MANUAL FOR THE PROASTweb APPLICATION[1]

12 January 2023

This manual should enable you to independently perform BMD analysis on your own data. In case you encounter any technical issues, or if you have questions on the BMD approach implemented in PROAST, you may contact us at proast@rivm.nl

[1] Suggested citation: RIVM (2023) Manual for the PROASTweb application, version 70.1. National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands. Available from https://www.rivm.nl/en/proast

Contents

RIVM (2023) Manual for the PROASTweb application, version 70.1. Bilthoven, The Netherlands.

Open the PROASTweb application
Go to https://proastweb.rivm.nl/

Page: Overview (before analysis)

After reading and closing the introduction text and disclaimer, you will see the overview page (Figure 1), containing a list of the analyses you have done today, which is empty when first starting.

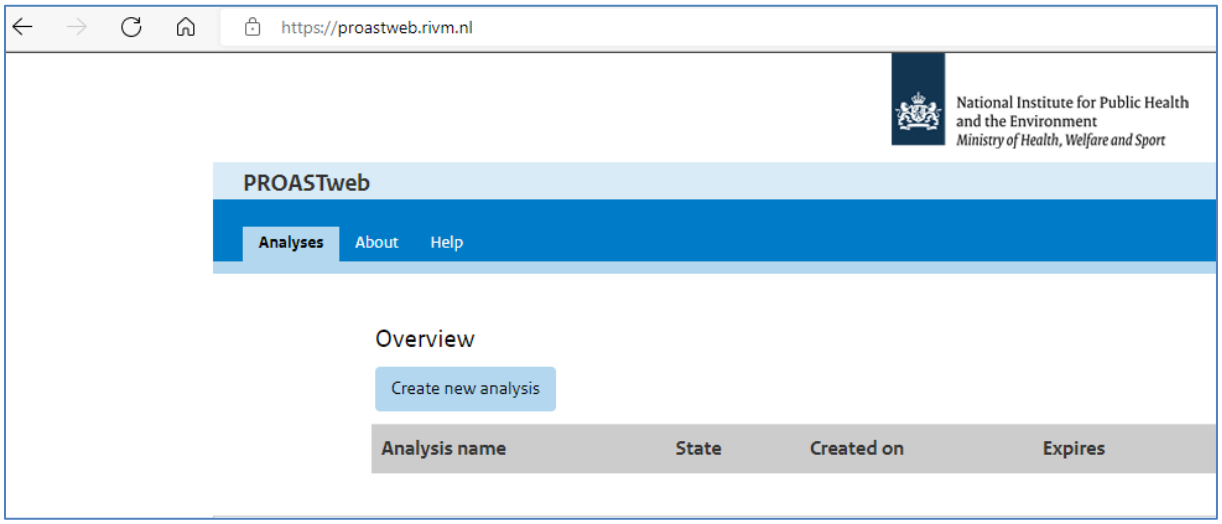

Figure 1. Overview page of PROAST web application.

Click on *Create new analysis.* The Analysis page will now be opened.

Page: Analysis

On this page (Figure 2) you can upload your data and provide a name for your analysis

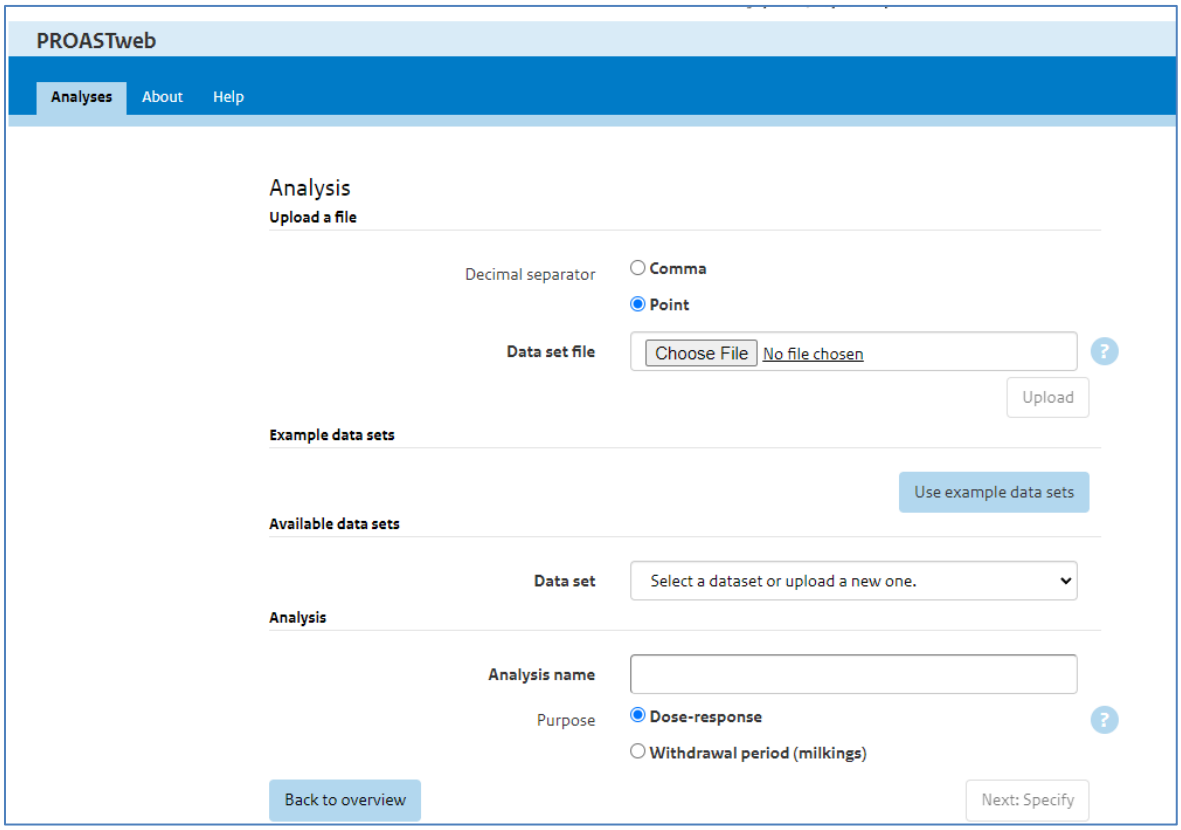

Figure 2. New page after clicking *Create new analysis.*

Section: Upload a file

Decimal separator

Check “comma” or “point”, to indicate which decimal separator is used in the dataset that you want to upload (this depends on the regional settings of your computer), see Figure 2.

Data set file

You can upload a dataset for your analysis. The help function, indicated by a question mark (next to *Choose file*) gives a brief description of the required format of a dataset. For a more detailed description on the requirements of the data format see Annex 1. Click on *Choose file* and browse to the dataset on you computer. Select the dataset, and then click on *Upload.* You will get a message that the data have been uploaded, and you will see the data at the bottom of the web page (you might need to scroll down).

Note: When you get an error message, this is most likely caused by a format error in the dataset. The help function, indicated by a question mark (, next to *Choose file*) gives a brief description of the required format of a dataset (for a more detailed description on the requirements of the data format see Annex 1). Check your dataset to see if the data format is correct.

Section: Example data sets

Besides uploading your own data, you can also use an example dataset, after clicking on *Use example data sets*.

Section: Available data sets

When you successfully uploaded your own dataset, the name of the updated dataset will appear in the field *Data set* of this section. If you have clicked on *Use example data sets*, then a list of available example data sets in PROASTweb will be available the field *Data set*, which will appear by clicking on the drop down symbol ( ). The available example data sets for BMD analysis are: dataset of continuous response, dataset of quantal response, and dataset of multiple continuous responses.

After uploading you own data or having selected an example data set, you can make a subset selection from the data (see Figure 3), by clicking on the drop down symbol ( ) next to the column header, and uncheck the values (or levels) you want to exclude. Cancellation of the selection of the subset(s) can be simply done by checking the previously excluded values (or levels).

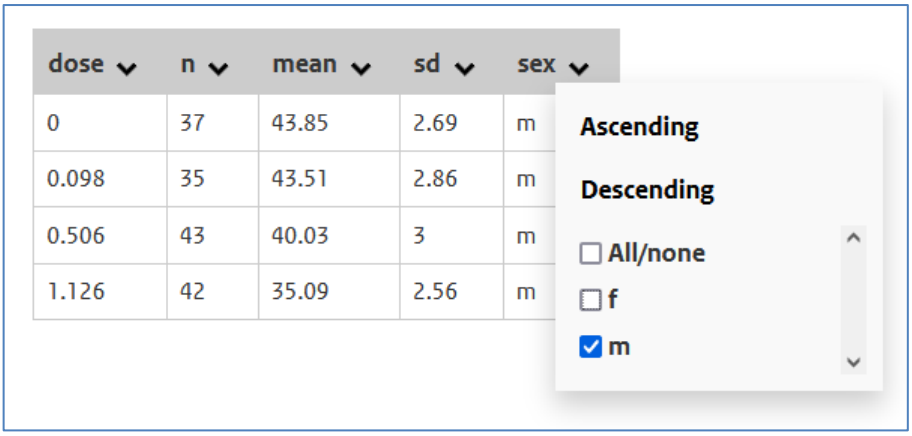

Figure 3. Illustration of making a subset selection from the data that were uploaded (in this example females were excluded from the data set).

Section: Analysis

Analysis name

Fill in a name for your analysis in the field *Analysis name*, which will be used in the overview of performed analyses on the (next) page called Overview.

Purpose

In this field choose the option *Dose-response*. More details (including the relevant manual) about the option *Withdrawal period (milkings)* can be found at https://www.rivm.nl/en/proast.

Click on *Next: Specify*

Page: Specify

On the page Specify, you can indicate which column in your dataset represents the dose, the response, and other settings, depending on the situation. Note that the question marks may help you with some more information.

Section: Analysis general

Dose column(s)

In the first field named *Dose column(s)* you need to indicate which column should be regarded as the dose (see Figure 4). By clicking on the empty field or the downward arrow on the right side ( ) a drop-down menu appears to make a selection.

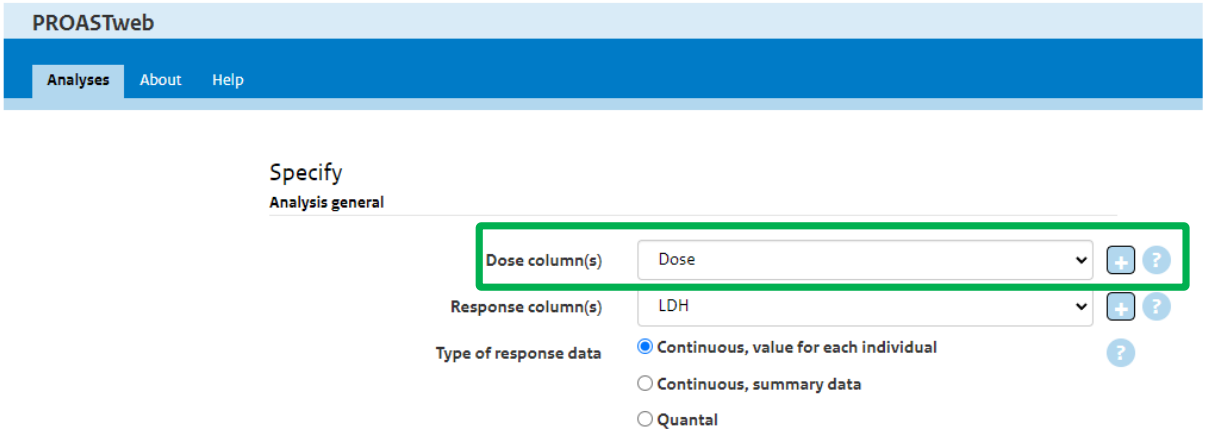

Figure 4. Illustration of selecting the dose column for the dose-response analysis. In this example the column Dose in the dataset is selected.

Response column(s)

In the field *Response column(s)* you need to indicate which column should be regarded as the response (see Figure 5).

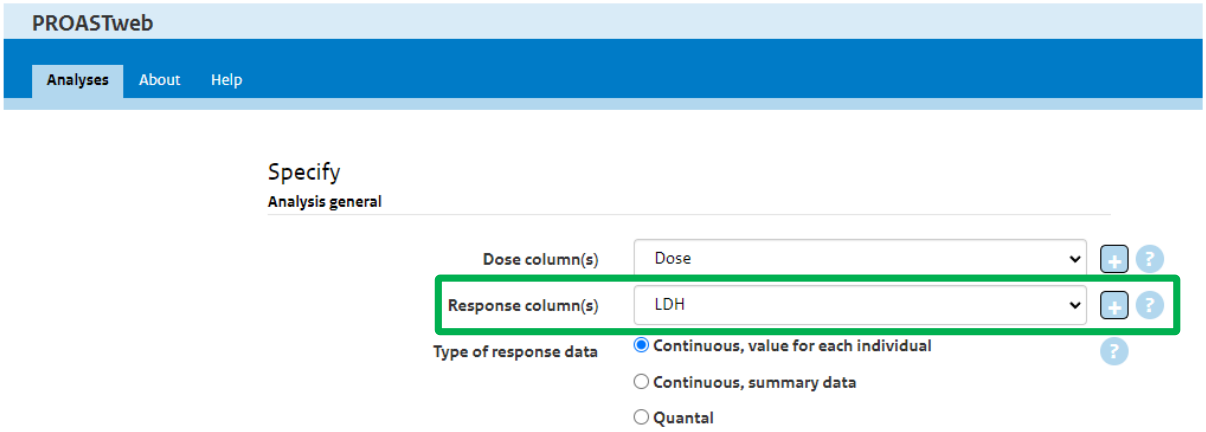

Figure 5. Illustration of selecting the response column for the dose-response analysis, in this example a column named LDH is selected for the analysis.

Type of response data

The options of the type of response data are (Figure 6) continuous value for each individual, continuous summary data, and quantal (or dichotomous).

Section: Dispersion measure

If you choose continuous summary data, then an extra section *Dispersion measure* needs to be filled out, including the field *Dispersion measure column(s)*, and the field *Relating to* which indicates whether the chosen column(s) relate to the standard deviation (SD) or standard error of the mean (SE).

Meanwhile if either continuous summary or quantal is selected for the *Type of response data*, also the information on group size need to be provided, this is indicated in the field *Group size column(s)* in the section *Other settings*.

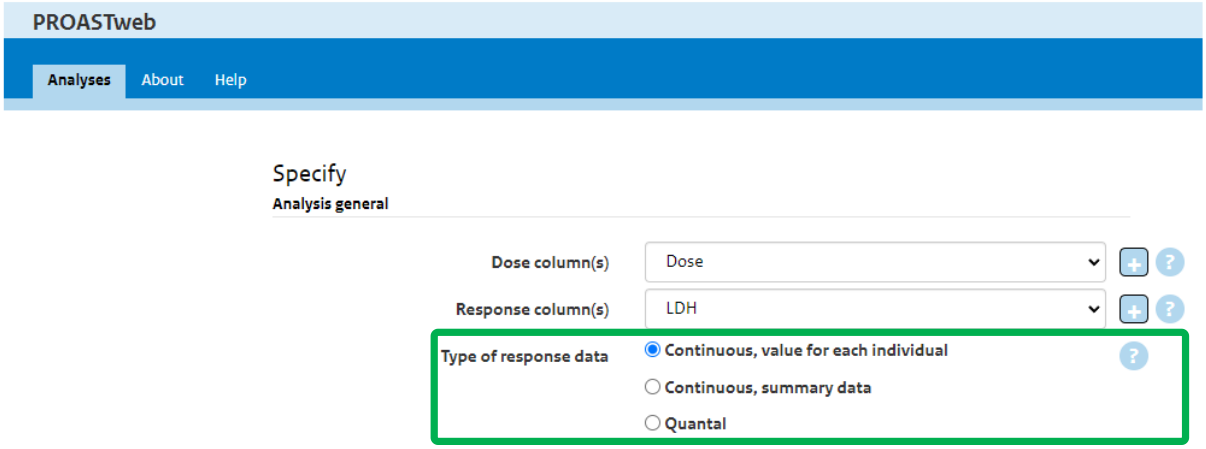

Figure 6. The options for choosing the type of response data in PROASTweb.

Section: Litter effect

PROASTweb allows for a dose-response analysis where the individual observations per dose group are not independent. Such data is often indicated as clustered data, nested data or litter data. For example, such data are obtained from developmental studies where the observations related to the embryos or the pups may show correlations within litters (litter effects). In this case the option *Litter effect* should be ticked.

Litter grouping column

When you analyze clustered individual continuous data and the option *Litter effect* is ticked, you also need to select in the field *Litter grouping column* a column which indicates the factor defining the clustering (the litter effect).

Number of bootstrap runs

When you analyze clustered continuous data (both individual and summary), the number of bootstrap runs have to be provided in the field *Number of bootstrap runs*. For clustered continuous data, the BMD confidence interval of each individual model is assessed by the bootstrap method in PROASTweb. The higher the number, the more accurate the BMD confidence interval will be. However, the calculations may take quite some time, and it is recommended not to use an overly large number of bootstrap runs.

In case of continuous litter effects model averaging is not implemented. Instead, the BMD confidence intervals of both the exponential and Hill models are calculated.

Section: Other settings

Group size column(s)

In case of continuous summary data and quantal, you need to provide a column with group size in the filed *Group size column(s)*.

Covariate column (optional)

In PROASTweb, the user may appoint a factor (e.g. sex) to be handled as a covariate in the model. PROAST will then explore if particular parameters in the model should receive different values, depending on the level of the covariate (e.g. males vs. females of the covariate sex). For instance, the parameter reflecting the background response (e.g. body weight in the controls) may differ between subgroups (e.g. males and females). Further, one subgroup may be more sensitive to the compound than another (reflected by different values for the parameter b of the dose-response models), or differences in residual variation (continuous data only) may exist. The covariate can be indicated in the field *Covariate column (optional).*

BMR (CES)

For the BMD analysis, the value of BMR (or CES, which is specifically used for continuous data) needs to be specified. In the field *BMR (CES)* the value of BMR needs to be filled in as a fraction, e.g. a BMR of 10% change in response or extra risk is entered as 0.10.

Continuous dose–response data are explicitly measured as gradually changing (in each individual). For example, the measure of organ weights, enzyme concentrations in blood, and large counts, e.g. of red blood cells, are continuous data. For continuous data, the metric for the BMR can be defined in various ways. The preferred option, implemented in PROASTweb, is to define the BMR as a relative (percent) change (or relative deviation) in mean response as compared to the mean background response:

$$BMR = \frac{|f(0) - f(BMD)|}{f(0)}$$

Where f(0) is the estimated background response and f(BMD) is the response at dose BMD. The BMD associated with such a BMR does not depend on the background response and the within-group variation. Therefore, BMDs derived from the same endpoint and corresponding to a BMR defined in this way can be compared (and extrapolated) among different studies, species and (sub)populations.

For quantal data, the BMR is defined in terms of an increase in the incidence of the lesion or response scored, compared with the background incidence. In PROASTweb, the metrics for reflecting such an increase is the extra risk:

$$BMR(extra\ risk) = \frac{P(BMD) - P(0)}{1 - P(0)}$$

Where P(BMD) is the incidence at the BMD and P(0) the incidence in the control. 1 – P(0) is therefore the non-affected fraction of the population.

Model averaging

Indicate whether to apply the model averaging approach by ticking this option (if apply model averaging) or leave the option unticked (if no). If model averaging is selected, the number of bootstrap runs should be given. The default number of bootstrap runs is 200.

If model averaging is not selected the BMD confidence intervals of the exponential and Hill models are calculated in case of continuous data, and of the suit of 7 quantal models in case of quantal data.

AIC criterion

For the purpose of comparing the fits of different models, the AIC (Akaike information criterion) is a convenient criterion as it directly integrates the log-likelihood and the number of model parameters in one single value. The default AIC criterion is 2, which is recommended by the EFSA Scientific Committee (for more details see the EFSA guidance 2017, doi: 10.2903/j.efsa.2017).

When you have finalized the specifications, click on *Run*. This will bring you (back) to the **Overview** page (see Figure 7).

Analyzing mixtures

If you want to analyze a mixture study, click on the  on the right hand side of the dose window (see Figure 4), and select the columns for the compounds you want to include. You can select multiple dose columns by holding the Ctrl key and meanwhile clicking each dose column. Do not forget to remove data from compounds you do not want to include in the mixture analysis (using subset selection in the uploaded data). For analyzing a mixture study, a column is needed in the dataset to denote if the row relates to a single chemical or a mixture (i.e. the single dose and mixture column). This column needs to be selected in the field *Single doses and mixture column*. For an example of the mixture dataset please see Annex 1.

Analyzing a batch of multiple endpoints

If you want to analyze multiple endpoints in one single run, you first need to choose in the field *Type of response data* (see Figure 6). Note that the option of analyzing multiple endpoints is not yet implemented for quantal data. Click on the  on the right hand side of the field *Response column(s)* (see Figure 5). Select multiple dose columns by holding the Ctrl key and meanwhile clicking each response column. In the case of continuous summary data you also need to specify in the field *Dispersion measure column(s)* the associated columns with the SD or SE. Similarly, you need to specify the associated columns with the group sizes for each endpoint selected. However, it is possible to select one column for the group sizes, in case they are the same for all endpoints. The order of the dispersion measure and group size columns in the uploaded dataset needs to be in the same order as the response means to ensure that the correct combination of mean, SD/SEM and group size are analyzed. To prevent errors, it is recommended to order the columns in triplets (response, SD or SEM, and N) for each endpoint, or in duplets (response and SD or SEM) if there is only one column for the group sizes.

Page: Overview (after starting an analysis)

After clicking *Run* on page *Specify* you are redirected to the *Overview* page. The analysis you just indicated is now visible, with *State* "Queued", followed by "InProgress" and "Finished".

When the analysis you started takes some time, you can in the meantime create and run a new analysis using the same or a newly uploaded dataset. The jobs that are running will just continue, and the results will be maintained for each analysis name that you specified. The analysis results will be stored for 24 hours after the analysis is finished.

> Note: PROASTweb is a publicly free tool, therefore in order to maximize the use of the resources supporting PROASTweb, analysis that runs longer than 8 hours will be terminated and erased from the server. For intensive analysis that requires rather high demand of computational capacity, it is recommended to use the (R-standalone) menu version of PROAST, which can be downloaded from https://www.rivm.nl/en/proast.

On the **Overview** page (see Figure 7), the third icon on the right of "*Finished*" contains "Report analysis results" (Figure 7).

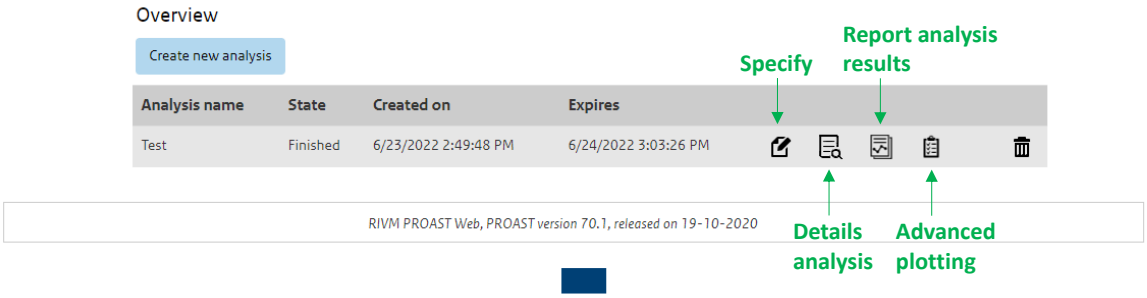

Figure 7. Overview page of PROASTweb application with a performed analysis named 'Test'. In this example the analysis is finished (i.e. *State* is "*Finished*"). On the right side of "*Finished*", there are four icons, namely Specify, Details analysis, Report analysis results, and Advanced plotting.

Page: Report

Clicking on the *Report* icon opens a new HTML page with a summary of the results, including the input values, graphical output, fitted models, and BMD confidence interval. When model averaging was chosen, the model weights are reported as well. And in this case the BMD confidence interval is based on model averaging. When model averaging was not applied, the lowest BMDL and highest BMDU from the applied set of models will be used to define the BMD confidence interval. To save your results you can copy-paste the content of the *Report* webpage to a Word document or print the webpage as pdf.

> Note: when copying to Word, sometimes the figure does not come along, in that case right click on the figure (without selecting it) and select *copy.* Then, in Word, click on *paste* (under *Home,* see upper left corner), and next on *Paste special*. In the window that appears select *Device Independent Bitmap*. Now the figure should show up.

Page: Details

The second icon *Details analysis* on the overview page contains some additional output of the analysis (the Console output). The Console output may contain important information on:

1) Significance test in case of a litter analysis. For a dose-response analysis where the individual observations per dose group are not independent, PROASTweb will fit the model, as if there are no litter effects, and after that estimate the within and between litter (or other subgroup) variances. Both variance components are reported in the Console window, followed by a P-value, which indicates whether the litter effect is statistically significant.

2) Outlier test. For the analysis of individual continuous data, PROASTweb will perform an outlier test (based on Grubb's test). The results of this test can be found in the Console output in this page.

3) Warning that the AIC of the best model (minimum AIC) is more than two units larger than that of the full model. The full model is used as a reference in the goodness-of-fit test, i.e. the AIC associated with the full model is used as a reference to the AICs of the fitted models, and lead to an alert if the AIC of the full model is substantially smaller than all other models. When the AIC of the best model (minimum AIC) is more than two units larger than that of the full model, there should be an alert that

a problem in the data might be indicated, in particular when the difference is much larger than two units (e.g. > 5).

Page: Specify

The first icon *Specify* can be used if you want to change the specification related to the same dataset. Note that the name of the analysis will not change, so that you will loose the previous results from the overview. If you want to keep these results, then click the *Create new analysis* button and redo the analysis with different specifications.

Page: Advanced plotting

The fourth icon *Advanced plotting* can be used for adjusting the format of the plot. When you click on the icon, you will see a page with various fields that allow you to change various plot characteristics (model type, axis scaling, axis limits, plot contents, titles, and style). After having clicked the *apply* button, you can find your adjusted plot in the report (third icon of the overview page).

ANNEX 1. Data format, examples for various data types

Data format

The required format of the data is as usual for statistical analysis of dose-response data. In the datasheet each column needs a header. The header should be one single string, without spaces or special characters[2], but dots and underscores are allowed. Use brief strings (say, no more than three characters without spaces or special characters) for the levels of the additional factors (potential covariates), as they may be used in the graphical output.

In its simplest form, a dose-response dataset consists of one column with the doses and one 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., male or female), 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, or be used to make a sub-selection of the dataset. 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 WebApp (as illustrated in this manual).

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 (empty cells in the datasheet are not allowed). Missing values should be indicated by NA. Values below the limit of quantification (LOQ) should be given as 0 (PROAST will then ask to provide a value for the LOQ). Conversely, zeros will be read as a value below the LOQ.

It is required to save the datasheet in a tab delimited text file or a comma delimited file (in Excel: use “save as”, go to “save as type”, and select Text (tab delimited)(*.txt) or CSV (Comma delimited)(*.csv)), and import the text or csv file in the PROAST web application. Below are some examples of different types of dose-response datasets.

Examples of dataset

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	NA	9.31	m
0	117.7	6.88	1.68	9.08	m

[2] ASCII punctuation and symbols; https://en.wikipedia.org/wiki/Basic_Latin_(Unicode_block)

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	NA	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
etc.					

Continuous summary 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
etc.				

Mixture dataset (continuous individual data)

For datasets from mixture studies, the doses for each single chemical are entered in a separate column. Meanwhile the mixture dataset should include a column indicating for each row which single chemical was applied, or that a mixture was applied (i.e. the single dose and mixture column). In the example below continuous individual data are shown. Chem1 and chem2 represent two chemicals. When continuous summary data or quantal data are obtained from a mixture study the format of the dataset is similar, i.e. instead of one dose column multiple dose columns are needed and a column indicating the single dose and mixture needs to be added.

Chem1-dose	Chem2-dose	Single.dose.and.mixture	Response

0	0	chem1	8029000
0	0	chem1	6675200
0	0	chem1	6144133
0	0	chem1	7710500
0.01	0	chem1	6905033
0.01	0	chem1	11172000
0.01	0	chem1	6534500
0.01	0	chem1	9646000
etc.			
0	0.02	chem2	9789500
0	0.02	chem2	8564500
0	0.02	chem2	7042000
0	0.02	chem2	5351500
etc.			
0.02	0.02	mix	8092000
0.02	0.02	mix	5680500
0.02	0.02	mix	7045500
0.02	0.02	mix	9464000
0.04	0.04	mix	7154000
etc.			

Quantal grouped data (incidences).

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

dose	animals_with_tumors	N	sex
0	0	20	m
0.5	2	18	m
3.5	5	19	m
25	6	20	m
50	12	18	m
0	1	20	f
0.5	0	19	f
3.5	3	20	f
25	7	19	f
etc.			

Individual continuous data with litter effects

For the individual continuous data with litter effects, it needs to be indicated which column represents the litter factor. In the example below the column 'dam_ID' represents the litter factor. For example, dam number 16 had 5 pups of which the body weights are reported in separate rows. Note that each dose group can contain multiple dams. Columns can be added to indicate subgroups (e.g. sex of the pup).

dose	dam_ID	foetalBW
0	16	4.17
0	16	4.28
0	16	3.94
0	16	4.3
0	16	3.94
0	25	4.23
0	25	4.21
0	25	4.42
0	30	5.03
0	30	4.88
0	30	4.76
270	40	4.25
270	40	4.21
270	40	4.18
270	40	4.28
270	40	3.71
270	64	4.31
270	64	4.56
270	64	4.4
270	64	4.4
270	64	4.55
270	85	3.97
270	85	4.11
270	85	4.47
270	85	4.33
270	85	4.22
350	7	4.86
350	7	4.31
350	7	4.78
etc.		

Continuous summary data and quantal data with litter effects

With clustered continuous summary data and quantal data there is no need to have a column defining the clusters, as they are already defined by the summary statistics (just like in clustered quantal data, see next example), each row of the table below indicates a cluster (or dam or litter).

dose	mean_foetalBW	sd	N*
0	4.126	0.177	5
0	4.287	0.116	3
0	4.890	0.135	3
0	4.567	0.233	7
0	4.423	0.212	3
0	4.680	0.113	2
270	4.126	0.236	5
270	4.444	0.108	5
270	4.159	0.224	8
270	3.817	0.400	3
350	3.908	0.155	5
350	4.810	0.270	4
350	4.562	0.136	6
350	3.796	0.312	5
350	4.129	0.243	9
350	4.747	0.159	7
350	4.623	0.302	4
etc.			

*N=number of pups in the litter

Quantal data with litter effects

dose	nr_born	dead_pups
0	11	0
0	11	1
0	12	0
0	11	0
0	14	0
0	12	0
0	13	2
0	13	0
0	14	1
0	13	0
2	8	3
2	14	0
2	10	3
2	13	0
2	14	1
2	9	0
2	13	0
2	12	0
6	12	0
6	13	1

6	10	4
6	13	2
6	9	4
etc.		