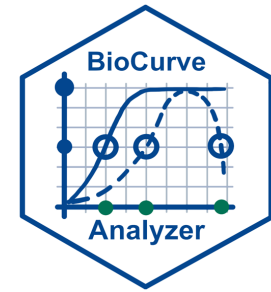


# Tutorial

## – Biocurve Analyze

In this tutorial, we will guide you through the process of analyzing sample data using Biocurve Analyzer. To begin, please install the application by following the instructions available on our GitHub page. The app features a navigation bar at the top, which includes three main tabs to assist users in completing the analysis step-by-step (see Figure 1). Each main tab consists of two panels: the side panel for input and the main panel for output.



BioCurve Analyzer

Welcome

Step 1: Input data

Step 2: ED<sub>50</sub>/T<sub>50</sub> Estimation

Step 3: Generate plot

Main functions of the app:

1. Analyze both the dose-response data and the time-to-event data.
2. Most popular models used to describe the data are provided as candidates, and it also helps to select the best model.
3. Calculate the ED<sub>50</sub>/T<sub>50</sub> values from the curves with diverse patterns.
4. Generate dose-response or time-to-event curves with customized appearance.
5. All the dataframes, figures, and a report can be downloaded.

References:

Wickham H (2014) Tidy Data. *Journal of Statistical Software*, Articles 59: 1–23

Ritz C, Baty F, Streibig JC, Gerhard D (2015) Dose-Response Analysis Using R. *PLoS One*. 10(12)

Reed LJ, Muench H (1938) A simple method of estimating fifty percent endpoints. *Am J Epidemiol* 27: 493–497

Ramakrishnan MA (2016) Determination of 50% endpoint titer using a simple formula. *World J Virol*. 5: 85–86

Serra A. Et al. (2020) BMDx: a graphical Shiny application to perform Benchmark Dose analysis for transcriptomics data. *Bioinformatics* 36: 2932–2933

Ritz C, Pipper CB, Streibig JC (2013) Analysis of germination data from agricultural experiments. *Eur J Agron* 45: 1–6

Onofri A, Mesgaran MB, Ritz C (2022) A unified framework for the analysis of germination, emergence, and other time-to-event data in weed science. *Weed Sci* 70: 259–271

Vaidya, A.S. et al. (2019) Dynamic control of plant water use using designed ABA receptor agonists. *Science*, 366(6464)

Eckhardt J, et al. (2024) Robotic Imaging and Machine Learning Analysis of Seed Germination: Dissecting the Influence of ABA and DOG1 on Germination Uniformity. *Plant Biology*

The logo for Biocurve Analyzer is a blue hexagon. Inside the hexagon, there is a graph with a blue curve and several data points. The text "BioCurve" is at the top and "Analyzer" is at the bottom of the hexagon.

Navigation bar with four tabs.

Figure 1 Welcome Tab

## Step1: Input Data

In this tab, users must first select the data type: either dose-response or time-to-event. Following that, the raw data needs to be submitted in a tidy format, with the columns arranged in a specified order. Users have three options to upload their data, as adapted from a previously published application [1]. Each data type requires specific information, and the necessary columns, along with their order, can be found by clicking the question mark next to the input data options. Additionally, users can download sample data to familiarize themselves with the required input format. To illustrate the functionality of this shiny app, we analyzed two datasets from our previously published papers [2–4]. Once the data is prepared, click the “Go” button to display it as a table and a simple line plot on the main panel (Figure 2).

BioCurve Analyzer   Welcome   **Step 1: Input data**   Step 2: ED<sub>50</sub>/T<sub>50</sub> Estimation   Step 3: Generate plot

**Step1: Select the data type.**

Data type:

☒ Dose-response data

☐ Time-to-event data

**Step2: Upload the input data.**

Input data:

The data should be in [tidy format](#) with column order as the follows:

**Factors** (if applicable, e.g. Protein, Compound, *et al.*; try to **avoid using underscore** in the name), **Biphasic**, **Concentration**, **Replicate**, **Response**.

Click the question mark to see the detailed requirements

Load the sample data

Download SampleData

Note: The sample data is from a previously published paper. (Vaidya, A.S. et al, *Science*, 2019)

Go

**Three ways to input data**

Upload the file

Choose File:

Note: The first row is used as header.

Browse... No file selected

Paste data

Paste data below:

Note: The first row is used as header.

Separator:

☒ Comma

☐ Tab

☐ Semicolon

Clear data

**Step3: Click on “GO” to upload the data and view the data as a table and line plot on the right.**

**Data table**

Is the curve biphasic? Y/N

Search:

Protein	Compound	Biphasic	Con_nM	Replicate	PP2C_Activity_%
AtPYR1	ABA	N	10000	1	5.44
AtPYR1	ABA	N	1000	1	10.97
AtPYR1	ABA	N	333.333	1	31.88
AtPYR1	ABA	N	111.111	1	65.01
AtPYR1	ABA	N	37.037	1	83
AtPYR1	ABA	N	12.346	1	95.53
AtPYR1	ABA	N	4.115	1	98.21
AtPYR1	ABA	N	0	1	102.16
AtPYR1	ABA	N	10000	2	5.71
AtPYR1	ABA	N	1000	2	10.55

Showing 1 to 10 of 96 entries

Previous 1 2 3 4 5 ... 10 Next

Click to see the format requirements

**Line plot**

**Figure 2 Input Data Tab**  
(Demonstrated by using dose-response data as the input data type)

## Step2: ED<sub>50</sub>/T<sub>50</sub> Estimation

In this tab, users can estimate the ED<sub>50</sub>/T<sub>50</sub> from their data by employing the best-fit model (see Figure 3). For both dose-response and time-to-event data, users should select a list of candidate models and determine the appropriate criterion for identifying the best-fit model. Additionally, users can specify constraints for the upper and lower limits of each model to better suit their data.

For dose-response data, there are several considerations to address before calculating the ED<sub>50</sub> values. The application enables users to estimate both relative and absolute ED<sub>50</sub> values, with three established methods available for value estimation. Furthermore, users can evaluate the models using the selected statistical tests. The final step involves specifying the minimum dose for plotting, as the concentration must be log-transformed. This step is optional, with the default method used being the one from the *drc* package. The differences between models, types of ED<sub>50</sub>, three ED<sub>50</sub> estimation methods, and model assessment methods are available for the users by clicking on the questions mark at the top right corner.

## Step3: Generate plot

In this step, users can generate the final figure using default formatting settings in one tab, while also having the option to customize the figure by modifying labels, legends, fonts, and line colors (see Figure 4).

First, users should define the layout, especially if multiple variables are involved in their experiments. They can choose how to present their data, either by displaying all replicates as a dot plot or by plotting the mean value of each group with the corresponding standard deviation (SD) represented as error bars.

If some of the data do not conform to any of the models chosen in the previous step, there is an additional option available. Users can either return to the previous step to adjust the candidate model list or proceed to plot the data as needed. The application allows for either a simple line plot that connects the mean values or fitting the data to the Loess model for visualization. In such cases, a message box will alert users about the extra step required when creating the figure. The results of both options are displayed in Figure 4, making this function particularly useful for handling messy data or inactive chemicals in experiments.

Once all selections are made, users can click on “Generate plot” to view the plot in the main panel. They can choose whether to display the ED<sub>50</sub>/T<sub>50</sub> on the plot and have further customization options available. The figures, data frames, and a comprehensive report can be downloaded from the app. Additional details and interpretations of the report are provided in Figure 5.

### Step1: Select the candidate models and set the constraints.

Select the candidate models:

- Monotonic Curves

Lower

Upper

☒ Log-logistic (4 parms)
 

Not Fixed

Not Fixed

☒ Log-logistic (5 parms)
 

Not Fixed

Not Fixed

☒ Weibull I
 

Not Fixed

Not Fixed

☒ Weibull II
 

Not Fixed

Not Fixed

### Step2: Choose the Criterion for model selection.

Criteria for model selection:

Akaike's Information Criterion

### Step3: Choose the methods and type of ED<sub>50</sub> you want to estimate.

Methods & Type of ED<sub>50</sub>:

- Type of ED<sub>50</sub>:

☒ Absolute
 

Three published methods

☐ Relative

- ED estimation method:

Ritz-Gerhard Method

☐ Include Reed-and-Muench Method

### Step4: Select the tests and the *p*-values for model assessment.

Model assessment methods:

☒ Lack-of-fit test
 

p-value

0.05

☒ Neill's test
 

0.05

☒ No effect test
 

0.05

☒ Parameters ≠ 0
 

0.05

### Step5: Specify the minimum dose in plotting.

Minimum dose:

default

**Note:** Please specify the minimum dose to achieve a logarithmic scale visual effect, applicable only when the minimum dose is zero. The default is the base-10 value corresponding to the rounded minimum log<sub>10</sub> value of all positive doses, as recommended by the *drc* package.

Calculate

## INPUT PANEL

Six models for dose-response data;  
Four models for time-to-event data.

Four criteria for dose-response data;  
two criteria for time-to-event data.

## OUTPUT PANEL

### ED<sub>50</sub> Estimation Table

Show 10 entries

Search:

	Protein	Compound	Response at ED50	ED50 Mean	ED50 SE	ED50 Lower Bound	ED50 Upper Bound
1	AtPYL2	ABA	50.00	36.52	2.77	30.77	42.28
2	AtPYL2	OP	50.00	7.52	0.40	6.68	8.36
3	AtPYR1	ABA	50.00	161.46	20.18	119.36	203.55
4	AtPYR1	OP	50.00	11.93	0.75	10.36	13.50

Showing 1 to 4 of 4 entries

Previous 1 Next

### Model Assessment Results

Show 10 entries

Search:

	Protein	Compound	Model	Lack-of-fit test	Neill's test	No effect test	Parameters ≠ 0
1	AtPYL2	ABA	Log-logistic (4 paras)	6.38E-02 *	3.19E-01 *	0.00E+00 *	Significant
2	AtPYL2	OP	Weibull I	4.48E-01 *	7.08E-01 *	0.00E+00 *	Significant
3	AtPYR1	ABA	Weibull I	9.29E-01 *	9.49E-01 *	0.00E+00 *	Non-significant
4	AtPYR1	OP	Weibull II	5.06E-01 *	7.27E-01 *	0.00E+00 *	Significant

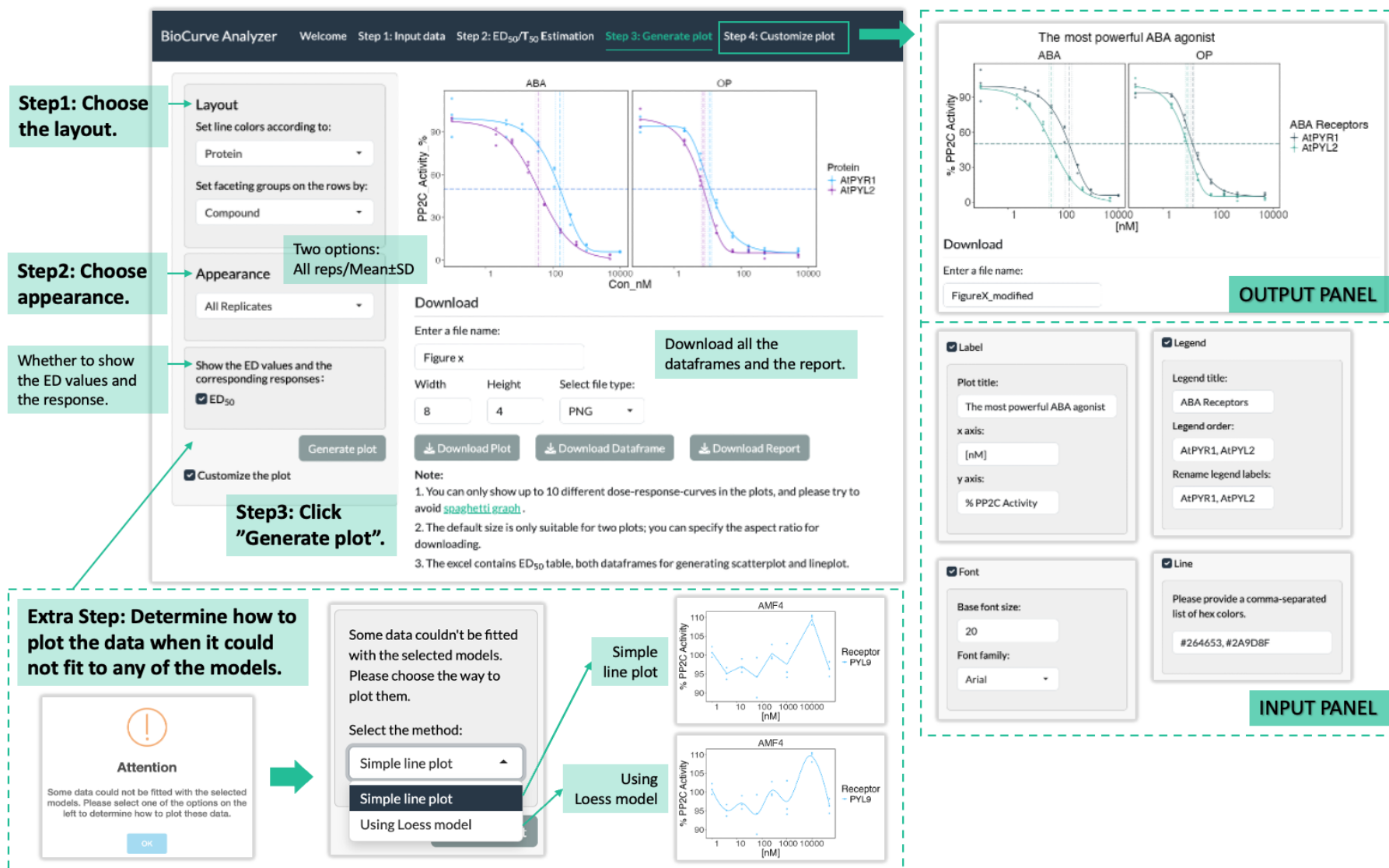
Showing 1 to 4 of 4 entries

Previous 1 Next

**Note:**  
To prevent over-fitting, we highly recommend you to choose the models based on your own experiment setup, if you let the app to choose the best models for you, the best models reported are selected from the *drc* analysis (Ritz C, et al., *PLoS One*, 2015), based on the criteria you choose on the left.

**Reference:**  
Ritz C, Baty F, Streibig JC, Gerhard D (2015) Dose-Response Analysis Using R. *PLoS One*. 10(12)  
Serra A. Et al. (2020) BMDx: a graphical Shiny application to perform Benchmark Dose analysis for transcriptomics data. *Bioinformatics* 36: 2932–2933  
Ramakrishnan MA (2016) Determination of 50% endpoint titer using a simple formula. *World J Virol*. 5: 85–86

**Figure 3 ED<sub>50</sub> Estimation Tab**  
(Demonstrated by using dose-response data as the input data type)



**Figure 4 Plot generation**  
(Demonstrated by using dose-response data as the input)

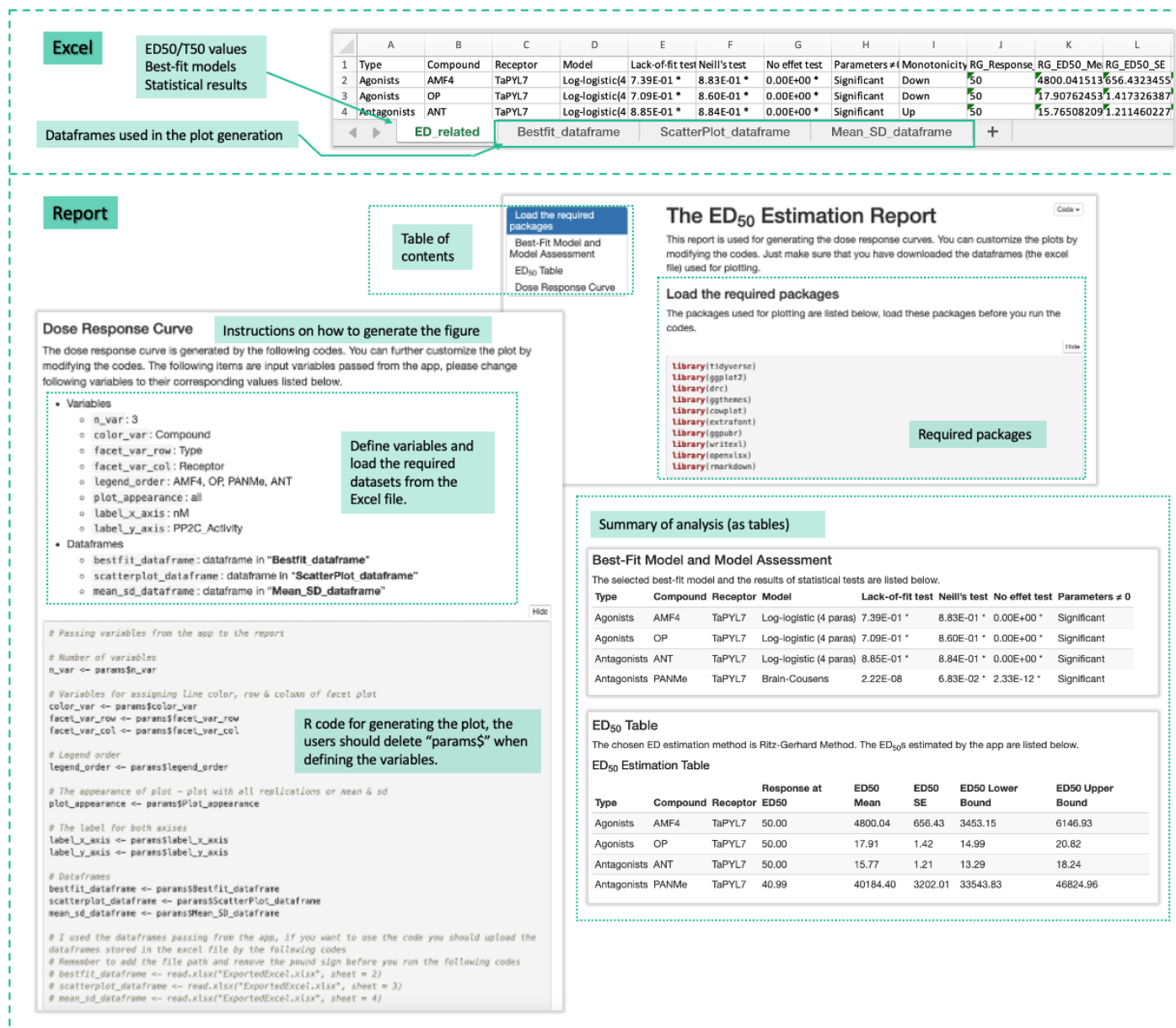


Figure 5 Example of output files  
(Demonstrated by using dose-response data as the input data type)

## References:

1. Spitzer M, Wildenhain J, Rappsilber J, Tyers M. BoxPlotR: a web tool for generation of box plots. *Nat Methods*. 2014;11:121–2.
2. Vaidya AS, Helander JDM, Peterson FC, Elzinga D, Dejonghe W, Kaundal A, et al. Dynamic control of plant water use using designed ABA receptor agonists. *Science*. 2019;366:eaaw8848.
3. Vaidya AS, Peterson FC, Eckhardt J, Xing Z, Park S-Y, Dejonghe W, et al. Click-to-lead design of a picomolar ABA receptor antagonist with potent activity in vivo. *Proc Natl Acad Sci U S A* [Internet]. 2021;118. Available from: <http://dx.doi.org/10.1073/pnas.2108281118>
4. Eckhardt J, Xing Z, Subramanian V, Vaidya A, Cutler S. Robotic Imaging and Machine Learning Analysis of Seed Germination: Dissecting the Influence of ABA and DOG1 on Germination Uniformity [Internet]. *Plant Biology*. bioRxiv; 2024. Available from: <https://www.biorxiv.org/content/10.1101/2024.05.10.593629v1.full.pdf>