



 @ImmPortDB

# Webinar Series

Aug 30, 2022

Moderator: Kate Hypes, ImmPort

## ImmPort Resources and Tutorials: An Introduction to Leveraging ImmPort Shared Data

ImmPort: An Overview

Sanchita Bhattacharya, Science Program Lead, ImmPort

Tutorials for Data Exploration

John Campbell, ImmPort Software Development Team

Q&A



University of California  
San Francisco



# ImmPort : An Overview

Sanchita Bhattacharya  
Bioinformatics Project Leader  
Science Program Lead, ImmPort  
Bakar Computational Health Sciences Institute  
University of California, San Francisco

# ImmPort data portal was developed to collect and share research and clinical trials data from NIAID/DAIT funded researchers



## Private Data

Upload Data  
Upload Templates  
Help



## Shared Data

Data Model  
Search/Download  
Gene Lists



## Data Analysis

Analysis Workflow  
Automated Clustering  
Tutorials



## Resources

Tutorials  
Documentation  
Publications

## FAIR Principles

Make your data:

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable

### Findable

- Descriptive metadata
- Persistent

### Accessible

- Determining what to share
- Participant consent and risk management
- Access status

### Interoperable

- XML standards
- Data Documentation Initiative
- CDISC

### Reusable

- Rights and licence models
- Permitted and non-permitted use

<http://datafairport.org/>

# ImmPort redistributes data from major NIAID-funded programs and more



Human Immunology Project Consortium COVID-19 - ImmPort

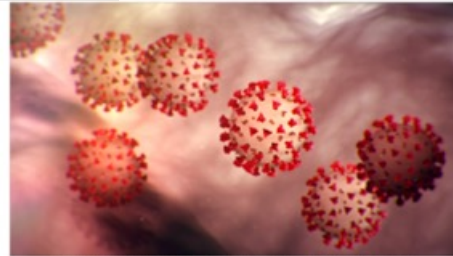


Image CDC Coronavirus Disease 2019 – COVID 19, Alissa Eckert

Coronavirus Disease 19

Powered by ImmPort

**Immunophenotyping Assessment in a COVID-19 Cohort (IMPACC)**  
**Multisystem Inflammatory Syndrome in Children (MIS-C)**

- Impact of Initial Influenza Exposure on Immunity in Infants (U01)
- Atopic Dermatitis Research Network (ADRN)
- Population Genetics Analysis Program
- Protective Immunity for Special Populations
- HLA Region Genomics in Immune-mediated Diseases
- Modeling Immunity for Biodefense
- Reagent Development for Innate Immune Receptors
- Adjuvant Development Program
- Immunity in Neonates and Infants
- Asthma and Allergic Diseases Cooperative Research Centers
- HLA and KIR Region Genomics in Immune-Mediated Diseases
- Cooperative Study Group for Autoimmune Disease Prevention
- Immunobiology of Xenotransplantation
- Centers for Medical Countermeasures against Radiation Consortium
- Inner City Asthma Consortium
- Systems Approach to Immunity and Inflammation
- Innate Immune Receptors and Adjuvant Discovery Program
- Maintenance of Macaque Specific Pathogen-Free Breeding Colonies
- Non-human Primate Transplantation Tolerance Cooperative Study Group
- Consortium for Food Allergy Research
- Development of Sample Sparing Assays for Monitoring Immune Responses (U24)
- Asthma and Allergic Diseases Clinical Research Consortium (AACRC)
- The Clinical Islet Transplantation (CIT) Consortium
- Autoimmunity Centers of Excellence (ACE)



LEVERAGING BIG DATA FOR PRETERM BIRTH RESEARCH



BILL & MELINDA GATES foundation



Powered by ImmPort

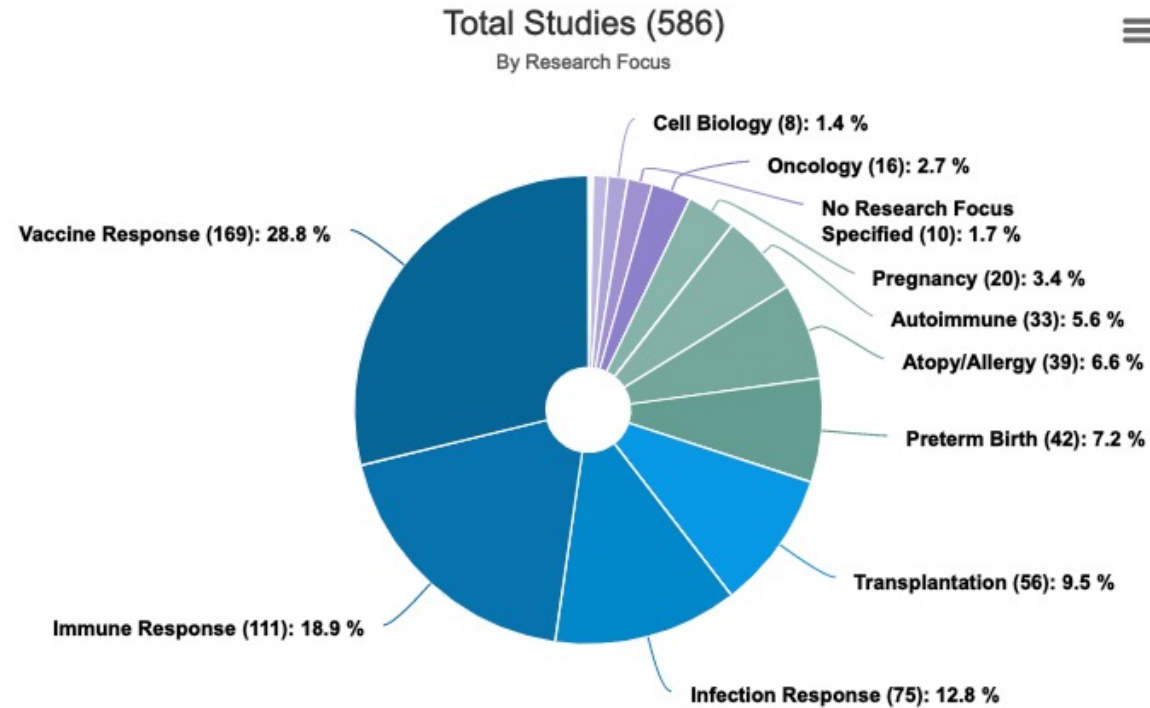


# Focus Areas and Data Types in ImmPort

## Data Summary: Release 44, June 2022

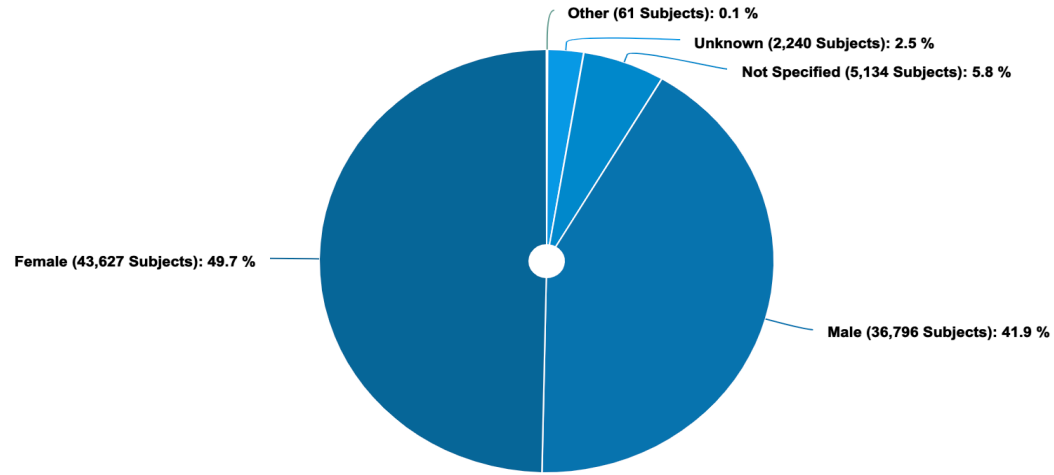
Click on the counts with  icon to visualize the count breakdown

Studies	 586
Subjects	 83513
Diseases	 130
Protocols	1810
Experiments	 2146
Total Results	 6304584
Elisa Results	349680
Elispot Results	71132
Flow Cytometry Results	4779083
PCR Results	127538
HAI Results	24219
HLA Typing Results	149342
Luminex/MBAA Results	730814
Neutralizing Antibody Results	23390
GEO/SRA/Other	49386
Lab Tests	1268118
Total Test Panels	6742

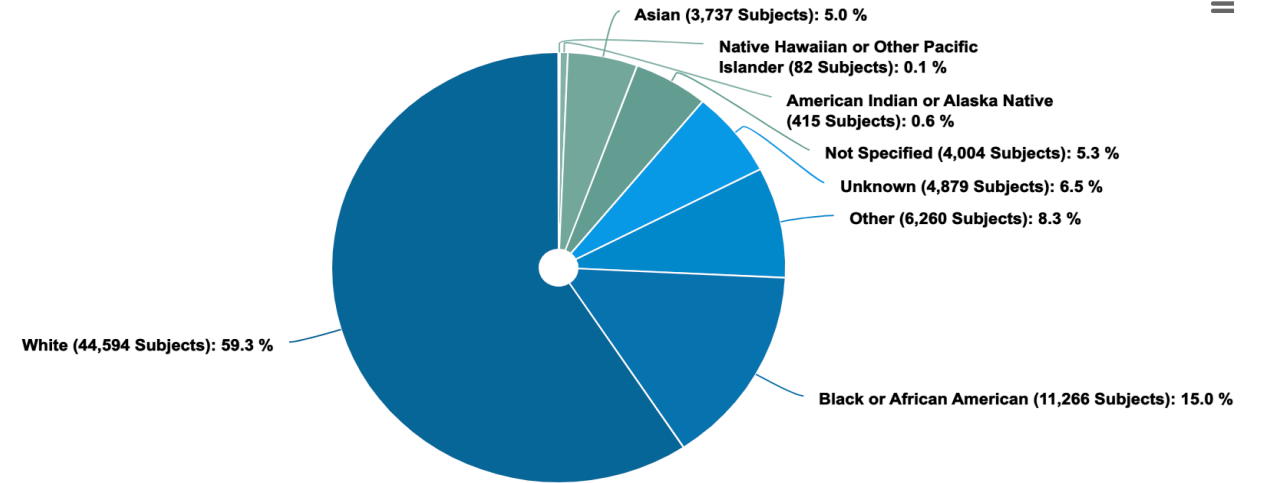


Immunology Studies- 422  
Clinical Trials- 164

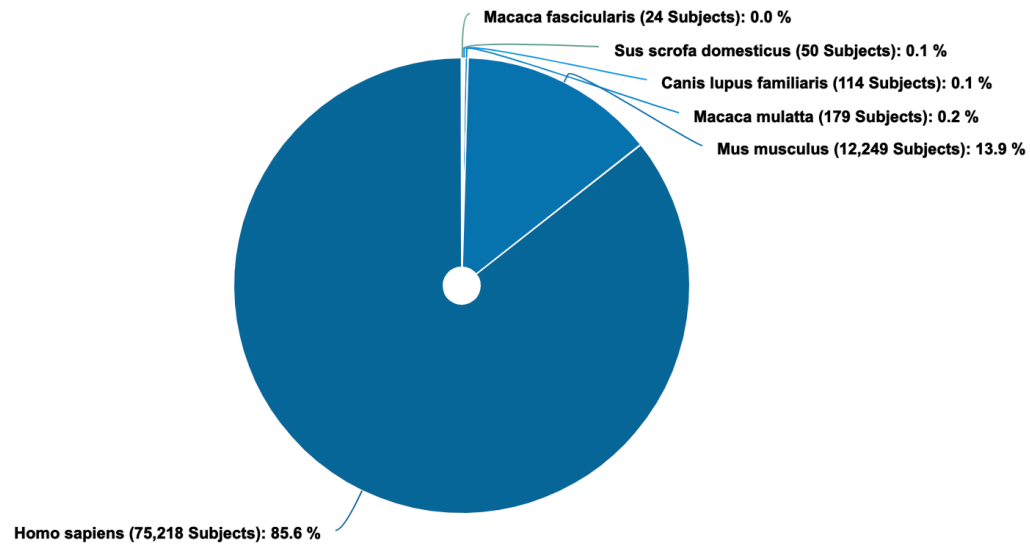
# Gender



# Race



# Species



# Data Ingestion

## Data Templates

basic_study_design	Version 2.24
Please do not delete or edit this column	

study
User_Def_Study_ID
Brief_Title
Official_Title
Study_Type
Brief_Description
Detailed_Description
Hypothesis
Objectives
Endpoints
Sponsoring_Organization

1	A	experiments	A	biosamples
2	A	Please do not delete or edit this column	B	flow_cytometry
				Version

	A	B	C	D	E	F	G	H	I	J
1	subjectshuman	Version 2.24								
2	Please do not delete or edit this column									
3	Column Name	Subject User-Defined ID	Subject Treatment Protocol ID(s)	Gender	Age	Age Unit	Age Event	Age Event Specify	Subject Phenotype	Study Arm Or Cohort ID(s)

adverseevents	Schema Version 2.28	Content
---------------	---------------------	---------

1	A	B	C
	labtests	Schema Version 2.28	Content Version 2.2
2	Please do not delete or edit this column		
3	Column Name	Substance Merge User-Defined ID	Subject ID
4			

Column Name	Biological Sample ID	Lab Panel ID	Study ID	Protocol ID(s)	Source Subject ID	Planned Visit ID	Biological Sample Type	Biological Sample Sub-Type	Biological Sample Name	Biological Sample Description	Study Time Collected	Study Time Collected Unit	Study Time TO Event	Study Time TO Event Specify	Lab Panel Name	Result Separator Column	Lab Test User-Defined ID	Lab Test Name	Result Value	Result Unit
-------------	----------------------	--------------	----------	----------------	-------------------	------------------	------------------------	----------------------------	------------------------	-------------------------------	----------------------	---------------------------	---------------------	-----------------------------	----------------	-------------------------	--------------------------	---------------	--------------	-------------

Honorific
-----------

1	A	B	C
	assessments	Schema Version 2.28	Content Version 2.28
2	Please do not delete or edit this column		

7	Column Name	Assessment Panel ID	Study ID	Assessment Name	Assessment Type	Is Standard	Status	CRF File Names	Result Separator Column	Assessment Component User-Defined ID	Planned Visit ID	Subject ID	Component Name	Study Day	Study Day From Partial Date	Age At Onset
---	-------------	---------------------	----------	-----------------	-----------------	-------------	--------	----------------	-------------------------	--------------------------------------	------------------	------------	----------------	-----------	-----------------------------	--------------

# Online Submission: Study Registration Wizard

Registration Home

## Study Registration ID 1038 for Workspace ID 1941

- 1 Basic Info
- 2 Categorization
- 3 Condition or Disease
- 4 Arm Or Cohort
- 5 Inclusion/Exclusion
- 6 Personnel
- 7 Planned Visit
- 8 Protocol
- 9 Publications/Study Links Optional
- 10 Study Files Optional
- 11 Finish

User Defined ID/Submitter ID \*  
study\_1038

Should be unique within the selected workspace. (Cannot be updated in Edit Mode) 10 / 150



Brief Title/Study Title \*



0 / 250

Study Official Title

0 / 500



Sponsoring Organization \*



0 / 250

Brief Description \*



0 / 4000

Detailed Description \*



Endpoints/Output Measurements/Clinical Assessments \*



Objectives





# Cohort Discovery Tool

ImmPort shares data on COVID-19 and other infectious diseases from NIAID and other funding agencies. Contact ImmPort\_Helpdesk@immport.org for any data sharing or uploading questions. Check out the ImmPort links. COVID-19 studies Influenza studies Respiratory-like illnesses studies Viral infectious diseases studies

This is a beta version of ImmPort's latest search engine with additional tools to search study and subject data. We welcome your Feedback

**Filters**

**Study Facets**

- Research Focus
- Condition or Disease
- Biosample Type
- Assay Methods
- Program Name
- Study Accession

**Subject Demographics**

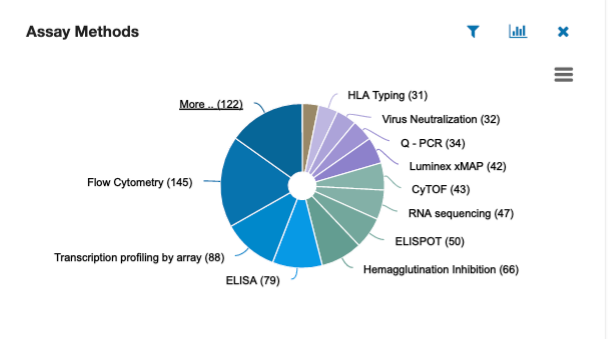
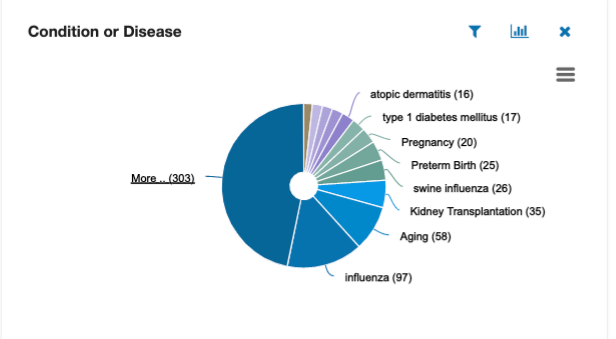
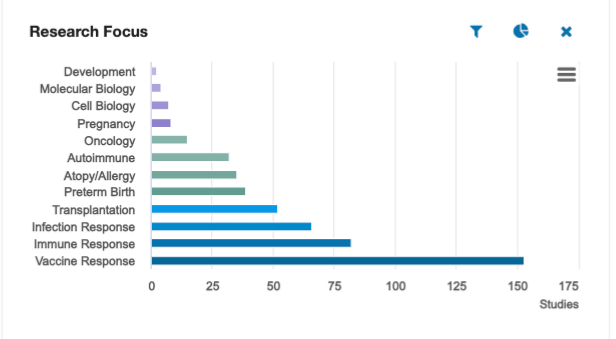
- Age
- Ethnicity
- Gender
- Race
- Species

**Miscellaneous**

- Is Clinical Trial
- Has Lab Tests
- Has Assessments Tests

Search free text: e.g. influenza (minimum 3 characters) [Q] [x Reset] [Bar Chart] [API]

**Studies(495)** Subjects (81908)



Found 495 studies in 66 ms 7 columns selected [Download All Studies] [Export]

Study	Title	Pubmed Id	Research Focus	Assay Methods Count (samples)	Latest Release Version	Latest Release Date
SDY283	Exploration of molecular mechanism(s) for TLR ligand response differences between adults and neonates		Immune Response	ELISA (948)	DR11	2014-08-07
SDY1688	Computational discovery of therapeutic candidates for preventing preterm birth	32051340	Preterm Birth		DR37	2020-12-16
SDY197	Responses to Influenza Vaccination in Systemic Lupus Year 2 2006-2007	21598235	Vaccine Response	ELISA (126) ELISPOT (252) Flow Cytometry (2331) Hemagglutination Inhibition (188) Western Blot (47)	DR22	2017-06-16

# Data Download

## Data Browser

[Data Browser](#)

ImmPort data browser allows users to download ImmPort data by individual file, directory, or study. The data browser uses a software tool called [Aspera Connect](#) to transfer files from ImmPort to users. Here are the [Instructions to install Aspera Connect on your browser](#).

+ Content Listings

[Browse Shared Data](#) > SDY1

\* Archive files are disabled for download in this view, please navigate into the archive directory to download them.

[Download](#) [Clear](#)

<input type="checkbox"/>	Name ^	Size	Last Modified
<input type="checkbox"/>	<input type="text"/>		
<input type="checkbox"/>	Protocols (5 files)	1.27 MB	Feb 9, 2016 12:17 PM
<input type="checkbox"/>	ResultFiles (11215 files)	3.86 GB	Jun 12, 2015 8:57 AM
<input type="checkbox"/>	SDY1-DR38_MySQL.zip	11.47 MB	Feb 23, 2021 15:11 PM
<input type="checkbox"/>	SDY1-DR38_Subject_2_Flow_cytometry_result.txt	9.70 MB	Feb 23, 2021 15:11 PM
<input type="checkbox"/>	SDY1-DR38_Tab.zip	10.31 MB	Feb 23, 2021 15:11 PM
<input type="checkbox"/>	SDY1-DR38_manifest.txt	701.76 KB	Feb 23, 2021 15:11 PM
<input type="checkbox"/>	SDY1-DR38_summary.txt	60.00 bytes	Feb 23, 2021 15:11 PM
<input type="checkbox"/>	SDY1-DR38_table_count.txt	2.31 KB	Feb 23, 2021 15:11 PM
<input type="checkbox"/>	StudyFiles (5 files)	43.97 MB	Apr 20, 2017 8:12 AM
<input type="checkbox"/>	* archive (170 files)	6.57 GB	Feb 23, 2021 15:10 PM

Showing 1 to 10 of 10 records

## API

Settings Help

Launcher Welcome.ipynb Download\_FCS\_Files\_File FOCIS\_2020\_SDY824\_An Download\_HAI\_Results\_DX Python 3

### Download FCS Files - File Download API

For this tutorial, we will first use the API to find FCS files for SDY212, then we will download a sample set of the raw FCS data files data using the File Download API.

In this tutorial we will use the <https://api.immport.org/data/query/result/filePath> endpoint to demonstrate retrieving FCS files for SDY212. The **result/filePath** endpoint will return meta data about the files, plus the file location of the within the shared data directory structure. We will then use the file path location to download the FCS file. To start, we will only be filtering the data based on the studyAccession, armName and measurementTechnique, many additional filters can be applied. Details on using the API's and the endpoints are available [here](#).

We recommend reviewing [DownloadingImmPortData](#) for an overview to download and API options.

### Python Environment

```
[1]: import sys
import os
import json
import requests
import pandas as pd
from io import StringIO

from pathlib import Path
from dotenv import load_dotenv

# We will be using the ImmPort API helper methods to access the API.
# If you would like to see the source code for this module
# enter "?import_api" in an empty notebook cell

sys.path.append('../common')
import immport_api as api
import immport_util as iu
```

```
[2]: ###?api
```

### Load Environment Variables

```
[3]: env_path = Path('../common')/'env'
load_dotenv("../common/.env")
OUTPUT_DIR = os.environ.get("OUTPUT_DIR")
API_ENDPOINT_BASE_URL = os.environ.get("API_ENDPOINT_BASE_URL")
```

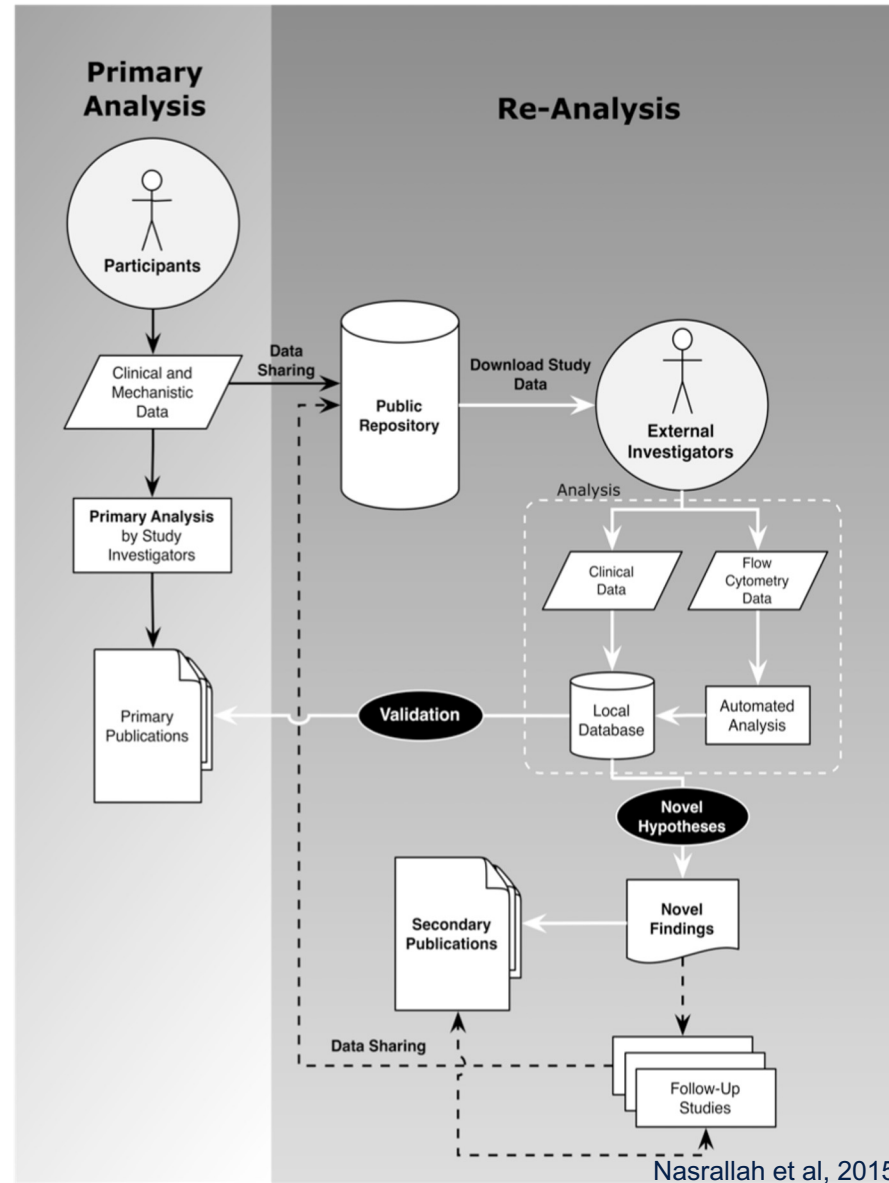
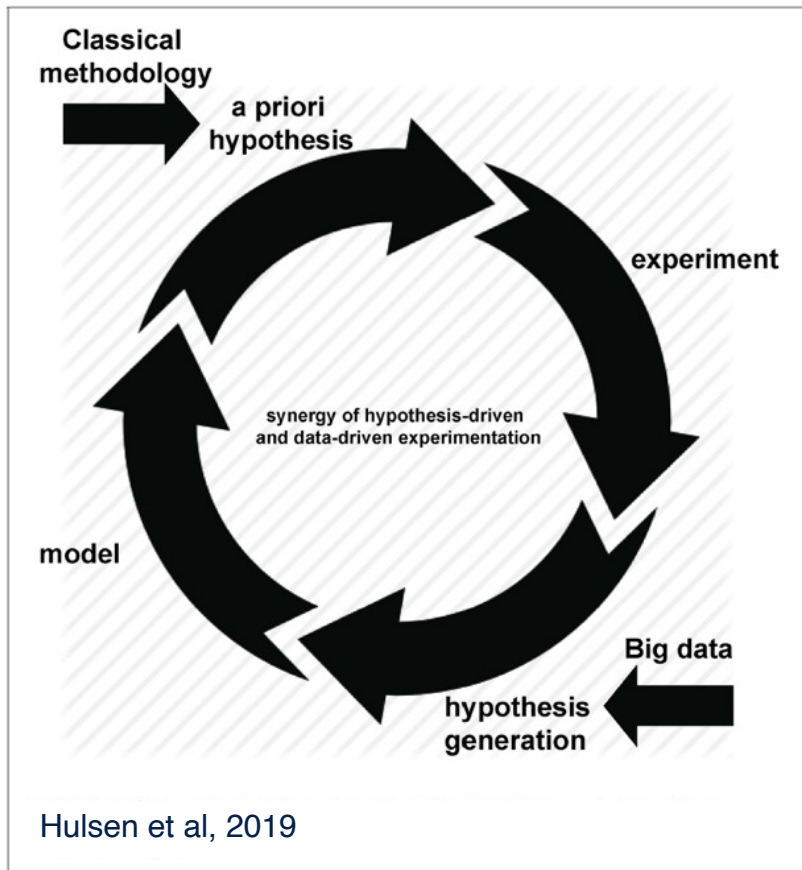
### User Specific Variables

Edit the code to use your ImmPort username and password

```
[4]: user_name = 'REPLACE_USERNAME'
user_password = 'REPLACE_PASSWORD'
```

Make API call

# Benefits of Open-Access Immunological Data



**R**eproducibility  
**R**e-Analyze  
**R**epurpose

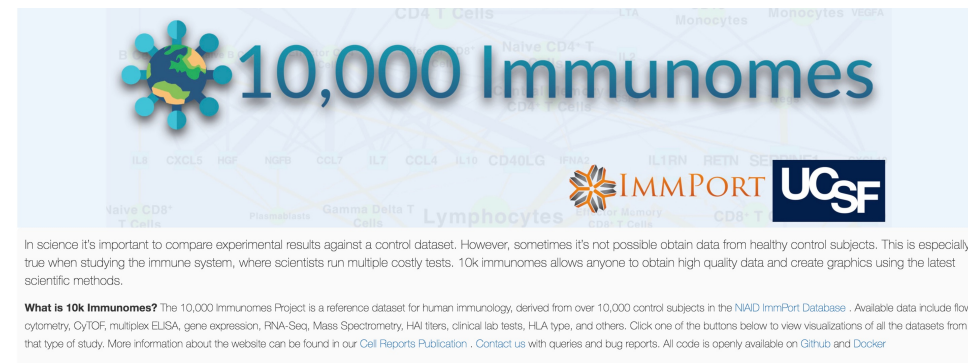
## Resource

### The 10,000 Immunomes Project: Building a Resource for Human Immunology

Kelly A. Zalocusky<sup>1, 2</sup>, Matthew J. Kan<sup>1, 2</sup>, Zicheng Hu<sup>1, 2</sup>, Patrick Dunn<sup>3</sup>, Elizabeth Thomson<sup>3</sup>, Jeffrey Wiser<sup>3</sup>, Sanchita Bhattacharya<sup>1, 2, 4</sup>, Atul J. Butte<sup>1, 2, 4, 5</sup> ✉

## 10kimmunomes.org

- Large, diverse, cleaned reference dataset for human immunology
- Interactive data visualization
- Custom control cohorts and standardized data download



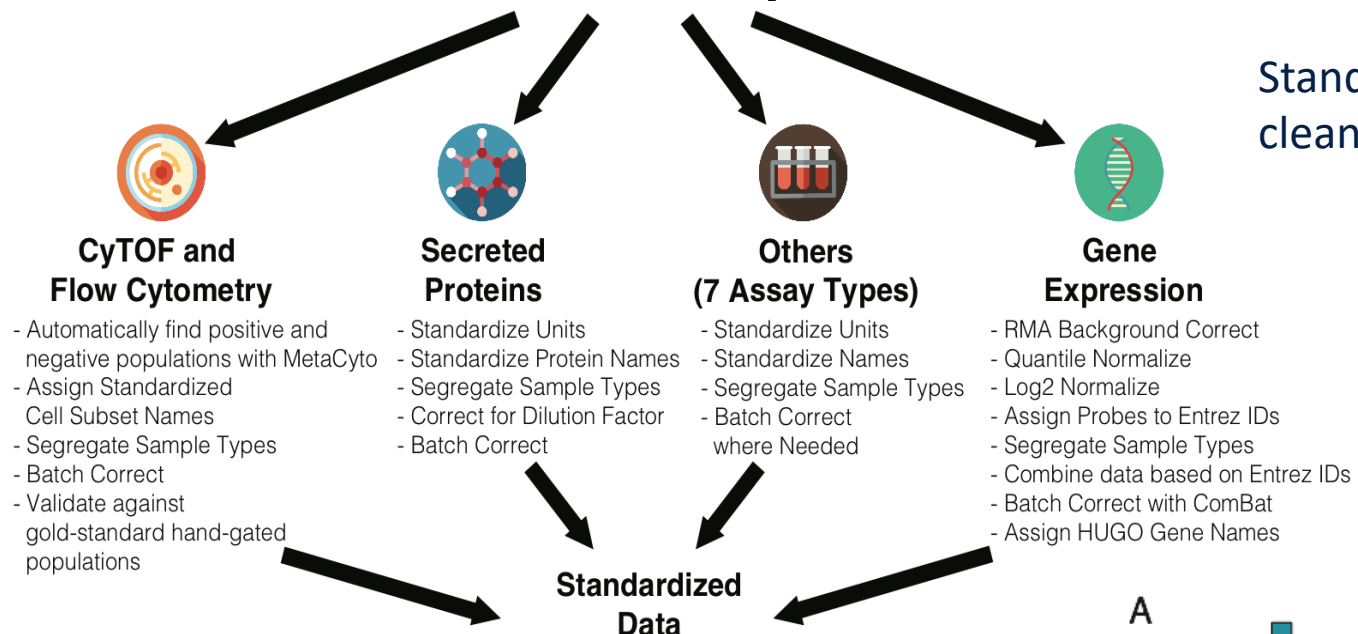
Data from 242 studies and 44,775 subjects was collected from the NAID Immunology Data and Analysis Portal, ImmPort. It includes flow cytometry, CyTOF, mRNA expression, secreted protein levels, clinical lab tests, HAI titers, HLA type, and others. We hand curated the entire contents of ImmPort to filter for normal

**Data available in the 10,000 Immunomes Project**

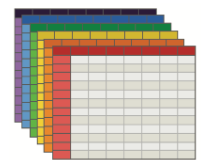
Total Samples	42117
Total Distinct Subjects	10344

MEASUREMENT	SUBJEC
<b>Secreted Proteins</b>	4835
<i>ELISA</i>	4035
<i>Multiplex ELISA</i>	1286
<b>Virus Titer</b>	3609
<i>Virus Neutralization Titer</i>	2265
<i>HAI Titer</i>	1344
<b>Clinical Lab Tests</b>	2639
<i>Complete Blood Count</i>	1684
<i>Comprehensive Metabolic Panel</i>	664
<i>Fasting Lipid Profile</i>	664
<b>Questionnaire</b>	1422
<b>Cytometry</b>	1415
<i>Flow Cytometry (PBMC)</i>	907
<i>CytoF (PBMC)</i>	583
<i>Flow Cytometry (Whole Blood)</i>	164
<b>HLA Type</b>	1093
<b>Gene Expression Array</b>	476
<i>Whole Blood</i>	311
<i>PBMC</i>	165

**85 Studies  
10,344 Subjects  
42,000+ Samples**



**Standardized Data**

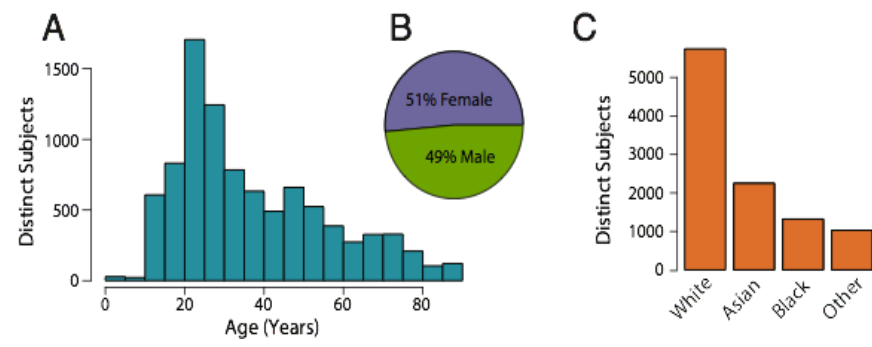
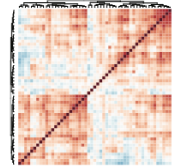


**Data Dissemination**

Web Interface for Data Visualization and Download



**Data Repurposing and Meta-analysis**



## Immunoassays

Immunoassays are a useful tool for obtaining precise measurements on the concentration of molecules, bacteria, and viruses inside of a sample. This technique relies on the strong bonding affinity between an antibody and its corresponding antigen. Specific antibodies are chosen to 'grab' the particle of interest. Immunoassays can also be used to measure the relative effectiveness of an antibody on a specific virus of interest. We provide immunoassays from ELISA, Multiplex ELISA, HAI Titer, and Virus Neutralization studies.

-  Transcriptomics
-  Proteomics
-  Immunoassays
-  Lab Tests

**Select Data Type:**

Normalized       Outliers

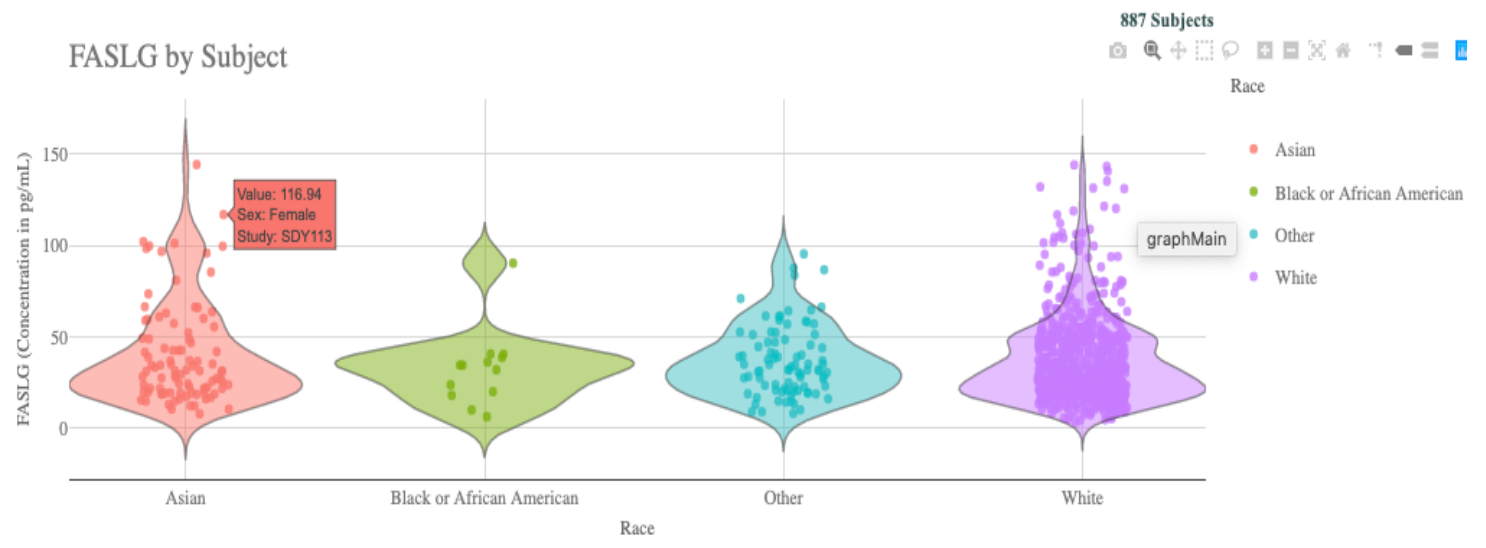
**Select Analyte:**

**Age Range:**

**Ethnicities:**  
 White  
 Black or African American  
 Asian  
 Other

**Sex:**  
 Female     Male

**Plot By:**  
 Age & Sex  
 Ethnicity  
 Study



**Download**

 Image
 Plot Data
 All Data

\* All Data is dataset's raw and formatted files



# IMMPORT

## Resources

Your site for documentation,  
training, and reference materials

# Tutorials and Links



## Using ImmPort APIs

ImmPort Application Program Interfaces can be used to query the Shared Data content, and download result files.



## ImmPort and ImmPort Data

Here are some resources to help understand the ImmPort data model, explore data in ImmPort and get started with your own re-analysis.

```
111000 101010 10101
001010 011101 10110
011010 110101 10001
```

## Basics of Programming

Useful links and tutorials exploring basics of programming, with a focus on Python, Jupyter Notebooks and R.



## Data Analysis

You have data, now what? Here are some data analysis tutorials in Python and R based on ImmPort data.



## Using ImmPort Galaxy

Still missing the old ImmPort Data Analysis interface? Here are some tips that might help you get used to the new ImmPort Data Analysis platform, ImmPort Galaxy.



## Preparing Data Upload

Uploading data to ImmPort doesn't have to be a challenge. Here are some resources to help you get started.

# Acknowledgements



Atul Butte, PI  
Director, Bakar Computational  
Health Sciences Institute, UCSF

## UCSF

Zicheng Hu  
Boris Oskotsky

## Peraton

Henry Schaefer  
Kate Hypes  
Zullinel Trilla-Flores  
John Campbell  
Srinivas Chepuri  
Ruth Monteiro  
Zhiping Gu  
Elizabeth Thomson  
Morgan Crafts

## Data Providers

## Funding Support

### NIAID

National Institute of Allergy and Infectious Diseases (NIAID)

Quan Chen  
Anupama Gururaj  
Dawei Lin

National Institutes of Health (NIH)  
Health and Human Services (HHS)  
Contract #: HHSN316201200036W