## Introduction

The Collaborative Informatics and Neuroimaging Suite Toolkit for Anonymous Computation, User Interface (COINSTAC) is software created to foster collaborative research by removing large barriers to traditional data-centric collaboration approaches. It is a desktop application that can be run on the major three operating systems (Windows, Mac, and Linux).

This document demonstrates how to run an analysis on COINSTAC. You will be guided through how to run a Single Shot Regression using the platform.

## Run an Analysis

### Log In

After you start COINSTAC, the first page you’ll see is the log-in screen, shown below. Enter your username and password and click the “LOG IN” button. If you are doing a test or demonstration, a username and password will be provided to you.

A screenshot of a cell phone

Description automatically generated

After you log in, you will see the COINSTAC home screen, as shown below.

A screenshot of a cell phone

Description automatically generated

### Create a Consortium

The next step is to create a consortium. A consortium is a group of users who run a decentralized analysis together. Click “Consortia” in the sidebar menu on the left, and you will be taken to the Consortia page. On this page, there is a card for each consortium containing its name, description, active pipeline, list of owner(s) and members, a button to view details about the consortium, and perhaps a button to join the consortium, as shown below.

A screenshot of a cell phone

Description automatically generated

If you are using COINSTAC in a group where you are not leading the analysis, you can find the consortium you want to join in the list and click the “JOIN CONSORTIUM” button or wait to be invited by the consortium owner.

If you are using COINSTAC by yourself or you want to lead the analysis, then click the Add Consortium button A picture containing object, first-aid kit

Description automatically generated, and you will be taken to the Consortium Edit page with the ABOUT tab open. Enter the name of your consortium and a short description. Then add all the people to the consortium who you would like to participate in the analysis. If you only want to run an analysis by yourself, then you don’t have to add yourself. Once you have entered all this information, click the SAVE button.

A screenshot of a cell phone

Description automatically generated

### Create a Pipeline

After you click the SAVE button on the ABOUT tab, you will be directed to the PIPELINES tab, as shown below. If you already have a pipeline, click OWNED PIPELINES and select your pipeline. If you do not already have a pipeline, click on the NEW PIPELINE + button.

A screenshot of a cell phone

Description automatically generated

You will be directed to the Pipeline Creation page, shown below. As the consortium owner and pipeline creator, you define the model that will be used in the decentralized analysis. You choose the type of analysis (the computation) and then specify what variables are in the model.

First, you should fill out the name and description of the pipeline. Next, add a step to the pipeline by clicking the ADD COMPUTATION STEP button. You can select any computation you would like, but in this tutorial, we select Single Shot Regression.

A screenshot of a cell phone

Description automatically generated

Click on the computation (Single Shot Regression) to define the model. You will generally see both an Input and Output section. The Input section lets you define your model, and the Output section lets you know the format of the results. For Single Shot Regression, you will see covariates, data, and lambda under Input Parameters. Covariates are the independent variables, data are the dependent variables, and lambda is the regularization parameter for the ridge regression. The Outputs section shows you what to expect in the results when the computation completes. In this example, you’ll see a list (array) of regression results, each of which has the region of interest, the global statistics of the shared model across sites, and the local statistics of your site’s model.

### A screenshot of a cell phone Description automatically generated

In this example, I’ve added two covariates, isControl and age. The former is a Boolean variable that indicates whether the subject is a control, and the latter is a number that indicates the age of the subject. For isControl, I clicked on the DATA TYPE button and chose BOOLEAN. Then, I clicked on the DATA SOURCE button and chose FILE. Finally, I entered the name of the variable as “isControl.” I followed a similar procedure for age. The data (dependent variables) are brain volumes that are outputs of the Freesurfer software, namely, the 3rd ventricle and the brain stem. I’ve also set the regularization parameter lambda to 0, which means no regularization will occur.

Note that the selection of multiple dependent variables does not create a single multivariate regression model. Instead, one model is created for every dependent variable chosen. For example, with the above parameters, we have the following simplified[[1]](#footnote-1) models:

### A screenshot of a cell phone Description automatically generated

### After you have entered all the details about the pipeline, press the SAVE PIPELINE button near the top right corner of the screen. At any time, you can view your pipeline by clicking on Pipelines in the sidebar on the left. As shown below, you will see your pipeline in the list and can either view details of your pipeline or delete it.

### A screenshot of a cell phone Description automatically generated

### Set Active Pipeline

Next, you must attach the pipeline you just created to your consortium. Click Consortia in the sidebar, find your consortium, and click the SET ACTIVE PIPELINE button, as shown below.

A screenshot of a cell phone

Description automatically generated

### Map Data

Now that you have created your consortium and pipeline, you and everyone else in your consortium has to map their data to the variables in the model you created. Users who have not joined the consortium must do so at this point. After joining the consortium, click on Maps in the sidebar. You will be taken to the Maps page, where you can see your consortium. Click on the MAP DATA TO CONSORTIA button.

A screenshot of a cell phone

Description automatically generated

You will be directed to the page shown below, which has the variables you specified in your pipeline on the left and a file selection area on the right.

The regression computation requires a CSV file containing covariates for each subject and names of local files, each of which is a text file listing brain volumes for each subject.

To save time, we encourage you to use our set of test data for initial testing purposes. Please download test data [here](https://github.com/MRN-Code/coinstac/releases/download/v4.0.2/freesurfer-test-data.zip). After downloading this file, extract the contents to a folder on your computer.

Click on the ADD FILES GROUP button on the right. You will see a dialog box that asks you to choose a file. Choose a CSV file in the abovementioned format to continue. If you are using the test data, click on any of the site folders and then click on the CSV file in that folder.

A screenshot of a cell phone

Description automatically generated

After you have chosen your file, the right side of the screen will be populated by information from the file, including name, date, extension, file path, and first row (header), as shown below. You will also see rounded buttons for each of the local variables (column headers) found in your CSV file. You can either click and drag each local variable to a variable in the model, or you can click the AUTO MAP button, which will automatically link your column headers to the model variables that have the closest names. This page allows for variation among sites in data preparation. In other words, every site can give the columns in their data files different names and can still participate in the analysis.

After you have mapped your data, the AUTO MAP button will display SAVE. Click it again to save your mapping. It will then change to BACK TO CONSORTIA button, and if you click it again, you will be taken to the Consortia page.

### A screenshot of a cell phone Description automatically generated

### Start Pipeline

After everyone has mapped their data, the consortium owner can then start the pipeline by going to the Consortia page, finding their consortium in the list, and clicking the START PIPELINE button.

A screenshot of a cell phone

Description automatically generated

### View Progress

To view the progress of the pipeline, click on Home in the sidebar. You will see a card for the pipeline you just started. It will list the name of the consortium and name of the pipeline at the top, along with a stopwatch that keeps track of time the pipeline has been running. Near the top of the card, you’ll also see the status, start date, and clients. Within that card, you’ll see one or two cards, one for the state of the pipeline on your computer, and one for the state of the pipeline on the COINSTAC remote server. Inside each of those cards, you’ll see the mode, state, iteration, and step count. At the bottom of the pipeline card, you’ll see a VIEW PIPELINE button that lets you see details of the pipeline and a STOP PIPELINE button that lets you stop the pipeline.

### A screenshot of a cell phone Description automatically generatedView Results

After your pipeline has completed, you can view the results by clicking on Results in the sidebar. You will be able to see statistics on both your local model and the global model for each dependent variable you chose.

The images below show what an example results page looks like. The first screen, shows two tables, the global statistics and local statistics for the 3rd ventricle. The coefficient, p-value, and t-statistic for each model coefficient are presented in the table, along with the degrees of freedom and the value of R2. The second screen shows the remainder of the second table and two more tables on the local and global statistics for the brain stem. If there were more than one member in this example, there would also be results for additional sites for both the 3rd ventricle and the brain stem.

If you would like, you can also download these tables to a CSV by clicking the DOWNLOAD TO CSV button.

A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated

1. This is a simplified model to give the reader a basic idea about the model that the software generates. The real model is decentralized and thus more complicated. For more details, see [Plis et al. 2016](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4990563/). [↑](#footnote-ref-1)