Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D1362Q

Title: Gene Expression and Biomarker Utility in Postmortem Samples

Creator: Scott Nicholson - **ORCID:** 0000-0002-2201-744X

Affiliation: Civil Aerospace Medical Institute (faa.gov)

Principal Investigator: Scott Nicholson

Data Manager: Scott Nicholson

Project Administrator: Scott Nicholson

Funder: Federal Aviation Administration (faa.gov)

Funding opportunity number: na

Grant: na

Template: Federal Aviation Administration (FAA) Data Management Plan (DMP) Template v1.1

Project abstract:

This study is intended to identify genetic biomarkers associated with consumption of cannabis in order to expand thresholds of detection and ability to detect use of drugs that are difficult to assay with traditional biochemistry-based toxicology assays. This project will produce a method of genetics-based detection of drug use in sample specimens used for traditional biochemical-based toxicology, thereby expanding the ability to detect use of drugs in toxicology samples. This project will expand the FAA's ability to assay toxicology samples and will produce biomarkers of THC use for use in future work.

Start date: 10-01-2019

End date: 04-26-2024

Last modified: 04-05-2024

Copyright information:

Gene Expression and Biomarker Utility in Postmortem Samples

Persistent Link:
https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003546.v1.p1

Change Log:NEW DMP 2023_01_13

NA

0. Dataset and Contact Information:

Project: Gene Expression and Biomarker Utility in Postmortem Samples

FAA line-of-business: AAM-600, Aerospace Medical Research Division

Email: scott.nicholson@FAA.GOV

Project number: a11j.am.11

RNAseq data in the form of fastq files will be produced during this research.

Other data, including aviation accident data, drug detection data, and other data will be in the form of .txt., .xlxs, .csv, or other standard formats.

All data will be de-identified.

The purpose of this research is to 1) determine how useful aviation accident autopsy samples are in detecting gene expression, and 2) what RNA biomarker of cannabis use can be found in such samples. The data and study findings will be published in a publicly available format, to be determines at a later date.

All data produced during this study will be in a text format, as described above.

Data will be collected from observations (autopsy data), toxicological analyses, and RNA sequencing.

Samples were collected from 10/1/2019 until 07/01/2022.

Data from these samples was collected during this period.

RNAseq data will be collected throughout 2023.

Existing autopsy, toxicological, and investigatory data will be collected from the FAA ToxDB system.

The data may be used in scientific research, such as medical or addiction research, and also by policymakers, industry researchers, or other unanticipated users.

This data may be useful for medical and addiction research,

Data from this study will be shared through a limited-access government database, and made available for legitimate research purposes.

Other data will be shared publicly through a technical report.

FAA AAM-600

NA

Sequence data will be in .fastq format

Other data will be in standard .txt, .doc, .docx, or .xlsx format.

All are open formats

na

Original data will be stored without modification.

Interim data versioning will be reflected by file naming including dates, version numbers, or iteration numbers.

Final data will be marked 'final'.

NA

The data generated here will be described in technical reports and using metadata standards consistent with NCBI repository requirements.

Standard NCBI metadata schema will be employed.

NA

Standard text editors may be employed to read the raw data.

The raw data is not amenable to direct viewing, summaries of the data and its significance will be made available in technical report(s).

There are a number of open-source quality control software packages, such as multiQC and fastQC, that will be used to assess data quality.

Technical reports including non-identifying raw and summarized data will be made available through Office of Aerospace Medicine reports and other publicly-accessible technical publications.

Raw sequence data will be made available to researchers through restricted-access NIH repositories designed to house that data.

All data involved in this study was de-identified at the site of collection.

RNA sequence data will be treated as sensitive data, and will be released in the restricted-access dbGaP NIH repository as described previously.

Non-sequence data will not be restricted.

Sequence data will be available to researchers who apply for access through the dbGaP data use committee.

Unless otherwise noted, the data described in this DMP is generated and managed by the Federal Aviation Administration. The data are in the public domain, and may be re-used without restriction.

Unless otherwise noted (e.g., data is partially proprietary by an external entity, where intellectual property is shared), this data is required to be made available in open, machine-readable formats, while continuing to ensure privacy and security in accordance with the OPEN Government Data Act, which is Title II of the Foundations for Evidence-Based Policymaking Act.

NA

NA

Unless otherwise noted, there is not a need for the data in this DMP to be licensed for reuse, redistribution, and/or its derivative products.

Unless otherwise noted, the data described in this DMP will be uploaded to the FAA's Enterprise Information Management (EIM) through the [FAA Data Governance Center.](#) This is the internal FAA landing page and access point to EIM uploaded datasets and processes. Here the metadata is curated and validated for quality and accuracy. The FAA Data Steward enters metadata and verifies quality and accuracy before publishing to data.faa.gov, which is the FAA's clearinghouse site for publicly available FAA data and managed and hosted by the FAA's, IT Shared Services organization - Chief Data Office, see https://catalog.data.faa.gov/about for more information.

Technical reports and non-sequence data generated in this project will be placed in the National Transportation library

Sequence data generated during this project will be placed in the NIH dbGaP repository

https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003546.v1.p1

Data and all research products (e.g., reports) are expected to be submitted within the period-of-performance of the research, which is planned to conclude 09/30/2024

Data will be stored in a secure cloud environment and on secure drives within CAMI.

Off-site data storage will be provided within a secure cloud environment.

Small data files will be stored in duplicate on separate secure drives.

Data will be stored in controlled-access sites, access will not be granted to individuals not involved in the research.

Original data will be stored in multiple secure locations.

Unless otherwise noted, the long term storage of the data described in this DMP will persist indefinitely in the FAA's Enterprise Information Management (EIM) platform following standard government policies and best practices.

Raw sequence data will be stored on the NIH dbGaP repository.

dbGaP employs persistent identifiers.

https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003546.v1.p1

NIH dbGaP is a government-maintained repository that meets all federal requirements and serves as the repository of record for human sequence data.

This data management plan was created to meet the requirements enumerated in the U.S. Department of Transportation's "Plan to Increase Public Access to the Results of Federally-Funded Scientific Research" Version 1.1 << https://doi.org/10.21949/1520559 >> and guidelines suggested by the DOT Public Access website << https://doi.org/10.21949/1503647 >>, in effect and current as of 2023_01_13

Planned Research Outputs

Data paper - "Gene Expression and Biomarker Utility in Postmortem Samples"

Technical report describing project

https://doi.org/10.21949/1529631

Dataset - "Gene expression and biomarker utility in postmortem samples RNAseq set"

RNAseq dataset located at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs003546.v1.p1

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?
Gene Expression and Biomarker Utility in Postmorte ...	Data paper	2024-09-29	Open	NCBI	1 TB	Creative Commons Attribution 4.0 International	None specified	No	No
Gene expression and biomarker utility in postmorte ...	Dataset	2024-02-07	Restricted	NCBI		Creative Commons Attribution 4.0 International	None specified	No	No