Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D1WK99

Title: Development of Cutibacterium-specific immunoassays to identify true Cutibacterium acnes infections.

Creator: William Mccoy iv - **ORCID:** 0000-0001-5115-3793

Affiliation: Washington University in St. Louis (wustl.edu)

Principal Investigator: William Howard McCoy IV, MD, PhD

Data Manager: William Howard McCoy IV, MD, PhD

Project Administrator: William Howard McCoy IV, MD, PhD

Funder: National Institute of Arthritis and Musculoskeletal and Skin Diseases (niams.nih.gov)

Funding opportunity number: PAR-22-119

Grant: https://grants.nih.gov/grants/guide/pa-files/PAR-22-119.html

Template: NIH-Default DMSP

Project abstract:

More than one million infections occur yearly in the USA due to an indwelling medical device (IMD), like a catheter, pacemaker, mechanical heart valve, cerebrospinal fluid (CSF) shunt, or artificial joint. These infections incur significant medical costs, morbidity, and mortality (>25% for mechanical heart valves), and microbes that live on or in the human body are the most often cause of these infections. Infected devices frequently must be replaced while undergoing antibiotic therapy due to the presence of microbial biofilms that shelter microbes from antibiotics. Strategies to prevent these infections (standard surgical skin preparation with topical anti-microbial agents like ChloraPrep, DuraPrep, and povidone-iodine scrub; prophylactic antibiotics) do not effectively control Cutibacterium acnes, a common cause of IMD infections. C. acnes is a normal inhabitant of human skin, where it grows as a biofilm. Widespread antibiotic use for acne vulgaris has led to significant C. acnes antibiotic resistance. C. acnes is now the most common cause of shoulder periprosthetic joint infection (PJI) and causes a high percentage of sternotomy (34%) and CSF shunt infections (6%). The indolent nature of C. acnes infections (minimal/no erythema, drainage, fever; normal labs) frequently delays their workup. It then takes 1-2 weeks to culture C. acnes anaerobically. True C. acnes infections are also difficult to differentiate from environmental contaminants. An assay is needed to quickly identify C. acnes-specific growth in humans. Our lab has generated antibodies that recognize a potential C. acnes growth biomarker and used them to develop assays. We propose to determine what (AIM 1) our antibodies bind on the biomarker, (AIM 2) type of C. acnes growth produces this biomarker, and (AIM 3) effect ligand-binding has on our antibodies' ability to detect this biomarker. Our study will help to develop the first diagnostics specific for C. acnes growth, and these assays will allow for rapid (<24 hours) identification of true C. acnes infections leading to improved patient outcomes (decreased morbidity/mortality) and lower healthcare costs.

Start date: 04-01-2024

End date: 03-31-2026

Last modified: 01-19-2024

Copyright information:

Development of Cutibacterium-specific immunoassays to identify true Cutibacterium acnes infections.

This study will generate genomic data, but it will only be used for control purposes. Therefore, this project is **not within the National Institutes of Health (NIH) Genomic Data Sharing (GDS) policy's intended scope.**

Some of the studies in AIM 1 will utilize specimens collected from patients who have provided informed consent (IRB ID#s: 201412158, 202302188) or a waiver of consent (IRB ID: 201203126) through one of three ongoing institutional IRB-approved protocols (IRB ID#s: 201412158, 202302188, 201203126). These specimens are not collected specifically for this study, random identifiers are generated for all patients in the datasets, and the study team will not have access to the subject identifiers linked to the specimens. Therefore, **this study is not considered human subjects research.**

This project will produce the following data types and amounts.

Table 1: Data types and amount

Specific Aim	Subaim	Data Type	Data/File Format	File Size*	File Number*	Total Data
1	A,B	Plasmid Sequencing**	.fasta	<10MB	<20 files	<200MB
1	A,B	Protein Gel	.TIFF	<10MB	<100 files	<1000MB
1	A,B	Western Blot (WB)	.TIFF	<10MB	<10 files	<100MB
1	A-C	ELISA	.csv	<10MB	<10 files	<100MB
2	A,B	Microbe Sequencing**	.fasta	<10MB	<10 files	<100MB
2	A,B	Bacterial Growth	.csv	<10MB	<100 files	<1000MB
2	A,B	ELISA	.csv	<10MB	<10 files	<100MB
3	A,B	Protein Gel	.TIFF	<10MB	<100 files	<1000MB
3	A,B	Western Blot (WB)	.TIFF	<10MB	<10 files	<100MB
3	A,B	Chromatography	.csv	<10MB	<100 files	<1000MB
3	B	Electron Microscopy	.TIFF	<50MB	<100 files	<5000MB

*Approximate values. **Sequencing to confirm the sequence of plasmid or reference microbe strain.

The total size of the data collected is projected to be ≤10 GB.

AIM 1A, 1B, 2, 3 - Data from Non-human Sources: This data includes all data generated from plasmids (e.g., sequencing), recombinant proteins (e.g., gels, WBs, ELISA, chromatography, electron microscopy), and reference microbial strains (e.g., sequencing, bacterial growth, gels, WBs, ELISA).

AIM 1C - ELISA Data from Human Biofluids/Tissues: Participants (IRB ID# 201412158, N=6; IRB ID# 201203126, N=24, IRB ID# 202302188, N≥70,) have been or will be recruited by ongoing studies at Washington University in St. Louis. The informed consent provided by these participants, or a waiver of consent, includes using their clinical samples in future studies, of which this proposal is one such study. These specimens were not and will not be collected specifically for this study. Details of human samples for ELISA are described in **Table 2**.

Table 2: Human Samples Summary

IRB#	Subject Group	Subject#	Sample Type (number)	Total Samples
201412158	Dermatology Patients	6	Skin swab (2/subject)	12
201203126	Infants/Children with Neurological Injury	24	Cerebrospinal fluid (1/subject)	24
202302188	Shoulder Arthroplasty Patients	≥70	Synovial fluid (1/subject)	≥70
202302188	Shoulder Arthroplasty Patients	≥70	Shoulder tissue* (3/subject)	≥210

*One of sample from each location: subcutaneous tissue, joint capsule, bone/cartilage.

The total number of subjects in this study is projected to be ≥100.

The total number of human samples in this study is projected to be ≥316.

All datasets described in **Table 1** will be preserved and shared through Digital Commons@Becker to enable validation of research results and accelerate future research directions.

For all datasets that will be deposited into Digital Commons@Becker, we will submit a README file to capture rich metadata and associated documents, such as data dictionaries and study protocols, along with datasets to facilitate the interpretation and reuse of the data.

The list of metadata that will be captured in a README file includes the following.

- General information: title, authors with ORCID iDs, organization (ROR), funder information (Funder Registry), award number, date of data collection, location of data collection, contextual description of the data
- Sharing/access information: licenses/restrictions placed on the data
- Data & file overview: list of file names and the relationship between files
- Methodological information: description of methods used for collecting/generation of data and processing the data, instrument and software-specific information to interpret the data, standards, and calibration information if appropriate, description of any quality-assurance procedures performed on the data, people involved with sample collection, processing, analysis and/or submission.
- Data-specific Information: number of variables, number of rows, variable list for each dataset

All data will be made available in open file formats (see **Table 1**), which can be accessed using publicly available software, and will not require the use of specialized tools to be accessed or manipulated. For example, fasta files can be opened using a text editor software such as Notepad and Microsoft Word, csv files can be opened using any spreadsheet program such as Google Sheets and Microsoft Excel, and TIFF files can be opened with an image file viewer such as Image J.

We will use the community standards and ontologies that are widely accepted in the fields such as CDISC, LONIC, standards on FAIRsharing.org and convert all data files into open file formats to enable interoperability.

All datasets described in **Table 1** will be preserved and shared through Digital Commons@Becker. Digital Commons@Becker is an institutional repository administered and maintained by the Bernard Becker Medical Library at Washington University in St. Louis and is a generalist repository that accepts all types of data.

DigitalCommons@Becker provides searchable study-level metadata for dataset discovery and assigns Digital Object Identifiers (DOIs) as persistent unique identifiers, and has a robust preservation plan to ensure long-term access. Data will be discoverable online through a standard web search of the study-level metadata as well as the DOI assigned to the dataset.

All scientific data generated from this project will be made available as soon as possible, and no later than the time of an associated publication or the end of the funding period, whichever comes first. Datasets archived in Digital Commons@Becker will be made available perpetually, and plans exist in instances of repository host change.

Data from Non-human Sources: There are no anticipated factors or limitations that will affect the access, distribution, or reuse of this scientific data generated by the proposal.

ELISA Data from Human Biofluids/Tissues: Human subjects' data will be shared within IRB bounds. All human subject data will be thoroughly reviewed, and all personally identifiable information (PII) will be removed. Informed consent for these ongoing studies at Washington University in St. Louis allows broad sharing, and so there are no anticipated factors or limitations that will affect the access, distribution, or reuse of the scientific data generated by the proposal. All human participants who provide informed consent (IRB ID#s: 201412158, 202302188) will consent to broad data sharing. For ELISA data generated from samples obtained using a waiver of consent (IRB ID: 201203126), demographic data will only be shared in aggregate, and the metadata for this ELISA data will only include infection status and the organism identified by culture (e.g., infected CSF: *Cutibacterium acnes*).

All data derived from non-human sources will be made available with a Creative Commons license with unlimited download. All

data derived from human samples, even though they are de-identified, will be made available via controlled access to the approved users who sign the Data Use Agreement (DUA) to ensure the requesters have a legitimate reason for access, agree not to attempt re-identification of human participants and not to distribute data to unauthorized users.

Samples collected from human participants will use randomly generated numbers and the study team will not have access to the subject identifiers linked to the specimens (de-identified). This research is not considered human subjects research and does not contain PII (Personally Identifiable Information). Therefore, no additional steps will be needed to ensure the confidentiality of human participants.

Principal Investigator William H. McCoy IV will be responsible for the day-to-day oversight of data management and sharing activities. The plan will be reviewed annually to ensure adherence and identify possible updates. Dr. McCoy will ensure that the data management and sharing practice follows the FAIR data principles. Dr. McCoy will report the DMS related activities outlined in this DMS plan in the annual RPPR and request approval for a revised plan if there is any deviation from the approved DMS plan. At the project conclusion, the final progress report will summarize how the DMS objectives were fulfilled and provide links to the shared dataset(s).

Planned Research Outputs

Image - "Gel images"

Images of protein gels

Image - "Western blot images"

Dataset - "ELISA microplate data from non-human sources"

Dataset - "Bacterial growth data"

Dataset - "Chromatography trace data"

Image - "Electron microscopy images"

Dataset - "Sequencing data"

Dataset - "ELISA microplate data from human samples"

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?
Gel images	Image	2026-03-31	Open	Digital Commons@Becker	2 GB	Creative Commons Attribution 4.0 International	None specified	No	No
Western blot images	Image	2026-03-31	Open	Digital Commons@Becker	200 MB	Creative Commons Attribution 4.0 International	None specified	No	No
ELISA microplate data from non-human sources	Dataset	2026-03-31	Open	Digital Commons@Becker	100 MB	Creative Commons Attribution 4.0 International	None specified	No	No
Bacterial growth data	Dataset	2026-03-31	Open	Digital Commons@Becker	1,000 MB	Creative Commons Attribution 4.0 International	None specified	No	No
Chromatography trace data	Dataset	2026-03-31	Open	Digital Commons@Becker	1,000 MB	Creative Commons Attribution 4.0 International	None specified	No	No
Electron microscopy images	Image	2026-03-31	Open	Digital Commons@Becker	5 GB	Creative Commons Attribution 4.0 International	None specified	No	No
Sequencing data	Dataset	2026-03-31	Open	Digital Commons@Becker	300 MB	Creative Commons Attribution 4.0 International	None specified	No	No
ELISA microplate data from human samples	Dataset	2026-03-31	Restricted	Digital Commons@Becker	100 MB	Custom Data Use Agreements/Terms of Use	None specified	No	No