AoU Guide: Cohorts & Dataframes for Condition Association Template

UPLOAD THE TEMPLATE TO THE WORKSPACE BEFORE BEGINNING!

Introduction

Welcome to the AoU guide tailored for the Condition Association Template! This step-by-step manual will walk you through the process of setting up specific cohorts and designing data frames using the AoU CohortBuilder and DatasetBuilder. By mastering these steps, you'll be perfectly equipped to harness the full potential of the Condition Association Template, enabling you to explore intriguing associations between any two conditions you select.

Outline of Instructions

Step 1: Define Your Conditions

- **Condition 1**: The primary condition you're interested in examining.
- **Condition 2**: An additional condition whose potential relation with the primary condition intrigues you.

Step 2: Set Up Cohorts

- **Case Cohort**: **Includes** individuals diagnosed with your primary condition (condition 1).
- **Control Cohort**: **Excludes** individuals with your primary condition (condition 1).
- **Condition Cohort**: **Includes** individuals diagnosed with the associated condition (condition 2).

Step 3: Create Data Frames

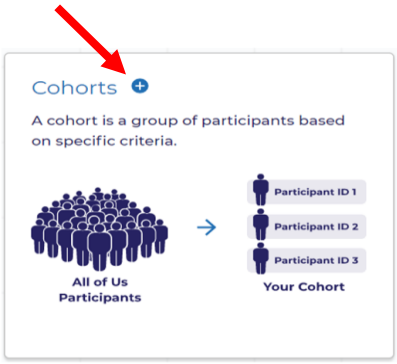
- For the case and control cohorts, you'll need data frames with the following variables:
    - **Using the DatasetBuilder**:
        1. **person_id**
        2. **date_of_birth**
        3. **race**
        4. **ethnicity**
        5. **sex_at_birth**

- Inside the template the following variables are auto-generated with pre-written code:
    - **Generated in Code Template:**
        1. **predictor**
        2. **outcome**
        3. **age**

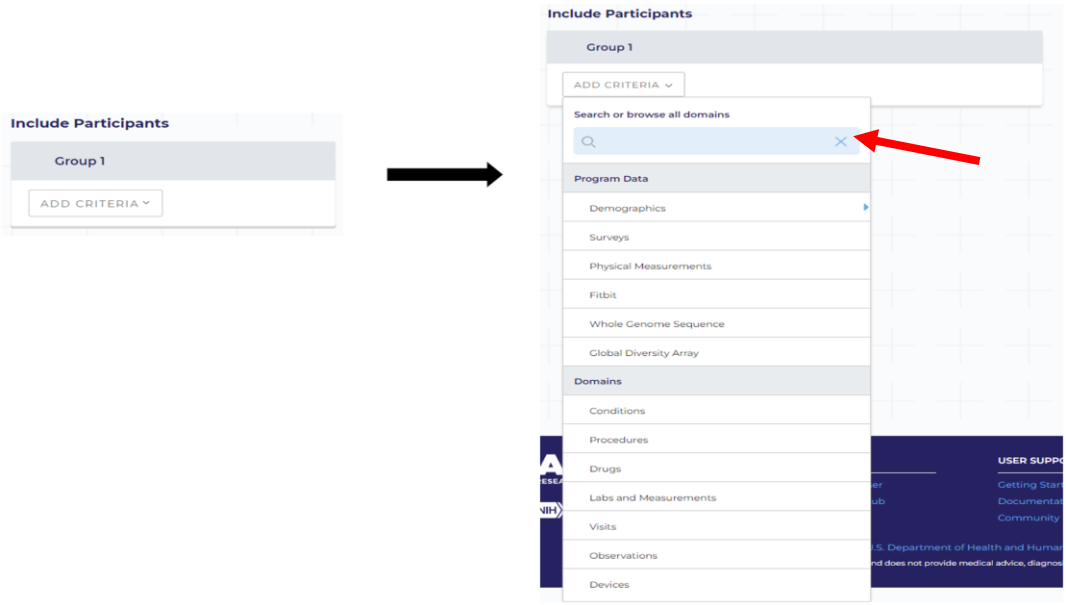
Define Conditions & Set Up Cohorts

Creating the case cohort:

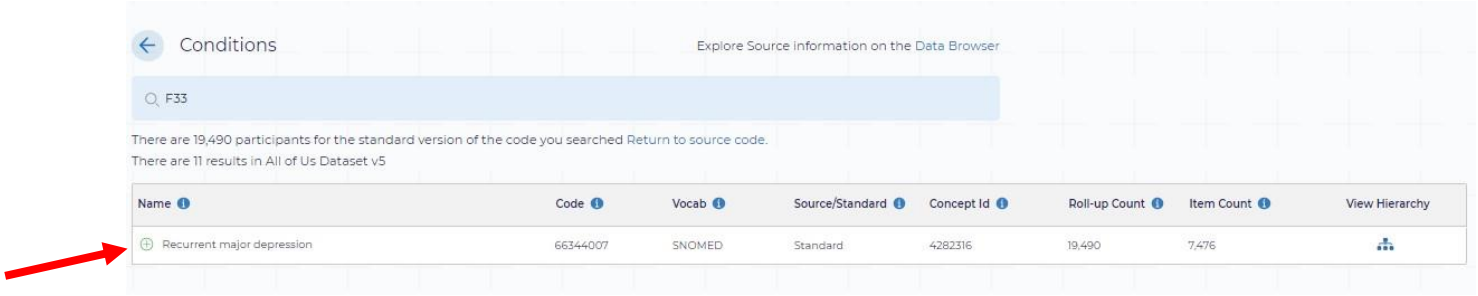
1. Go to the CohortBuilder in the All of Us Researcher Workbench.



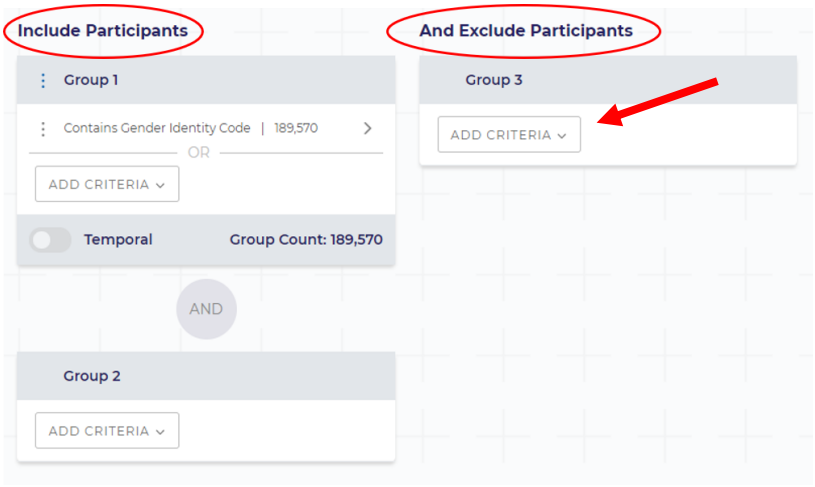
2. Under **Group 1**, search for **condition 1**.



3. Add **condition 1** to the inclusion criteria.



4. Go to exclusion criteria under **Group 3**, then select **Demographics**.
    a. For the variables **race**, **ethnicity**, and **sex_at_birth**, select the values you **do not** want included in the analysis (for example, you can exclude those who answered "Skip").
    b. DO NOT put the exclusion criteria for all variables under one group, or it will be interpreted as an 'or' statement.



5. Name and save the cohort.

Creating the control cohort:

1. Go to the CohortBuilder in the All of Us Researcher Workbench.
2. Under **Group 1**, select **Demographics,** then select **Age**.
    a. This will be a dummy inclusion criteria, so we can access the exclusion criteria.
    b. Keep the age range as it is and save the selection.
    c. The exclusion criteria should now be available.
3. Under **Group 3,** add **condition 1** to the exclusion criteria.
4. Go to exclusion criteria under **Group 4**, then select **Demographics**.
    a. For the variables **race**, **ethnicity**, and **sex_at_birth**, select the values you **do not** want included in the analysis (for example, you can exclude those who answered "Skip").
    b. **DO NOT** put the exclusion criteria for all variables under one group, or it will be interpreted as an 'or' statement.
5. Name and save the cohort.

Creating the condition cohort:

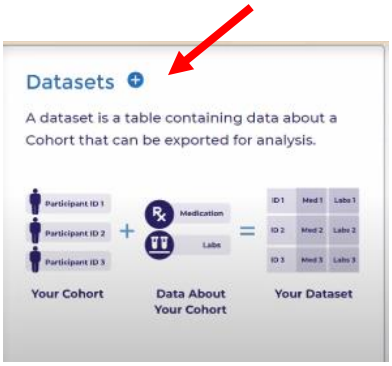
1. Go to the CohortBuilder in the All of Us Researcher Workbench.
2. Under **Group 1**, search for **condition 2**.
3. Add **condition 2** to the inclusion criteria.
4. Name and save the cohort.

Set Up & Export Data Frames

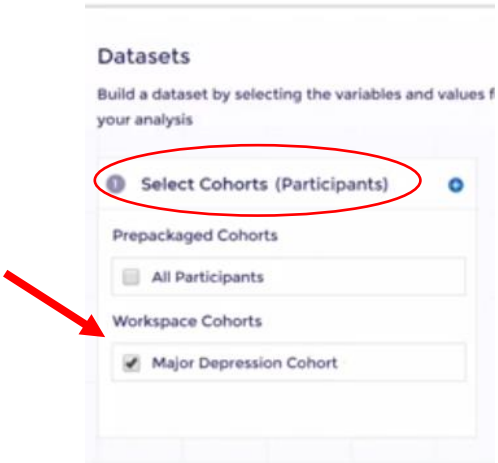
Please note that these data frames will be imported to the bottom of the template. You will have to move the cells up to the top of the notebook so they are the first things to run.

Creating a data frame for the case cohort:

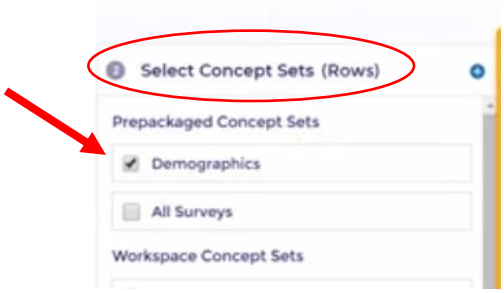
1) Go to the DatasetBuilder in the Researcher Workbench.



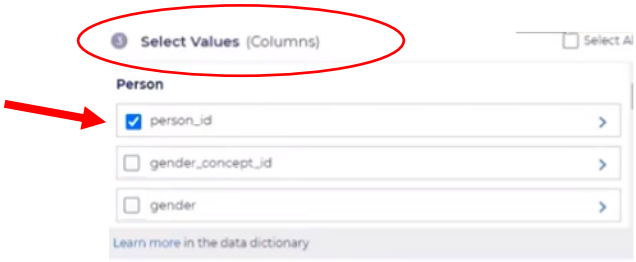
2) Under **Select Cohorts**, check the box next to the case cohort you created.



3) Under **Select Concept Sets**, check the prepackaged **Demographics** concept set.



4) Under **Select Values**, check the boxes for **person_id**, **sex_at_birth**, **date_of_birth**, **race**, and **ethnicity**. You **must** import these values for the dataset to work in the template.



5) Select **Create Dataset** and name your data frame. Select **Save**.

6) Select **Analyze**, choose the existing notebook "ConditionAssociation-R-Empty.ipynb" then select **Export**.

Creating a data frame for the control cohort:

1) Under **Select Cohorts**, check the box next to the control cohort you created

2) Under **Select Concept Sets**, check the prepackaged **Demographics** concept set

3) Under **Select Values (Columns)**, check the boxes for **person_id**, **sex_at_birth**, **date_of_birth**, **race**, and **ethnicity**. You **must** import these values for the dataset to work in the template.

4) Select **Create Dataset** and name your data frame. Select **Save**.

5) Select **Analyze**, choose the existing notebook "ConditionAssociation-R-Empty.ipynb" then select **Export**.

Creating a data frame for the condition cohort:

1) Under **Select Cohorts**, check the box next to the condition cohort you created.

2) Under **Select Concept Sets**, check the prepackaged **Demographics** concept set.

3) Under **Select Values (Columns)**, check the boxes for **person_id**. You **must** import this value for the dataset to work in the template.

4) Select **Create Dataset** and name your data frame. Select **Save**.

5) Select **Analyze**, choose the existing notebook "ConditionAssociation-R-Empty.ipynb" then select **Export**.