SMS – Determining if there is a medical vs. human factors/experience risk tradeoff in pilots operating under Basic Med 17.6

Persistent Link

Include the persistent identifier (PID) that is associated with the dataset.

Persistent Link:

https://doi.org/10.21949/1528566

Recommended Citation

The recommendeded data citation to be used when citing the dataset.

Recommended Citation:
U.S. Department of Transportation, Bureau of Transportation Statistics. (2023).

Medical versus Human Factors mishap risk tradeoff dataset (OC 17.6).

Change Log

Table of Contents

Optional table of contents included here, in order to better organize the DMP.

CONTENTS:

6. Policies Affecting this Data Management Plan

0. Dataset and Contact Information

Please provide the following information:

- Name of the dataset or project for which data is being collected.
- Name of the FAA Line-Of-Business/Office for which the associated dataset is being generated.
- Email for the FAA Line-Of-Business/Office (key field).
- If applicable and as reference, project number, contract number, or other number used to link this DMP.

0. Dataset and Contact Information:

The name of the data dictionary/table: "BasicMed and 3rd Class Medical and Human Factors Mishap Rates".

The FAA line of business is Aerospace Medicine (AAM) of Aviation Safety (AVS).

The project is managed by AAM-600, the Medical Research Division of CAMI, the Civil Aerospace Medical Institute.

1. Data Description

Name the data, data collection project, or data producing program. Provide high level narrative.

The dataset provides metadata, initial values, and sources of those values for mishap rates by cause type (human factors or medical) by age for BasicMEd and 3rd medical certification class pilots in addition to output variable values such as summary by age decile.

Describe the purpose of your research and whether results will be documented in a published document or report. How will it be used?

Presentation given to FAA AAM Safety Management Council.

Describe the data that will be generated in terms of nature and scale (e.g., numerical data, image data, text sequences, video, audio, database, modeling data, source code, etc.).

Numerical data in table form.

Describe methods for creating the data (e.g., simulated; observed; experimental; software; physical collections; sensors; satellite; enforcement activities; researcher-generated databases, tables, and/or spreadsheets; instrument generated digital data output such as images and video; etc).

Researcher-generated databases and open source government (DOT) data sources.

Describe the period of time over which the data will be collected and frequency at which it will be updated.

2016 through 2019.

If using existing data, describe the relationship between the data you are collecting and existing data.

The data used is an extract and aggregation of open source available government data with PII information removed.

Describe potential users of the data and the expected manner in which they may use it.

It can be used to derive accident rates for 3rd Class and BasicMed certified pilots.

Discuss the potential value of having the data available not only to your institution but also for the public, e.g., might be renewed interest and value in reanalyzing the data with updated and more universally comparable metrics or recently developed analytical methods.

Later data may be used to extend these data and results.

State clearly if data can be shared publicly or not. If you request permission not to make data publicly accessible, explain rationale for lack of public access.

Choose what is applicable or delete otherwise (add if necessary, not limited to list):

1. Confidentiality. The data includes content personally identifying information (PII) that could be used to identify research participants and cannot be de-identified sufficiently to guarantee confidentiality of the participants' identities. Examples include:

- Media (e.g. photographs, video, audio) that could be used to identify participants
- Internal Review Board (IRB) protected data
- Health Insurance Portability and Accountability Act of 1996 (HIPAA) protected data
- System of Records Notices (SORNs) requirements (e.g. Paperwork Reduction Act (PRA)

related data retention schedules)
- Data that have not been or cannot be deidentified
- Data collected under the approval of an Institutional Review Board covered by [49 Code of Federal Regulations (CFR), Part 11, Protection of Human Subjects](#) and [Protection of Human Subjects (FAA Order 9500.25B)](#), where data cannot be de-identified sufficiently to meet the requirements of the CFR and the FAA Order.

2. Classified. The data for this project are classified to protect national security under……(FAA order, etc) and cannot be released publicly. Examples include:

- Data collected from research utilizing air traffic controllers as participants is covered by National Air Traffic Controllers Association (NATCA), and generally prohibits public release of project data
- Collaborative projects with military organizations (e.g. Tinker Air Force Base)

3. Bargaining unit agreement. This project was conducted under an agreement with a bargaining unit or union organization and is governed by the requirements of that agreement which prohibits public release of data pertaining to the project. Examples include:

- Identification of experimental subjects
- Performance of experimental subjects

4. Operational data. The data for this project include data that are used by the organization to conduct official business and are governed by FAA rules regarding the release of such information. Examples include:

- Information related solely to the internal personnel rules and practices of the organization such as employee performance data, sensitive agency infrastructure, employee selection and hiring process data

5. Proprietary. The data for this project are owned and governed by a private organization pursuant to agreement with the FAA. Examples include:

- data generated in the development of proprietary software or hardware information
- data that would give a company or companies a competitive advantage or disadvantage
- trade secrets
- commercial or financial information that is confidential or privileged
- Data that are not releasable per Memorandum of Understanding (MOU).

6. Federal Prohibition. Information that is prohibited from disclosure by federal law (refer to specific law).

Indicate the party responsible for managing the data.

Unless otherwise noted, refer to "Section 0: Dataset and Contact Information", the FAA line-of-business (LOB) is identified, which is responsible for generating the data, and is also responsible for managing the data initially, and by default long-term, the FAA's Enterprise Information Management (EIM) will manage and catalog the data. Refer to the [FAA Data Governance Center](#), this is landing page and access point to EIM uploaded datasets.

Describe how you will check for adherence to this data management plan.

Unless otherwise noted, refer to "Section 0: Dataset and Contact Information," the FAA line-of-business (LOB) is identified, which is responsible for generating the data, and is also responsible for managing the internal project management processes to ensure adherence to the published data management plan (DMP). Details of the particular FAA LOB's DMP adherence processes can be provided on-demand. Typical processes require management review and sign-off at project start and close-out.

2. Standards Employed

List in what format(s) the data will be collected. Indicate if they are open or proprietary.

2. Standards Employed:

Unless otherwise noted, this FAA research project has descriptive project data posted in https://rip.trb.org/ at project launch and while under development and https://researchhub.bts.gov/ database beyond. These databases have published standards. The project's metadata will be posted in [Catalog.Data.Faa.Gov](#). This catalog follows the DCAT-US Schema v1.1 (Project Open Data Metadata Schema) https://resources.data.gov/schemas/dcat-us/v1.1/ – a set of required fields (Title, Description, Tags, Last Update, Publisher, Contact Name, etc.) for every data set displayed on Catalog.Data.FAA.gov.

If you are using proprietary data formats, discuss your rationale for using those standards and formats.

NA

Describe how versions of data be signified and/or controlled.

Unless otherwise noted, refer to "Section 0: Dataset and Contact Information", the FAA line-of-business (LOB) is identified, which is responsible for generating the data, and is also responsible for version control initially. Once uploaded by default upon project completion and long-term, the data is managed by FAA's Enterprise Information Management (EIM), which also applies configuration control on dataset versions. Refer to the [FAA Data Governance Center.](#) This is the internal FAA landing page and access point to EIM uploaded datasets and processes.

If the file format(s) you are using is(are) not standard to your field, describe how you will document the alternative you are using.

Unless otherwise noted, this project's metadata will describe the data and formats whether open source or non-standard to the particular field domain of the research. if additional description is required for non-standard formats, the researcher will list references and/or definitions here.

List what documentation you will be creating in order to make the data understandable by other researchers.

Unless otherwise noted, this project's metadata will describe the data and formats and by default should be understandable by other researchers and on the FAA's Enterprise Information Management (EIM), which requires application of published standards like DCAT-US Schema v1.1 (Project Open Data Metadata Schema) https://resources.data.gov/schemas/dcat-us/v1.1/ – a set of required fields (Title, Description, Tags, Last Update, Publisher, Contact Name, etc.). Most data sets use open standard and common formats (e.g., CSV, XML, JSON) and if not, described in this DMP.

Indicate what metadata schema you are using to describe the data. If the metadata schema is not one standard for your field, discuss your rationale for using that scheme.

This project's metadata and associated data schema is posted with its data on the FAA's Enterprise Information Management (EIM), which requires application of published standards like DCAT-US Schema v1.1 (Project Open Data Metadata Schema) https://resources.data.gov/schemas/dcat-us/v1.1/ – a set of required fields (Title, Description, Tags, Last Update, Publisher, Contact Name, etc.).

Describe how will the metadata be managed and stored.

Unless otherwise noted, refer to "Section 0: Dataset and Contact Information", the FAA line-of-business (LOB) is identified, which is responsible for generating the data, and is also responsible for generating the metadata. Once uploaded by default upon project completion and long-term, the data and its associated metadata is managed by FAA's Enterprise Information Management (EIM). Refer to the FAA Data Governance Center. This is the internal FAA landing page and access point to EIM uploaded datasets and processes.

Indicate what tools or software is required to read or view the data.

Unless otherwise noted, open data formats are used as much as possible. If not possible, the researcher shall list proprietary data formats and associated tools and software required to read/view the data here. Citations to the required tools and software would be included.

Describe your quality control measures.

Refer to "Section 0: Dataset and Contact Information", the FAA line-of-business (LOB) identified is responsible for managing quality control standards in the data generation and initial creation of the associated metadata. Once uploaded by default upon project completion and long-term, the data and its associated metadata is managed by FAA's Enterprise Information Management (EIM). Refer to the [FAA Data Governance Center.](#) This is the internal FAA landing page and access point to EIM uploaded datasets and processes. Thus, all data uploaded to the EIM platform follows the quality control measures set forth in managing FAA datasets, where EIM states "FAA Data Stewards publish data thru the FAA Data Governance Center hosted and managed by the FAA Chief Data Office. 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."

3. Access Policies

Describe what data will be publicly shared, how data files will be shared, and how others will access them.

Consistent with the previous two sections, "Section 1. Data Description" and "Section 2. Standards Employed," the default long term storage and access location for the data documented in this DMP is the FAA's Enterprise Information Management (EIM). Refer to the [FAA Data Governance Center.](#), which is the internal FAA landing page and access point to EIM uploaded datasets and processes.

Indicate whether the data contain private or confidential information. If so:

- **Discuss how will you guard against disclosure of identities and/or confidential business information.**
- **List what processes you will follow to provide informed consent to participants.**
- **State the party responsible for protecting the data.**

For this DMP's particular dataset, "Section 1. Data Description" Question #9 the researcher already delineated if the data was able to be publicly accessible or restricted and reason associated with the latter. Restricted data is identified upon upload to the FAA's Enterprise Information Management (EIM) and details of the protections in place are documented at the [FAA Data Governance Center](#), which is the internal FAA landing page and access point to EIM uploaded datasets and processes.

If applicable, describe how you will deidentify your data before sharing. If not:

- **Identify what restrictions on access and use you will place on the data.**
- **Discuss additional steps, if any you will use to protect privacy and confidentiality.**

Unless otherwise noted, all concerns and mitigations associated with the need to deidentify fields in the data are addressed in the previous question.

4. Re-Use, Redistribution, and Derivative Products Policies

Name who has the right to manage the data.

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.

Indicate who holds the intellectual property rights to the data.

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.

List any copyrights to the data. If so, indicate who owns them.

Unless otherwise noted, there is no shared copyrights on the data described in this DMP.

Discuss any rights that are transferred to a data archive.

There are no rights transferred to the permanent archive or repository to accompany this dataset described in this DMP.

Describe how your data will be licensed for reuse, redistribution, and derivative products.

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.

5. Archiving and Preservation Plans

Discuss how you intend to archive your data and where (include URL).

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.

Indicate the approximate time period between data collection and submission to the archive.

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 mm/dd/yyyy.

Identify where data will be stored prior to being sent to an archive.

Unless otherwise noted, the permanent archive of the data described in this DMP shall be uploaded, stored, and managed permanently by the FAA's Enterprise Information Management (EIM) platform. However, until the upload upon completion of the project or at a convenient time before, the data will reside locally by the researcher. Refer to "Section 0: Dataset and Contact Information", the FAA line-of-business (LOB) is identified, which is responsible for generating the data, and is also responsible for managing the data initially.

Describe how back-up, disaster recovery, off-site data storage, and other redundant storage strategies will be used to ensure the data's security and integrity, initially and for the long-term.

Unless otherwise noted, the data described in this DMP shall be uploaded, stored, and managed permanently by the FAA's Enterprise Information Management (EIM) platform. This platform is managed and hosted by the FAA's, IT Shared Services organization - Chief Data Office and all back-up, disaster recovery, off-site data storage, and other redundant storage strategies are managed internally by this office and adhering to all FAA mission support policies. For more information and details on these processes, see FAA EIM Platform or contact the FAA line-of-business (LOB) that is identified in "Section 0: Dataset and Contact Information," which is responsible for generating the data.

Describe how data will be protected from accidental or malicious modification or deletion prior to receipt by the archive.

Unless otherwise noted, the data described in this DMP will initially (prior to receipt into the FAA's Enterprise Information Management (EIM) platform) be generated and managed by the FAA line-of-business (LOB), identified in "Section 0: Dataset and Contact Information." The FAA LOB will maintain (3) copies of the data within protected and monitored FAA government servers, facilities, and cloud platforms.

Indicate how long the chosen archive will retain the data.

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.

Indicate if the chosen archive employs, or allows for the recording of, persistent identifiers

linked to the data.

Unless otherwise noted for the FAA researchers in this DMP, the persistent identifiers can only be linked to the Catalog.Data.faa.gov, which provides access to metadata. Access to the research data itself currently requires secure access, including a secure government credentialed sign-on, referred to as MyAccess. This is a role based security profile and intrusion detection monitoring policy to maintain a secure boundary for the EIM Platform that hosts the data.

Discuss how your chosen data repository meets the criteria outlined on the [Guidelines for Evaluating Repositories for Conformance with the DOT Public Access Plan](#) page.

Unless otherwise noted, the data described in this DMP shall be uploaded, stored, and managed permanently by the FAA's Enterprise Information Management (EIM) platform. The EIM Platform is an FAA-developed, cloud-based, big data platform that consists of two key items: (1) "Data Mall" – this is a large repository for FAA data. It is organized and catalogued for easy access, but safeguarded to preserve its integrity and protect data from unauthorized access. And (2) an "App Mall" – this is a collection of curated technologies and tools to enable FAA personnel to transform data into information. For more information, see [FAA EIM Platform](#). The platform's DATA.FAA.GOV 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. It is public gateway to the Enterprise Information Management (EIM) platform that is dedicated to managing data and information to improve efficiency, reduce costs, promote transparency, and enable business insight across the FAA. Thus, this FAA repository meets all the criteria outlined in the DOT Public Access Plan above.

6. Policies Affecting this Data Management Plan

Include policies that the data management plan was created to meet, such as the DOT public access plan.

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 Month(Write out) Day(XX), Year(XXXX).

Planned Research Outputs

Dataset - "OC 17.6 Dataset"

An Excel spreadsheet containing two worksheets: a data table and a (metadata) dictionary.

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?
OC 17.6 Dataset	Dataset	2023-10-29	Open	National Transportation Library	1 MB	Creative Commons Attribution 4.0 International	None specified	No	No