivation therapy.

Based on ethical and legal considerations, only the following data produced in the course of the project will be preserved and shared: aggregated and anonymized biometric data (step count, lean mass, fat mass), clinical frailty assessment scores, aggregated patient-reported outcomes on pain and fatigue, and summary statistics of Serious Adverse Events (SAEs). Individual participant identifiers and any other sensitive or personally identifiable information will not be shared to protect participant privacy and comply with regulatory standards. The final dataset will include aggregated biometric data from wearable devices, clinical frailty assessment scores, aggregated patient-reported outcomes on pain and fatigue, clinical/cancer/treatment related information and summary statistics of Serious Adverse Events (SAEs). Appropriate measures such as data aggregation, removal of all personally identifiable information, and use of unique coded identifiers will be used for data de-identification and sharing, and informed consent forms will reflect those plans. The rationale for sharing specified subsets of data is to contribute to the broader scientific understanding of the relationships between physical activity, body composition, clinical frailty, and treatment outcomes in older adults with metastatic hormone-sensitive prostate cancer (mHSPC), while ensuring the utmost protection of participant privacy and adherence to ethical and legal standards governing human subjects research. Sharing this data allows for validation of the findings, fosters ease of data reuse, collaborative research, and supports the broader dissemination of knowledge which could ultimately contribute to improved clinical care and outcomes for this population.

To facilitate interpretation of the data, a comprehensive data dictionary, metadata, documentation, statistical analysis plans, study protocols, and data collection instruments will be created, shared, and associated with the relevant datasets. In addition to the aggregated and anonymized data being shared openly, the researcher will share the statistical analysis code, analytic plans, and any transformation algorithms used during data processing. Documentation and support materials will be formatted to align with the [clinicaltrials.gov](https://www.clinicaltrials.gov) Protocol Registration Data Elements. The data dictionary will provide detailed definitions and explanations for each variable within the datasets, including the methods used for data collection. Metadata will encompass information regarding the data collection process, data cleaning and transformation procedures, and the methodologies employed for statistical analysis. Study protocols will detail the design, methodology, and ethical considerations of the study. Data collection instruments such as the assessment forms, survey questionnaires, and the software interfaces for

wearable devices will also be shared to provide clarity on how the data was initially captured. All these elements will be compiled and organized in a structured manner, and a README file will be generated to provide an overview of the data package, explaining the organization of the datasets, metadata, and documentation, and offering guidance on how to accurately interpret and utilize the data.

Patient-Reported Outcome (PRO) and biometric data will be made available in a proprietary format, which requires the use of specialized tools, such as the EMA Wellness app and associated software for data access and manipulation. These tools are fee-based, proprietary software available from the EMA Wellness team. The EMA Wellness app and software are essential for comprehensively managing and analyzing the collected data, ensuring its integrity and accuracy while providing a user-friendly interface for researchers and a HIPAA compliant data flow and storage that has been vetted by our institutions IT security. The longevity of these proprietary tools is expected to be sustained for the duration of the research project and beyond, with a commitment from the EMA Wellness team for ongoing support and updates and deletion of data upon study completion. Alternative access to the data will be provided by exporting files to a nonproprietary format such as CSV for limited use and reuse by the investigators. This format allows for broad accessibility and can be manipulated with common, freely available software tools like Excel or R, ensuring that the data remains accessible to the wider research community. Links to file viewers or open-source software alternatives for accessing and manipulating the CSV data will be provided, along with documentation to facilitate understanding and use of the data in this alternative format.

Data will be stored in common and open formats, such as CSV for our Patient-Reported Outcome (PRO) data and biometric data. Information needed to make use of this data [e.g., the meaning of variable names, codes, information about missing data, other metadata, etc.] along with references to the sources of those standardized names and metadata items will be included wherever applicable. Formal standards for biometric data have not yet been widely adopted. However, our data and other materials will be structured and described according to best practices which are as follows: adherence to a well-defined data dictionary, maintaining a comprehensive metadata record, and ensuring data consistency through rigorous validation and quality control processes by our study team's investigators.

Specialized data such as Electronic Patient-Reported Outcomes (ePROs), step count, and body composition data collected through the EMA Wellness app and paired Samsung Galaxy watch will be stored in EMA Wellness's proprietary repository, which is structured for secure and HIPAA-compliant data storage and sharing, validated by our IT security department. This de-identified data will be shared with the investigators who will extract and store this data using RedCap Case Report Forms that are stored on UChicago Center for Research Informatics Servers

On the other hand, clinical data including demographic, clinical/cancer/treatment, geriatric assessment, and treatment outcomes data (serious adverse events) will be stored using RedCap on the CRI cloud, known for its robust data preservation and sharing mechanisms, ensuring long-term access and discoverability.

Clinical data including demographic, clinical/cancer/treatment, geriatric assessment, and treatment outcomes data (e.g. serious adverse events) will be stored in the CRI cloud. The CRI Cloud Repository offers comprehensive metadata, persistent identifiers such as DOIs (Digital Object Identifiers), and ensures long-term data accessibility. This repository is sustained by The University of Chicago, and the dataset(s) are accessible UChicago Biological Science Division Faculty who have at least 4 Tera Bytes free and available through a specified request process ensuring controlled data dissemination.

The EMA Wellness's proprietary repository provides metadata, persistent identifiers such as DOIs, and long-term access. EMA Wellness supports this repository and dataset(s) are available under a specified Master Services Agreement with the University of Chicago Medical Center, which ensures HIPAA-compliant access and sharing.

Shared data generated from this project will be made available upon request by interested parties no the funding period's end. The duration of preservation and sharing of the data will be a minimum of 10 years after the end of the funding period.

Due to ethical and legal considerations, access, distribution, and reuse of the resulting scientific data will be limited and overseen by a designated committee. The process entails a formal request submission to the committee, which will evaluate each request on a case-by-case basis, ensuring alignment with the University of Chicago's data sharing policies and compliance with NIH guidelines. The committee's mandate extends to monitoring the distribution and reuse of the data, thereby safeguarding the integrity of the data and the privacy of individuals represented therein. This structured approach facilitates a controlled dissemination of the data, while abiding by the prevailing ethical, legal, and institutional frameworks.

Given the sensitive nature of the dataset, data will be made available in the CRI Cloud data repository, which restricts access to the data to qualified investigators with an appropriate research question and an approved Data Use Agreement (DUA). Data can be accessed through a formal request process, which entails submission of a detailed research proposal and the requisite Data Use Agreement to the designated review committee. Upon approval, investigators will be granted access to the data via a secure online platform, ensuring both the integrity of the data and compliance with established ethical and legal frameworks.

In order to ensure participant consent for data sharing, IRB paperwork and informed consent documents will include language describing plans for data management and sharing of data, describing the motivation for sharing, and explaining that personal identifying information will be removed.

To protect participant privacy and confidentiality, shared data will be de-identified using the methods in compliance with HIPAA guidelines. This will include the removal or modification of any information that could be used to identify individuals, such as names, geographical information, and unique identifiers.

Moreover, access to the data will be governed by an approval process, where only qualified investigators with a valid research question and an approved Data Use Agreement (DUA) will be granted access. This controlled access mechanism, alongside the robust de-identification process, fortifies the safeguarding of participant privacy, rights, and confidentiality, fostering a conducive and ethically sound environment for the conduct and dissemination of human-centric research.

Lead PI Nabil Mir MD, ORCID: 0000-0003-0635-3985, will be responsible for the day-to-day oversight of lab/team data management activities and data sharing. Broader issues of DMS Plan compliance oversight and reporting will be handled by the PI and Co-I team as part of general campus(es) stewardship, reporting, and compliance processes. The compliance with the Plan will be reviewed on regularly by a designated Data Safety Monitoring Committee within the institution. The committee will evaluate adherence to the data management and sharing procedures outlined in the DMS Plan, assess the effectiveness of de-identification methods, and ensure that data sharing activities align with informed consent agreements. Any non-compliance issues identified will be addressed promptly, and corrective measures will be implemented as needed to ensure ongoing compliance with the DMS plan.

Planned Research Outputs

Dataset - "Biometric and Clinical Dataset of Older Adults with Metastatic Prostate Cancer "

This dataset is a culmination of a multicenter, prospective cohort study on 120 older adults with metastatic Prostate Cancer (mPCa), spanning 12 months. It intricately captures the patients' health trajectory amid intensified androgen deprivation therapy. The biometric data, encompassing daily step count, lean and fat mass, are measured using Samsung Galaxy Watch 5 and EMA Wellness software, presenting a longitudinal view of physical activity and body composition. Clinical frailty assessments, conducted at baseline, 3, 6, 9, and 12-month intervals, include Cancer Aging Risk Group (CARG) Score, Geriatric-8 (G8), Karnofsky Performance Status Scale (KPS), and Katz Activities of Daily Living (ADL), among others. The dataset also records treatment regimens, dosages, frequencies, and the occurrence and classification of Serious Adverse Events (SAEs). Patient-reported outcomes (PROs) of pain and fatigue, graded per the PRO-CTCAE measures, are meticulously logged. Demographic and cancer-specific data are included to provide a comprehensive background. The dataset is temporally structured, allowing for a nuanced longitudinal analysis. It serves as a vital resource for delving into the nexus between biometric indicators and clinical frailty, shedding light on treatment impacts, and investigating the predictive efficacy of biometric measures concerning adverse events in geriatric oncology. Data accessibility is managed through secure, designated repositories with controlled access to uphold privacy and facilitate collaborative research. Accompanying metadata and persistent identifiers enhance discoverability and consistent access. Preservation and sharing protocols align with the funding institution's guidelines and ethical boards' directives, ensuring the dataset's long-term accessibility and utility for the broader scientific community. This dataset is poised to significantly contribute to geriatric oncology, unraveling complex relationships between physical activity, body composition, clinical frailty, and treatment outcomes in older adults with mPCa.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Biometric and Clinical Dataset of Older Adults wit ...	Dataset	Unspecified	Open	None specified		None specified	None specified	No	No