MANUAL FOR PROAST GUI

Table of Contents

PRELIMINARIES

Installation of required software

- The R software needs to be installed on your computer (free download from www.cran.r-project.org).
  Note 1: if you are asked to select a cran mirror, simply select one near to you. A cran mirror is a public server from where you can download files.
  Note 2: when you do not have administrator rights on your computer, the installation may fail. In that case you may try installing again and, in the window "Select Installation Location", browse to a folder where you have writing rights (e.g., My Documents) on the hard disk of your computer. A memory stick is also possible, but it will make working in R slower.
- The PROAST package ("proast65.5.zip") needs to be installed within R. *Make sure you are online*. After opening R, click on "packages", next on "Install package(s) from local files" (see Fig. 1). Browse to the folder where you have stored the PROAST package, and click on it. Note that the PROAST package comes as a zip file, and should ***not*** be unzipped. (Mac users should install proast65.0.tgz).
- The GUI version of PROAST needs some auxiliary packages. These can be installed by copying the following line into the Console window of R:
  `> install.packages(c('assertive', 'gWidgetstcltk', 'hwriter'))`
  Alternatively, you could install these packages one-by-one by clicking on Install package(s) in the R window (see Fig. 1)
  Note that installing these packages may take a couple of minutes, depending on your computer.
- Now, load the PROAST package, by clicking on "packages" and then "Load package".
- Finally, check if your installation was successful by typing
  `> g.proast()`
  in the Console window.

If you have problems with installing PROAST:
It might occur that you get error messages when trying to install or load PROAST, due to problems with the automatic installation of auxiliary packages, specifically related to your computer system. Read the errror message, and determine the (first) package mentioned that could not be found. Then, click on "Install package(s)" (see Fig. 1), and you will get the list of all available R packages. Find the packages you are missing (e.g. "assertive"), and click on it. It will now be installed from the R website.

Note: If you get an error message related to assertive, you probably have an earlier version of assertive installed in your R library. You may solve this as follows. First type
`> remove.packages("assertive")`
and then re-install it:
`> install.packages("assertive")`
and next
`> install.packages(c('gWidgetstcltk', 'hwriter'))`

The PROAST package has a version number, so that you could reproduce results from earlier versions. It is recommended to always use the latest version, however, and only use earlier versions in specific circumstances (see Annex 2 for how to switch between versions).

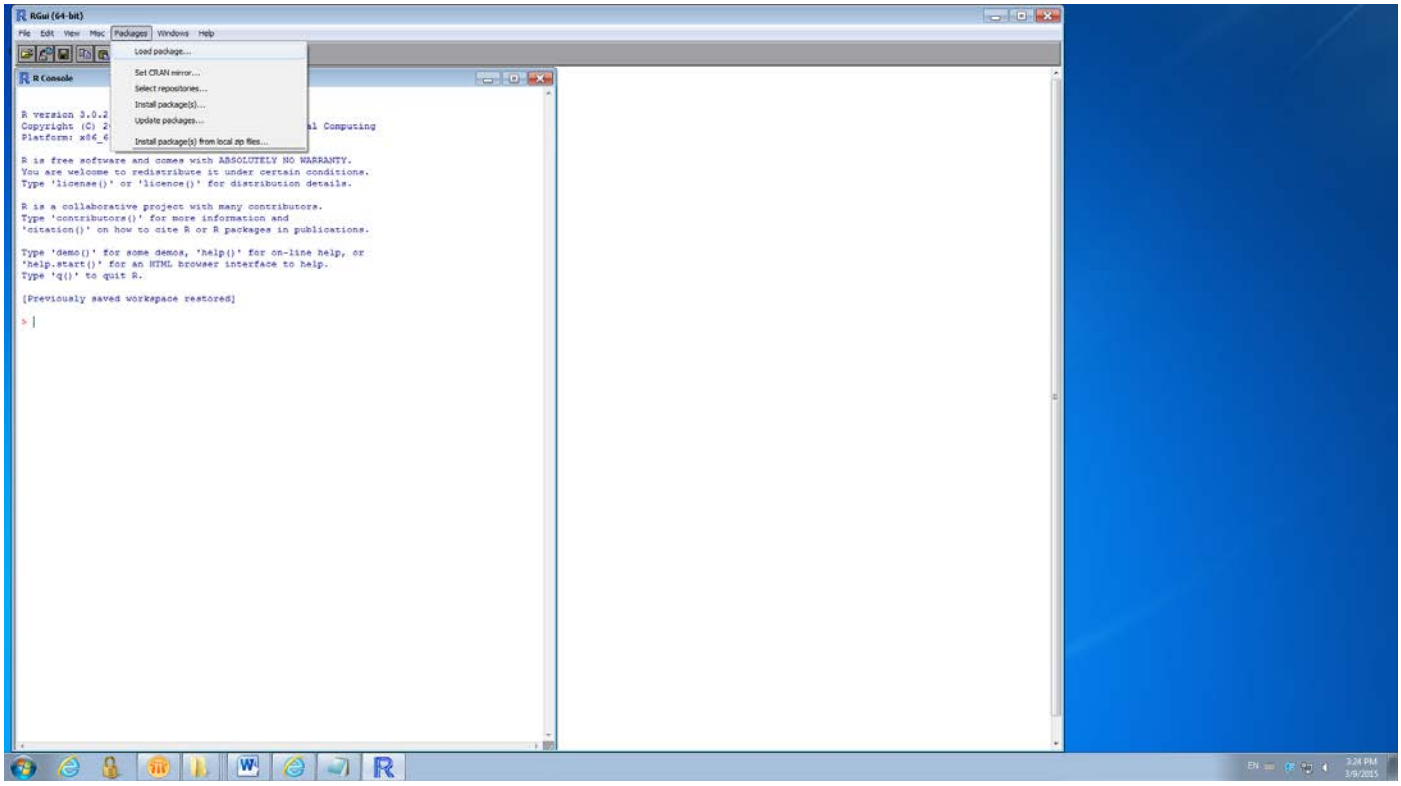


Fig. 1. The preferred setting of the R windows, and installing packages. The PROAST package can be installed by clicking on “Install package(s) from local files…” in the scroll down menu under “packages”. If necessary, auxiliary packages can be installed by clicking on “Install package(s)…”.

Working in R

When working in R, you are working in an environment that can only be read by R. This is called the Workspace. The Workspace can be safed as a file that contains the data to be used and all the results created during an R session. Once you have save a Workspace, you can re-open it (double click on it) to continue the analyses starting from the point where you saved it. When you have various projects, you can create various Workspaces, to prevent confusion and keep oversight of your data and associated results. An R-Workspace has the extension .Rdata, actually the default name is just .Rdata.

An R-Workspace can be regarded as analogous to a Word document, which you first create, then you save it, and next you may re-open it to continue working on it. Just like the content of a Word document can only be read by opening it in Word, the content of an R-Workspace can only be read by opening it in R.

The following steps are recommended when you want to create a new R workspace:

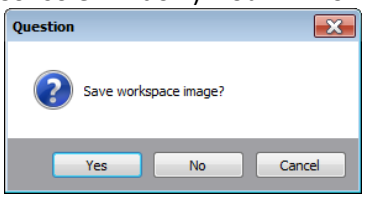
1. *Define Working directory.* The Working directory is an existing or newly created folder that contains the relevant data (related to some defined problem or project). The use of different folders is just to make it easier for you to find back a specific dataset or analysis at a later point in time. The datasets need to have a specific format, as will be described below.
2. *Open R.* Open the R-GUI by double clicking on the R icon on your desktop (or in the start menu of your computer). In the upper left corner of the R-GUI click on “file”, and next on

"Change dir …". Browse to the folder (the working directory) with the relevant data and click "OK".

3. *Create Workspace.* In the R Console window type:
   ```
   > getwd()
   ```
   which means: get Working Directory. (Note that the first > is the prompt already present in the Console window). You should now see the path of the Working Directory that you selected in step 2. When you type
   ```
   > dir()
   ```
   you will see the files in your current Working Directory. When you type
   ```
   > ls()
   ```
   you will see the content of the current R Workspace (when you have not yet done anything it will be empty).

   To illustrate how the R-Workspace stores new results, do the following simple exercise. In the Console window (see Fig. 1) type
   ```
   > aa <- 5
   ```
   which creates an object called aa, with content the number 5. Type
   ```
   > ls()
   ```
   and you will see aa as one of the objects in the list. Type
   ```
   > aa
   ```
   and you will see the content of aa.

   Similarly, during an R session you can create all sorts of "objects" (usually more meaningful, like the results of an analysis), which will be kept in the Workspace. If you save the Workspace during or at the end of the session (see next step), all objects will remain, and will be available when re-opening the Workspace.
4. *Save the Workspace.* Close the PROAST-GUI (click on the red cross in the upper right corner), and then close the R-GUI (click on the red cross in the upper right corner; or type q() in the Console window). You will now get the question:

   Here, answer Yes. This will save the workspace in the current working directory with name .Rdata.
5. *Re-open the Workspace.* Browse to the folder where you saved the Workspace. When the Workspace is shown as a big blue R, this means the Windows recognizes it as an R application, and you can simply double click on it. This re-opens your last session. Type ls() to check if the object aa still exists.
   - If Windows on your machine does not recognize the workspace as an R application, you may right click on saved Workspace (.Rdata) and select Rgui.exe (see folder R/R-3.2.0/bin/i386) as the associated program. If that does not work, repeat step 2, and then click File/Load workspace, to load your last saved workspace.
   - When you re-open the workspace, it might happen that you get a welcome message from PROAST, without having loaded PROAST (this will happen if you closed R without having closed the PROAST GUI window first). To solve this type
     ```
     > rm(.proast.wizard)
     ```
     safe the workspace and re-open it. Then the message should no longer appear. Further, it will take less time to open the workspace.

Names of Workspaces
When you save the Workspace at the end of the session, as in step 4 above, the Workspace will be saved under the name .Rdata (i.e., only the extension). During the session you can also save the Workspace by clicking File/Save Workspace. In that case, you could provide a name in front of the extension .Rdata. However, as long as you keep different projects organized in different folders, each with its own Workspace, different names for the Workspaces are not really needed. Actually, using various Workspaces in the same folder is somewhat risky, as they might be overwritten unintendedly (after step 4 above).
Regularly save your Workspace
When you are doing something unexpected, it might happen that the R session crashes. Therefore, it is recommendable to save your workspace during a session every now and then, in particular if you did a calculation taking a substantial amount of time.
When you want to re-open the session by double clicking .Rdata after an irregular ending of the session, it may happen that R opens very slowly, and then shows you the PROAST message, even though you did not yet load the package. This should be solved as soon as you have ended a session with the PROAST GUI in a regular fashion, i.e. first close the PROAST GUI, and next the R GUI, and save the workspace after that.

GUI does not respond
If the GUI does not respond (after some unexpected action), you may be able to re-activate it by typing g.proast() once more, even though the GUI already exists.

STARTING THE PROAST GUI

Create a .Rdata workspace in the folder with the datasets that you are going to analyse, as just described. Open the workspace.
Arrange the R windows as shown in Fig. 1.

The left part is the Console window. It has two functions: it shows the non-graphical output from the PROAST analyses, and it can be used for specific commands (in R language; some examples are given below).
The middle part is used for graphical output
The right part will be used for the PROAST GUI (see Fig. 3).

Every time you opened the Workspace, the PROAST needs to be loaded by clicking on “Load package” (see Fig. 1).
In the Console window, type

```
> g.proast()
```

(see Fig. 2). This will result in creating the PROAST GUI; it may take some time as a number of auxiliary packages need to be loaded first.

Place the PROAST GUI outside the R window, to make sure that it remains visible at any time during the session (see Fig. 3).

In case of problems with loading required packages:
It might occur that you get error messages when required packages are loaded. Read the errror message, and determine the (first) package mentioned that could not be found. Then, click on “Install package(s) … ” (see Fig. 1), and you will get the list of all R packages available on the R website. Find the packages you are missing (e.g. “assertive”), and click on it. It will now be installed from the R website. Again, type g.proast() to launch the PROAST GUI.

NOTE: whenever you want to stop the PROAST session, first close the PROAST GUI before closing the R GUI, otherwise re-opening the .Rdata workspace will be slower.

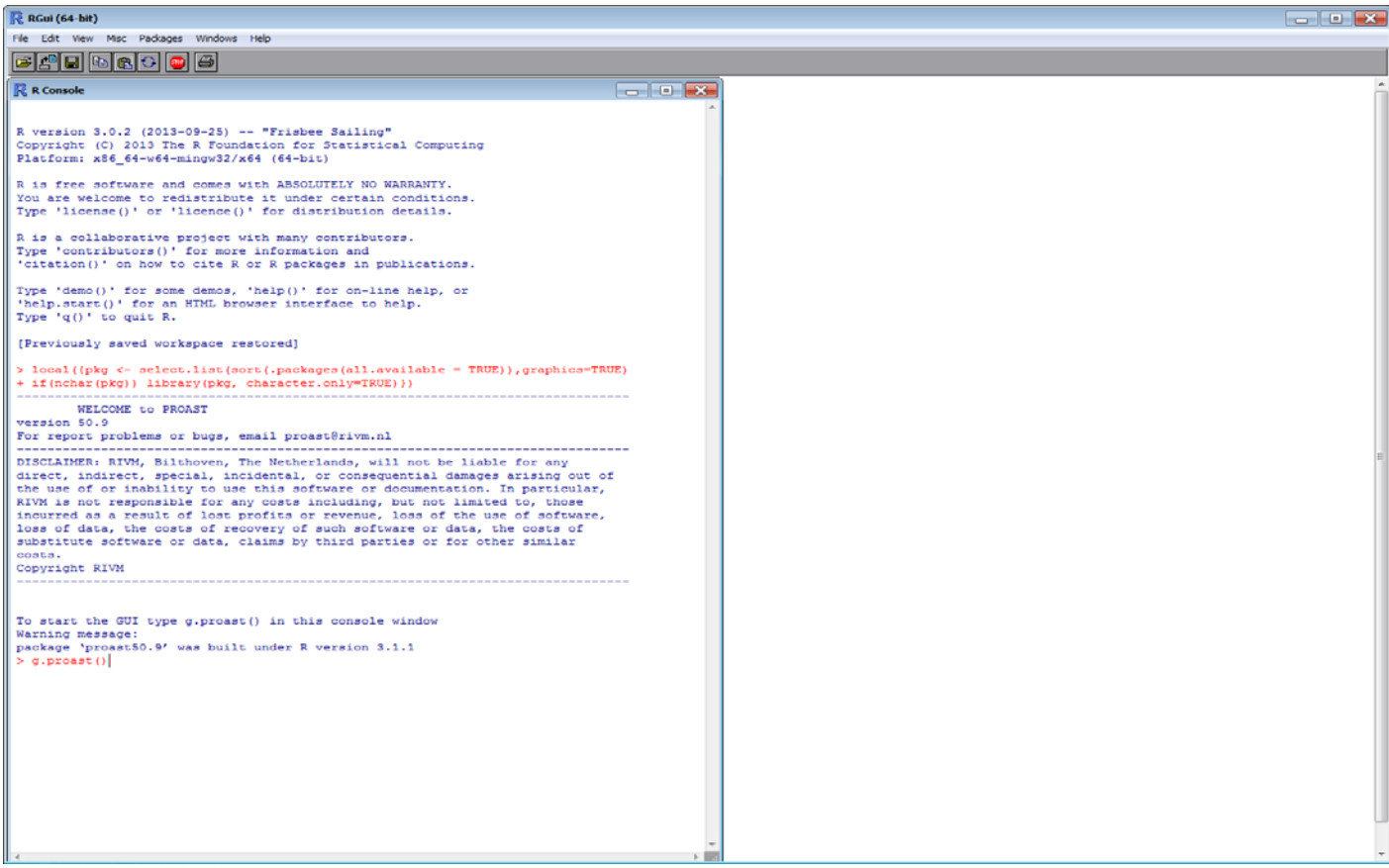


Fig. 2. Start the PROAST GUI by typing g.proast() in the Console window of R.

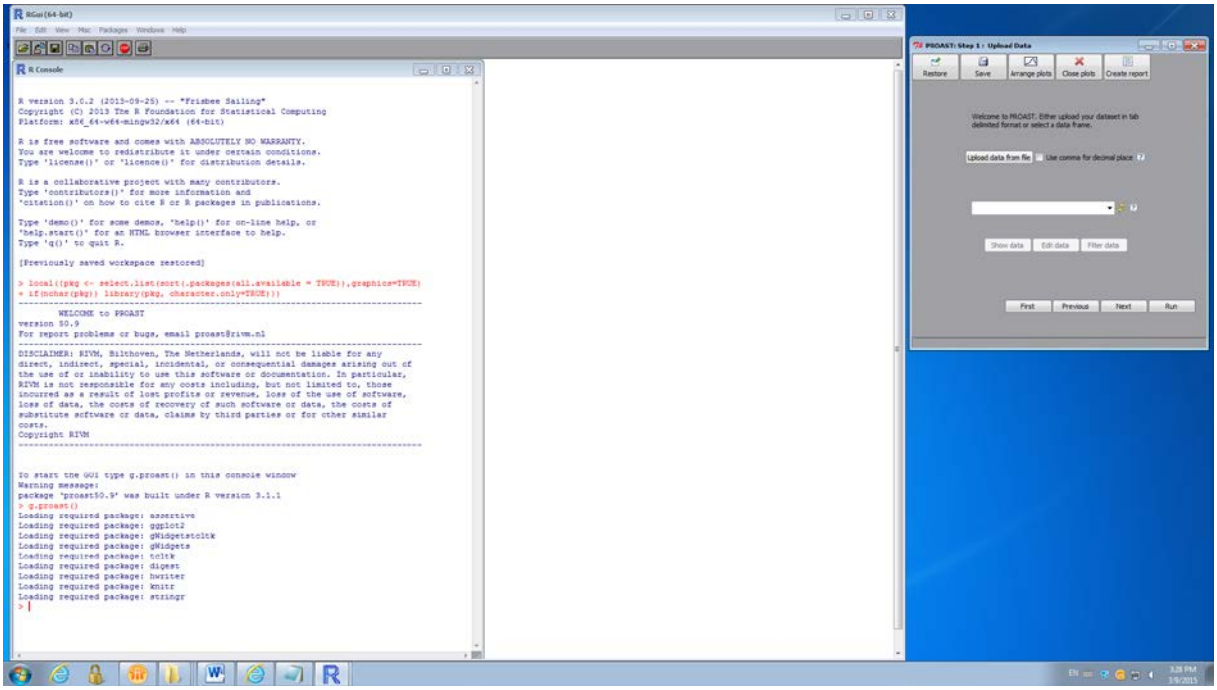


Fig. 3. The PROAST GUI is placed outside the R window, to keep it visible and accessible at all time.

EXAMPLE ANALYSIS

When going through the manual for the first time, you might consider to use the dataset below (foetal weights) to get a first impression of the PROAST GUI and as a first exercise.

Enter (copy-paste) the following dataset in an Excel sheet, and save it as a text file (tab delimited), preferably in the directory where you have created your .Rdata workspace, under a convenient name, e.g. foetw.txt.

dose	sex	foetal_w_mn	SD	n
0	m	4.437	0.373	271
270	m	4.339	0.340	115
350	m	4.369	0.338	264
450	m	4.234	0.351	115
580	m	4.185	0.341	109
750	m	3.843	0.462	182
970	m	3.481	0.439	27
1250	m	3.230	0.099	2

Next, go through the six steps in the GUI as described below, focussing on the sections with headers followed by (*), and skipping the paragraphs on covariates.

RUNNING THE GUI (*)

You are now ready to run the GUI. You will go through 6 steps, which will be discussed consecutively. After having completed a particular step click on “Next”. You can always go back when changing your mind (click on “Previous”). Clicking on a question mark will give some information on the related issue. Note that windows that you are trying to create may not show up. Usually they are hidden behind other windows already opened on your screen. I that case right-click on the blue R on the bottom bar of your screen.

STEP 1: Define the dataset (*)

Step 1 is used for importing and defining the dataset to be analysed (see Fig. 4). When you create a dataset for analysis by PROAST, it needs to obey certain format rules.

Data format

The required format of the data is as usual for statistical analysis. In its simplest form it consists of a column with the doses and a column with the responses, or multiple columns with various responses (endpoints or toxicological parameters). Each row may relate to an individual animal (or some other experimental unit), or to a group of animals, such as a dose group. With grouped data, an additional column is needed with the group sizes. This holds for both continuous data (e.g. mean body weights) and for quantal data (number of animals with some effect). For continuous grouped data (such as mean body weights) yet another column is required: the SDs or SEMs related to the means. Additional columns may be added, indicating particular characteristics of the data in that row, such as sex (e.g., m or f), a study label (e.g., study 1, study 2, study 3), or any other factor of interest. These columns may be used as covariates in the analysis (see step 5), or be used to make a sub-selection of the dataset (in step 1, see below). It is recommendable to compose one larger dataset rather than many small datasets. This opens the option to compare dose-response relationships among the subgroups defined by any covariate column, while sub-selections of smaller datasets can always be made in the GUI (see below). For examples of various datasets, see Annex 1.

Note that it is not allowed to add additional information to the datasheet, such as a title, or additional information at the end. If you wish to do so, you may create a copy with the additional information. However, the sheet that is intended to be imported in PROAST should not have any additional fields with information. Each cell of the table needs to be filled with one single value or text string. Missing values should be indicated by NA. Values below the limit of quantification (LOQ) should be given as 0 (the PROAST GUI will then ask to provide a value for the LOQ).

When you have created the dataset in Excel, it can be directly imported by the PROAST GUI (if it is the first sheet). However, it is recommended to save the datasheet in a text file (in Excel: use “save as”, go to “save as type”, and select Text (tab delimited), (*.txt)), and import the text file in the PROAST GUI.

Importing the data in the GUI (*)

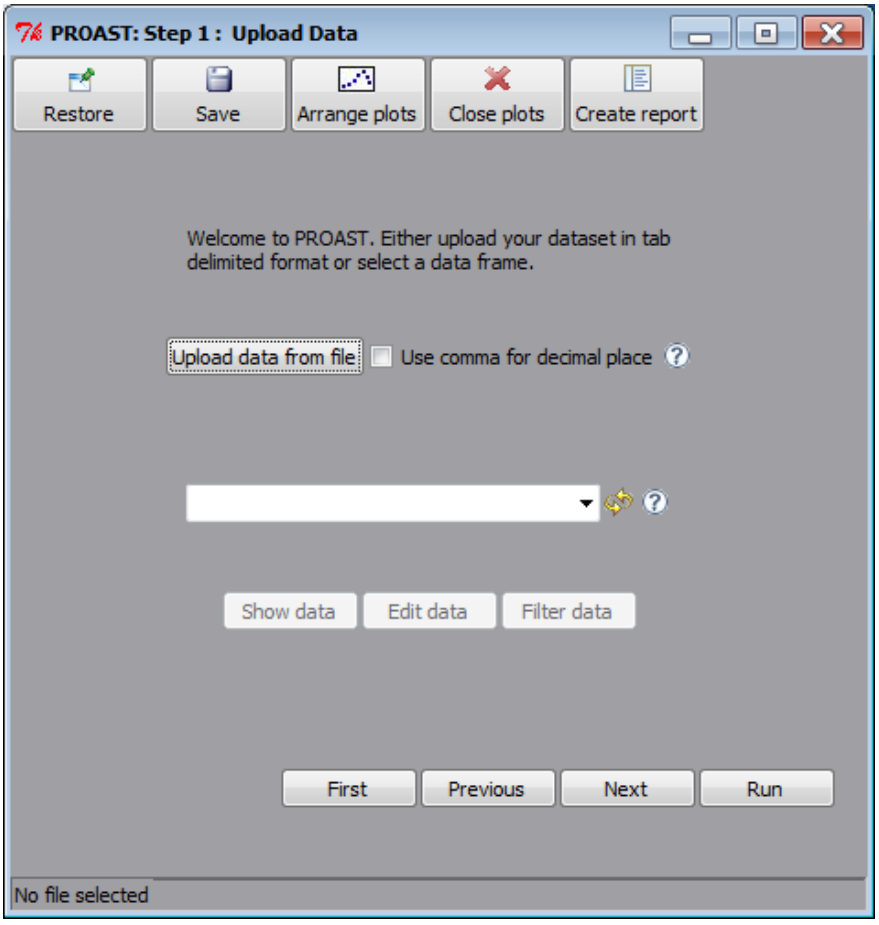
The most important button in step 1 is “Upload data from file”. After clicking this button, navigate to the directory where you have located the dataset you want to analyse (if needed), select the file and click on “open”.

Note: R assumes that a dot is used as the decimal marker. If the regional settings on your machine use a comma, you could tick off the box “Use comma for decimal place”. Or, you could change your regional settings.

Once the data are imported, you can click on “Show data”, and inspect the data just imported.

Note: When the data do not appear, check if they are hidden behind other windows currently opened on your screen (right click on the R in the bottom bar of your screen).

If needed, you can change the data manually (click on button “Edit data”).



Making a subset selection

If you want to make a subset selection from the dataset (e.g. males only), click on “Filter data”. This will open another window, called “Filter the dataset”, see lower right corner of Fig. 4.

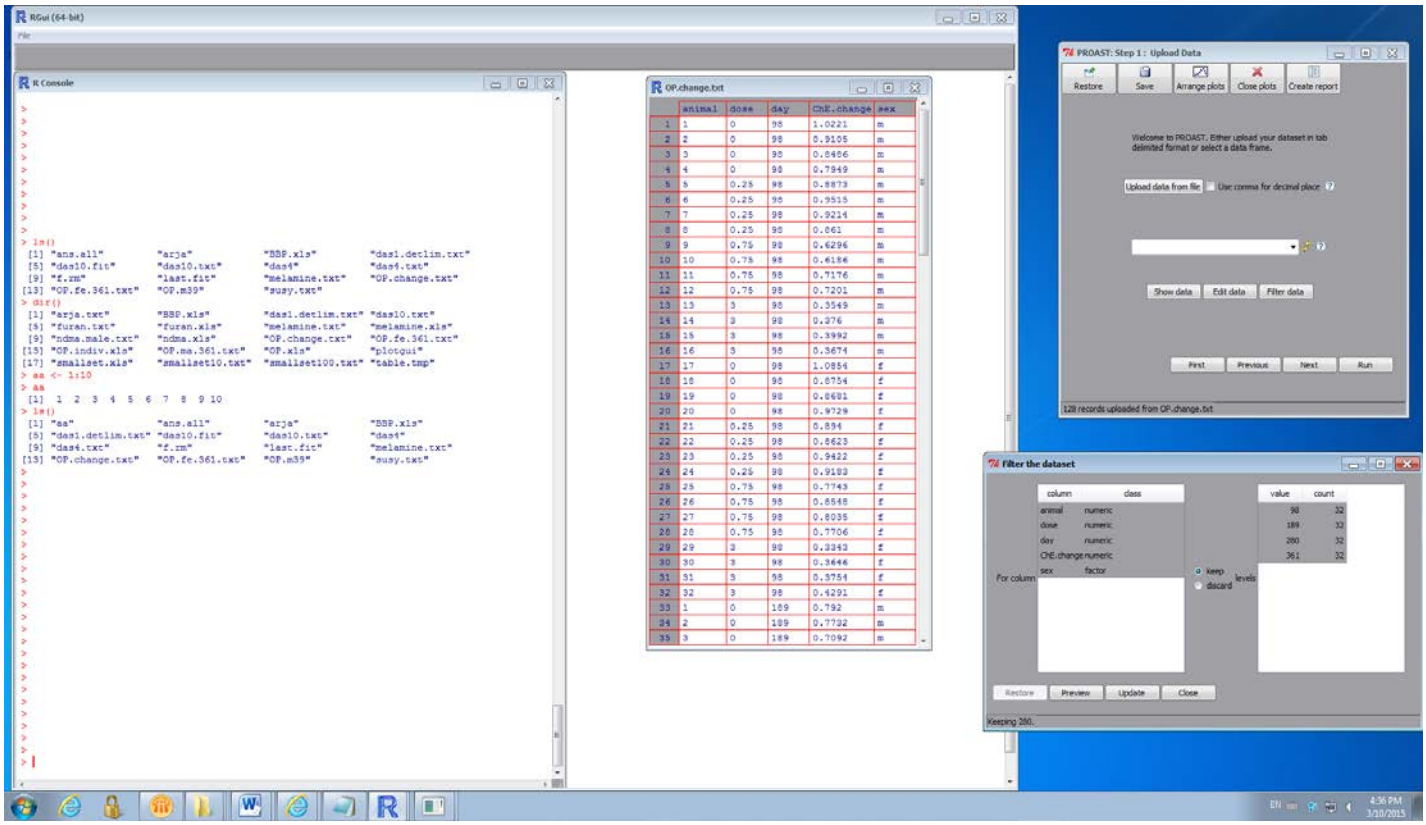


Fig. 4. After clicking “Filter data” in step 1, the window “Filter dataset” will appear.

In this new window, you can select a factor for which you want to make a selection. For instance, in Fig. 4 the factor ‘day’ is selected (just click on it), which then produces the levels of that factor on the right side under “value” (under “count” you will see the frequency for each of them). When you now click on one value, e.g. 280, and then click on update, the window changes as in Fig. 5. Click on close, then close the window with the data, and re-open it (click on “show data”). You will now see that the desired subselection has been made.

Note 1: this window can also be used to remove outliers, by ticking off “discard” rather than “select”.

Note 2: If you want to select more than one level, keep the control key press while selecting levels.

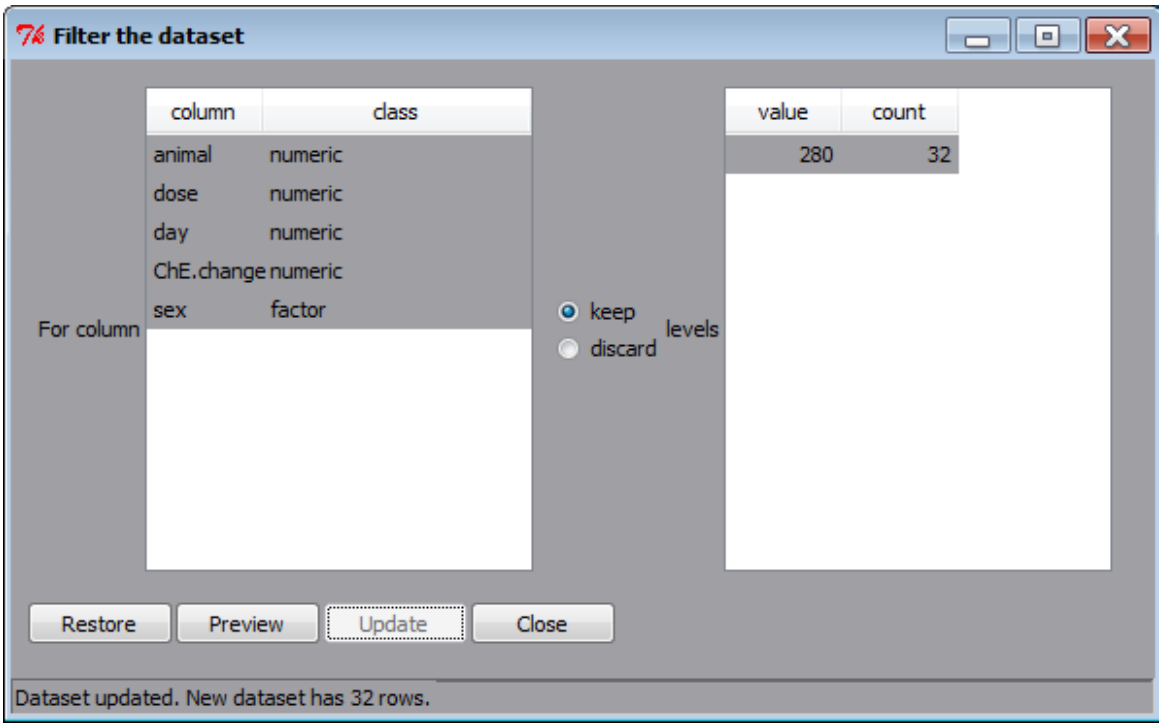
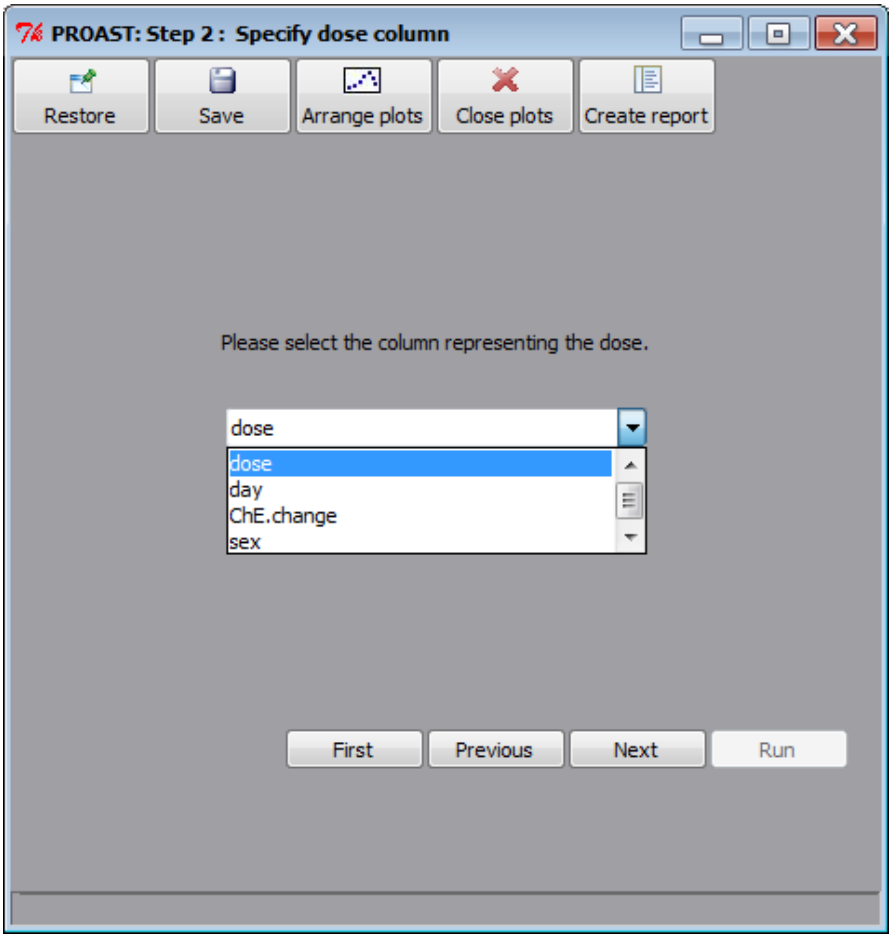


Fig. 5. Select day 280 from the dataset.

STEP 2: Define the dose (*)

Step 2 is used for selecting the column with the dose information. Click on the downward arrow on the right hand side of the small window in the middle, and you will see the names of the columns in your dataset. When one of the columns is called “dose” this will already be selected by PROAST.

Note: it may happen that you do not see all column names, use the scroll bar to check that.



STEP 3: Define the response (*)

First, select the column with the response information (just as you did with dose).
Next, you need to indicate what type of data the response relates to. To do so, you need to decide:

- Are the response data continuous or quantal?
- In case of continuous: individual data or summary data?
- Are there potential litter effects?

When there are no litter effects, select the appropriate data type from the first three options. For continuous data, you may have an observation in each individual animal (first option), or just the means for each dose group (option 2: groups of animals). In the latter case, the data they need to include SDs (or SEMs) and group sizes.
In the case of quantal data, the data are provided per dose (number of animals with the lesion in that dose group, out of the total number in that group), so they may be considered as data on groups of animals as well.
When there are litter effects, choose the fourth or the fifth option, depending on having continuous or quantal data.
You might also have summary data for a continuous response with litter effects. This option is not in the list. However, in this case you may choose option 1: by taking the mean of each litter, and consider that as individual data, the litter effect has disappeared, and the means can be considered as individual observations per dam. Note that this should be regarded as an approximate analysis.

> Note: the menu version offers the option to analyse the data as summary data with litter effects, but the results will often be similar.

When you have individual continuous data that contain zeros, you need to enter a value of the limit of quantification (LOQ). PROAST will show the lowest nonzero value of your response data (see Fig. 6). If you have no specific LOQ information, you may use the lowest nonzero value as the LOQ.

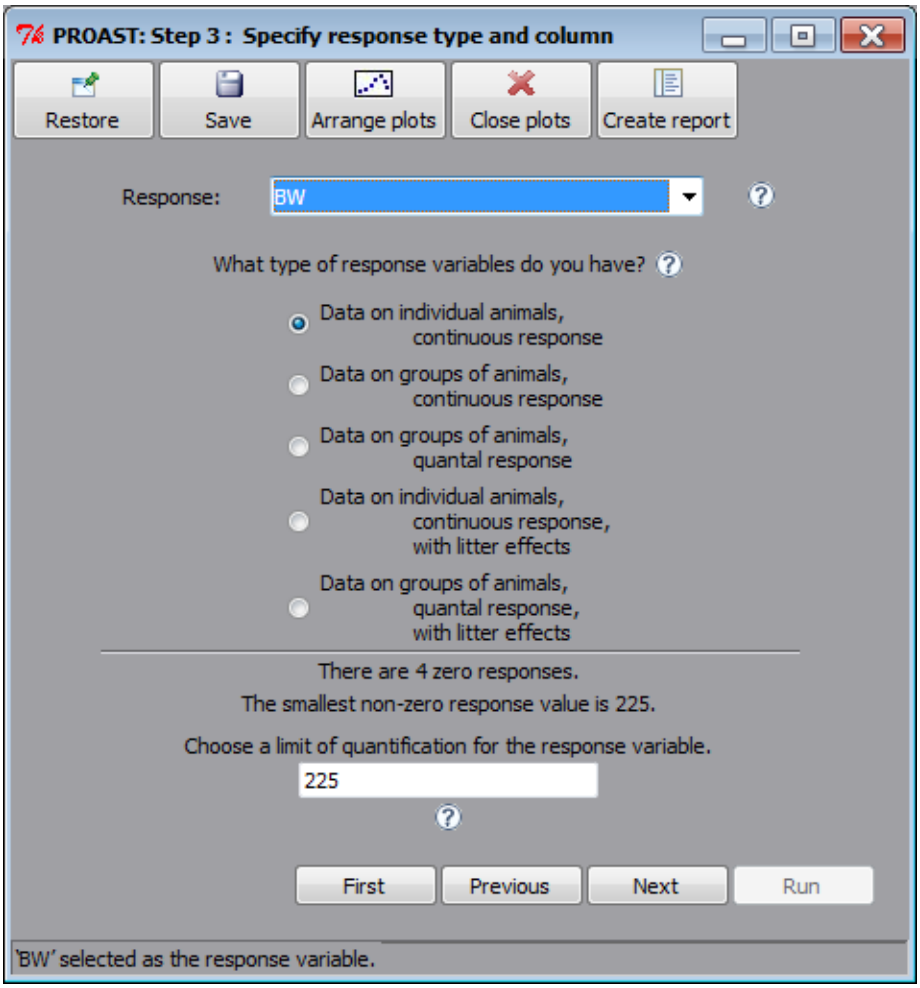


Fig. 6. When individual continuous data contain zeros, PROAST reports the lowest nonzero response value, and uses this value as a default for the LOQ. If you have information on the LOQ, you may enter that value.

STEP 4: Specify additional information for response (*)

Step 4 depends on the type of response selected in step 3.

Individual continuous data

When you selected individual continuous data, no further information is needed, and step 4 will be skipped.

Continuous summary data (*)

When you selected continuous summary data, you need to indicate the column with the SDs (or SEMs). Do not forget to tick off SD or SEM, as this will make a difference for the output. Further, you need to indicate the column with the group sizes. See Fig. 7 for an illustration.

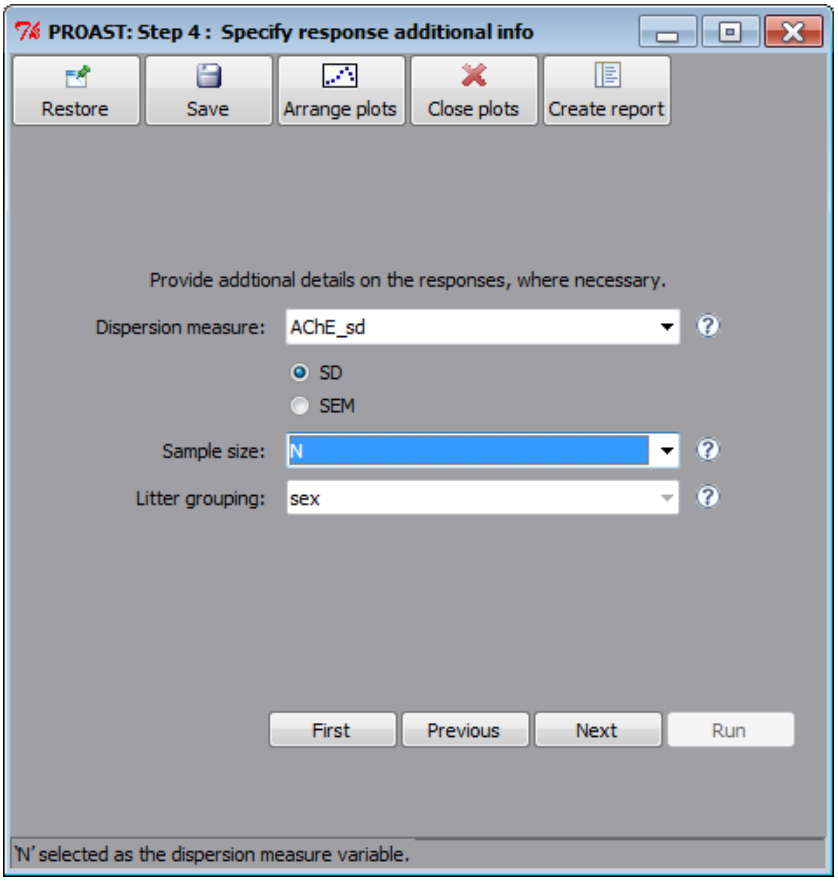


Fig. 7. Additional information needed in step 4 when the responses are continuous summary data: dispersion measure and sample size. You also need to indicate if you have SDs or SEMs in the datasheet. (Note that litter grouping is grey, it is not relevant in this case).

Quantal data

With quantal data you need to indicate the column with the group sizes.

Data with litter effects

When you have (continuous or quantal) data with litter effects, you need to indicate which column represents the litter factor. See Fig. 8 for an illustration.

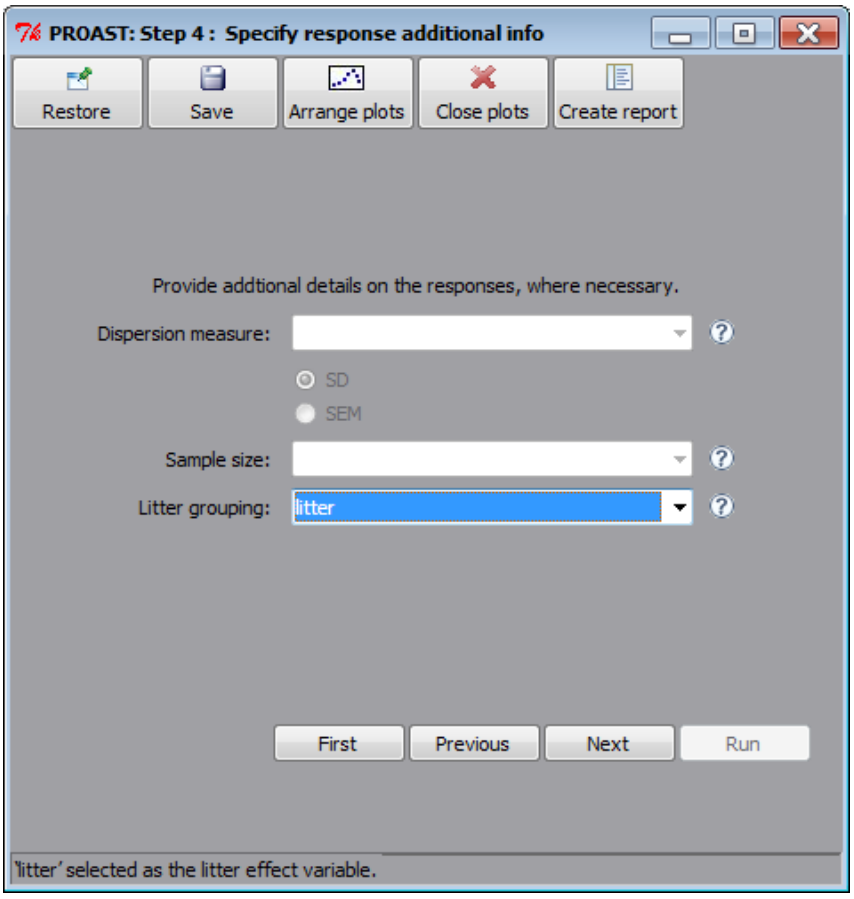


Fig. 8. Data with litter effects, the column representing the litters need to be indicated after “Litter grouping”.

STEP 5: Specify a covariate

In step 5 you may choose a covariate, defining subgroups in your dataset for which you want to compare the dose-responses. For example, when you have data for both males and females, you may select sex as a covariate. PROAST will then examine if the dose-response differs significantly between males and females, and if so, in what sense. For example, they might differ in background response but be similar otherwise, or they might differ in sensitivity to the compound, or both. For continuous data, the within-group variation is also compared among subgroups. See Fig. 9 for an illustration.

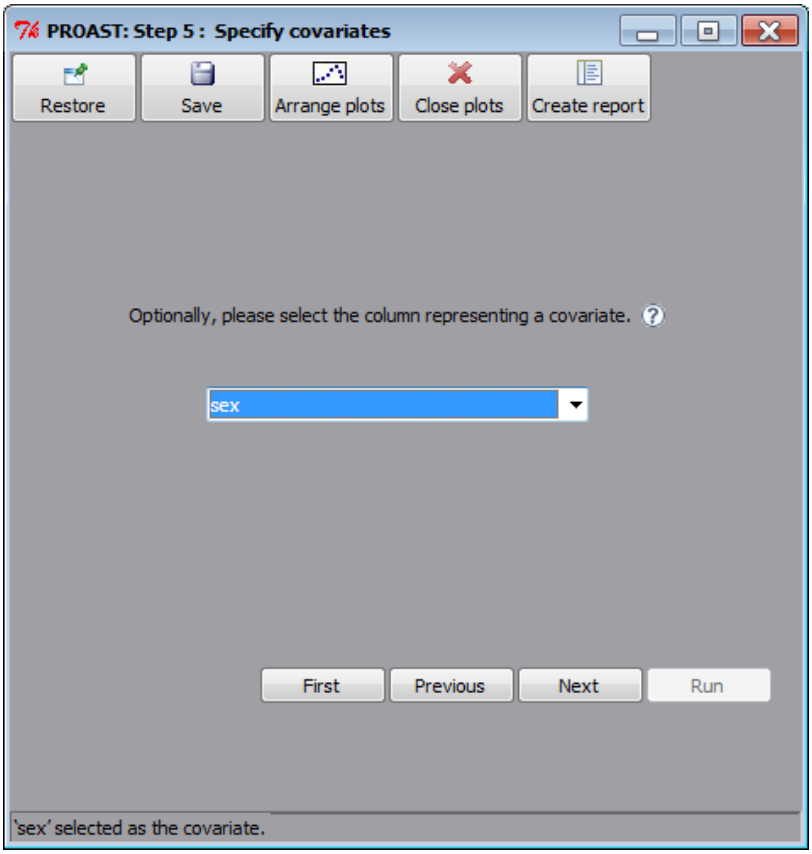


Fig. 9. Step 5 is used for specifying a covariate. A covariate may also be omitted: select <no covariate>.

STEP 6: Running the BMD analysis (*)

In step 6 you can give the command to start the analysis. Before doing so, you have the opportunity to make some final decisions.

You may want to deviate from the default BMR value. In the PROAST GUI, the default values are 5% for continuous data, and 10% (extra risk) for quantal data. If you want to specify other values, they can be entered in the first window of step 6 (see Fig. 10).
You could also enter the value 0 here. In the case of continuous data a BMR of zero will tell PROAST to omit calculation of the confidence intervals (which shortens calculation time, in particular when the covariate has many levels). In the case of quantal data, a BMR of zero will make PROAST to calculate BMDs related to a 50% response (i.e. the BMD is the ED50).

In continuous data, there are two options of selecting a model from the (exponential and Hill) families of models. The default option is to fit models 3 and 5 from each family of models, and select the one with lowest AIC for calculating the BMD confidence interval (according to EFSA 2016). There is also an option to select model 5 without considering the AIC of model 3. This second option may be useful in specific cases.

In quantal data, a suite of models is fitted, and an AIC criterion is used to distinguish the better from the poorer models. In particular, the default criterion (EFSA 2016) is a difference of 2 AIC units: the ones with AIC more than two units higher than the lowest AIC found for this dataset are discarded. The user can change the default critical difference in AIC of 2 units to another value, if deemed appropriate. (see Fig. 10). This may be needed if all or most models are rejected by the default of 2 AIC units (as this normally results from anomalies in the data, such as outliers).

The button “Advanced plotting options” may be used after fitting the model (see below). So, you are now ready to click on “Run”.

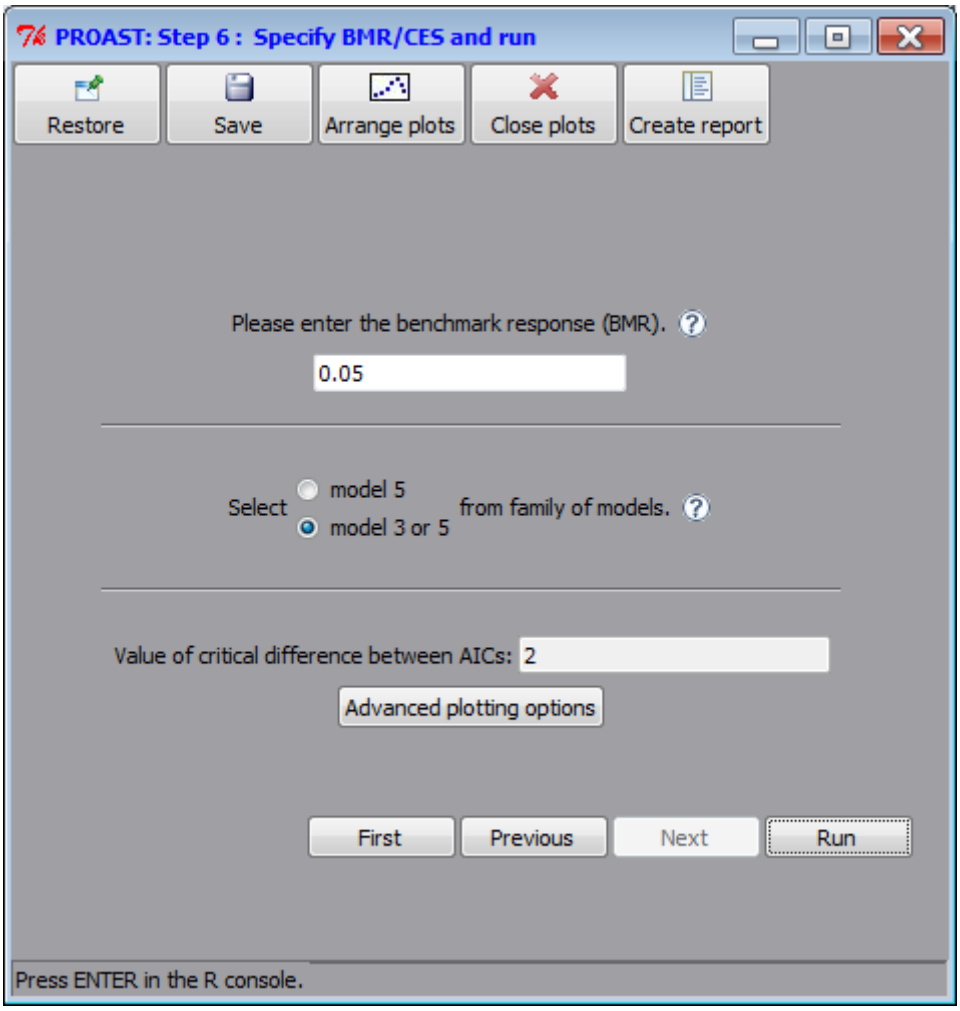


Fig. 10. In step 6, check the BMR.

PROAST output (*)

The PROAST output consists of numerical output in the Console window, and graphical output in the form of dose-response plots. These are discussed for continuous and for quantal data separately.

Output for continuous data: graphical (*)

Various graphical windows are opened during the analysis, and the final one summarizes the most important information. It consists of two plots, the left one showing the fitted exponential, the right one showing the fitted Hill model (see Fig. 11). Each plot has a legend on the right hand side, with the most important numerical results, as explained in Table 1.

Note 1: A graphical window can be copied and pasted into a Word document (right click in the plot; select Windows metafile), which may be your report platform. If you want a higher resolution, click on File in the R window, select Save as, and choose, e.g., JPEG 100%.

Note 2: It might be that the legend on the right hand side of the plot does not fit within the plot frame; this can usually be solved by enlarging the graphical window a bit.

Note 3: Just before the last plot, PROAST created two other graphical windows. They show plots of the nested models that were consecutively fitted. This information might be of interest occasionally, but normally they can be ignored.

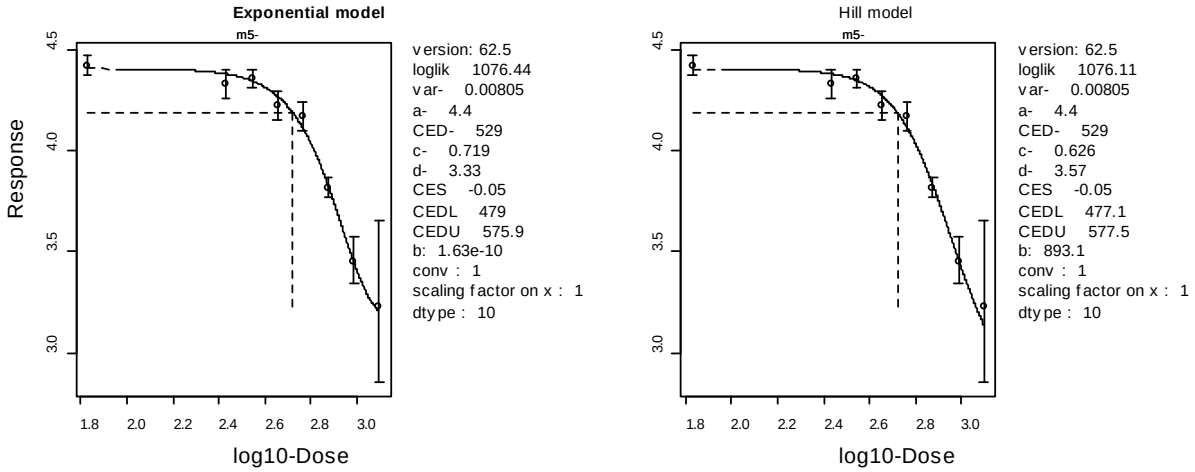


Fig. 11. Final graphical output with continuous data, in this case group means with confidence intervals. The left panel shows the results for the exponential model, and the right plot for the Hill model. The marks (circles) indicate the (geometric) means at each dose. The horizontal dashed line represents the BMR, the vertical line the BMD for the fitted curve. Note that the log of zero equals minus infinite, so the controls are plotted on an arbitrary (but low) location on the log-dose scale; for that reason the fitted curve at the lower end is represented as a dashed line.

version	the version of PROAST used
loglik	the log-likelihood value associated with the best fit (i.e. the curve shown)
var	The within-group variance (related to the natural log- responses)
a	The value of *a* (background response) according to the fitted model
CED	The CED (Critical Effect Size), or, equivalently, the BMD associated with the specified percent change in mean response
c	The value of *c* (maximum response) according to the fitted model
d	The value of *d* (steepness) according to the fitted model
CES	The value of the Critical Effect Size, which is the BMR in terms of a percent change in mean response.
CEDL	The lower confidence bound of the CED confidence interval
CEDU	The upper confidence bound of the CED confidence interval
b	The value of *b* (potency) according to the fitted model (in this case, one value for each sex)
conv	Did the fit algorithm converge? Yes if 1, no if 0.
scaling factor on x	always 1 in the GUI version (only in the menu version this value can be adjusted)
dtype	Data type, see table 2 for explanation
covariate (fig. 12)	Name of the column in the datasheet that was used as a covariate

Table 1. Explanation of the legend in the graphical output for the case of continuous data (see Fig. 11)

Type of data	Value of dtype
Individual continuous data	1
Continuous summary data	10
Quantal data	4
Individual continuous data with litter effects	5
Quantal data with litter effects	6

Table 2. Summary of datatypes possible in the PROAST GUI, with the associated values for dtype, as used in the legends of the PROAST plots (see Fig. 11).

When you had included a covariate in your analysis, PROAST will test if there are any differences between the subgroups defined by the covariate. The parameters that are found to differ among the subgroups will be reported by adding the name of the subgroups to the relevant parameter. As an example, consider Fig. 12.

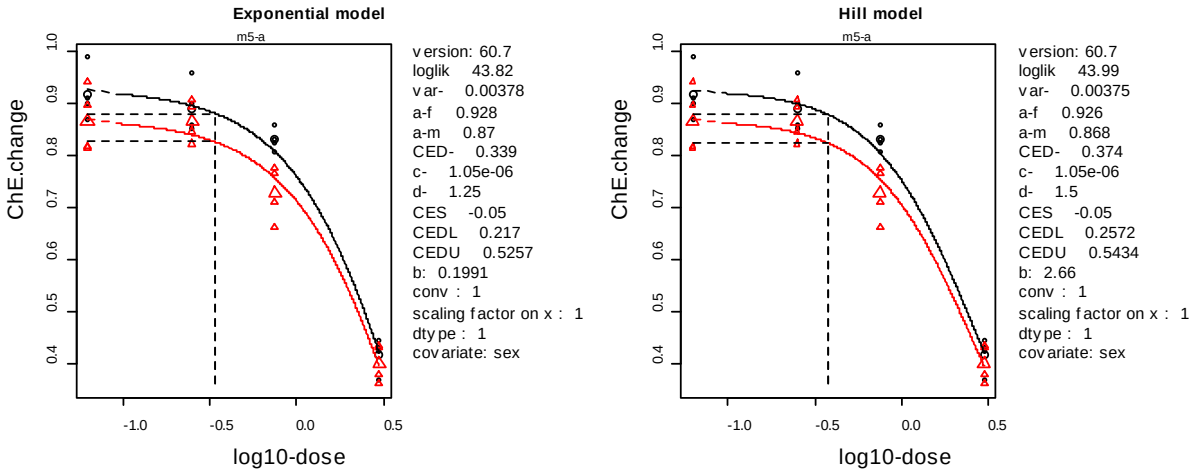


Fig. 12. Final graphical output with continuous data with sex as a covariate. The CED and its confidence interval are in this case the same for males and females, but the background responses differ, indicated by the extensions –m and –f. The larger marks in the plot indicate the (geometric) means, the smaller marks the individual observations. The triangles (and the lower curve) relate to the males, the circles (and the upper curve) to the females (this can be deduced from the fact that a-m is smaller than a-f).

Output for continuous data: numerical (*)

PROAST also shows some numerical results in the Console window, which partly overlaps the graphical output, and partly provides some additional information. It consecutively shows the analysis based on the exponential model, and that based on the Hill model. In each of these two analyses PROAST fits a series of sub-models with different numbers of parameters in it. Different numbers of parameters can be achieved in two ways. First, some of the model parameters may simply be omitted, as illustrated in Table 3. Second, some of the model parameters may be assumed to depend on subgroups in the dataset (e.g. sex), so that the number of parameters increases. By combining both options, a large list of possible models may arise. PROAST aims to compare the sub-models with varying numbers of parameters in an efficient way (i.e. not all options are fitted), and then selects one of them as being an optimal description of the data at hand. As already discussed, in the preferred approach models 3 and models 5 (possibly with covariates) are fitted, and from those models the one with lowest AIC is selected.
In the middle panel of step 6 (see Fig. 1) you can also indicate to select model 5 anyway, without considering the AIC from model 3.
When you have indicated to include a covariate in the analysis, PROAST selects the covariate model (i.e. a model with at least one parameter being dependent on the covariate) with the lowest AIC. In cases where alternative covariate models (i.e. with other parameters being dependent on the covariate) differ less than 2 units from the model with lowest AIC, this is indicated in the Console window, by saying that those other models are plausible as well.

In the PROAST output, the models in the nested families of models are denoted as in the first column of Table 3. In the case of a covariate, the parameter(s) that were assumed to depend on the covariate are added to it. For example, m3-abv means the (exponential or Hill) model that includes the parameters *a*, *b*, and *d*, while parameters *a*, *b*, and *var* are assumed to depend on the covariate.

In the Console window you can see which models were fitted, together with the number of parameters in that model, and the associated log-likelihood value (and whether the fit algorithm converged). Next, it is reported which model is selected, and for that model the parameter values are reported, as well as the BMD confidence interval. Once this procedure has been completed for the exponential model, it is similarly performed for the Hill model.
Finally, you will see the lowest BMDL and highest BMDU (combining the selected exponential and Hill model), for each subgroup separately, if applicable.

Table 3. The exponential and Hill families of models. The interpretation of the parameters is, briefly, as follows: *a* = background response, *b* = potency, *c* = maximum fold-change in response, *d* = "steepness" of the curve. Note that omitting parameters is achieved by fixing them at a specific value. The remaining parameters are "free" parameters, estimated by fitting the model to data. In the PROAST versions later than 62.1 model 4 has been omitted while model 2 is only used for getting start values for the higher models.

	Exponential family	*Hill family*	*fixed parameters*
m1	y = a		b = 0, c = 0, d = 1
m2	$y = a \exp(bx)$	$y = a [1 - x/(b+x)]$	c = 0, d = 1
m3	$y = a \exp(bx^d)$	$y = a [1 - x^d/(b^d+x^d)]$	c = 0
m4	$y = a [c – (c-1)\exp(- bx)]$	$y = a [1 + (c-1)x/(b+x)]$	d = 1
m5	$y = a [c – (c-1)\exp(- bx^d)]$	$y = a [1 + (c-1)x^d/(b^d+x^d)]$	

Output for quantal data: graphical

At the end of the analysis of a quantal dataset, you will see that the graphical window contains eight subplots, each representing another model fitted to the dataset. The first six models are the usually recommended list of models for quantal data (except the probit model, which frequently leads to numerical problems). The last two models are so-called latent variable models (LVMs), the first based on the exponential, the second on the Hill model. See the menu version of PROAST for a discussion of these models.

When you have included a covariate in the analysis, each of the eight models if fitted without covariate, then with covariate dependent parameter *a* (the background response), then with covariate dependent parameter *b* (the potency parameter), and finally with both parameters dependent on the covariate. From these four models PROAST selects the one with lowest AIC. For that model, PROAST calculates the BMD confidence interval, but only for the most sensitive subgroup (if applicable).
When any of the other models differ less than 2 AIC units from the best model, this is reported in the Console window.

After the analysis is completed, PROAST creates a plot with each model fitted to the data, as illustrated in Fig. 13. In the section "Advanced plotting options" below it is shown how the plot can be changed, and how a single plot related to one selected model can be generated, together with a legend, just like in the plots for continuous data.

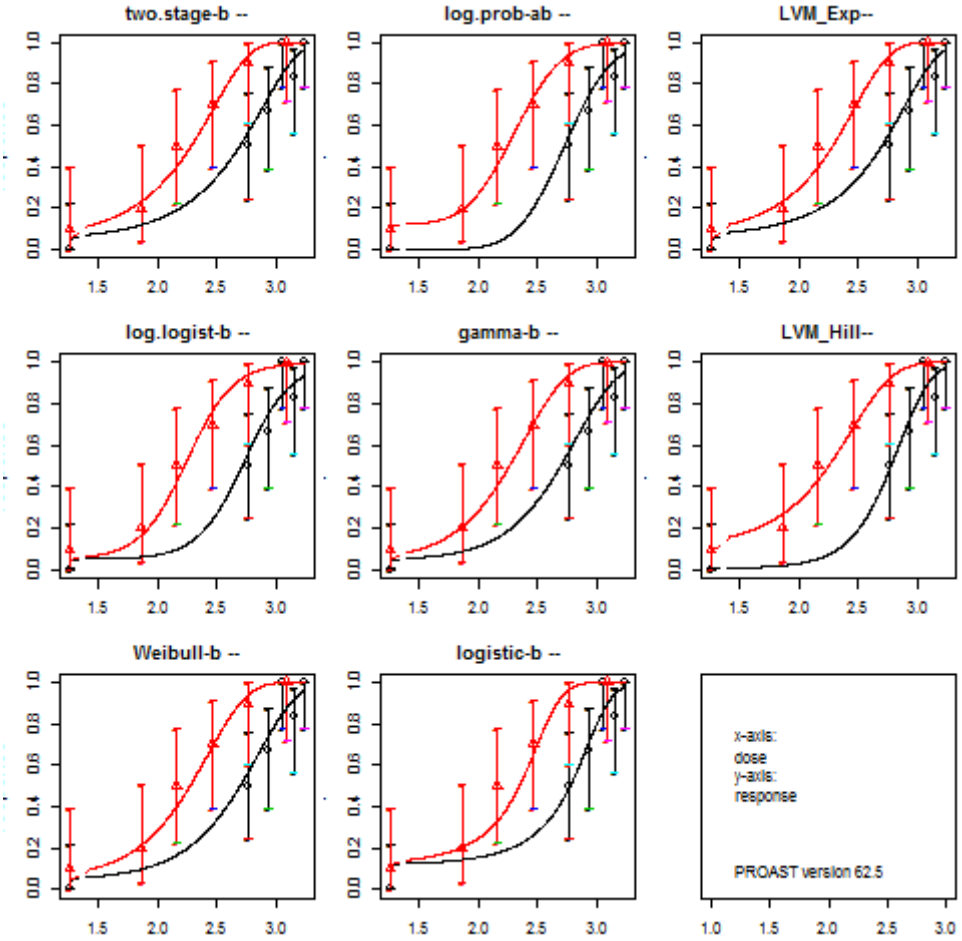


Fig. 13. Illustration of the graphical output from an analysis with quantal data. In this case, 8 different models are fitted, with two subgroups that differed in parameter *b* in most models (as indicated in the headings). The vertical whispers reflect the confidence intervals for the observed responses.

Output for quantal data: numerical

In the Console window you will see an analysis with the exponential and Hill latent variable models (LVMs), which is similar to that for continuous data.
Next, you see the message that the other models are fitted, and that the associated BMD confidence intervals are calculated.
Finally, a summary table is produced, with the results for each model, in particular, it includes the following columns:

- Model name, with selected sub-model
- Number of parameters in the model
- Value of the log-likelihood
- Whether or not the model is accepted according to critical AIC difference
- The BMDL, i.e., the lower confidence bound of the BMD
- The BMDU, i.e. the upper confidence bound of the BMD
- The value of the BMD, i.e., the MLE (maximum likelihood estimate), or the “best” estimate of the BMD.

When you used a covariate in the analysis, there are two additional columns: one providing the model parameter(s) that were determined to differ among subgroups, and one that indicates which level of the covariate was the most sensitive subgroup according to that model. The BMD confidence interval that is reported in the table relates to that subgroup.

> Note: It is important to realize that the BMD could just as well be ignored. The information that counts is the BMD confidence interval, i.e. both the BMDL and the BMDU are important, as they indicate the range of plausible values that the (true) BMD might have. In particular, when the BMD interval is wide, the “most likely” value (the BMD estimate) is not any more important than other values that are just slightly less likely.

In the section “Reporting results” it will be shown how the summary table can be made available as an editable table in other applications, including Word.

ADVANCED PLOTTING OPTIONS

When you want to change the format of the final plot, click on “Advanced plotting options” in step 6. This will create a window as shown in Fig. 14.

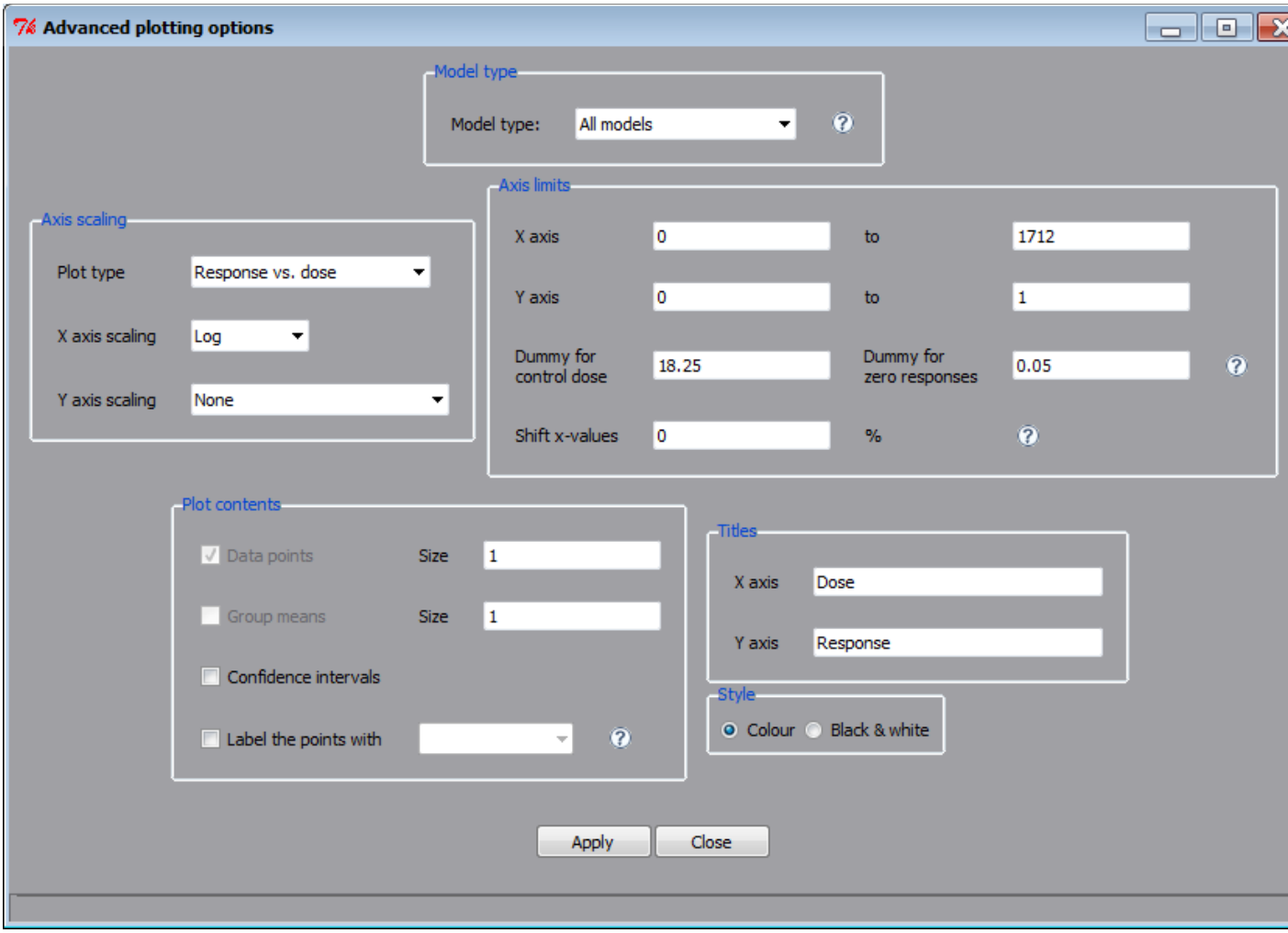


Fig. 14. The window for advanced plotting.

Upper panel

Here you can select the model you want to re-plot. When you choose “All models”, you will get a plot like Fig. 11 for continuous data, and like Fig. 13 for quantal data.
When you had continuous data, you can select the exponential model, or the Hill model. When you had quantal data, you can choose one of the eight models listed.

Left middle panel

Plot type

Next to plotting response against dose, you may, in the case of individual continuous data, also plot the residuals (i.e. the data corrected for the dose-response, or, the distance of the individual observations to the fitted dose-response curve, on log-scale). There are two options. You may want to see the QQ-plot of the residuals (see Fig. 15), to check the assumption that the data are lognormally distributed. Or, you may choose to plot the residuals against dose, to visually check the

assumption that the within group variance (for the log-responses) does not show a clear trend with dose.

Note: when plotting the residuals, you first need to select either the exponential or the Hill model, as the residuals depend on the model.

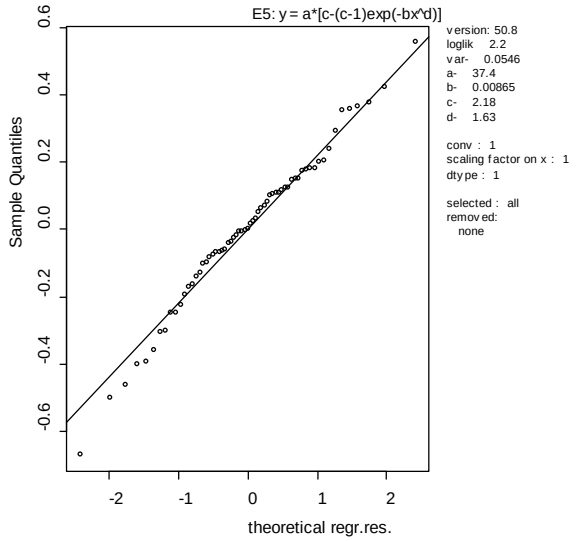


Fig. 15. Example of a QQ-plot of the regression residuals, which are the distances of all data points to the fitted curve (on log-scale). Since in PROAST the within group variation is assumed to be constant, the regression residuals can be considered as a single sample from the same distribution. That distribution is assumed to be lognormal in PROAST. Instead of creating a histogram of the regression residuals the log-normality assumption can be better checked by creating a QQ-plot, which plots the ordered residuals against the theoretical quantiles of the distribution. If the residuals are a sample from the assumed (in this case lognormal) distribution the points should approximately lie on a straight line. The line drawn in the plot is based on the sample variance of the residuals. Note that the points in the tails may deviate from a straight line due to sampling errors. If they deviate more than statistically expected, PROAST provides a message that outliers were found.

X-axis scaling

You can change the scale of the x-axis (default: log-scale) into original dose-scale, or square root-dose scale (the latter for continuous data only).

Similarly, you can change the scale of the y-axis, into log-response scale for all data types, or into sqrt-response scale for continuous data, or into arcsine-sqrt-scale for quantal data.

Right middle panel

Here, you can change the lower and upper bounds of the x-axis and the y-axis.

When you have selected a log-scale (for either the x- or y-axis) PROAST will enter a dummy value for plotting zeros (if applicable; note that the log of zero equals minus infinity). This value may be changed by the user.

When you have various subgroups in the dataset considered, it may happen that data from different subgroups are plotted at the same dose, hampering visual distinction. This may be solved by entering a value after “Shift x-values”, e.g. 80. In this way, the data from each consecutive subgroups are

shifted in the horizontal direction by a distance equaling a fraction 1/80 of the total range of the x-axis.

Lower left panel

When the dose-response data relate to individual continuous data, you can choose here if you want to plot the group means and/or the individual data, and change the size of the marks (for individual and group data points separately). If you plot both the individual observations and the group means, the marks of the group means should be larger to make them visible.

When group means or group incidences are plotted you can choose to add the confidence intervals or to omit them.

> Note: quantal data (incidences) are in fact group data, but in the plotting window they are considered as the data points. The reason is that in the case of litter effects you may want to plot the means of the incidences over all nests, which will then be considered as the group means in the advanced plotting window.

You can identify any data points by the option called “Label the points with”. If you tick this off, you can select one of the columns in your datasheet as the factor for labelling data points in the plot. After clicking apply, left click near to a data point that you want to identify, and the relevant value of the labelling factor will appear in your plot. You can repeat this several times, and right click in the plot window to stop the labeling action.

Lower right panel

Here you can change the text along the two axes, or make a black-and-white plot rather than a coloured plot.

Final note

It is recommended to close the advanced plotting window if no more changes are needed.

SAVING AND RESTORING RESULTS

At the left top of each window there are two buttons, called Restore and Save.
The Save button may be used once you have finalized a particular analysis, and you want to save the results for other or later uses. It may be helpful to give all saved results a name with the same string at the end, e.g., BW.fit and RBC.fit, where "fit" indicates that these objects are saved results.

For example, you might have copied the final plot by copy-paste into you Word document, but when you want to publish it you need a better quality. Then you could open the relevant Workspace, open the PROAST GUI, and click the button Restore. This will open a new window (if not visible, check the big blue R in the bottom bar of your screen for an item called "Load saved results"), and select the object you need. This will produce a plot with the data and fitted model. If you want, you could change the plot (using advanced plotting options), and then use the *Save as* option under *File* (left upper corner of R window). Here you will find some other formats for storing pictures (possibly needed to achieve a better solution, e.g. for publication).

As another example, you might want to do additional analyses at a later point in time (i.e., in a new session), e.g, with another value of the BMR. In that case, restoring the saved object will let you go through the steps with the previous selections are already filled in, and you only need to check them (or change them, as desired).

The number of potential additional analyses that you could do within the GUI is not as extensive as in the menu version of PROAST. If you wish to apply one of those other options, in the Console window type:
`> f.proast( , BW.fit)`
where *BW.fit* is any object that you had saved earlier. Or, just after the end of a particular session you can directly type (without saving)
`> f.proast(, ans.all)`
where ans.all is the generic temporary name of the current results.

Options in the menu version of PROAST not available in the GUI include:

- Include covariates to the shape parameters of the models
- Include more than one single covariate (related to different parameters)
- Choose other transformations on the response variable for continuous data (square root or no transformation)
- Analyze a (large) number of endpoints in one single PROAST run, resulting in a plot with BMD intervals for each endpoint
- Pairwise testing of dose-groups against the controls
- Adjust the start values of the parameters (in case of local optima, e.g.)
- Adjust the parameter constraints
- Use another factor than the covariate in the model for plotting the data with different marks
- Include a right censoring value (as opposed to the left censoring LOQ)
- Minimize the sum of products rather than the sum of squares when the data for both axes are subject to error/variation
- Relax the conditions in the fit algorithm, to decrease the calculation time in cases where the number of parameters to be estimated is large
- A larger collection of available models (including some non-monotone models)
- The option to calculate confidence intervals for any parameters in the fitted model
- The option to derive confidence intervals for RPFs (relative potency factors) or comparable parameters

- Fitting dose-response models to mixture data, based on dose addition, making clear if dose-addition applies or not over the whole dose-range
- Analyse response as a function of time (e.g. body weights), and then select a parameter characterizing that curve (e.g. growth curve) in a dose-response analysis
- Plot subgroups with their fitted curves separately, in particular useful with datasets consisting on multiple subgroups
- Make plots where the curves for different subgroups are scaled to the same background response
- Calculate BMD confidence intervals using parametric bootstrapping for any datatype
- Analyze ordinal (e.g. histopathological) data, and derive BMD confidence intervals
- Analyze binary data (i.e., yes or no lesion in each animal or experimental unit)
- Calculate withdrawal periods for veterinary medicines

See the “PROAST MANUAL menu” for a description of the options in the menu version of PROAST.

REPORTING RESULTS *

After a PROAST analysis you can create a summary of the results in an HTML page. Simply click the button Create report. You can copy-paste the HTML page into Word, and edit it as desired.

REFERENCES

EFSA (2009). European Food Safety Authority. Guidance of the scientific committee on use of the benchmark dose approach in risk assessment. EFSA J 1150:1–72.

US EPA (2012) Benchmark Dose Technical Guidance.

http://www.epa.gov/raf/publications/benchmarkdose.htm

ANNEX 1. Data format, examples for various data types.

Each column needs a header. The header should be one single string, without spaces of special characters, but dots and underscores are allowed.
Empty cells in the datasheet are not allowed. Missing values should be indicated as NA.
Use brief strings (say, no more than three characters) for the levels of the additional factors (potential covariates), as they may be used in the graphical output.
An observation below the limit of quantification (LOQ) should be indicated by a zero. Conversely, zeros will be read as a value below the LOQ.

Continuous individual data.

Here, various endpoints are combined in the same datasheet, with one additional covariate (sex).

dose	BWabs	relKidney	relThym	relBrain	sex
0	104.2	7.34	1.34	8.74	m
0	106.1	6.54	1.61	9.31	m
0	117.7	6.88	1.68	9.08	m
0	112.4	7.35	1.32	8.32	m
0	109.4	7.11	NA	9.86	m
75	99.5	7.62	1.67	9.26	m
75	102.3	7.07	1.89	9.46	m
75	106.1	7.67	1.92	9.71	m
75	98.9	7.11	1.74	9.28	m
75	118.7	7.78	1.86	9.65	m
300	107.5	NA	1.49	9.13	m
300	111.2	8.81	1.88	8.84	m
300	108.8	8.47	2.07	9.45	m
300	97.2	8.15	1.57	9.17	m
300	90.2	8.61	2.11	10.9	m
0	55.2	7.62	2.29	12.28	f
0	52.5	7.29	1.69	13.24	f
0	57.3	7.21	2.14	12.35	f
0	57.9	7.01	2.2	12.06	f
0	52.9	6.7	1.63	12.07	f

Continuous grouped data (summary statistics).

For summary statistics (mean observations) it is required to have a column with SDs (or SEMs), and a column with the group sizes (here called: N).

dose	mean.bw	sd.bw	N	sex
0	704	124.7	33	m
0.5	739	140.5	35	m
3.5	742	97.7	40	m
25	646	119.4	41	m
50	572	97	49	m
0	496	105.7	37	f
0.5	477	132.6	33	f
3.5	480	106.8	32	f
25	402	106.8	27	f
50	361	81.1	20	f

Quantal grouped data (incidences).

In quantal data the number of animals with a given lesion should be accompanied by the group size.

dose	animals_with_tumors	N	sex
0	0	20	1
0.5	2	18	1
3.5	5	19	1
25	6	20	1
50	12	18	1
0	1	20	2
0.5	0	19	2
3.5	3	20	2
25	7	19	2
50	10	19	2

Individual continuous data with litter effects.

dose	sex	exposure_duration	dam	foetalBW
0	1	2	2	4.11
0	1	2	2	4.22
0	1	2	2	4.27
0	1	2	2	4.26
0	1	2	2	4.03
0	1	2	2	4.13
0	1	2	2	4.07
0	1	2	2	4.27
0	1	2	2	4.21
0	1	2	2	4.03
0	2	2	2	4.17
0	2	2	2	3.93
0	2	2	2	4.12
750	1	1	6	3.96
750	2	1	6	3.97
750	2	1	6	4.05
750	2	1	6	3.39
750	2	1	6	3.27
750	2	1	6	3.59
750	2	1	6	3.55
350	1	1	7	4.86
350	1	1	7	4.31
350	1	1	7	4.78
350	1	1	7	4.62
350	1	1	7	4.63
350	1	1	7	4.53
350	1	1	7	4.3
350	1	1	7	4.26
350	1	1	7	4.44
350	2	1	7	3.87
350	2	1	7	4.3
350	2	1	7	3.92
350	2	1	7	4.16
350	2	1	7	3.95
580	1	1	9	3.62
580	1	1	9	4.32
580	1	1	9	4.18
580	2	1	9	3.84
580	2	1	9	3.86
580	2	1	9	3.64

ANNEX 2. Some basic R commands

Some basic R commands where discussed at the end of section “Starting the PROAST GUI”. This annex provides some more, as they may be helpful, in particular for more regular users.

R can use the information available from other folders or packages, but only when they are “attached”. The first in the list of attached folders is your working directory. When you type

```
> search()
```

you will see all attached folders and packages. The first is called “.GlobalEnv”, which is in fact your current working directory. Some of the packages in this list are attached automatically when you open R. Others have been attached after starting the PROAST GUI, as they are required for various tasks. And, of course, you will see the PROAST package (if you have loaded it).

Switching to another PROAST version.
It might happen that you want to change to an earlier PROAST version (for instance, when the current version resulted in a problem that did not occur in an earlier version, or when a numerical result seems to differ from an earlier version). After having established the position of the PROAST package in the list (by using the search command), first type

```
> detach(10)
```

where 10 is position in the list with the PROAST package (or use another number if appropriate). When you now type search() it should be gone. You are now ready to load another PROAST package (possibly, you may first need to install it). If you type search() again, you should now see the other version.

Organizing the objects in your Workspace.
When your Workspace is getting crowded, it may be helpful to remove objects that are no longer needed. This can be done by typing

```
> rm(name)
```

where *name* is any object to be removed.
If you want to rename an object, simply type

```
> newname <- oldname
```

and next

```
> rm(oldname).
```