User Manual of "Biomarker Explorer" R-Shiny Web App

Analytical Solution Groups

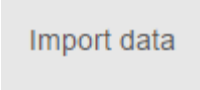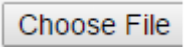Contact: Qinghua.song@gilead.com

1. Description:

This R-shiny web app was developed at Gilead Sciences, Inc. for biomarker statisticians and scientists to generate graphs and tables for biomarker exploratory analysis including time profiling and association, in an interactive platform. The app can be accessed via http://rshiny.gilead.com/dev/biomarker_explorer/.

The app consists of multiple navigation-bar pages with each fulfilling distinct purposes as described below

- **Import data**: the user uploads the data file in this page.
- **Time profiling**: the user explores the pattern of certain variable of interest over time through some most commonly used types of graphs.
- **Association**: the user explores the distribution of one variable or association relationship between two variables in this page.
- **Output results**: the user prepares the parameter specifications of output graph in a foot and title (TNF) file of Excel format and obtains all the desired graphs in a zipped file.

2. User Direction:

Step 1. Data File Upload

- Click **Import data** on the navigation bar on the top of the platform to import data.

- Click the **Choose File** button to select and upload user's data file onto the web-based platform with file extension being one of the following:
  - RData, R saved workspace data file. Note that only one data object is allowed in the saved workspace for the platform to correctly identify the data to use
  - CSV/TSV, delimiter-separated files
  - XLS/XLSX, Excel file
  - TXT, plain text file
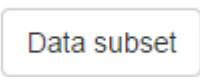 - sas7bdat, SAS data file

To ensure that the application runs correctly, the data file the user uploads must satisfy the following requirements (You may click the link "Need input data template?" in

Choose file to import (Need input data template?) to download an input data template):

- ❖ In data frame format
- ❖ The following columns are required
  - o STUDYID (study id)
  - o USUBJID (subject id)
  - o PARAM (biomarker names)
  - o AVAL (biomarker measurements)
  - o BASE (biomarker baseline measurements)
  - o CHG (biomarker change from baseline measurements)
  - o PCHG (biomarker percentage change from baseline measurements)
  - o PRCHG (biomarker proportion of baseline measurements)
  - o XTICKER (numeric visit time, e.g. 6)
  - o XTICKERL (nominal visits containing time unit, e.g. 'Week 6')
- ❖ The following columns are optional
  - o UNIT (biomarker measurement unit)
- ❖ Missing entries should either be left empty or marked '**NA**'

Once the uploading is complete, a button [⊞ View top 10 rows] will appear on top of the main panel. Click it to view top 10 rows or more.

Step 1.1 Data filtering

Once the data uploading is complete, a button [Data subset] will appear on top of the main panel. Click it to open a modal window to perform data filtering as follows.

- Choose a column name from the drop-down list of **Variable 1**
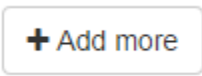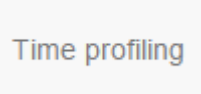- Corresponding, a numeric slider input or a drop-down list will appear depending on the continuity of the selected column in **Variable 1**. Rows will be filtered according to the interval determined by the numeric slider input or the categories by the drop-down input list.
- Click [+ Add more] to apply more filtering criterions or [✔ Done] to finish the data filtering process.

- Click **⟳ Reset** to clear all applied data filtering criterions.

Step 2 Visiualization

Step 2. 1 Time profiling

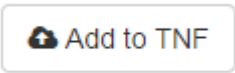- Click **Time profiling** on the navigation bar on the top of the platform.
- Click the "Study" bar to select a specific Study or Studies of interest. If the user wants to compare time-profiling patterns among multiple studies, he/she can leave this blank and choose "STUDYID" in the "Group" select input.
- Click the "Biomarker" bar to select the Biomarker of interest. One Biomarker can be selected at each.
- Click the "Y variable" to select the one of the biomarker-dependent Y variables in the time profiling graph:
  - AVAL
  - BASE
  - CHG
  - PCHG
  - PRCHG
- Click "Graph Type" to select one of the graph types:

  - Spaghetti plot
  - Mean + SE
  - Mean + SD
  - Median + IQR
  - Boxplot

Once the "Study", "Biomarker", "Y Variable" and "Graph Type" are specified, a graph will be displayed correspondingly on the main panel. Please note that data of all studies will be displayed if the "Study" input is left blank.

- Click "Group" to select a group variable (such as "SEX", "DISEASE TYPE", etc) over which the plots can be plotted. This is optional.
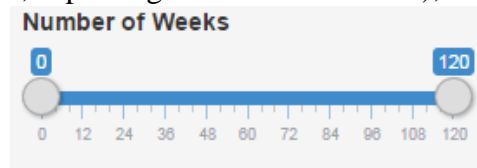
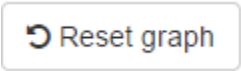Once the user is satisfied with the shown time-profiling graph and would like to save it to the final outputs, the user can click **☁ Add to TNF** button. Consequently, the current specification of the graph will be added as a new row to the TNF file, which can be easily previewed or downloaded in "Output results" page.

There are several ways that the user can interact with the static plot generated by specifying the below input widgets.
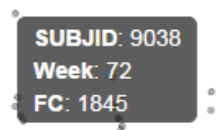
- **Zooming:** typically instead of the entire scope of the graph, one is interested in certain sub-region. This can be easily accomplished in one of the following two manners.
  - o **Slider bar:** Move the end point in the slider bar below, "Number of XX" (XX can be Weeks, Days, and etc., depending on the units of time), to change the range of X axis
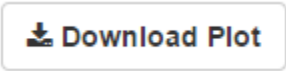

  - o **Brush:** Move the mouse on the plot, click, hold and drag, to zoom in to the region to be focused. To go back, double click on the graph or click beneath the graph.
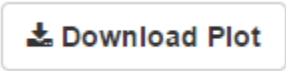- **Tooltip:** when the mouse hovers over a data point, a small semi-transparent panel will appear at the lower right of the mouse displaying crucial information of the data point.
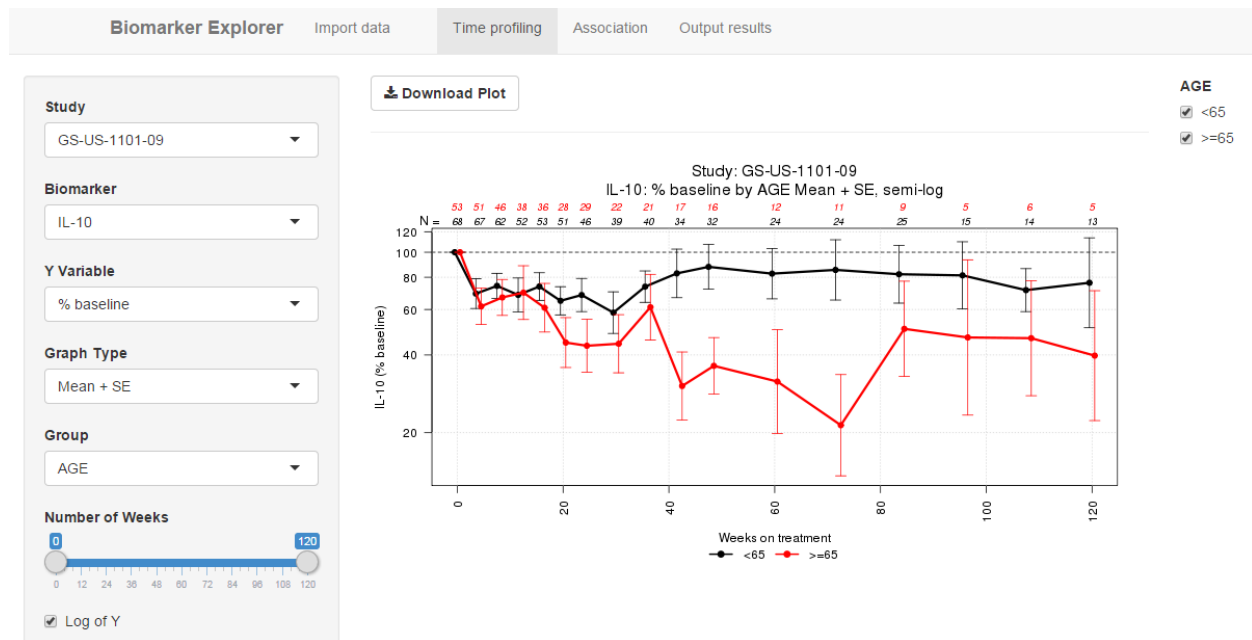

- **Log scale:** click "log of Y" checkbox, then the scale of Y will be transformed into semi-logarithm.

Finally, when the user is satisfied with the graph and wants to keep a physical copy, he/she can

- Click the on the left top of graph to download the plot into a pdf file. Note: If you'd like to download the plot in a specific area, you could use the brush tool, zoom in and download the plot before resetting the graph.

Example 1: Mean+SE plot (Graph Type=Mean + SE) of the % baseline (Y Variable= % baseline) of Biomarker IL-10 (Biomarker=IL-10), grouped by age (Group=AGE), semi-log

Step 2. 2 Association

- Click **Association** on the navigation bar on the top of the platform.
- Click the type of Association from:

Association type

- ● Continuous vs Continuous
- ○ Continuous vs Categorical
- ○ Categorical vs Categorical

   - o For "Continuous vs Continuous", a scatter plot of continuous Variable X vs continuous Variable Y will be generated
   - o For "Continuous vs Categorical", a box plot of continuous Variable Y vs categorical Variable X will be generated
   - o For "Categorical vs Categorical", a contingency table of categorical Variable Y vs categorical Variable X will be generated

- Click "Study" bar to select a specific Study or Studies of interest. If the user wants to compare time-profiling patterns among multiple studies, he/she can leave this blank and choose "STUDYID" in the "Group" select input.
- Click the "Variable X" bar to select Variable X;
- Click the "Variable Y" bar to select Variable Y
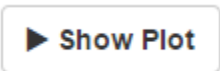
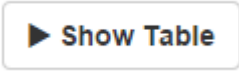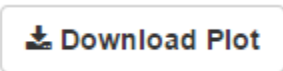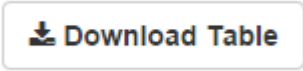Once Variable X or Variable Y is selected, a check box

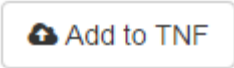☐ Is variable X biomarker-dependent?   or   ☐ Is variable Y biomarker-dependent?   will appear, check the box
if Variable X or Variable Y is biomarker dependent. Example of biomarker dependent
variables (variables whose values are related to biomarkers) are

- o AVAL: Measurement of Biomarker value
- o BASE
- o CHG
- o PCHG
- o PRCHG

- Click "Biomarker X" or "Biomarker Y" and select one biomarker for each if "Variable
  X" or "Variable Y" is biomarker dependent
- Click "Visit X" and "Visit Y" to select Visits for association graph/table. One or
  multiple visits can be selected

- Click   ▶ Show Plot   or   ▶ Show Table   on left top of main panel to generate
  an association graph or table, respectively.

- Click the   ⬇ Download Plot   or   ⬇ Download Table   on the left top of graph
  to download the plot or table into a pdf file.

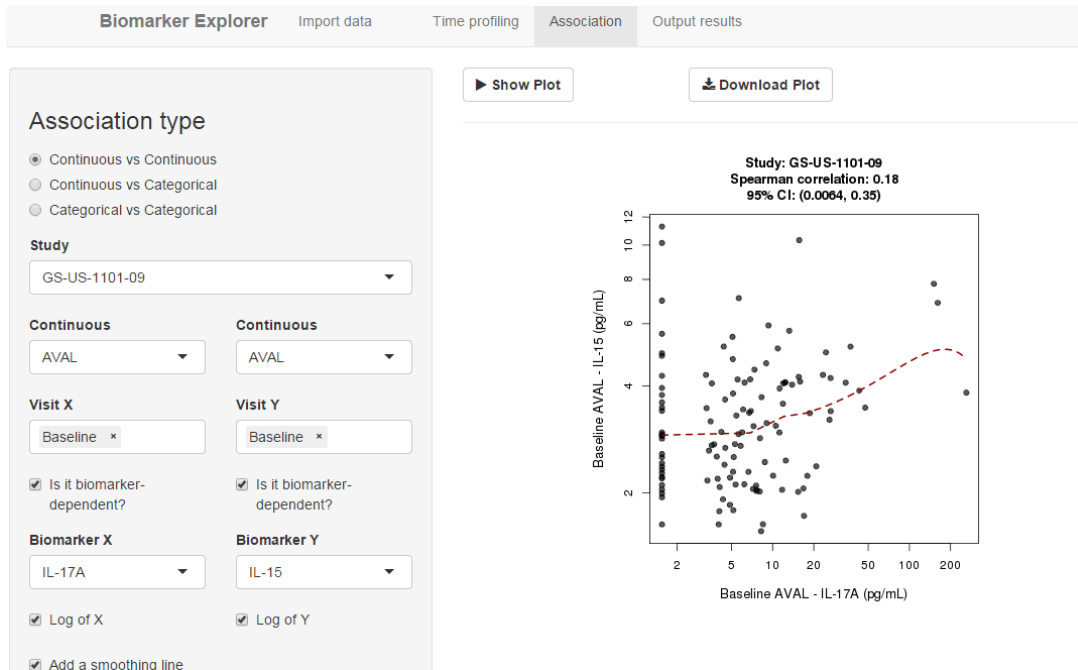Once the user is satisfied with the shown association graph or table and would like to save it to

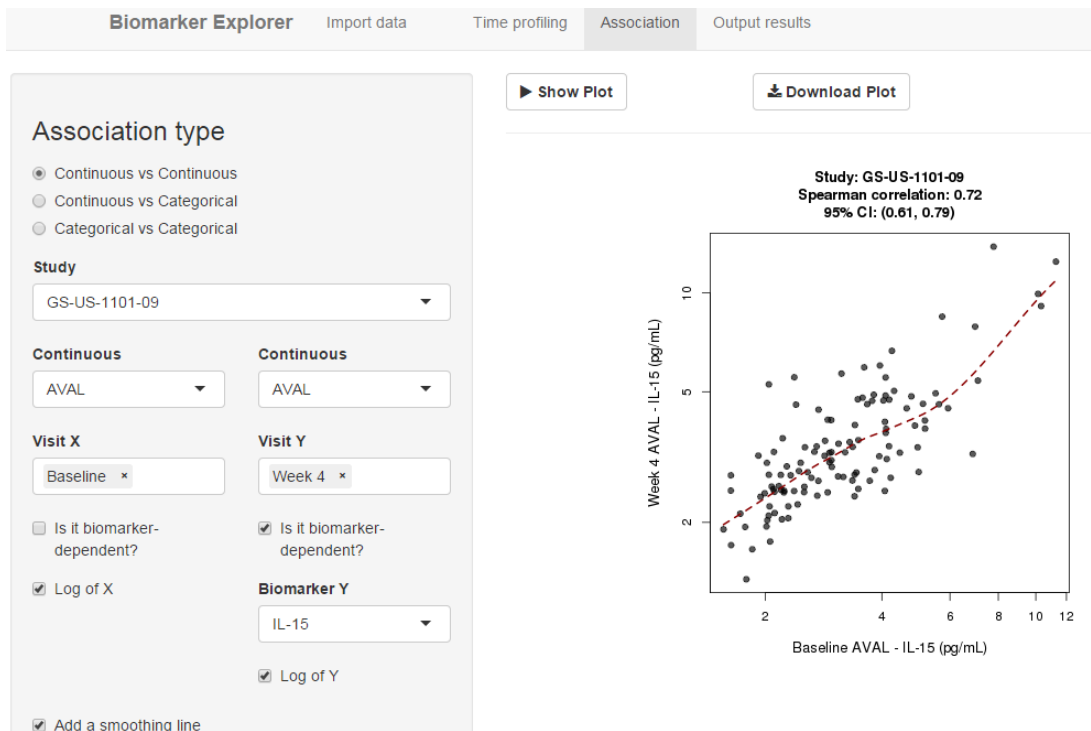the final outputs, the user can click   ⬆ Add to TNF   button. Consequently, the current
specification of the graph or table will be added as a new row to the TNF file, which can be
easily previewed or downloaded in "Output results" page.

Example 2.1 (Continuous vs Continuous): Association of Biomarker IL-15 (Variable Y=AVAL, Biomarker=IL-15) vs Biomarker IL-17A (Variable X=AVAL, Biomarker=IL-17A) at Baseline (Visit Y =Baseline, Visit X=Baseline) with log on X and Y, adding a smoothing line
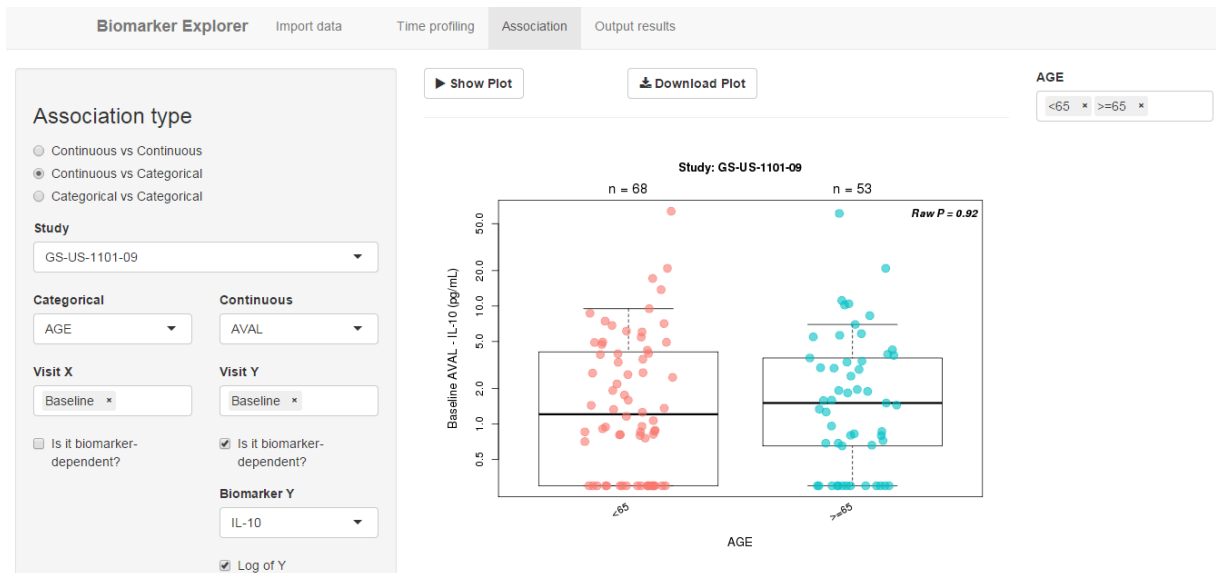

Example 2.2 (Continuous vs Continuous): Association of Biomarker IL-15 (Variable Y=AVAL, Biomarker=IL-15) at Week 4 (Visit Y=Week 4) vs. Biomarker IL-15 (Variable X=AVAL, Biomarker=IL-15) at Baseline (Visit X=Baseline) with log on X and Y, adding a smoothing line:

Example 2.3 (Continuous vs Continuous, multiple time points): Association of Biomarker IL-15 (Variable Y=AVAL, Biomarker=IL-15) vs Biomarker IL-17A (Variable X=AVAL, Biomarker=IL-17A) at Baseline and Week 4 (Visit Y =Baseline, Week4; Visit X=Baseline, Week4) with log on X and Y, adding a smoothing line:

Example 2.4 (Continuous vs Categorical): Association of Biomarker IL 10 (Continuou=AVAL, Biomarker=IL-10) vs age (Variable X=AGE) at Baseline (Visit X=Baseline; Visit Y=Baseline):

Example2. 5 (Categorical vs Categorical): Association of age (Variable X=AGE) vs Disease Type (Variable Y=Disease_type) at Baseline (Visit X=Baseline; Visit Y=Baseline):
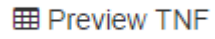

Step 3. Output Results

Step 3.1 Upload TNF file

- Click **Output results** on the navigation bar on the top of the platform.
- Click **Choose File** button to select and upload a TNF file. The standardized format of TNF files for Biomarker Explorer Rshiny tool can be found by clicking the link "Need input data template?" in **Choose file to import (Need input data template?)** to download a template.
- Once the TNF file is uploaded, a button **Output results** will appear on the top left of main panel. Click it to save the graphs and/or tables in a zipped file to the directed folder.

Step 3.2 Preview / Download TNF file

Once the user clicked **Add to TNF** button in either time-profiling or association page, a TNF file in Excel format will be created and constantly updated at the back end that saves the specification of all the graphs or table the user would like to output. The app provides an easy way to preview or download this TNF file which can be directed uploaded in Step 3.1

- To preview the TNF file, click **⊞ Preview TNF** button. A tab panel will be automatically created. A "Time Profiling" tab will be created if the user has saved any time-profiling graph specifications. An "Association" tab will be created if the user has save any association graph or table specifications.

- To download the TNF file, click **⤓ Download TNF** button. An Excel file containing all the saved graph and/or table specifications will be downloaded to the directory specified by the user.