

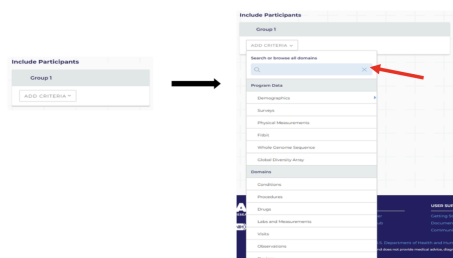
# DATA SET BUILDER\_AOU

## Creating the Case Cohort:

1. Go to the cohort builder within the AOU workbench.



2. Under Group 1 Search for sickling disorder due to hemoglobin S.



3. Go to group 2 and exclusion criteria and update with the following information:

A screenshot of the 'Sickle Cell with Hemo Labs' cohort builder interface. The interface is divided into several sections: 'Include Participants', 'And Exclude Participants', and 'Results by'. The 'Include Participants' section shows two groups: Group 1 with criteria 'Sickling disorder due to hemoglobin S' and Group 2 with criteria 'Hemoglobin [Massvolume] in Blood (Any value)'. The 'And Exclude Participants' section shows three groups: Group 3 with criteria 'Race - None Indicated Race - Skip Race - I Prefer Not To Answer', Group 4 with criteria 'Ethnicity - Skip Ethnicity - Prefer Not To Answer Ethnicity - No Matching Concept', and Group 5 with criteria 'Sex Assigned At Birth - Skip Sex Assigned At Birth - I Prefer Not To Answer Sex Assigned At Birth - None Sex Assigned At Birth - Unknown'. The 'Results by' section shows a bar chart of the total count (571) and a breakdown by race and age at CDB.

4. Preview your dataset

# DATA SET BUILDER\_AOU

Time for access renewal.  
23 days remaining.

Get started

Time for Controlled Tier access  
renewal. 23 days remaining.

Get started

DATA

ANALYSIS

ABOUT

All of Us Controlled Tier Dataset

These resources will get you started using the genomic data. Genomic data paths, genomic data formatting, and how to build a cohort.  
You can always learn more in the User Support Hub section on Genomics.

X

Datasets - Person, Survey, Demo, Zip, Labs

Build a dataset by selecting the variables and values for one or more of your cohorts. Then export the completed dataset to Notebooks where you can perform your analysis.

Select Cohorts (Participants)

Prepackaged Cohorts

☐ All Participants

Workspace Cohorts

☒ Sickle Cell with Hemo Labs

Select Concept Sets (Rows)

Prepackaged Concept Sets

☒ Demographics

☒ All Surveys

☒ The Basics

☒ Lifestyle

Select Values (Columns)

☒ Deselect All

Person

☒ person\_id

☒ gender\_concept\_id

☒ gender

Learn more in the data dictionary

Preview Dataset

A visualization of your data table based on concept sets and values you selected above.  
Once complete, export for analysis.

View Preview Table

Person	Survey	Zip_code_socioeconomic							
person_id	gender_concept_id	gender	date_of_birth	race_concept_id	race	ethnicity_concept_id	ethnicity	sex_at_birth_concept_id	sex_at_birth
1050502	45880669	Male	1938/06/15 00:00:00	8527	White	38003563	Hispanic or Latino	45880669	Male
1162746	45878463	Female	1938/06/15 00:00:00	8536	Black or African American	38003564	Not Hispanic or Latino	45878463	Female
7646687	45878463	Female	2002/06/15 00:00:00	8536	Black or African American	38003564	Not Hispanic or Latino	45878463	Female
6845346	45880669	Male	1935/06/15 00:00:00	8527	White	38003564	Not Hispanic or Latino	45880669	Male
2716495	45880669	Male	2002/06/15 00:00:00	8536	Black or African American	38003564	Not Hispanic or Latino	45880669	Male
1489704	45880669	Male	1963/06/15 00:00:00	8536	Black or African American	38003564	Not Hispanic or Latino	45880669	Male
821718	45880669	Male	1963/06/15 00:00:00	8536	Black or African American	38003564	Not Hispanic or Latino	45880669	Male
					Black or African				

SAVE DATASET

ANALYZE

Click analyze

Time for access renewal.  
23 days remaining.

Get started

Time for Controlled Tier access  
renewal. 23 days remaining.

Get started

DATA

ANALYSIS

ABOUT

All of Us Controlled Tier Dataset

These resources will get you started using the genomic data. Genomic data paths, genomic data formatting, and how to build a cohort.  
You can always learn more in the User Support Hub section on Genomics.

X

Your Analyses

ANALYZE ALL ANALYSES

Application	Name	Last Modified Time	Last Modified By
No results found			

Analyze Data

Select an application

FullYear

CLOSE

NEXT

Time for access renewal.  
23 days remaining.

Get started

Time for Controlled Tier access  
renewal. 23 days remaining.

Get started

DATA

ANALYSIS

ABOUT

All of Us Controlled Tier Dataset

These resources will get you started using the genomic data. Genomic data paths, genomic data formatting, and how to build a cohort.  
You can always learn more in the User Support Hub section on Genomics.

X

Your Analyses

ANALYZE ALL ANALYSES

Application	Name	Last Modified Time	Last Modified By
No results found			

Analyze Data

Select an application

FullYear

CLOSE

NEXT

Time for access renewal.  
23 days remaining.

Get started

Time for Controlled Tier access  
renewal. 23 days remaining.

Get started

DATA

ANALYSIS

ABOUT

All of Us Controlled Tier Dataset

These resources will get you started using the genomic data. Genomic data paths, genomic data formatting, and how to build a cohort.  
You can always learn more in the User Support Hub section on Genomics.

X

Your Analyses

ANALYZE ALL ANALYSES

Application	Name	Last Modified Time	Last Modified By
No results found			

Analyze Data

Select an application

FullYear

CLOSE

NEXT

Time for access renewal.  
23 days remaining.

Get started

Time for Controlled Tier access  
renewal. 23 days remaining.

Get started

DATA

ANALYSIS

ABOUT

All of Us Controlled Tier Dataset

These resources will get you started using the genomic data. Genomic data paths, genomic data formatting, and how to build a cohort.  
You can always learn more in the User Support Hub section on Genomics.

X

Your Analyses

ANALYZE ALL ANALYSES

Application	Name	Last Modified Time	Last Modified By
No results found			

Analyze Data

Select an application

FullYear

CLOSE

NEXT

# DATA SET BUILDER\_AOU

**All of Us**  
Research Program

ACU Sickle Cell Disease\_BCM

Time for access renewal  
An app is missing.

Time for Connected Tier access  
 renewal, 20 days remaining.

Get Connected

DATA

ANALYZES

ABOUT

88 of 100 Connected Tier Datasets (1)

These datasets will get you started using the genomic data, genomic data tools, genomic data formatting, and how to build a cohort. You can always learn more in the case Support from datasets on Genomics.

Your Analyses

NO RESULTS FOUND

Application	Name (1)	Last Modified Time (1)	Last Modified By (1)
	Search Name		

No results found

New Jupyter Notebook

Name:

Programming Language:  

Python 3

Julia

BACK

CANCEL

CREATE NOTEBOOK

**All of Us**  
Research Program

Quick Links

[Home](#)  
[About](#)  
[Genomic Resources](#)  
[Research Data](#)  
[How to Connect](#)

Need Support? Ask

[Contact Us](#)  
[Helpdesk](#)  
[Feedback](#)  
[Privacy Policy](#)

For all programming and data files, please refer to the [Data and Code Access Policy](#) for more information.

For all research and data files, please refer to the [Data and Code Access Policy](#) for more information.

These resources will get you started using the genomic data: Genomic data paths, genomic data formatting, and how to build a cohort. You can always learn more in the User Support Hub section on Genomics.

## Creating New Notebook:

CANCEL



Initializing notebook  
server, may take up  
to 5 minutes



Authenticating with  
the notebook server



Creating the new  
notebook



Redirecting to the  
notebook server



You are prohibited from taking screenshots or attempting in any way to remove participant-level data from the workbench.

You are also prohibited from publishing or otherwise distributing any data or aggregate statistics corresponding to fewer than 20 participants unless expressly permitted by our data use policies.

For more information, please see our [Data Use Policies](#).

Click Edit