

PERIOPERATIVE PREDICTORS OF CHRONIC POST-SURGICAL PAIN AFTER KNEE ARTHROPLASTY (PROSPER)

A Data Management Plan created using DMP Assistant

Creator: Behnam Sadeghirad

Principal Investigator: Behnam Sadeghirad

Contributor: Lawrence Mbuagbaw, PROSPER initiative members

Affiliation: Other Organization

Funder: The Canadian Institutes of Health Research (CIHR)

Template: Alliance Simplified Template (Funding Application Stage)

ORCID iD: 0000-0001-9422-5232

Project abstract:

PROSPER seeks to build an international repository of individual patient data for perioperative prognostic factors of chronic pain after Total Knee Arthroplasty (TKA), to develop a prognostic prediction model and to validate the risk prediction algorithm. Our objectives are:

1. To develop an international data repository of patient-level data from randomized and non-randomized studies of patients undergoing Total Knee Arthroplasty.
2. To perform a systematic review and IPD meta-analysis of randomized and non-randomized studies to investigate perioperative prognostic factors of chronic pain after Total Knee Arthroplasty.
3. To develop a risk prediction model for chronic post-surgical pain in patients who undergo Total Knee Arthroplasty and assess internal and external validity and robustness of the risk prediction algorithm.

Identifier: 17995

Start date: 01-10-2024

End date: 01-10-2035

Last modified: 26-12-2025

Grant number / URL: <https://mac-prosper.ca>

PERIOPERATIVE PREDICTORS OF CHRONIC POST-SURGICAL PAIN AFTER KNEE ARTHROPLASTY (PROSPER)

INTRODUCTORY GUIDANCE

Introductory Guidance

This Data Management Plan (DMP) is developed to guide the research team through systematic management of data generated and collected for PROSPER project. This project involves consolidating anonymized de-identified patient-level datasets from research studies to investigate prognostic factors associated with chronic pain after knee replacement surgery. The purpose of this DMP is to establish clear protocols and procedures for managing data across all phases of the research lifecycle, from initial data acquisition and validation through storage, analysis, and eventual archiving or sharing. Given the complex nature of this project involving international collaborators, this plan addresses critical data management challenges including data standardization, harmonization of outcome definitions and measurement scales across studies, handling of missing data (both within-study and between-study missingness), and ensuring appropriate anonymization while maintaining data integrity. The DMP provides guidance on documentation standards, quality assurance procedures, metadata capture requirements, and security measures needed to protect sensitive health information across multiple jurisdictions. This plan also facilitates collaboration with international research partners by clearly communicating expectations around data formats, variable naming conventions, and documentation requirements. All team members should familiarize themselves with the procedures outlined in this plan and ensure adherence to established protocols throughout the research process.

PLAN QUESTIONS

What considerations will you take into account with respect to ethical, legal, or commercial issues?

Describe any applicable ethical, legal, or commercial considerations related to your project and data. This includes research involving Indigenous communities and knowledges, human subjects, legal and commercial considerations/agreements, partnerships or data with a high level of risk associated with it

Anonymized and de-identified data will be stored in password-protected XLSX format for all relevant variables. All file names will include a description of the file contents with the created/modified date and accompanied with metadata files (protocol for this study and our standard operating procedures for data modification and merging, roles and responsibilities regarding data capture, data review and data validation). Only investigators and analysts will have access through password protected computers. Collected data will be retained for the period agreed to in the Data Sharing Agreements (DSA) after open-access publication of our findings – as per our Data Sharing Agreement applications and data transfer agreements. We

will follow privacy and confidentiality protections consistent with applicable federal, provincial, and local laws, regulations, and policies.

All data transfer and management comply with the DSA established with participating research institutions and with approval from the Hamilton Integrated Research Ethics Board. Each collaborating institution provides de-identified patient data according to standardized anonymization guidelines that remove personal identifiers (names, contact information, medical record numbers) while retaining clinically relevant information as listed in our published protocol (available from project website). The Provider institutions retain nominal files linking study IDs to patient information, ensuring that original data cannot be re-identified at McMaster University. Our team will use the data exclusively for the stated research purpose and will not share, disclose, or re-use the data beyond the scope of this individual patient data meta-analysis without explicit written permission from Provider institutions. If identifiable data is submitted to the McMaster team, it will be securely deleted and the Provider contacted to revise their submission.

Data safeguarding procedures include restricted access protocols limiting data access to authorized study team members with demonstrated need-to-know, secure storage on encrypted university servers with regular backups, and prohibition against using the data to contact or re-identify study participants. The data will be securely destroyed according to the study protocol timeline. For long-term data stewardship, analyzed de-identified results and study outputs will be retained at McMaster University and may be published, presented, or made available through peer-reviewed journals and institutional repositories. The intellectual property and any tools or models developed from the analysis—including artificial intelligence-informed predictive algorithms—remain the intellectual property of McMaster University as the Recipient Institution, with Provider Investigators acknowledged as co-authors according to academic standards. Preliminary findings will be shared confidentially with Data Providers before publication, with confidentiality obligations binding all parties until formal publication. All team members acknowledge their responsibility to comply with the terms of the Data Sharing Agreements and institutional privacy policies.

What data will you collect or otherwise bring into your project under this plan?

Describe the data that will be collected, generated, and/or acquired.

We will include randomized trials and prospective observational studies (e.g., cohorts, longitudinal or before-after studies, time-series) that: (1) enrolled adults undergoing primary elective unilateral or bilateral TKA, (2) assessed perioperative risk factors for chronic post-surgical pain (including but not limited to any of the factors listed in Table 1), and (3) measured knee pain or reported the incidence of chronic pain at least 3 months after surgery. We will exclude studies of revision or partial knee replacement surgeries and studies with more than 40% missing participants data for pain outcome or predictors listed in Table below. Our outcomes of interest included be moderate-to-severe CPSP (pain score ≥ 4 on a 10cm VAS) and physical functioning. For the purpose of data collection, we will consider continuous or discrete pain intensity and physical functioning measurement at 3-month post-TKA and any timepoint after that. We will prioritize the collection of continuous variables for pain measurements at two or more time points. We will collect and use a binary variable for chronic pain at a certain time point only if a continuous or discrete pain measurement is not available. Potentially eligible perioperative prognostic factors include:

Socio-demographic patient characteristics

age, sex, gender, body-mass Index (BMI), smoking and drinking habits, marital status, living status (living alone, home care, long-term/assisted home, race and ethnicity, education, occupation/employment status, and litigation/insurance

Preoperative psychological and clinical risk factors

preoperative pain catastrophizing, preoperative opioid consumption, preoperative knee pain duration and intensity, chronic pain from previous surgery, preoperative physical and emotional functioning, disease type (rheumatoid arthritis vs osteoarthritis), previous knee arthroscopy, Diabetes, comorbidities affecting mobility, patellar grind and crepitus, knee alignment, state of anterior cruciate ligament (ACL), preoperative Range of Motion, and disease severity (e.g., Kellgren and Lawrence grading), psychological well-being (including anxiety, depression, PTSD, poor coping strategies)

Procedure-related risk factors

operation side, type of procedure, type of implant, duration of procedure, type of anesthesia, opioid use

Post-operative risk factors

acute post-operative pain, post-operative pain management and medications, mobility, mechanical or surgical complications, duration of hospital stay, infection, duration of postoperative opioid use, postoperative support (e.g., physiotherapy, psychological support, social support)

How will you document data for future re-use or validation?

Describe how you will document your data to ensure that it is easily read and interpreted correctly throughout the research process.

If applicable, specify any data and/or metadata standards that are being used to support your research project.

We will create a complete data documentation package that includes a master data dictionary, study codebook, data handling manual, and standardized operating procedures (SOPs) that detail all data management processes. The data dictionary will provide structured, machine-readable documentation of all variables in the master dataset, including variable names, definitions, data types, measurement units, valid value ranges, and any transformations applied during data harmonization. The study codebook will offer a human-readable companion document summarizing variable descriptions, outcome definitions, prognostic factor categories, and coding schemes used across all included studies.

All data documentation will follow established metadata standards to support interoperability and future data sharing. We will adopt the FAIR data principles (Findable, Accessible, Interoperable, Reusable) and utilize metadata standards compatible with open, machine-readable formats. Additionally, we will document all data preparation and analysis processes through comprehensive README files, annotated analysis scripts (Stata Do file with detailed comments), and audit trails documenting all data cleaning, variable creation, and analytical decisions made at each stage. Data quality assurance procedures, including validation checks, missing data assessments, and outlier identification, will be documented alongside resolution strategies used.

We will assign dedicated personnel with expertise in data management and biostatistics to oversee documentation quality. The complete documentation package will be retained alongside the archived dataset and made available to Provider institutions for validation purposes and to support any future secondary analyses or meta-research questions, ensuring our work can be independently verified and built upon by other researchers.

How will data be stored, accessed and worked with?

Describe both where and how data will be stored, accessed, and worked with during the active phases of your research including as applicable:

- **all versions of data (e.g., raw, master, analytic)**
- **All activities (e.g., data collection, processing, analysis, dissemination)**
- **All software and platforms**
- **Who requires access, including security measures (e.g., Investigators, research staff, collaborators, partners)**
- **How data will be backed up to prevent data loss**

The master dataset will be stored on secure, encrypted servers maintained by McMaster University's Information Technology Services, with access restricted to authorized research team members. Raw de-identified data received from data Provider or their institutions will be stored separately from the master dataset and analytical datasets. We will maintain multiple versions of data throughout the project lifecycle: raw data files received from data Provider (archived and unchanged), the master dataset (consolidated and harmonized data from all studies), and analytic datasets (subsets used for specific analyses, version-controlled with documentation of all transformations applied).

Data will be accessed exclusively through McMaster University's secure systems with multi-factor authentication required for all logins. Access will be granted only to the core research team (Dr. Behnam Sadeghirad, biostatistician, data manager, and analysis staff) and to authorized collaborators at participating institutions who require access for specific analytical tasks, with access logs maintained and regularly reviewed. All data work will occur on secure university-managed computers; data will never be downloaded to personal devices or transferred via unsecured email. Regular backup verification will ensure data integrity and prevent data loss. Upon project completion, all data will be securely retained or destroyed according to Data Sharing Agreement requirements and institutional retention policies.

How will data be managed, discoverable, and accessible for the long term?

Describe plans for long-term management of your data after the active phases of your research have concluded including data deposit and sharing.

Consider and describe as applicable plans for:

- **all versions of data deposited (raw, master, analytic, published)**
- **All activities (e.g., curation, preservation, ethical compliance, publishing etc.)**
- **All software and platforms (e.g., data repositories)**

Clean collated data will be retained for the period agreed to in the DSA after open-access publication of our findings. After publication, the cleaned datasets will be deposited in McMaster Dataverse (<https://borealisdata.ca/dataverse/mcmaster>), a secure Canadian research data repository with multifactor authentication. McMaster Dataverse supports open discovery, management, sharing, and preservation of research data. Researchers outside of the participating author groups can access the datasets by proposing new projects by submitting a detailed proposal and statistical analysis plan (SAP). The proposal and SAP will be reviewed by a data management group and core PIs to ensure methods are valid and feasible. The SAP will

then be circulated among all collaborators for approval and they will retain the option to opt out and not participate or share data for the new sub-project. Researchers seeking to use PROSPER datasets will need to sign a Data Sharing Agreement. Applications from external parties will be entertained at the end of the project. Privacy and confidentiality protections will be consistent with applicable federal, provincial, and local laws, regulations, and policies.