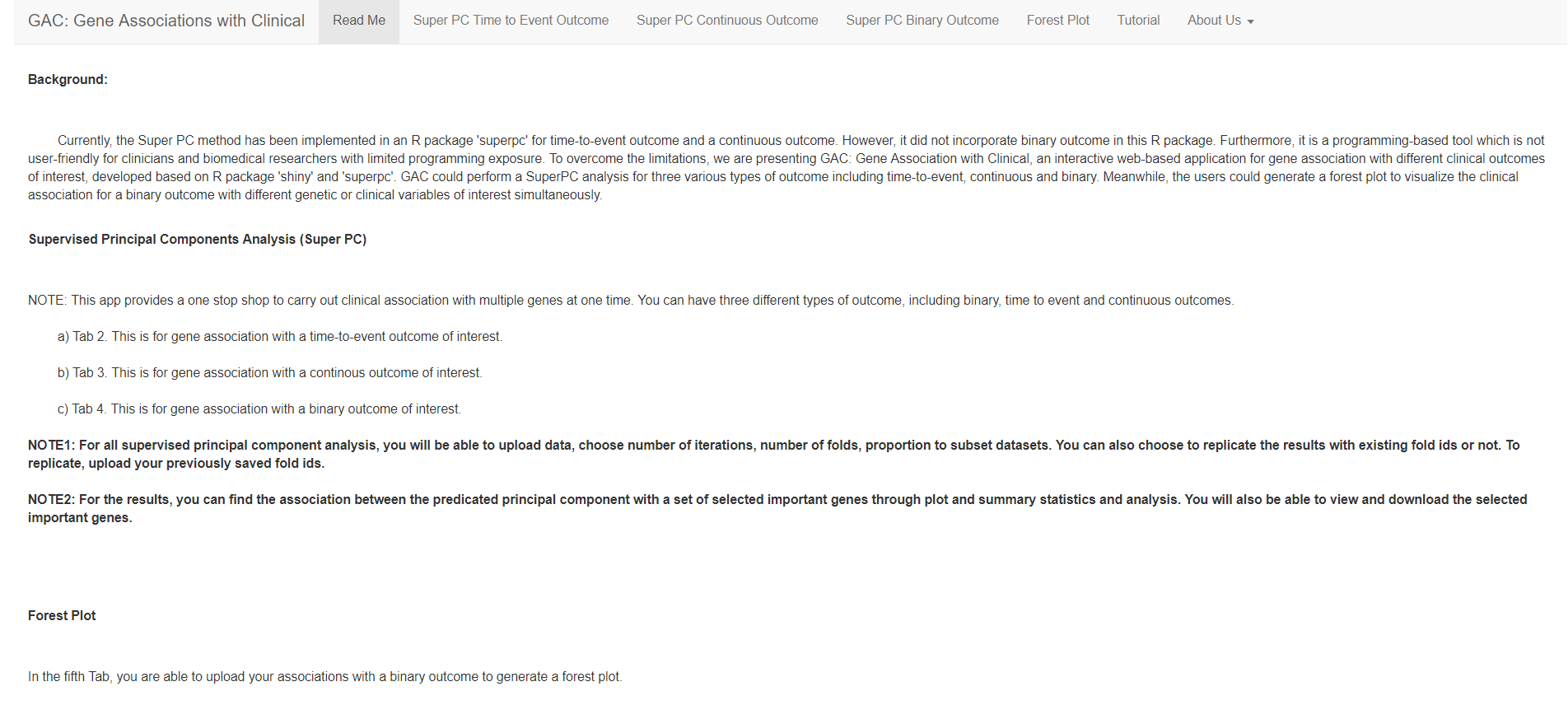
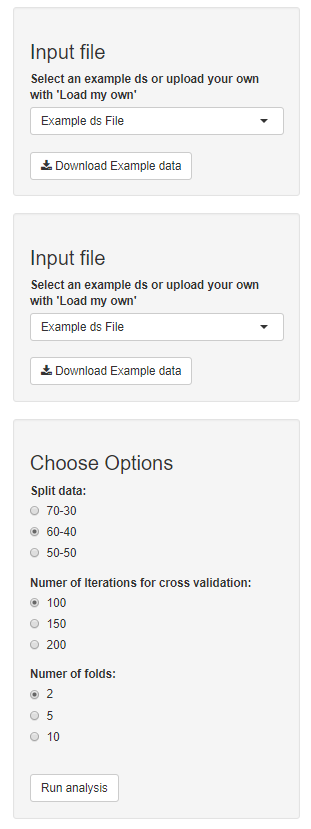
GAC Tutorial

**Tab 1: Read Me**



**Tab 2: Super PC Time to Event Outcome**





Step 4:

Hit button to run analysis after each change in option.

Step 3:

Choose data split into training and validation. Default uses 60-40 i.e. data is split into 40 % for training and remaining 60% for validation. You can also choose the number of interactions. For example purposes a smaller number 100 is used.

Also, select number of folds for cross validation. Default is 2.

Step 2:

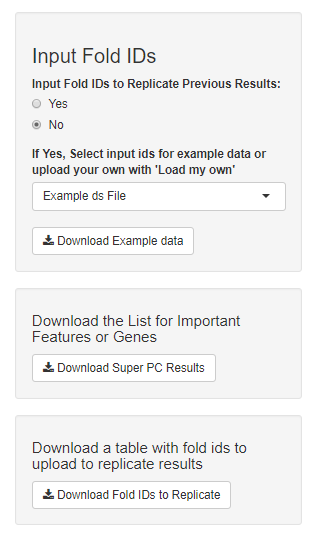
Select example ‘expression’ dataset or upload your own. The patients should be ordered in same order and should be exact same.

To view example data and format, use download button to view .csv file.

Step 1:

Select example ‘Time to event’ clinical dataset or upload your own.

To view example data and format, use download button to view .csv file.



Step 7 (Optional):

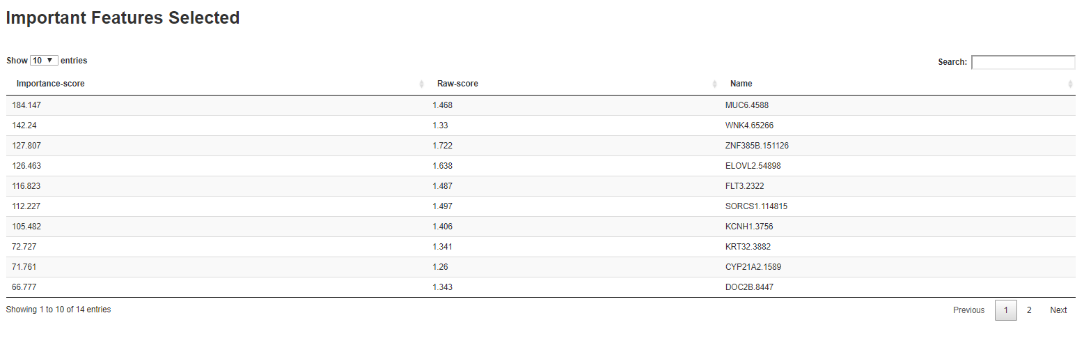
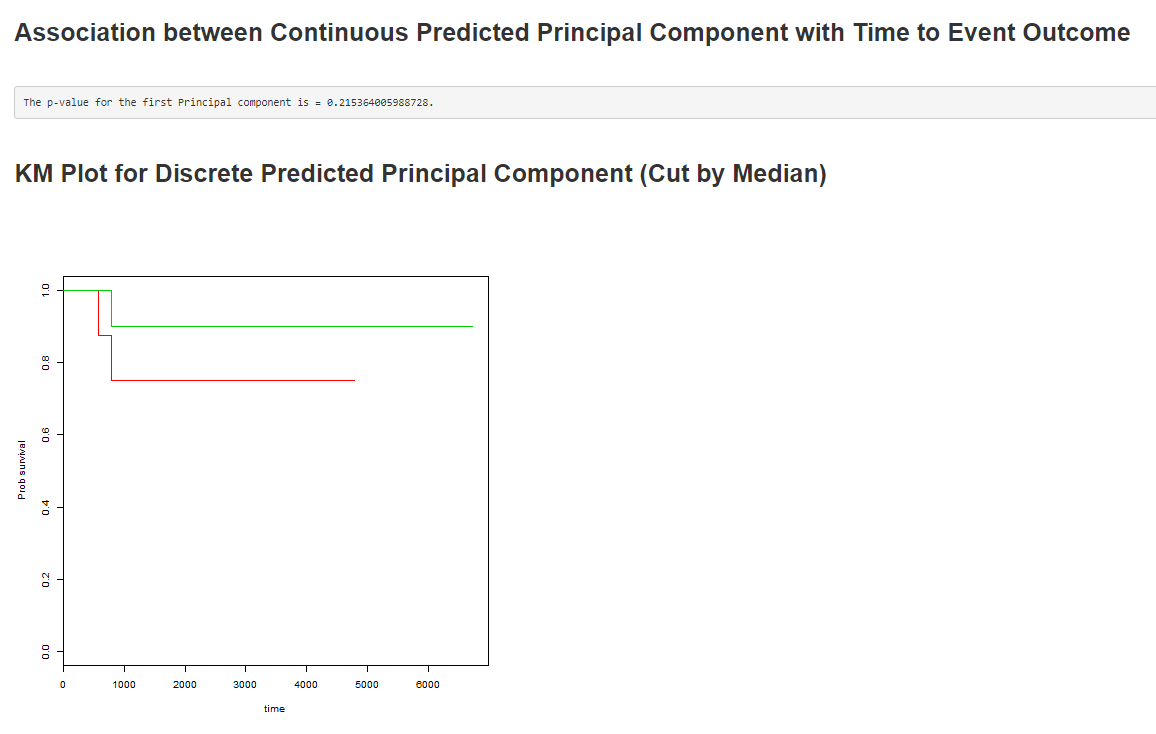
The user can download the fold-ids in csv format to replicate results at a later time. These can be uploaded in the input fold id section.

Step 6 (Optional):

The user can download the tabular results in .csv file.

Step 5 (Optional):

The user can replicate previously generated results by uploading fold-ids (download available below). When running analysis for first time, leave option to ‘No’.



List of all significant genes, in order of decreasing importance score are reported

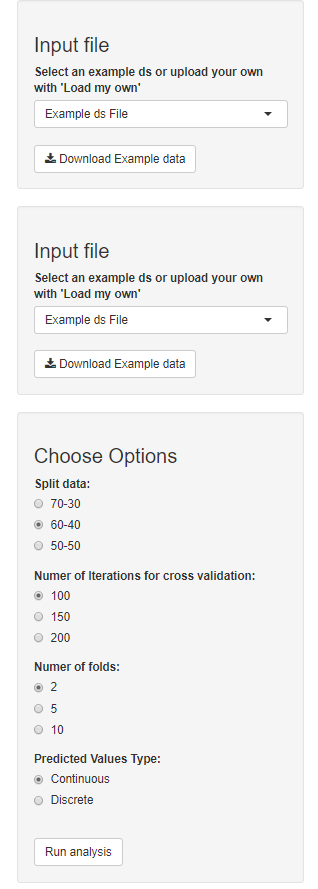
Main panel shows the results from the super PC analysis.

The p-value from 1st supervised PC is reported.

A discrete (categorical) predictor is created by cutting the predictor at its median to form two groups for the Kaplan-Meier analysis.

**Tab 3: Super PC Continuous Outcome**





Step 4:

Hit button to run analysis after each change in option.

Step 3:

Choose data split into training and validation. Default uses 60-40 i.e. data is split into 40 % for training and remaining 60% for validation. You can also choose the number of interactions. For example purposes a smaller number 100 is used.

Also, select number of folds for cross validation. Default is 2.

Additionally, choose to display predicted values as scatter plots (for continuous predictors with Pearson correlation’s) or boxplot (for binary groups cut off at median with t-test).

Step 2:

Select example ‘expression’ dataset or upload your own. The patients should be ordered in same order and should be exact same.

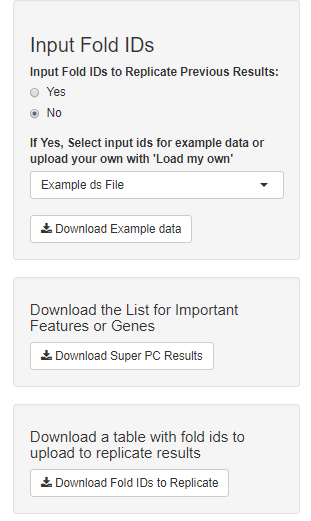
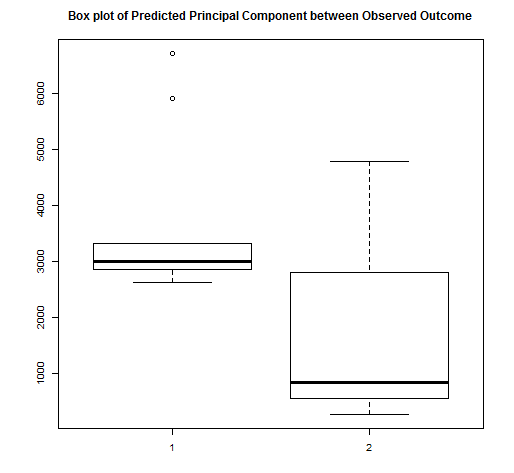
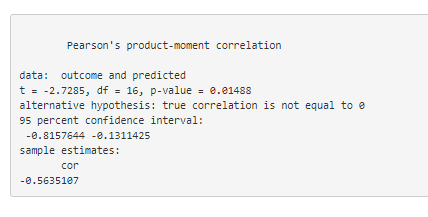
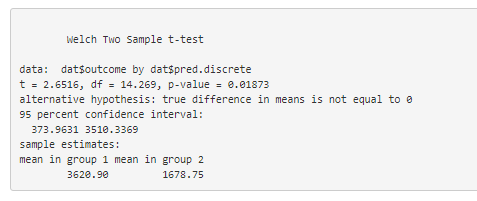
To view example data and format, use download button to view .csv file.

.

Step 1:

Select example ‘continuous’ outcome dataset or upload your own.

To view example data and format, use download button to view .csv file.

Step 7 (Optional):

The user can download the fold-ids in csv format to replicate results at a later time. These can be uploaded in the input fold id section.

Step 6 (Optional):

The user can download the tabular results in .csv file.

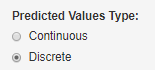
Step 5 (Optional):

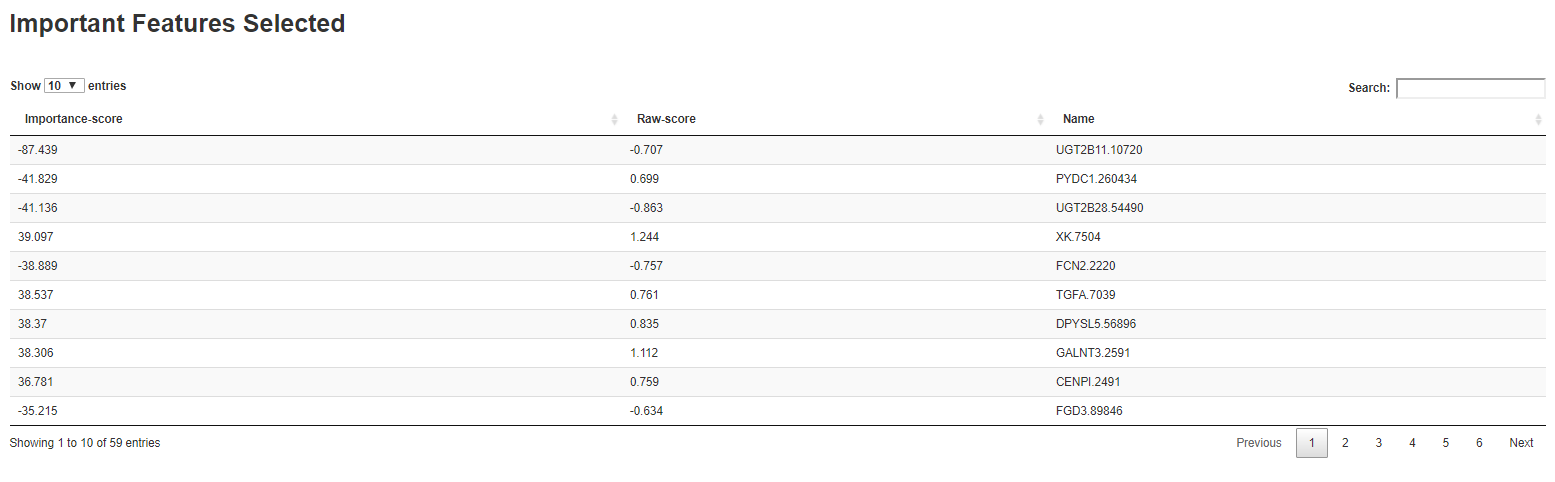
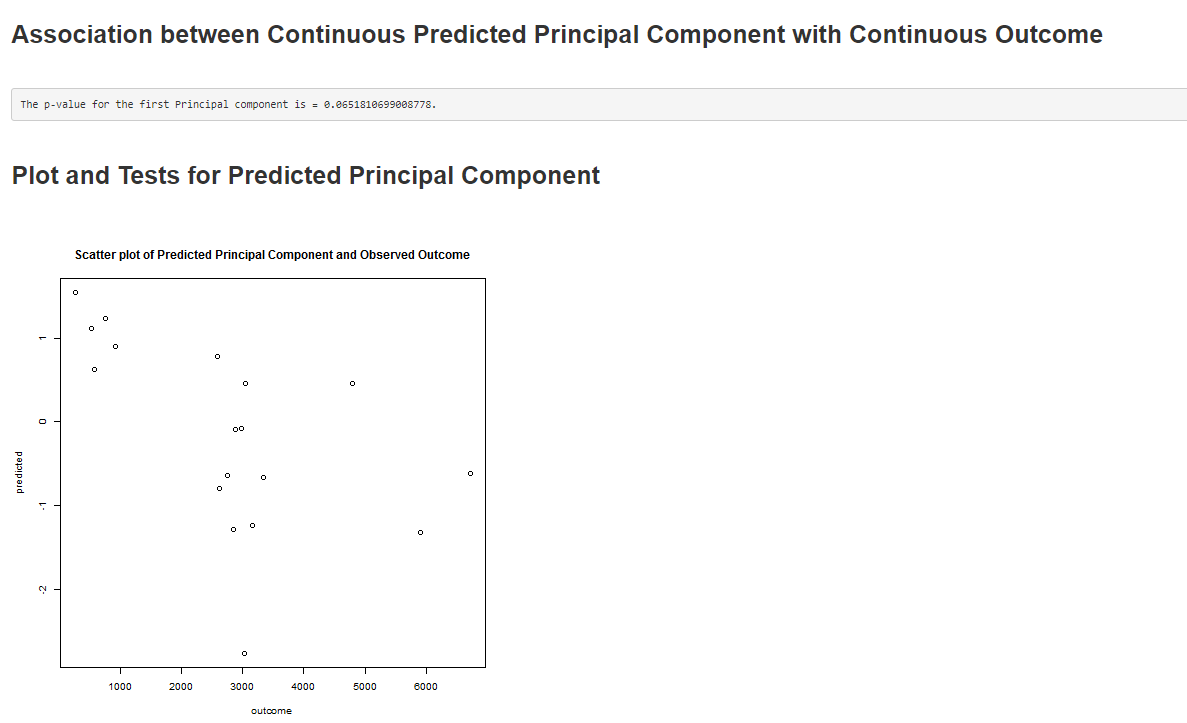
The user can replicate previously generated results by uploading fold-ids (download available below). When running analysis for first time, leave option to ‘No’.

Main panel shows the results from the super PC analysis.

The p-value from 1st supervised PC is reported.

Scatter plot showing correlation of continuous predictor with outcome using Pearson correlation is reported. Discrete predictors are created by cutting the predictor at its median to form two groups and reported as boxplots.

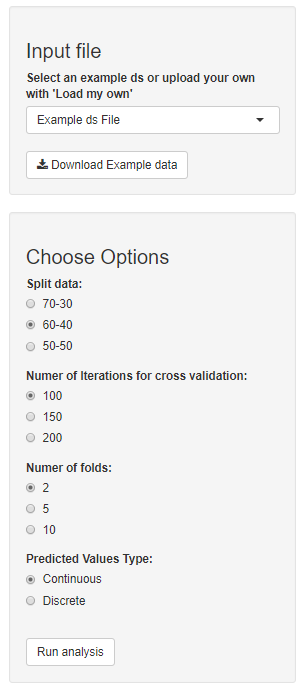


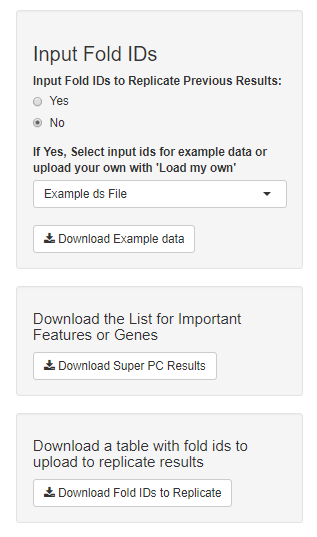
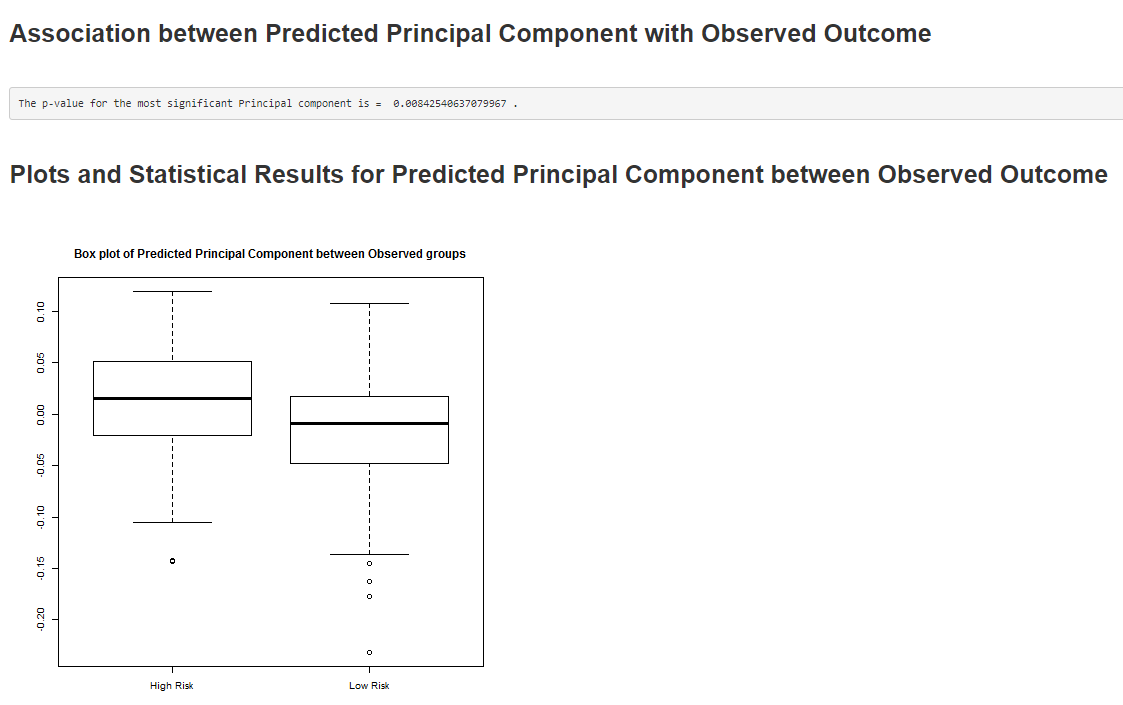
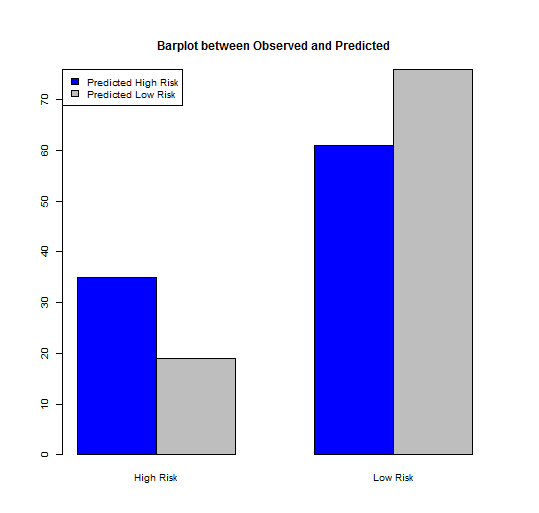
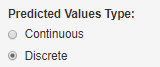
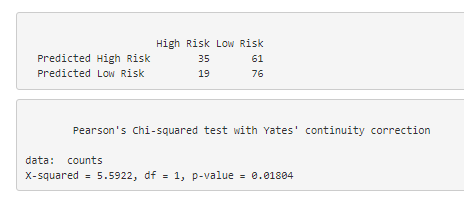
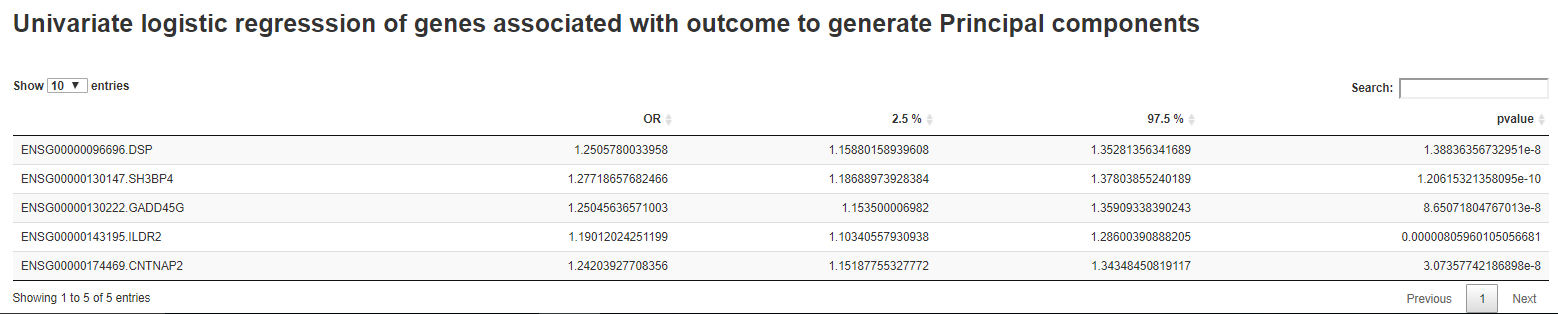
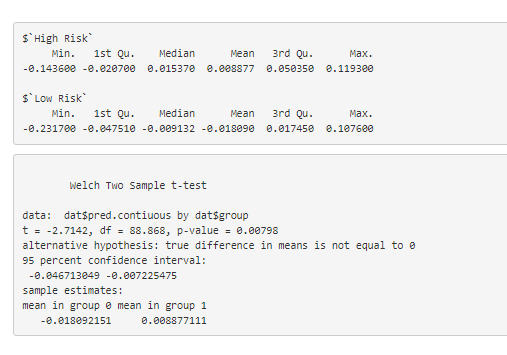
 

List of all significant genes, in order of decreasing importance score are reported

**Tab 4: Super PC Binary Outcome**

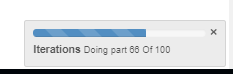




Step 3:

Hit button to run analysis after each change in option. A progress indicator may appear to the right corner of the page.



Step 4 (Optional):

The user can replicate previously generated results by uploading fold-ids (download available below). When running analysis for first time, leave option to ‘No’.

Step 6 (Optional):

The user can download the fold-ids in csv format to replicate results at a later time. These can be uploaded in the input fold id section.

Step 5 (Optional):

The user can download the tabular results in .csv file.

Step 2:

Choose data split into training and validation. Default uses 60-40 i.e. data is split into 40 % for training and remaining 60% for validation. You can also choose the number of interactions. For example purposes a smaller number 100 is used.

Also, select number of folds for cross validation. Default is 2.

Additionally, choose to display predicted values as boxplots (for continuous predictors with t-test) or bar plot (for discrete groups cut off at median with chisq test).

Step 1:

Select example ‘expression with binary outcome indicator’ dataset or upload your own.

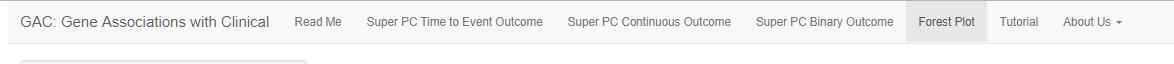
To view example data and format, use download button to view .csv file.

Main panel shows the results from the super PC analysis.

The p-value from 1st supervised PC is reported.

Scatter plot showing correlation of continuous predictor with outcome using Pearson correlation is reported. Discrete predictors are created by cutting the predictor at its median to form two groups and reported as boxplots.

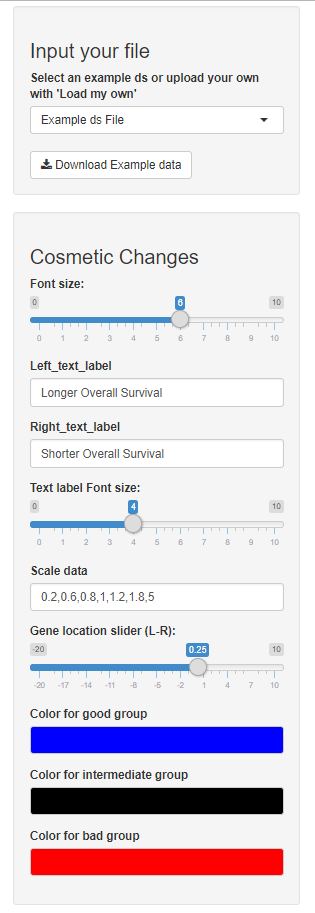
List of all significant genes, in order of decreasing importance score are reported

**Tab 4: Forest Plot**

Step 1:

Select example dataset or upload your own.

To view example data and format, use download button to view .csv file.



Step 8:

Change color of outcomes. Genes with HR and their 95% CI below 1 are coded blue (good group), overlapping 1 are coded black (the intermediate group) and those over 1 are coded red (the bad group).

Step 7:

Move IDs (genes) slightly to right or left to avoid overlap with the forest plot.

Step 4:

Choose x ticks to display below the horizontal line depending on dataset.

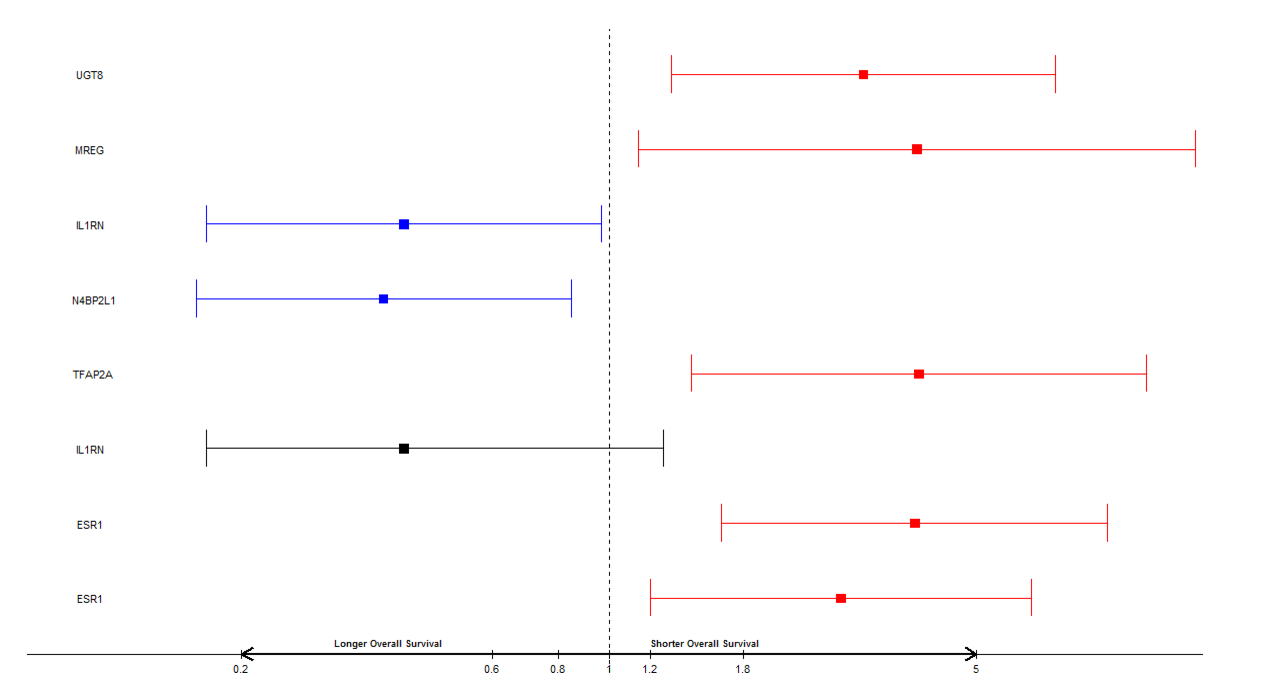
Step 3:

Change label for text to be displayed above the horizontal line to the left and right of the HR= 1 dotted line.

A slider bar to change font size for these labels is also available.

Step 2:

Change font size for the IDs (genes) displayed to the left of forest plot.



Reporting results from super PC analysis through forest plot. Genes with HR < 1 are coded in blue, those overlapping 1 are coded black and those greater than 1 are coded red.