EMORY
LIBRARIES

An Introduction to the Upcoming NIH Policy on Data Management and Sharing

Jen Doty, MSI, Research Data Librarian, Emory Libraries

Jeremy Kupsco, PhD, Research Informationist, Emory Libraries

DMS Policy Overview

Applicability

All research, funded or conducted in whole or in part by NIH, that results in the generation of "**scientific data**".

See Research Covered Under the Data Management & Sharing Policy

Requirements

- ✓ **Submission of DMS Plan** with all applications for funding
- ✓ **Compliance with the DMS Plan** approved by the funding NIH Institute, Center, or Office

DMS Policy: Scope

Applies to all research, funded or conducted in whole or in part by NIH, that results in the generation of "scientific data".

"Scientific data" is defined as:

> ***"the recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications."***

Exclusions from the DMS Policy

Scientific data do <u>not</u> include:

- Data **not** necessary for or of sufficient quality to validate and replicate research findings,
- Laboratory notebooks,
- Preliminary analyses,
- Completed case report forms,
- Drafts of scientific papers,
- Plans for future research,
- Peer reviews,
- Communications with colleagues, or
- Physical objects, (e.g., laboratory specimens)

Activities Subject to the DMS Policy

APPLIES TO…

All research generating scientific data, including but not limited to:

- Research Projects
- Certain Career Development Awards (Ks)
- Small Business SBIR/STTR
- Research Centers

DOES NOT APPLY TO…

research projects <u>not</u> generating scientific data or non-research projects, including but not limited to:

- Training (Ts)
- Fellowships (Fs)
- Construction (C06)
- Conference Grants (R13)
- Resources (Gs)
- Research-Related Infrastructure Programs (e.g., S06)

Compliance/Enforcement

- Extramural Awards: The Plan will become a Term and Condition of the Notice of Award. Failure to comply with the Terms and Conditions may result in an enforcement action, including additional special terms and conditions or termination of the award, and may affect future funding decisions.

- Questions will be added to Research Performance Progress Report (RPPR) to help determine compliance with Plan

Allowable Costs

- **Reasonable costs allowed in budget requests**
  - Curating data/developing supporting documentation
  - Preserving/sharing data through repositories
  - Local data management considerations

- **NOT considered data sharing costs**
  - Infrastructure costs typically included in indirect costs
  - Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)

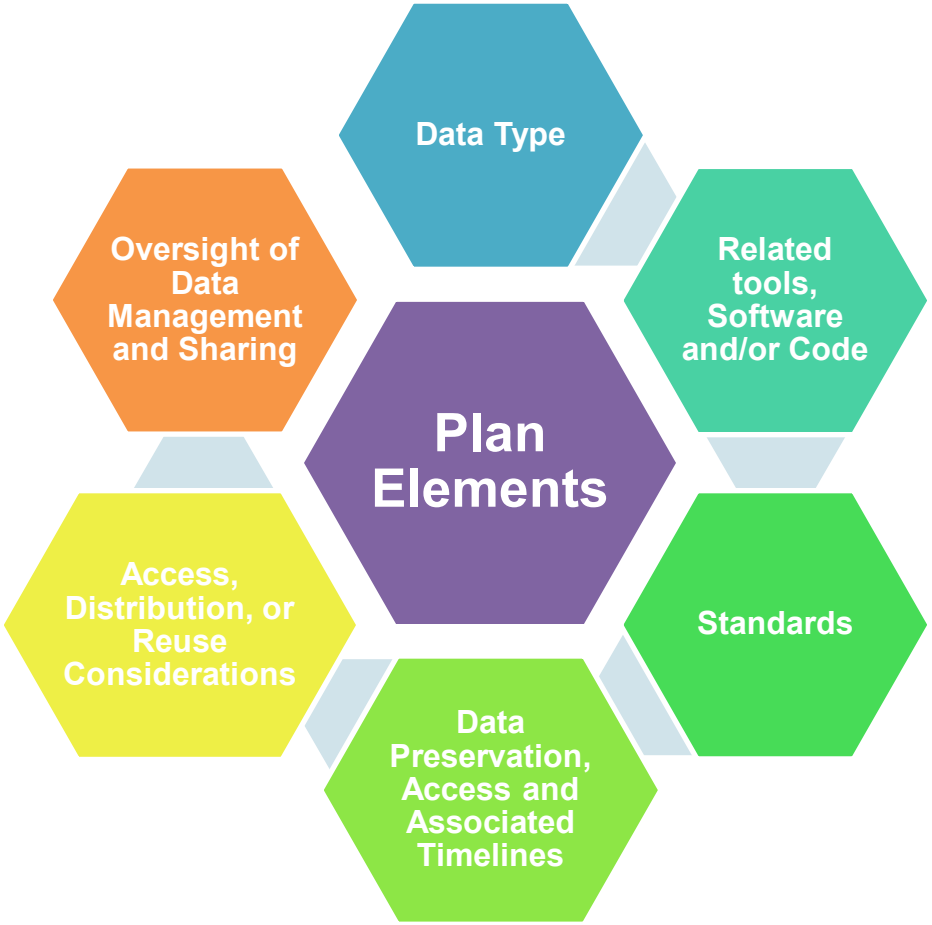
Important Links

- NIH FAQs: https://osp.od.nih.gov/scientific-sharing/nih-policy-for-data-management-and-sharing-faq/
- National Institute for Mental Health Example Data Sharing Plans:
  - https://www.nimh.nih.gov/funding/managing-your-grant/nimh-data-sharing-for-applicants-and-awardees#4
- DMPTool NIH Contest award winner:
  - https://dmptool.org/plans/74502/export.pdf?export[question_headings]=true

How the Library Can Help Researchers

DMPTool and Dataverse

Elements of a DMS Plan



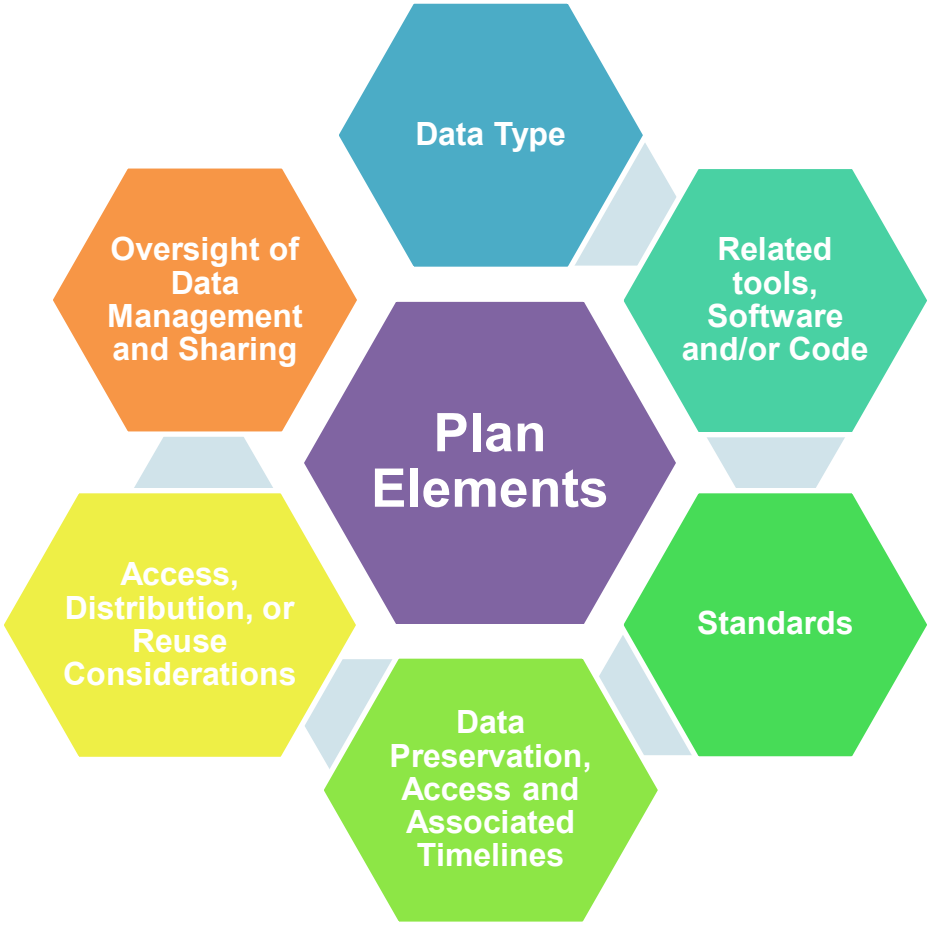
- **Data type**
  - Identifying data to be preserved and shared
- **Related tools, software, code**
  - Tools and software needed to access and manipulate data
- **Standards**
  - Standards to be applied to scientific data and metadata
- **Data preservation, access, timelines**
  - Repository to be used, persistent unique identifier, and when/ how long data will be available
- **Access, distribution, reuse considerations**
  - Description of factors for data access, distribution, or reuse
- **Oversight of data management and sharing**
  - Plan compliance will be monitored/ managed and by whom

See Writing a Data Management & Sharing Plan for details

Adapted from: NIH Webinar I: Understanding the New NIH Data Management and Sharing Policy

EMORY LIBRARIES

Format of a DMS Plan



- ✓ Plans should be no more than 2 pages in length
- ✓ Optional format page will be available from NIH forms, or use template from DMPTool (dmptool.org)

See Writing a Data Management & Sharing Plan for details

DMPTool

Build your Data Management Plan

Funder Requirements Public DMPs Help

dmptool.org

Language

Sign in / Sign up

Email address *

For SSO, use institutional address.

Continue

Problems signing in? Contact us.

77,861 Users

343 Participating Institutions

75,605 Plans

Latest News from DMPTool

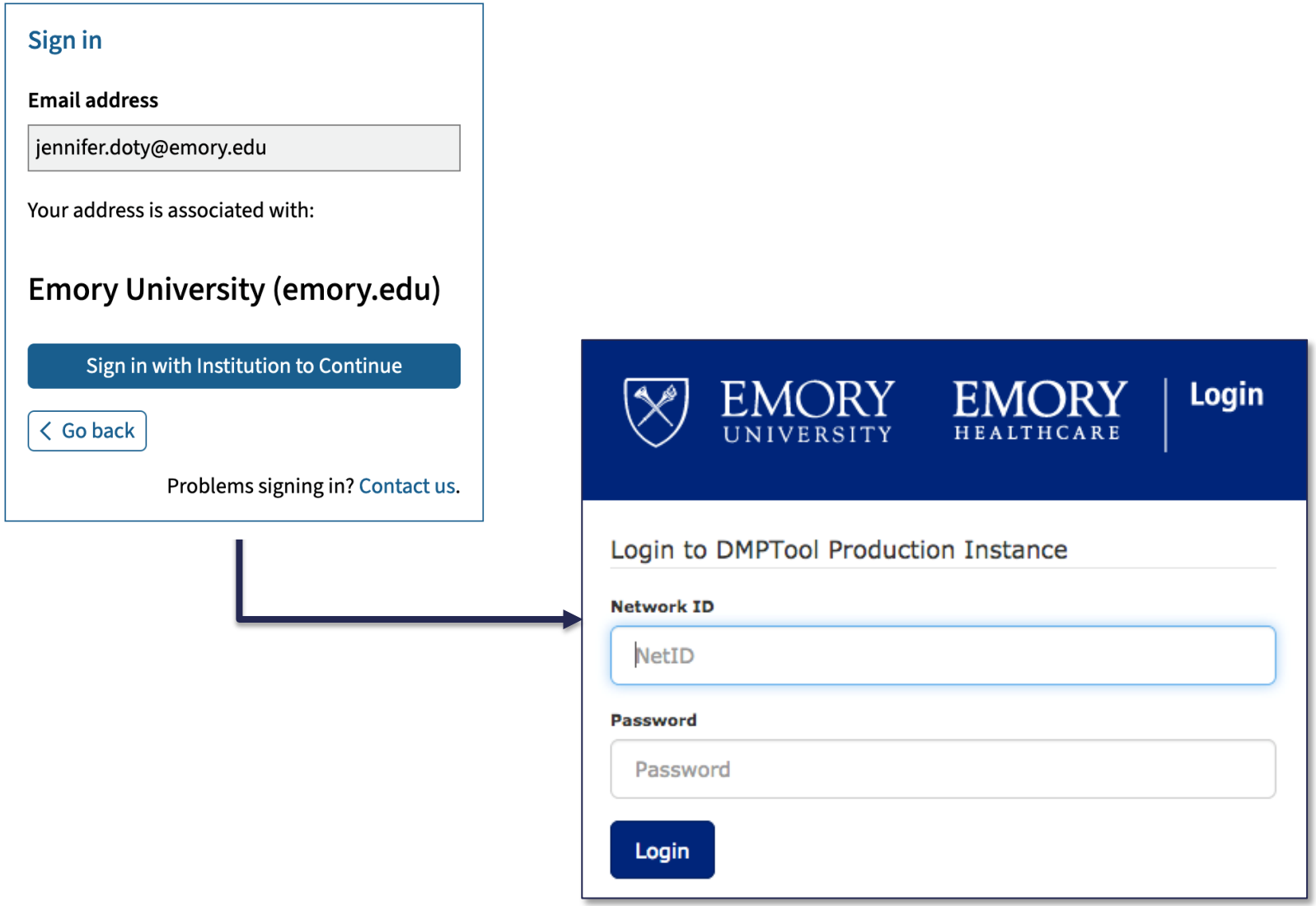
Things to know about the updated DMPTool website

View all news

University of California CDL California Digital Library

DMPTool is a service of the California Digital Library, a division of the University of California Office of the President. Version: v3.5.11

Single sign on with Emory NetID



Funder Requirements

Templates for data management plans are based on the specific requirements listed in funder policy documents. The DMPTool maintains these templates, however, researchers should always consult the program officers and policy documents directly for authoritative guidance. Sample plans are provided by a funder or another trusted party.

nih Search

Template Name	Download	Organization name	Last Updated	Funder Links	Create a new plan	Sample Plans (if available)
NIH-GDS: Genomic Data Sharing		National Institutes of Health (nih.gov)	09-20-2022	NIH Genomic Data Sharing Policy [PDF] NIH Public Access Plan [PDF] NIH GDS Policies and Supplemental Information Data Submission and Release Expectations Developing Genomic Data Sharing Plans		
NIH-GEN DMSP (Forthcoming 2023)		National Institutes of Health (nih.gov)	09-13-2022	Final NIH Policy for Data Management and Sharing		Elements of an NIH Data Management and Sharing Plan Selecting a Repository for Data Resulting from NIH-Supported Research Allowable Costs for Data Management and Sharing
NIH-GEN: Generic (Current until 2023)		National Institutes of Health (nih.gov)	10-25-2021	NIH Data Sharing Policy and Implementation Guidance NIH Public Access Plan [PDF] NIH Data Sharing Policies by Program		NIH Examples of Data Sharing Plans

Clear search results

Project Details | Collaborators | Write Plan | Research outputs | Download | Finalize / Publish

This plan is based on the "NIH-GEN DMSP (Forthcoming 2023) " template provided by National Institutes of Health (nih.gov) - (ver: 4, pub: 2022-09-13).

expand all | collapse all

0/13

Data Type (0 / 3)

Briefly describe the scientific data to be managed, preserved, and shared.

Types and amount of scientific data expected to be generated in the project: ***Summarize the types and estimated amount of scientific data expected to be generated in the project.***

Describe data in general terms that address the type and amount/size of scientific data expected to be collected and used in the project (e.g., 256-channel EEG data and fMRI images from ~50 research participants). Descriptions may indicate the data modality (e.g., imaging, genomic, mobile, survey), level of aggregation (e.g., individual, aggregated, summarized), and/or the degree of data processing that has occurred (i.e., how raw or processed the data will be)

B I

Save

NIH example answer

This project will produce _____ [Data type, e.g., imaging, sequencing, experimental measurements] data generated/obtained from _____ [e.g., instrument, method, survey, experiment, data repository]. Data will be collected from ___ [number] of research participants/specimens/experiments, generating ___ [number] datasets totaling approximately ___ [amount of data] in size. The following data files will be used or produced in the course of the project: _____ [list input data files, intermediate files, and final, post-processed files]. Raw data will be transformed by ____ [analysis, method] and the subsequent processed dataset used for statistical analysis. To protect research participant identities, _____ [e.g., individual, aggregated, summarized] data will be made available for sharing.

Comments & Guidance

Guidance | Comments

NIH | Emory | DMPTool

NIH Guidance

The final DMS Policy defines Scientific Data as: “The recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications. Scientific data do not include laboratory notebooks, preliminary analyses, completed case report forms, drafts of scientific papers, plans for future research, peer reviews, communications with colleagues, or physical objects, such as laboratory specimens.”

Even those scientific data not used to support a publication are considered scientific data and within the final DMS Policy’s scope. We understand that a lack of publication does not necessarily mean that the findings are null or negative; however, indicating that scientific data are defined independent of publication is sufficient to cover data underlying null or negative findings.

Additional Guidance

Research projects vary widely in the types of data produced. In this section, you will describe the categories, amounts, and degree of processing of your data.

Scientific data that will be preserved and shared, and the rationale for doing so: ***Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.***

Guidance | Comments

Sharing Data

Encourages use of established repositories

Depositing data in a quality repository generally improves the FAIRness of data – Findable, Accessible, Interoperable, Reusable

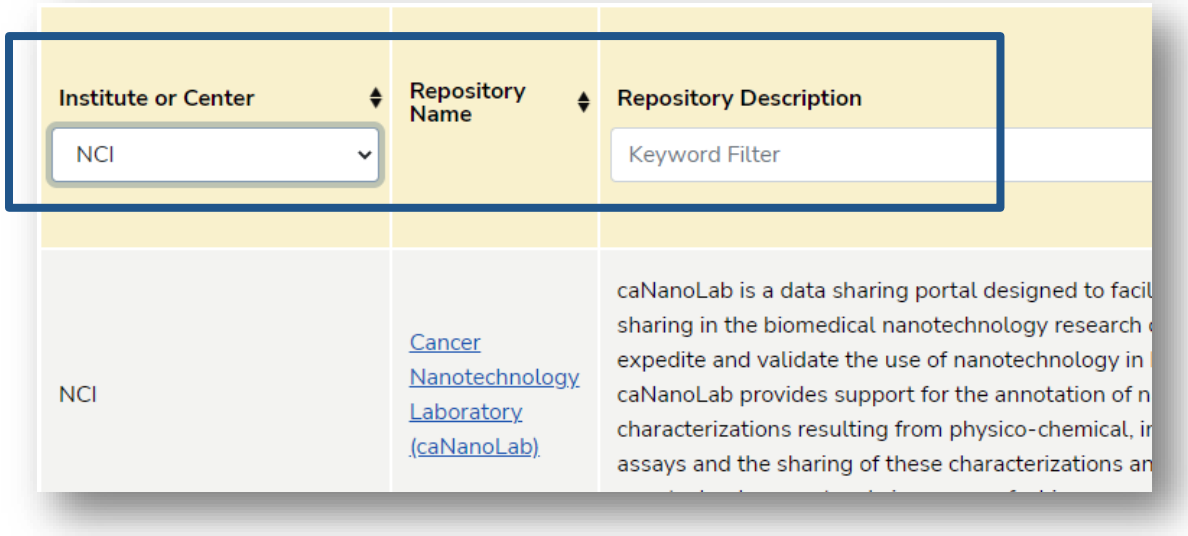
NIH ICs may designate specific data repository(ies)

See Selecting a Data Repository for details

Finding and Selecting a Repository: NIH & Other Resources

- NIH-Supported Repositories

- Filterable list of 70+ NIH Repositories



- Other Repository Resources

- Generalist repositories
- Nature's Data Repository Guidance
- Registry of Research Data Repositories

See Repositories for Sharing Scientific Data

Emory Dataverse

The Dataverse Project

Emory Dataverse

sco.library.emory.edu/dataverse

- **Dataverse** is Emory's open data repository, offered through a partnership between Emory and UNC’s Odum Institute.
- Data deposited with the Emory Dataverse is made available through a web-accessible repository at no cost to depositors or users.
- **Provides persistent access to your data**. Each dataset in Dataverse is assigned a Digital Object Identifier (DOI) for reliable citation and linking.

Emory Dataverse Deposit Policy

Digital, machine-readable data only

No data considered Internal, Confidential, or Restricted as defined by Emory policies

De-identified human subject data accepted

File size limit: 2 GB

Behaviors, movements, and transmission of droplet-mediated respiratory diseases during transcontinental airline flights

Vicki Stover Hertzberg[a,1,2], Howard Weiss[b,1], Lisa Elon[c], Wenpei Si[d], Sharon L. Norris[e], and The FlyHealthy Research Team[3]

[a]Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA 30322; [b]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30313; [c]Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322; [d]Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322; and [e]Boeing Commercial Airplanes, The Boeing Company, Bellevue, WA 98004

Edited by Burton H. Singer, University of Florida, Gainesville, FL, and approved February 13, 2018 (received for review June 30, 2017)

With over 3 billion airli
mission of infectious
concern. Over a dozen
infections have been d
conduit for the rapid s
pandemics. Despite sen
risks of transmission of respiratory viruses in an airplane cabin are unknown. Movements of passengers and crew may facilitate disease transmission. On 10 transcontinental US flights, we chronicled behaviors and movements of individuals in the economy cabin on single-aisle aircraft. We simulated transmission during flight based on these data. Our results indicate there is low probability of direct transmission to passengers not seated in close proximity to an infectious passenger. This data-driven, dynamic network transmission model of droplet-mediated respiratory disease is unique. To measure the true pathogen burden, our team collected 229 environmental samples during the flights. Although eight flights were during Influenza season, all qPCR assays for 18 common respiratory viruses were negative.

airplane transportation | infectious disease transmission | influenza | SARS | pandemic

With over 3 billion airline passengers annually, the inflight transmission of infectious diseases is an important global health concern (1). Over a dozen cases of inflight transmission

Data deposition: Data and software for the simulations are available at dx.doi.org/10.15139/S3/OOYETQ.

account for the spread due to transmission while in route, the evidence cited above notwithstanding. Recently, a model has been proposed that allows for transmission among passengers (15). However, this model assumes that passengers mix randomly. Very little is known about how passengers and crew (flight attendants) mix on airplanes, enabling infection transmission. Given the restraints of time periods when passengers and crew must be seated and the physical restraints of seating in an airplane, it is difficult to believe that random mixing of passengers occurs. We report here on our study of behaviors and movement of passengers and crew on 10 transcontinental flights on

Significance

With over 3 billion airline passengers annually, the inflight transmission of infectious diseases is an important global health concern. Over a dozen cases of inflight transmission of serious infections have been documented, and air travel can serve as a conduit for the rapid spread of newly emerging infections and pandemics. Despite sensational media stories,

EMORY

Emory Dataverse (Emory University) Dataverse is Emory's open data repository. Learn more about sharing your data with Dataverse.

Contact Share

FlyHealthyTM study

Version 1.2

Hertzberg, Vicki, 2018, "FlyHealthyTM study", https://doi.org/10.15139/S3/OOYETQ, UNC Dataverse, V1

Cite Dataset Learn about Data Citation Standards.

Dataset Metrics

221 Downloads

Description The dataset contains programs and data for contacts between passengers and crew on 1000 simulated flights, to be used to simulate infectious disease transmission on these flights. Two READ ME files describe how to use the other files here. (2017-11-07)

Subject Social Sciences

Keyword social networks, infectious diseases, influenza, pandemic

Related Publication Vicki Stover Hertzberg, Howard Weiss, Lisa Elon, Wenpei Si, Sharon L. Norris, The FlyHealthy Research Team. Behaviors, movements, and transmission of droplet-mediated respiratory diseases during transcontinental airline flights. Proceedings of the National Academy of Sciences (PNAS), March 19, 2018, 201711611. doi: 10.1073/pnas.1711611115

Files Metadata Terms Versions

Search this dataset... Find

Filter by

File Type: All Access: All

Sort

1 to 10 of 24 Files Download

aisle.sas7bdat

SAS System - 5.0 KB - Mar 1, 2018 - 12 Downloads

MD5: 23f0afbd59fc1f24bd925237da387d63

SAS data set called by program

Download

We are happy to come talk to your Faculty and answer their questions

Contact: jennifer.doty@emory.edu or jkupsco@emory.edu