CATEGORICAL REGRESSION (CATREG)

USER GUIDE

VERSION 3.1.0.7

Release Date: June 16, 2017

Developed for US EPA

United States
Environmental Protection
Agency

TABLE OF CONTENTS

1.0 INTRODUCTION

The U.S. Environmental Protection Agency (EPA) developed the Categorical Regression (CatReg) application as a tool to facilitate exposure-response analyses. The latest version can be obtained from the EPA CatReg website.

In general, CatReg should be viewed as a statistical tool for examining exposure-response-time relationships and addressing related questions. A thorough analysis may require numerous executions of CatReg, ideally guided by both toxicological and statistical considerations. The program's Windows interface and model options facilitate sensitivity analysis and produce numerous plots in addition to hypothesis-testing results.

Key features of CatReg include:

- The use of multiple independent variables to explain the response
- The ability to predict exposures related to various levels of effect severity
- Extensive plot features for the analysis of model results
- A user-friendly hypothesis testing interface that assists the user in:
    - Assessing and comparing how well models fit the data
    - Determining the feasibility of combining data from multiple studies by testing for differences across studies and the significance of covariates within single or pooled studies
    - Detecting outliers

CatReg's advantages over other dose-response analysis tools include its abilities to:

- **Use multiple independent variables to explain the response.** Most often, the independent variables modeled in CatReg are exposure (or dose) and duration of exposure. The latter allows for estimating risk at different exposure durations—even time points not included in the dataset being modeled.
- **Predict exposures related to various levels of effect severity.** Using categorical regression to evaluate different levels of severity is of particular use for developing emergency planning guidelines, which are defined in terms of the exposure limits needed to protect people from increasingly severe effects (e.g., protection from mild effects, severe effects, and death).
- **Combine multiple studies in a single analysis.** The meta-analytical use of categorical regression can be invaluable when individual studies provide only minimal useful dose-response data. Classifying toxic effects into severity categories provides a way to put different endpoints on a common scale for analysis and also allows the use of all types of data in the analysis. For example, individual animal studies investigating liver lesions and serum biochemistry can be combined with human studies reporting preclinical and clinical liver effects to provide an overall estimation of the dose-response for liver toxicity.

    Continuous, descriptive and categorical data can be used, in addition to incidence data, as long as these effects can be classified into severity categories (Appendix A describes an approach for converting continuous response information into severity categories).
- **Test hypotheses.** Finally, the CatReg 3.0 implementation of categorical regression offers valuable and practical plotting and hypothesis testing tools, combined with a user-friendly interface, to assist the user in determining whether data associated with variables that can sometimes influence a dose-response (e.g., sex, species, diet, age, genetic differences) should be evaluated separately or combined together into a more robust, data-rich multiple study meta-analysis.

This documentation provides instruction on how to use CatReg. However, the documentation does not address in detail CatReg concepts or guidance on CatReg methods. While the EPA

CatReg methods guidance has not been finalized at this time, every attempt has been made to make this software consistent with the most recent working draft guidance and discussions of the EPA Benchmark Dose Work group.

Until formal CatReg methods guidance is available, users of this software are strongly encouraged to review existing background material such as the CatReg Software User Manual: R-Version (EPA, 2006) before using this software.

> **Note:**   This user guide supersedes the 2006 User Manual with respect to software-specific instructions such as installation procedures and user interface features and options.

The U.S. Environmental Protection Agency's (EPA's) National Center for Environmental Assessment encourages the broad application of this software. In this document, however, EPA has chosen to focus on the application of the software to the assessment of adverse effects associated with acute inhalation exposure. The description of the CatReg input files for the software reflect this application. The user is free to modify the input fields to support other applications. Appendices B and C provide additional technical description and background for the statistical methods and models used by the program.

1.1   About CatReg

CatReg is a computer program developed to support toxicologists, risk assessors, and health scientists in conducting exposure-response or, when intake information is available, dose-response analyses, most often for controlled animal experiments.

"Exposure" or "intake" has two components:

- Exposure or intake level, indicated by a concentration or dose of the agent of interest

- Exposure or intake duration (time), when the dose varies within the data

"Response" refers to occurrence of a detrimental health effect of a user-defined level of severity.

More specifically, effects observed in toxicological studies are assigned to ordinal severity categories and associated with the exposure conditions (e.g., dose and time) under which the effects occurred. "Ordinal" here means that the categories have a natural ordering in terms of severity or strength of response, but the spacing between ordinal scores is not subject to direct interpretation.

For example, response might have four levels of severity coded as:

- 0 = "no adverse effect"

- 1 = "mild adverse effect"

- 2 = "moderate/severe effect"

- 3 = "lethal effect"

An ordinal response of 2 is higher than a response of 1, but the difference is not necessarily the same as the difference between 3 and 2. The simplest case is dichotomous response data, with just two severity levels, such as: 0 = "no adverse effect", 1 = "adverse effect".

If data are reported on a continuous scale, such as mean and standard error of respiratory rate depression, the user can distribute the total number of experimental subjects over the severity levels using a method discussed in Section 5.11.3 and Appendix A.

1.2   CatReg's Two Models

CatReg provides two basic models, with variations to be explained, to relate the probabilities of the different severity categories to exposure level and exposure duration, taking user-defined covariates into account (e.g., species, sex, target organ, etc.).

The parameters in the models are an intercept term and coefficients of dose and time, either of which may be log-transformed (to the base 10, denoted as "log," "log10," or "log10").

- Model 1, the cumulative odds model, allows the intercept term to vary with severity level, but not the coefficients of dose and time.

- Model 2, the unrestricted cumulative model, allows any of the parameters to vary with severity level.

The probability that a specified severity level or worse will occur increases as dose or time increases. The user can choose for either Model 1 or Model 2 to conform to the logistic, normal, or Gumbel cumulative probability distribution (see Appendix B and Appendix C).

There is a function (called the link function) in each case that transforms the probability for each severity level to a linear function of the unknown parameters, the format of a linear statistical model. The link functions are the logit, probit, and cloglog (complementary log-log) functions for the logistic, normal, and Gumbel cumulative probability distributions, respectively.

The parameter estimates and their statistical characteristics, including standard errors and significance levels, are routinely output by CatReg, along with an analysis of deviance table to assess model fit and a table of estimates of extra risk concentrations (ERC) (doses at which extra risk is a user-specified value) for 1-, 4-, 8-, and 24-hour exposure durations.

1.3   CatReg and Toxicology Data

CatReg was developed for, but is not limited to, meta-analysis of toxicology data from controlled animal studies. Meta-analysis refers to the analysis of data or results from multiple studies simultaneously. Meta-analysis becomes valuable when individual experiments are too narrow to address broad concerns.

For example, in acute inhalation risk assessment, it is important to investigate the combined effects of dose and time of exposure, but few published experiments vary both the dose and the time of exposure (Guth et al., 1997). By combining information from multiple studies, the contribution of both dose and time to toxicity can be estimated. Moreover, the combined analysis allows the analyst to investigate variation among experiments, an important benchmark for the level of model uncertainty.

Different exposure-response experiments may consider the same or different toxicological endpoints, and toxicological judgment is required to determine if, and when, two different endpoints, or gradations of the same endpoint, are of comparable severity.

A relatively simple example is analysis of mortality studies, with two severity levels: 0 = "not lethal", 1 = "lethal". The same endpoint is used for all studies and no intermediate degrees of health gradation are addressed.

A little more complicated example might involve a single health effect, or mode of action, but with more than one severity level corresponding to manifestations of progressive "stages" of development.

Where studies address dissimilar endpoints that may be the consequence of different modes of action, particular care needs to be exercised to decide if comparable severity levels can be assigned across studies. It may not be reasonable to include all studies in the same analysis.

For example, two toxicology experiments might report stages of anesthesia while another reports suppression of the shock-avoidance response. It might be the case that a toxicologist can confidently assign endpoints of the first two studies to comparable severity levels, but not be able to include the third study. In that case, one analysis could address the first two studies and a second analysis the third study, since the studies cannot all be put on a "toxicologically equivalent" severity scale for analysis together.

1.4   "Concentration" and "Dose" are Equivalent Terms in CatReg

You will see the terms "concentration," "dose," and "concentration or dose" used throughout this User Guide. (In the CatReg user interface, "Dose" is used throughout.)

CatReg can be used on oral or inhalation toxicity data, and as such "dose" and "concentration" are used interchangeably throughout this document.

Indeed, CatReg was originally developed for acute, inhalation data, hence the term "extra risk concentration" (ERC) is used in the program. However, in practice, the ERC values are similar, and can serve the same purpose as, benchmark dose (BMD) values traditionally used by EPA for the derivation of reference values and cancer slope factors.

2.0   INSTALLING CATREG

2.1   Step 1: Meet CatReg System Requirements

- 16 Megabytes of RAM
- Microsoft Windows 7 or higher
- Microsoft .NET 4.5 or later. If your computer is set up to allow or be notified of downloads of Microsoft software, then it is likely that that version of the .NET framework is already installed. Otherwise, you can download and install Microsoft's latest version of .NET for client desktops.
- R for Windows version 3.4 or later (either 32- or 64-bit).

Note: CatReg relies on the open source package R for Windows to perform all statistical calculations. For more information on R, refer to the R Project's web site.

2.2   Step 2: Install CatReg

You do not need to uninstall previous versions of CatReg. You can run this version of CatReg alongside earlier versions.

1. Download the CatReg31.msi Windows Installer file from the CatReg Web site.
2. Run the CatReg31.msi file. *Your system's antivirus program may warn you about the file*; if so, approve or allow the file so the install can continue.
3. Step through the install wizard. You can enter a different install directory or accept the default.
4. Navigate to the install directory and select the CatReg.exe file to start the program.

When installing the CatReg application folder, you should:

- Place the folder in the simplest, shortest directory possible. For most EPA users, the preferred path would be C:\Users\[EPA user's LAN ID]; for non-EPA users, this could be as simple as C:\.
- Have administrative rights for the folder.
- Ensure there are no special characters or spaces in the folder names.

If you have problems or concerns, please contact the CatReg development team using our Help Desk Form.

3.0   OVERVIEW OF A CATREG SESSION

The following steps illustrate one path through a CatReg session.

1. Open CatReg by selecting the CatReg.exe file.

2. From the menu bar, select File > New Analysis (for a new input file containing one or more datasets) or Open Analysis (to continue work from a previous session). The Analysis Screen opens, as shown in the following figure.

Figure 1.—Initial CatReg window.
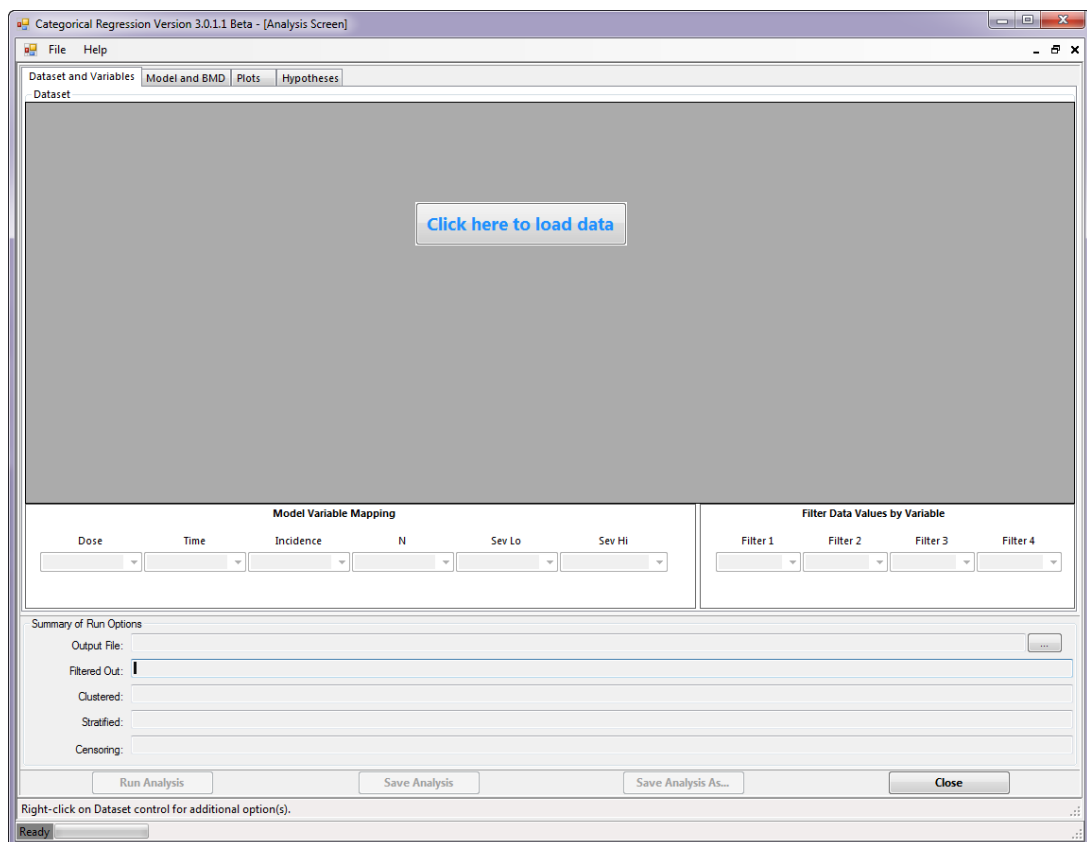
3. In the Analysis Screen, select the **Click here to load data** button. CatReg displays an Open dialog box; navigate to and select the CSV input file you want to open. CatReg displays the loaded dataset(s) and enables the dropdown lists.

Note: A CatReg input file can contain one or more sets of dose-response data.

Note: At this time, CatReg can only read files in CSV format.

CatReg automatically displays the data sorted into treatment groups and displays the dialog box shown in Figure 2. Click on **Select Grouping…** to group the data by other criteria.

Figure 2.—Grouping message box, with grouped data.
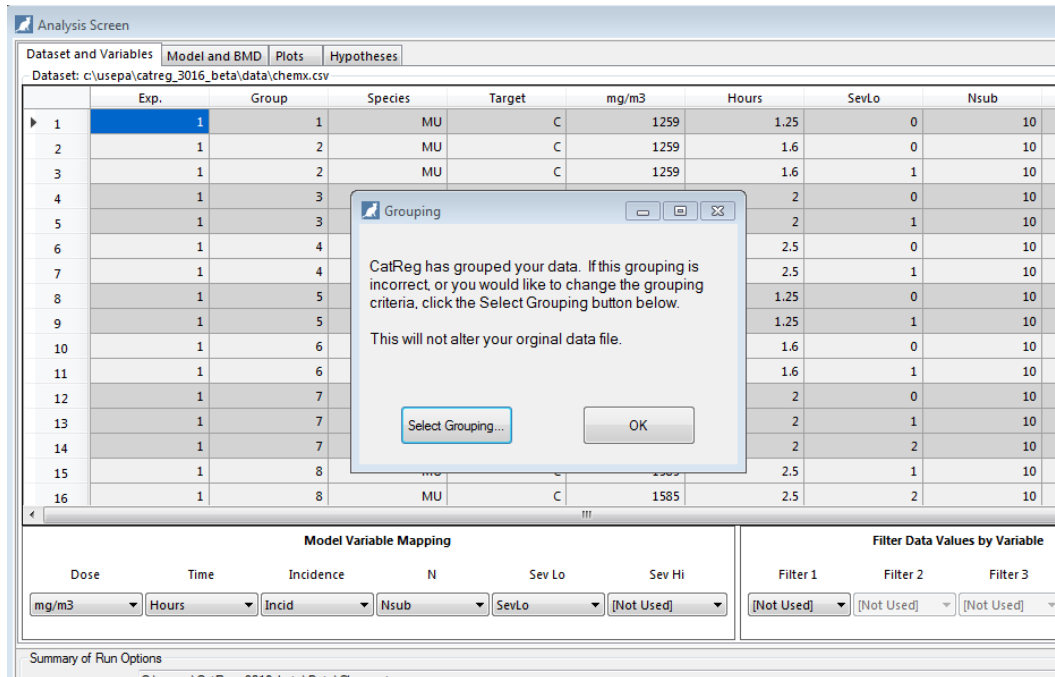
4. In the Dataset and Variables tab, you must map the dose-response data variables to CatReg's standardized model variable names. If you want to filter out specific values, select them here. Note that the summary fields at the bottom of the screen record the selections you make.

Note: Although you can select **Run Analysis** after specifying the model variables and any filtering, it is *highly recommended* that you review the Model and BMD tab settings first.

5. In the Model and BMD tab, specify the options that CatReg will use to calculate an exposure-response curve. All you need to do is select options from picklists, but understanding clustering and stratification will help inform your selections.

6. Select the **Run Analysis** button. CatReg opens a separate window to display the text-based results (.OTX file). By default, CatReg saves the .OTX file to the directory containing the input file used for the current analysis.

7. On the Plots tab (which is enabled after an analysis is run), select the plots you want generated for the analysis. Select the **Run Plots** button. CatReg opens a separate window to display the plots. You can copy and paste the plots into other graphics programs, such as PowerPoint or GnuPlot, to print or edit them. By default, CatReg saves all plots to the directory containing the input file used for the current analysis.

8. On the Hypotheses tab (which is enabled after an analysis is run), you can test exposure-response hypotheses for Intercept, Dose, and Time parameters. Select **Run Tests**. CatReg will display a separate window to display the text-based analysis (.ANX file). By default, the file is saved to CatReg's "OptionFiles" directory. Select the "Save Analysis As…" button to save to another directory.

4.0   INPUT FILE FORMATTING CHECKLIST

A CatReg input file contains dose-response data from one or more studies.

CatReg input files that conform to the following guidelines should provide consistent operation and results.

> **Note**   CatReg includes two sample input files in its Data subdirectory: Chemx.csv and ChemxHu.csv.

☐ The input file is a comma-delimited text file with a ".CSV" extension. Spaces are interpreted as characters. For example, ", MU," is distinct from ",MU,".

☐ The data and units must be capitalized correctly. Because R is case-sensitive, "mU," is different from "Mu,".

☐ The first row for each set of dose-response data contains variable names (columns). Subsequent rows contain data for the variables.

☐ Each dataset should contain the following columns that you will map to variables in CatReg. The columns can be in any order. See Table 3 in Section 5.9, "Data Requirements: Severity Levels" for more information.

  o Exposure concentration or dose (required)

  o Exposure time in hours (required unless times are equal; default=1)

  o Number of subjects in the treatment group (required)

  o Incidence of severity level or severity range for the record (required)

  o Lowest severity level for the record (required)

  o Highest severity level for the record (required for "censored" analysis if response spans more than one severity level; see Section 5.9, "Data Requirements: Severity Levels," for more information.)

☐ The following conditions apply for severity levels in the input file. See Section 5.9, "Data Requirements: Severity Levels," for more information.

  o The minimum number of severity levels is two (severity levels coded as 0 and 1, corresponding to absence or presence of an effect).

  o Severity levels must be specified as small, consecutive ordinal values (0, 1, 2, 3, 4).

  o It is not necessary to include rows that explicitly specify incidence=0.

  o Individual studies in the input file can have incidence=0 for a severity level.

  o Each severity level in the input file should have incidence > 0; that is, at least one study in the input file must have incidence > 0 for a given severity level..

  o If a severity level has zero incidence (whether explicitly indicated by row entries or not) across all studies in an input file, all severity levels in the input file must be renumbered to eliminate the zero incidence. For example, an input file has severity levels 0, 1, 2, and 3, but there is zero incidence for level 1. In this case, all the severity levels should be renumbered to 0, 1 (formerly level 2), and 2 (formerly level 3).

  o Filtering out all instances of a severity level is the same as having incidence=0 for that severity level for the whole input file.

  o For stratification, it is acceptable to have some strata where incidence=0 for a severity level.

☐ If both concentration/dose and time effects are to be modeled, then both C and T need to be varying in the data. (If T is constant, then a reduced model is fit so that T is dropped from the model.)

☐ A treatment group's incidence must sum to the treatment group size (Nsub).

- o If you use CatReg's default grouping method, then a treatment group's records must be consecutive in the dataset.
- o If you use CatReg's new grouping command to select column(s) that uniquely identify treatment groups, then the treatment group records can be in any order. See Section 5.4, "Managing Treatment Groups," for more information.
- o If a Group variable is present, CatReg will use it to sort and display treatment groups. Otherwise, CatReg determines treatment groups by reading records until the values of Incid sum to Nsub, then starting over with the next record.

5.0   DATASET AND VARIABLES TAB

After either starting a new analysis or opening a previous analysis file, CatReg opens its Analysis Screen.

Figure 3.—CatReg Analysis Screen with loaded dataset, striped treatment groups, and mapped variables.
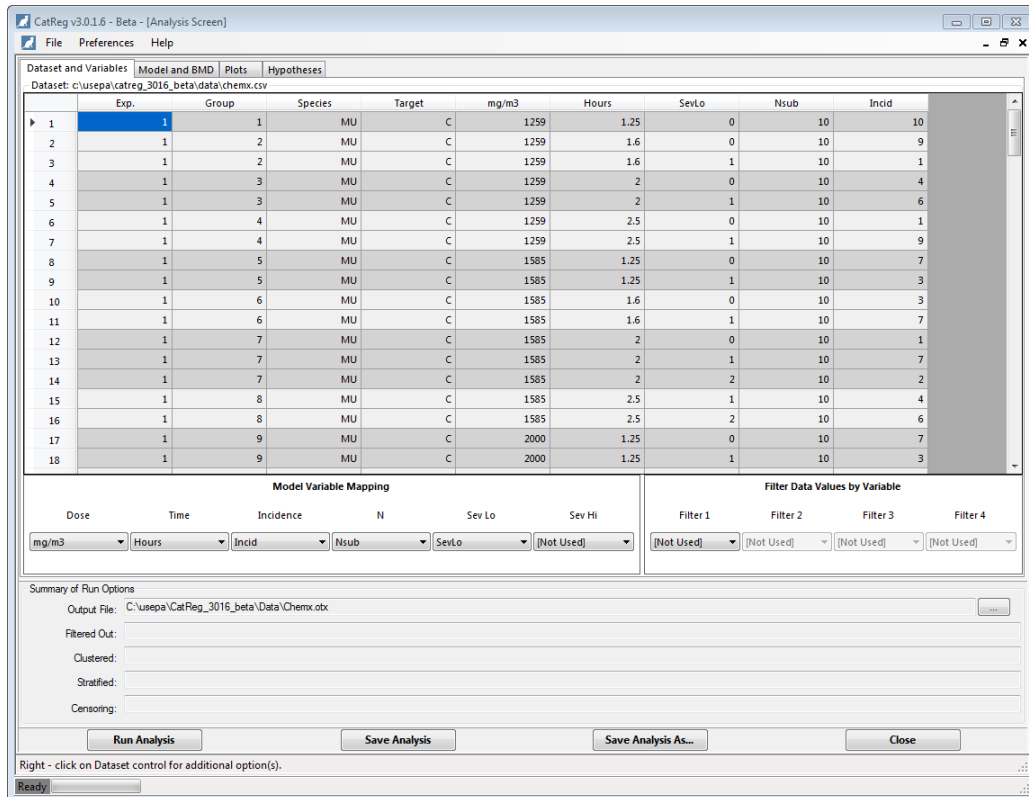
Table 1: *Selections Available on the Dataset and Variables tab*

Specifications	Description	Notes
Variables	Variables as defined in the dataset's column headers.	N/A
Dose	Specify the variable for exposure dose.	Required.
Time	Specify the variable for duration.	Exposure time in hours. Required unless all times are equal. (Default value = 1.)
Incidence	Specify the variable for incidence of severity level or severity range.	Required.
N	Specify the variable for number of subjects in a treatment group.	Required.
SevLo	Specify the variable for the lowest severity level.	Required.

Specifications	Description	Notes
SevHi	Specify the variable for the highest severity level.	Required if response spans more than one severity level. SevHi is required if the severity level for one or more records is entered as a range, but not otherwise. For example, if the severity level for a record is a range such as level 1 to level 2, then SevLo is entered as 1 and SevHi is entered as 2. Data records for which SevLo $\neq$ SevHi are referred as censored data. If only one severity level applies, e.g. level 1, then SevLo = SevHi = 1
Filter	Check to display a popup window showing values for the selected variable. Check the option boxes beside a value to remove (filter) it from the analysis. The summary fields at the bottom of the screen display the filtered values in the Filtered Out text box.	Use this option to fit the exposure-response curve to a subset of the data. One reason to filter data is to investigate how the fit changes when certain observations are excluded. For instance, a particular study may be suspect, and it may be desirable to compare the parameter estimates with and without the suspect study.

5.1   Loading an Input File

- *If no dataset is loaded*: Select File > New Analysis, and select the **Click here to load data** button.
- *If a dataset is already loaded*: right-click on the Dataset and Variables window and select Load Dataset… from the context menu.

5.2   Input Files Must be in CSV Format

Microsoft Excel spreadsheets can be used to construct an input file, but the file must be saved as a comma-delimited file with a ".CSV" extension, rather than as an Excel file with a ".XLS" extension.

Because CatReg assumes that data are separated by commas rather than by blank spaces, spaces are interpreted as characters and should be avoided unless intended to be part of the data. For instance, ", MU," is distinct from ",MU,". Because R is case-sensitive, "mU," is different from "Mu,".

5.3   Model Variable Mapping

For the Model Variable Mapping, select from the picklists the dose-response variables that correspond to the CatReg variables of Dose, Time, Incidence, etc. For Time and Sev Hi, [Not Used] is a valid choice.

CatReg automatically maps variables based on the loaded input file's column names.

a. **If default mapping fails for required variables**, then the error message shown in Figure 4 appears:

Figure 4.—Error message when CatReg cannot map required variables.
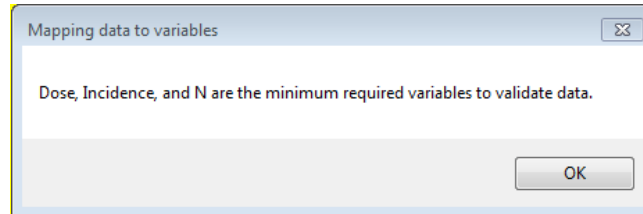
In addition, CatReg does not stripe the data into groups, as shown in Figure 5. For this case, you must manually map the required variables so CatReg can perform its default validation.

Figure 5.—Unstriped treatment groups mean CatReg cannot validate the input file data.

If the validation succeeds, the data will be striped into treatment groups, as shown in Figure 3. If you prefer a different grouping, right-click on the Dataset window and select "Regroup Dataset…" from the menu. See the next section for more information on managing treatment groups in CatReg.

b. **If default mapping is successful**, then a Grouping message box (see Figure 2) will appear.

Each time you change the Model Variable Mapping, and if all required variables are mapped, then CatReg performs its validation methods (see Section 5.5).

5.4   Managing Treatment Groups

CatReg attempts to automatically sort treatment groups in the loaded input file based on column names. Although CatReg displays sorted treatment groups in the Dataset and Variables tab, CatReg does not change or rewrite the input file on disk.

To save the grouped data to disk, right-click on the Analysis Screen and select Save Dataset As… You can choose to overwrite the existing input file or save the displayed grouped data to a new file.

5.4.1   CatReg's Default Grouping Algorithm

CatReg performs the following treatment-group checks on loaded dose-response groups to ensure that the required variables dose, incidence, and number of subjects are mapped:

1. The data is read row by row in the order it is provided (no sorting), and the sum of the incidence column is calculated.

2. The end of a group is determined if the sum of the incidences = number of subjects, or if the dose value or the time (optionally, if specified) value changes.

3. The grouping validation fails if for any group the incidence does not equal the number of subjects.

5.4.2 Custom Grouping

You can choose to proceed with CatReg's default grouping or you can specify a custom grouping.

To specify a custom grouping:

1. Right-click in the Dataset window and select Regroup Dataset… from the menu. The Grouping dialog box appears (Figure 6).

Figure 6.—Grouping dialog box.
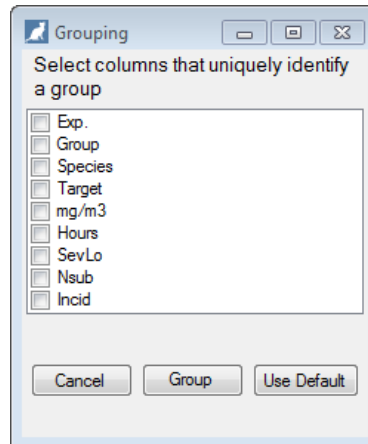
2. Select any number of column(s) that will uniquely identify groups in the dataset.

3. Select the **Group** button and CatReg will sort the groups based on the selected columns.

Behind the scenes, CatReg sorts the data provided (without altering the original .CSV file) based on the selected columns.

- The ordering of items in the "Grouping" dialog list determines the sorting order. For example, if "Exp" and "Group" are selected, then the data is sorted by "Exp" and then by "Group" for each value of "Exp". If all the values in a column are numeric, then sorting is done numerically. If the columns contain any non-numeric values, then sorting (including numbers) is done alphabetically

- CatReg reads the data row by row in the sorted order and calculates the sum of the incidence column.

- CatReg determines the end of a group if any values in the selected grouping columns change.

- Grouping validation fails if, for any group, the incidence total does not equal the number of subjects.

To save the regrouped data to disk, right-click on the Analysis Screen and select Save Dataset As… You can choose to overwrite the existing input file or save the regrouped data to a new file.

5.5 CatReg's Validation Criteria After Custom Grouping

CatReg uses the following criteria to validate the data if the grouping method is changed:

- **Grouping validation.** The data is checked according to the grouping algorithm used. See Section 5.4.1, "CatReg's Default Grouping Algorithm," for more information.

- **Numeric validation.** Certain variables are required to be numeric. CatReg validates that the mapped columns in the data corresponding to these variables do not contain any non-numeric values.

If validation fails:

- A separate dialog box appears for each failed validation method (grouping or numeric validation) with a list of the row numbers that failed validation (see Figure 7).

Figure 7.—Example of failed validation error message.
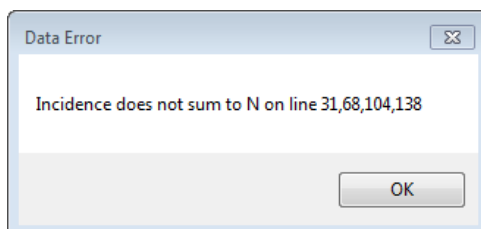
- A red "X" icon appears on the tool strip at the bottom left of the Analysis Screen with the message "Validation Failed" (see Figure 8). Click the icon to redisplay any failed validation messages.

Figure 8.—Validation Failed error.
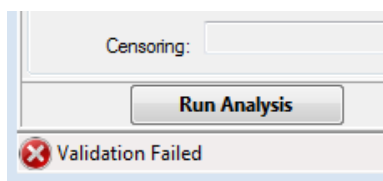
5.6   Filtering Data Values by Variable

From the dropdown lists, select the dose-response group data variables whose records you want to be removed (filtered) from the analysis, without removing them from the input file. If no variable is selected, "[Not Used]" appears. (See Figure 3.)

This option is used to fit the exposure-response curve to a subset of the data. One reason to filter data is to investigate how the fit changes when certain observations are excluded. For instance, a particular study may be suspect, and it may be desirable to compare the parameter estimates with and without the suspect study.

To filter data, select the variable of interest; CatReg will display all the values for that variable. Tick the box for the value you want to be removed, and select **OK**.

This option is also useful for scanning a particular field to see what values have been observed. Unexpected values may indicate a problem with the input file (e.g., errors in data entry).

5.7   Minimum Data Requirements

Certain minimal data requirements need to be satisfied to estimate the categorical regression model:

- There needs to be at least one response in each severity category. For example, if the input file contains a severity 4, then severity categories 3, 2, and 1 should each contain responses. If any categories in the input file contain no non-zero response, then categories will need to be combined. Refer to Section 5.9, "Data Requirements: Severity Levels," for more information.

- If both concentration/dose and time effects are to be modeled, then both C and T need to be varying in the data. If T is constant, then a reduced model is fit so that T is dropped from the model.

There are additional technical limitations on the complexity of the model relative to the data:

- The most obvious is that the model cannot include more parameters than the number of independent observations.

- Less obvious problems sometimes occur if certain variables are redundant. CatReg will display R error messages if there are redundant variables in the model. This means that at least one variable in the model can be expressed in terms of the others. Reducing the number of stratification variables usually will solve the problem.

- Sometimes, R will return a "failure to converge" message. This is an indication that there are too many variables in the model, and that the model needs to be simplified. This problem relates to the number of variables needed to completely isolate the different severity categories. The solution is to remove one or more variables from the model.

5.8   Data Requirements: Outcomes

When the source document for an experiment does not report the outcome for individual subjects, or otherwise report the incidence of different health effects, the data may not be suitable for CatReg. For example, a report of "mild" pathology for a treatment group might mean that a few or many in the group manifested that response or that the "mild" response was the most common, with both lesser and more severe effects also present in the group. In either case, there is not sufficient information to divide a treatment group into incidence of severity categories.

It is sometimes reasonable to represent a health outcome measured on a continuous scale as categorical data. Continuous data from acute studies, such as enzyme activities, tidal volume, respiratory rate, blood pressure, etc., often are reported as a mean value, with a measure of dispersion, such as the standard error or standard deviation, for each treatment group. To convert these data to severity levels for CatReg, each severity level needs to be equated to an interval of values on the continuous scale.

For example, if the full range of responses is 0 to 100, the user might decide to classify outcomes 0 to 20 as "no effect", 21 to 40 as a "mild adverse effect", 41 to 65 a "moderate adverse effect", and 66 to 100 "severe effect". The mean for a treatment group falls into a single severity level, but some of the individual responses of subjects in the group may have been dispersed over adjacent severity levels.

Knowing the mean and standard deviation (or standard error that can be converted to a standard deviation by multiplying by the square root of the size of the treatment group) and assuming a distribution for the continuous data (e.g., normal), an estimate can be made of the incidence at each severity level (see Appendix A for details). The estimated incidence figures need not be whole numbers, but must still sum to the total group size. Incidence estimation is not possible if the mean is reported without a measure of dispersion.

5.9   Data Requirements: Severity Levels

The same category system of severity levels must be used for all data in an input file. Considerable toxicological judgment may be required for classification of various health effects into severity levels and for achieving comparability across experiments. When that cannot be done for all the studies of interest, it may be necessary to group the studies into more than one input file. Classification judgments must be made systematically according to documented criteria.

The minimum number of severity levels is two (severity levels coded as 0 and 1, corresponding to absence or presence of an effect).

Suggested severity categories for a three-category classification are "no adverse effect", "adverse effect", and "lethal effect", coded as severity levels 0, 1, 2, respectively.

A four-category scheme might be "no adverse effect", "mild adverse effect", "moderate/severe effect", and "lethal effect", coded as 0, 1, 2, 3, respectively.

In some toxicology studies, it may not be possible to score all response data completely. Consider a four-category scoring system in which 0 = "no observable effect," 1 = "mild effect," 2 = "moderate effect," and 3 = "severe effect." Published data from an animal mortality study may not include nonlethal outcomes; therefore, the response score for an animal that survives is uncertain, or "censored." That score is known to be less than 3, but it is not known whether the score should be 0, 1, or 2. Such an observation is said to be "interval censored." Another situation where the response score may be interval censored is in combining data from experiments with different endpoints. For some endpoints, it may not be clear from the toxicology whether a specific response should be considered "mild" or "moderate." An interval censored analysis simply could report that the response is either 1 or 2, but the specific score is not known.

The ability to include partial information about the ordinal scores is one of the important features of CatReg. CatReg incorporates this type of partial information in an interval-censored analysis. In general, interval censoring occurs if the response is known only to lie in an interval of potential values. Such intervals are specified in CatReg by supplying the lower and upper limits of the known range for each observation. Suggested codes to indicate species and sex in an input file are provided in Table 2.

Table 2: *Recommended Codes for Species and Sex*

Species	Code	Sex	Code
Human	HU	Female	F
Rat	RT	Male	M
Mouse	MU	Both sexes	B
Rabbit	RB	N/A	N/A
Guinea pig	GP	N/A	N/A

Each column of the user input file is referred to as a data field, with the first record (row) being variable names and all subsequent rows containing data for the variables. CatReg requires information for four- six data fields, depending on the data. The names of these data fields and the corresponding default variable names that CatReg looks for in the user input file are shown in Table 3. For example, "conc" refers to a data field for exposure concentration and CatReg looks for the variable name "mg/m3" to identify that field, unless the variable name has been changed from the default. The default variable name for "conc" might be changed, for example, if the concentration or dose used in experiments is different from milligrams per cubic meter.

Using the default variable names in Table 3 as an example, the user input file must include data in each record (beyond the first that contains variable names) for variables "mg/m3", "Nsub", "Incid", and "SevLo". Data are also required for the variable "Hours" unless all exposure duration times are equal, in which case it can be omitted (CatReg uses 1 as the default value in that case). The variable SevHi is required if the severity level for one or more records is entered as a range, but not otherwise. For example, if the severity level for a record is a range such as level 1 to level 2, then SevLo is entered as 1 and SevHi is entered as 2. Data records for which SevLo ≠ SevHi are referred as censored data. If only one severity level applies, e.g. level 1, then SevLo = SevHi = 1.

Table 3: *Data Fields That May Be Required*

Data Field	Default Variable Name	Description
Conc	mg/m3	Exposure concentration or dose. Use of the human equivalent concentration or dose is recommended. Always required.
Time	Hours	Exposure time in hours. Required unless all times are equal. (Default value = 1.)
Nsub	Nsub	Number of subjects in a treatment group. Always required.

Data Field	Default Variable Name	Description
Incid	Incid	Incidence of severity level (or severity range) for the record. Always required.
Loscore	SevLo	Lowest severity level for the record. Always required.
Hiscore	SevHi	Highest severity level for the record. Required if response spans more than one severity level.

A separate record is entered for each severity level (or range of severity levels in the case of censored data) observed in a treatment group. For example, if the user determines three severity classifications for health effects, denoted as 0, 1, and 2, then the outcome for a treatment group is represented as an incidence for each severity level that is observed. To illustrate, a treatment group of size 10 might result in 3 subjects being classified at severity level 0, 4 at severity level 1, and 3 at severity level 2, which would require three consecutive records (consecutive rows of data) in the input file. The records for the treatment group must not only be consecutive but their incidence (values of Incid) must sum to the treatment group size (Nsub).

A severity level with no observations need not be entered as a separate record. For example, if a treatment group of size 10 had 6 subjects classified at severity level 0 and 4 at severity level 1, then only two records would be required to enter the data. The value of Nsub would be 10 in both records; the value of Incid would be 6 in one record and 4 in the other.

Variables in addition to those in Table 3 may be added to the input file at the user's discretion, either for use in execution of CatReg or to facilitate organizing and keeping track of the data. The user can refer to them in the same manner as variables in the required fields when using CatReg options. For example, one might want to add "strain" as a variable to distinguish between two strains of mice and have CatReg test whether their exposure-response curves are significantly different in some respect, or add "Ref.id" to record the source of the data, even if it is not used during the execution of CatReg.

5.10   Data Issues that Trigger Error Messages

CatReg attempts to display any R error messages in plain language.

Sometimes "NaN" (not a number) is given for some of the model deviance iterations. It is displayed when CatReg is searching for the solutions to the parameter estimates in early iterations of the program. CatReg tries to compute the deviances for those solutions and finds they are undefined (NaN). Once a solution is determined for runs like this, the following warning message may appear: `NAs produced in: log(likes)` (Na means "not available").

The message `Warning: Gamma hit its maximum bound!!!` may occur when the parameter γ (gamma) is being estimated. The smallest positive value of concentration or dose in an input file is used as a practical boundary on gamma. The estimate shown for gamma, and the other parameters in the summary table of estimates, are not maximum likelihood in that case, and the user is advised to consider the setup option that assumes the background risk is zero.

An error message occurs if a coefficient of concentration/dose or time is negative, e.g., `Error: TIME is negative! Estimates of coefficient parameters do not satisfy non-negativity constraint on the parameters. A negative estimate is evidence of no effect. This run will terminate.` The user needs to modify the run. The data may be indicating there is no effect or there may just be too many parameters in the model.

Similarly, the estimates of severity intercepts may violate the order constraint, resulting in a message such as `Sev 1 < Sev 2! Estimates of severity parameters do not satisfy constraint on order of parameters. Incorrectly ordered severity estimates is evidence of too many severity levels in the data. This run`

`will terminate.` The user needs to modify the dose-response group data by, for example, combining severity levels.

Most error messages at this point in a CatReg run result from an attempt to calculate too many parameters with too little data. This may generate error and /or warning messages from the optimization routine (used to find maximum likelihood estimates), such as: `Error in optim(transcoefs, fn = hfdeviance, gr = hfdeviance.grad, method = "BFGS",: initial value in vmmin is not finite`

In general, when errors occur during the calculation of parameter estimates, the user should change the link function if the cloglog link function was used or simplify the model and run it again. The model can be simplified by reducing or omitting stratification of parameters. If Model 2 (unrestricted cumulative model) is in use, the user can consider switching to Model 1 (cumulative odds model). Additionally, one can try different options for the scale of the explanatory variables Dose and Time. `Dumped` means that the CatReg analysis has been terminated because a solution could not be computed.

5.11   Input File Examples

Three varied examples of input files follow, described as experimental results for hypothetical chemicals named chemx, chemy, and chemz.

The data were generated by simulation, except the data for what is being called chemy that were constructed from actual rodent experiments.

- The input file for chemx is an example of four experiments, one each on the four combinations of species (RT and MU) and target organs (C and L).
- The input file for chemy has a more complicated structure and illustrates how a toxicologist might determine the severity levels.
- The input file for chemz is an illustration of converting a continuous response to severity categories for use in CatReg.

5.11.1   chemx.csv

Table 4 displays the first part of the input file for chemx.csv. Four experiments were conducted under identical exposure conditions, each consisting of 10 observations at each combination of four concentrations/dosages (mg/m3) and four times (Hours), for a total of 64 treatment groups. The dosages are 1259, 1585, 2000, and 2512 mg/m3; the times are 1.25, 1.6, 2.0, and 2.5 hours.

There are three severity levels: no adverse effect (SevLo = 0), mild adverse effect (SevLo = 1), moderate/severe effect (SevLo = 2). Two experiments are on mice (Species = MU) and two are on rats (Species = RT), with one of the two experiments on each species reporting effects on the central nervous system (Target = C) and the other reporting effects on the liver (Target = L). "Exp" denotes an experiment number, "Group" the treatment group within the experiment, "Nsub" the number of subjects in the treatment group, and "Incid" the incidence in the treatment group of the severity level (SevLo) shown in the record (row of data).

The names of data fields are user-specified, but the data fields associated with the Table 4 variable names mg/m3, SevLo, Nsub, and Incid are required. CatReg displays an error message (see Figure 4 in Section 5.3, "Model Variable Mapping") if any of them is missing. In this example, the exposure times vary so the variable Hours is included. SevHi would have been included as a variable if the severity level had been censored for one or more records (spanned more than one severity level). For example, to make the 10 subjects in the first record (row of data) classified as severity level 0-1, the variable SevHi would be added to the input file and the first record would remain unchanged except for SevHi = 1. In that case, subsequent records that are not censored would be given the same value for SevLo and SevHi. For example, the second record indicates that 9 subjects were classified at severity level 0. If SevHi were included as a variable, then SevHi would be set to 9 for that record, making SevLo = SevHi = 9.

It may be noted that variables Exp., Group, Species, and Target are names created by the user. Species and Target were included in this case to be able to distinguish between species and target organ in the data analysis, but other names could be used in their place. Exp. and Group were added by the user to facilitate record keeping. The variable Group is not required by CatReg but records for the same treatment group must be together in the input file, all with the common value of Nsub and values of Incid that sum to Nsub. Adding a variable such as Group provides a convenient check of the data for the user. CatReg determines treatment groups by reading records until the values of Incid sum to Nsub, then starting over with the next record.

Table 4: *Part of the Input File CHEMX.CSV*

Exp.	Group	Species	Target	mg/m3	Hours	SevLo	Nsub	Incid
1	1	MU	C	1259	1.25	0	10	10
1	2	MU	C	1259	1.6	0	10	9
1	2	MU	C	1259	1.6	1	10	1
1	3	MU	C	1259	2	0	10	4
1	3	MU	C	1259	2	1	10	6
1	4	MU	C	1259	2.5	0	10	1
1	4	MU	C	1259	2.5	1	10	9
1	5	MU	C	1585	1.25	0	10	7
1	5	MU	C	1585	1.25	1	10	3
1	6	MU	C	1585	1.6	0	10	3
1	6	MU	C	1585	1.6	1	10	7
1	7	MU	C	1585	2	0	10	1
1	7	MU	C	1585	2	1	10	7
1	7	MU	C	1585	2	2	10	2
1	8	MU	C	1585	2.5	1	10	4
1	8	MU	C	1585	2.5	2	10	6
1	9	MU	C	2000	1.25	0	10	7
1	9	MU	C	2000	1.25	1	10	3
1	10	MU	C	2000	1.6	0	10	2
1	10	MU	C	2000	1.6	1	10	5
1	10	MU	C	2000	1.6	2	10	3
1	11	MU	C	2000	2	1	10	6
1	11	MU	C	2000	2	2	10	4
1	12	MU	C	2000	2.5	1	10	4
1	12	MU	C	2000	2.5	2	10	6
1	13	MU	C	2512	1.25	1	10	9
1	13	MU	C	2512	1.25	2	10	1
1	14	MU	C	2512	1.6	1	10	5
1	14	MU	C	2512	1.6	2	10	5
1	15	MU	C	2512	2	1	10	2
1	15	MU	C	2512	2	2	10	8
1	16	MU	C	2512	2.5	2	10	10
2	1	RT	C	1259	1.25	0	10	10
2	2	RT	C	1259	1.6	0	10	9
2	2	RT	C	1259	1.6	1	10	1
2	3	RT	C	1259	2	0	10	9
2	3	RT	C	1259	2	1	10	1

5.11.2 chemy.csv

Table 5 is part of a larger input file that illustrates a more elaborate coding system and some other features not included in the preceding example, e.g. censoring, and provides a realistic example for discussion of toxicological judgment in severity classification. The available studies varied on the organ sites and endpoints examined and a four-category system of severity levels was implemented: no adverse effect (0), mild adverse effect (1), moderate/severe effect (2), and lethal effect (3). Again, CatReg expects to find variables mg/m3, Nsub, Incid, and SevLo, at a minimum, and Hours as well, if exposure time varies, and SevHi if there are any censored data.

Also (again) notice that a separate record is required for each different severity (or range of severity in the case of censored data) in a treatment group, and that Incid sums to Nsub for records in the same treatment group.

"Ref.id" (reference identification) is a number assigned by the user to the source of the information in the record. In the example, Ref.id = 20938 is the source of the material used to construct the first 10 data records shown. "Exp." identifies experiments within each Ref.id, numbered sequentially from 1; "Group" numbers treatment groups within each experiment (i.e., subjects alike with respect to all methods and materials variables); and "Nsub" is the number of subjects in each group. A separate record (row) is entered for each severity level (or range of severity levels) in a treatment group. "Marker" just numbers the data records. The severity levels are entered under "SevLo", the lowest possible severity level for the record, and "SevHi", the highest severity level for the record.

When SevLo ≠ SevHi, "y" is entered under "Censored"; "n" when SevLo = SevHi. In Table 5, the variable Censored has been added by the user to readily distinguish between records with a single severity level and those with a range of severity levels; it is not required by CatReg. The variable BestNum also has been added by the user to indicate the most likely severity level when SevLo ≠ SevHi; otherwise, the common value for SevLo and SevHi is entered for BestNum. To use the scores in the BestNum column, the default variable names SevLo and SevHi must both be changed to BestNum by editing the names in the input file.

The first reference (Ref.id = 20938) reported two experiments, one with mice (Species = MU) and one with rats (Species = RT), with both sexes in both experiments (Sex = B). There is a user-defined variable coded to indicate the target organ (e.g., Target = Resp) and the primary endpoint (Endpoint = Lethality), both of which were the same for both experiments. The first group (Group = 1) under the first experiment (Exp = 1) is for rats (Species = RT) of both sexes (Sex = B) exposed at a concentration or dose of 330 mg/m3 (mg/m3 = 330) for 6 h (Hours = 6). There were 26 rats at risk (Nsub = 26). The user-defined severity classification was the same for all 26: severity level 0 to 2 (SevLo = 0, SevHi = 2), with severity level 1 the best guess for a single severity level (BestNum = 1). Because the results variables were the same for all 26 subjects, only one record is needed to record the data for the whole treatment group.

The sixth and seventh records (Markers 6 and 7) are for the first treatment group (Group = 1) of the second experiment (Exp = 2) in Ref.id = 20938. Mice (Species = MU) of both sexes (Sex = B) were exposed to 360 mg/m3 (mg/m3 = 360) for 6 h (Hours = 6). Two records are required because there are two distinct severity classifications: one for 23 subjects (Incid = 23) with severity level 1 to 2 (SevLo = 1, SevHi = 2) and severity level 2 the best guess (BestNum = 2) and the other with three subjects at severity level 3 (SevLo = 3, SevHi = 3) and BestNum = 3. Six records in the example have different values for SevLo and SevHi, as indicated by Censored = y.

In Ref.id = 20938 of Table 5, the effect severity for dosages at which no subjects died were censored 0 to 2 (e.g., Marker 1) for no adverse effects to severe adverse effects, because the effects were unknown. Survivors from groups in which some subjects died (e.g., Marker 2) were assumed to have suffered adverse effects and were censored 1 to 2 because the effects could have been mild to severe. Survivors from groups in which most of the subjects died (e.g., Marker 5) were assumed to have suffered severe effects. For the 360-, 390-, 420-, and 460-mg/m3 ChemY exposures, there is one record for the number of subjects exhibiting lethal effects, and another for the number of subjects exhibiting nonlethal effects. Only one record was made for exposure to 330 mg/m3, because all subjects were assumed to exhibit effects of severity 0 to 2, and one record was made for exposure to 500 mg/m3, because all subjects died (severity 3).

Table 5: *Part of the Input File CHEMY.CSV*

Marker	Ref.id	Exp	Group	Species	Sex	mg/m3	Hours	Target	Endpoint	Nsub	Incid	BestNum	SevLo	SevHi	Censored
1	20938	1	1	RT	B	330	6	Resp	Lethality	26	26	1	0	2	y
2	20938	1	2	RT	B	390	6	Resp	Lethality	26	20	2	1	2	y
3	20938	1	2	RT	B	390	6	Resp	Lethality	26	6	3	3	3	n
4	20938	1	3	RT	B	460	6	Resp	Lethality	26	23	3	3	3	n
5	20938	1	3	RT	B	460	6	Resp	Lethality	26	3	2	2	2	n
6	20938	2	1	MU	B	360	6	Resp	Lethality	26	23	2	1	2	y
7	20938	2	1	MU	B	360	6	Resp	Lethality	26	3	3	3	3	n
8	20938	2	2	MU	B	420	6	Resp	Lethality	26	13	3	3	3	n
9	20938	2	2	MU	B	420	6	Resp	Lethality	26	13	2	1	2	y
10	20938	2	3	MU	B	500	6	Resp	Lethality	26	26	3	3	3	n
11	61831	1	1	RT	M	14	4	Resp	N lavage	12	12	0	0	0	n
12	61831	1	2	RT	M	278	4	Resp	N lavage	12	12	1	0	1	y
13	61831	1	3	RT	M	556	4	Resp	N lavage	12	12	2	1	2	y

Notes:

Marker - Record number.

Ref. id - Source identifier.

Exp - Experiment number within a source.

Group - Treatment group number (within an experiment).

Species - Species.

Sex - Sex.

mg/m3 - Exposure concentration or dose.

Hours - Exposure time.

Target - Target organ.

Endpoint - Toxic endpoint.

Nsub - Number of subjects in treatment group.

Incid - Number of animals responding.

BestNum - Analyst's best estimate of severity category for censored data; same as SevLo and SevHi for noncensored data.

SevLo - Lowest applicable severity level.

SevHi - Highest applicable severity level.

Censored - Severity level reported as a range if "y".

Ref.id = 61831 of the example illustrates a case where health effects are described for each treatment group, for which the severity levels can be decided, but incidence of the severity levels is unknown. These data are not suitable for CatReg because correct values for Incid are unknown. Nevertheless, one might want to test the sensitivity of the results of CatReg to different assumptions regarding the incidence values and it provides a useful example for discussion of deciding severity levels. It is included here primarily as an example of deciding severity levels from toxicological data. The data are from a study in which groups of 12 rats were exposed to 0, 14, 278, and 556 mg/m3 ChemY for 4 h. Incid = 12 is used in the example for illustration, which is equivalent to assuming that all 12 rats in each treatment group have the same severity classification shown for the group as a whole.

Data from this study were assigned to severity levels as follows. Groups of four rats were killed at 1, 20, and 44 h after 4-h exposure for the examination of biochemical indicators of injury and inflammatory response in the respiratory tract. Nasal lavage fluid was examined for lactate dehydrogenase (LDH), alkaline phosphatase (ALP), protein, and number of nucleated cells. Bronchoalveolar lavage (BAL) fluid was examined for activities of LDH, ALP, and □-glutamyl transpeptidase. All measurements were reported as means, plus or minus standard deviations. No changes in any parameters were noted among rats exposed to 14 mg/m3 ChemY. The only parameter significantly different from controls in nasal lavage fluid at all post-exposure time periods was increased cellularity, which was significant at the 556-mg/m3 exposure. In BAL fluid, LDH activity was elevated at 44 h post-exposure, and ALP activity was significantly decreased at 20 and 44 h after the 278-mg/m3 exposure. At all post-exposure times for the 556-mg/m3 exposure, protein concentration and LDH activity were elevated, but ALP was decreased.

Because a number of biochemical and cellular parameters were measured at several post-exposure periods, significant changes were considered to be adverse only if the changes were still significant at the last post-exposure measurement. In other words, reversible changes were classified as no-observed-adverse effects. Table 6 shows how these effects were categorized using the four-category severity scheme. Because no changes were noted after exposure to 14mg/m3 ChemY, effects were coded 0 for no adverse effect. Effects at the 278-mg/m3 ChemY exposure were estimated to range from severity 0 to 1 (no adverse effect to mild adverse effect) because significant changes in two biochemical parameters were observed. The adversity of those changes was uncertain but assumed to be less than severe. Effects caused by exposure to 556 mg/m3 ChemY were estimated to range from severity 1 to 2 (mild to severe adverse effect) because of changes in the activities of several enzymes and nasal cytopathology. No deaths occurred during this study, so category 3 (lethality) was not used.

Table 6: *Example of Severity Categorization for Nonlethal Effects*

Exposure Concentration or Dose (mg/m3 ChemY)	Statistically Significant Effects Reported	Severity Score	Censored
14	None	0	No
2748	↑LDH, ↓ALP	0-1	Yes
556	↑Protein, ↑LDH, ↓ALP, ↑cellularity, nasal cytopathology	1-2	Yes

5.11.3 chemz.csv

An artificial example of how continuous data might be coded is displayed in Table 7. It is assumed that the standard error was included along with the mean for each treatment group. In this case, the incidence of each severity in a group was estimated from the mean and standard error, which produces fractional subjects. The method to estimate incidence at each severity from group means and a measure of variability is described in Appendix A.

Male rats in treatment groups of 10 each were exposed to a toxicant at various dosages and times. Adverse effects occurred in the respiratory tract, with severity indicated by lung weight. For illustration, a four-category severity classification is used, with lung weight 0 to 20 classified as "no effect", 21 to 40 as "mild adverse effect", 41 to 65 as "moderate adverse effect", and 66 to 100 as "severe effect". The mean for a treatment group falls into a single severity level, but some of the individual responses of subjects in the group may have been dispersed over adjacent severity levels.

Table 7: *Part of the Input File CHEMZ.CSV*

Marker	Ref.id	Exp	Group	Species	Sex	mg/m3	Hours	Target	Endpoint	Nsub	Incid	BestNum	SevLo	SevHi	Censored
1	1	1	1	RT	M	330	2	Resp	Lung wt	10	1	0	0	0	n
2	1	1	2	RT	M	360	2	Resp	Lung wt	10	1	1	1	1	n
3	1	1	3	RT	M	390	2	Resp	Lung wt	10	1	1	1	2	y
4	1	1	4	RT	M	410	2	Resp	Lung wt	10	1	2	2	2	n
5	1	1	5	RT	M	460	2	Resp	Lung wt	10	1	2	2	2	n
6	1	2	1	RT	M	460	1	Resp	Lung wt	10	1	1	1	1	n
7	1	2	2	RT	M	510	1	Resp	Lung wt	10	1	1	1	1	n
8	1	2	3	RT	M	560	1	Resp	Lung wt	10	1	2	2	2	n
9	1	2	4	RT	M	610	1	Resp	Lung wt	10	1	2	2	2	n
10	1	3	1	RT	M	560	0.5	Resp	Lung wt	10	1	1	1	1	n
11	1	3	2	RT	M	610	0.5	Resp	Lung wt	10	1	1	1	2	y
12	1	3	3	RT	M	660	0.5	Resp	Lung wt	10	1	2	2	2	n
13	1	3	4	RT	M	710	0.5	Resp	Lung wt	10	1	2	2	2	n
14	2	1	1	RT	M	330	2	Resp	Lung wt	10	7.1	0	0	0	n
15	2	1	1	RT	M	330	2	Resp	Lung wt	10	2.9	0	1	1	n
16	2	1	2	RT	M	360	2	Resp	Lung wt	10	2.9	1	0	0	n
17	2	1	2	RT	M	360	2	Resp	Lung wt	10	7.1	1	1	1	n
18	2	1	3	RT	M	390	2	Resp	Lung wt	10	4.9	1	1	1	n
19	2	1	3	RT	M	390	2	Resp	Lung wt	10	5.1	1	2	2	n
20	2	1	4	RT	M	410	2	Resp	Lung wt	10	1.2	1	1	1	n
21	2	1	4	RT	M	410	2	Resp	Lung wt	10	8.8	2	2	2	n

6.0   MODEL AND BMD TAB

Figure 9.—Model and BMD tab.

The Model and BMD tab provides further opportunity to specify the options that CatReg will use to calculate an exposure-response curve. All you need to do is select or deselect option boxes or buttons, but having an understanding of clustering and stratification will help inform your selections.

See also Appendix C, for technical background discussions on exposure-response models and extra risk concentration (ERC).

6.1   Setting BMD Specifications for an Analysis

Table 8: *BMD Specification Options on the Model and BMD Tab*

Specifications	Description
Risk	Extra is the default risk type. Extra risk is the additional risk divided by the predicted proportion of animals that will not respond in the absence of exposure, 1 - P(0).
BMR	The response, generally expressed as in excess of background, at which a benchmark dose or concentration is desired. User input value (or default of .1000). must be a number >0 and <1.
Confidence Level (%)	The confidence level (default 0.95) associated with the statistical lower bound of BMD (BMDL) calculation. Confidence level must be a number >0 and <1.
Time	Exposure duration, in hours

6.2   Setting Model Specifications for an Analysis

Table 9: *Model Specification Options on the Model and BMD Tab*

Specifications	Available Options	Notes
Link Function	Options are: • Logit (default) • Probit • Cloglog	For more information, see page 33 .
Model Form	Options are: • Cumulative Odds (default) • Unrestricted Odds	For more information, see page 34.
Zero Background Response	Options are: • Yes (default) • No	For more information, see page 35.
Dose	Options are: • Log10 (default) • Linear	N/A
Time	Options are: • Log10 (default) • Linear	N/A
Worst Case Analysis	Options are: • Yes • No (default)	For more information, see page 35.

6.3   Running an Analysis

After specifying the parameters and options, select the **Run Analysis** button. CatReg will execute the analysis according to your selections.

CatReg opens a new window to display the output file specified in the Output File box at the bottom of the Analysis Screen. In the following screenshot, the output file's name is ChemxHu.otx.

Figure 10.—CatReg Output File.
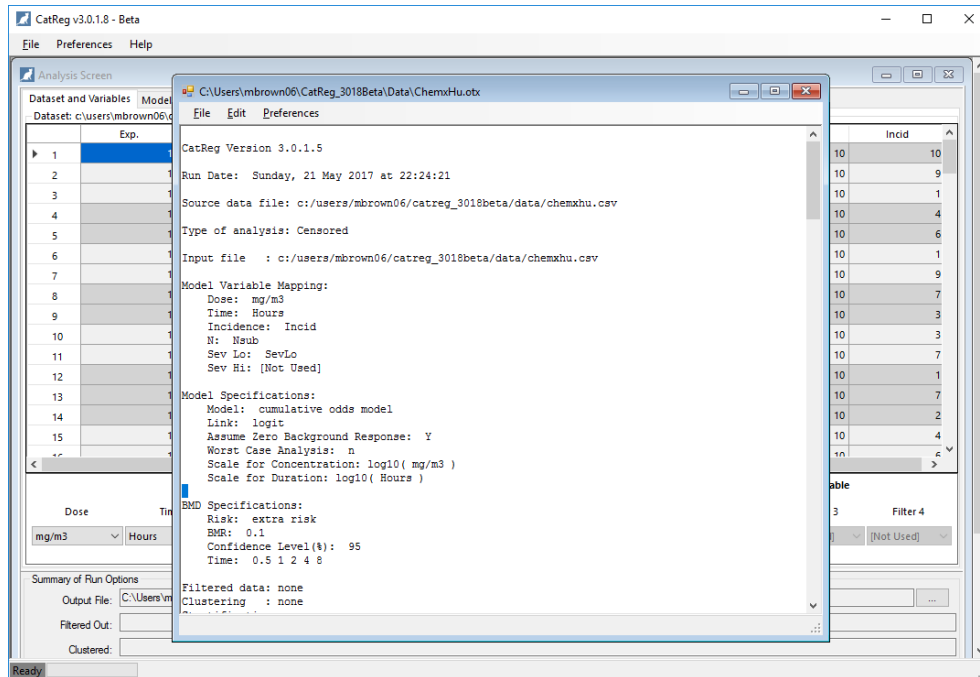
6.4   About the CatReg Output Files

CatReg's output file provides summary information such as the name of the input file, the setup options used, the table of coefficient estimates, an analysis of deviance table, and extra risk concentrations (ERCs) with upper and lower confidence bounds at exposure times of 1, 4, 8, and 24 hours.

CatReg's output file has the extension *.otx. Figure 24 in Section 9.0 contains a sample output file.

The .otx output file is a simple text file that CatReg automatically saves to the directory holding the currently loaded input file. From the text file window, you can save, print, edit, and otherwise manipulate the file's text. You can also set the file's font and line wrapping preferences.

CatReg by default suggests a name for the output file based on the currently loaded input file's name. For example, if the input file is named "ChemZ.csv," then CatReg will save the output file as "ChemZ.otx".

CatReg also saves the output results in .csv format to the directory holding the currently loaded input file. The .csv file name uses the input file's name and appends "_summary".

6.5   Changing the Output File's Name and Location

To change the name and location of the output file, select the **...** button beside the Output File field in the Analysis Screen's Summary of Run Options area. CatReg displays a Save As dialog box. Use this dialog to define a new location and filename.

By default, CatReg saves output files to the directory holding the currently loaded input file.

6.6   Stratifying

On the Model and BMD tab, select the variable(s) from the picklists to stratify by Intercept, Dose, or Time.

Stratification is a way of allowing one or more of the regression parameters (intercept, coefficient for concentration or dose ("mg/m3"), and coefficient of time ("Hours") to change when a specified variable changes value. For example, instead of assuming a common intercept parameter for three different species, stratification of intercept on the variable Species adds two more intercept parameters so there is one for each species. Stratification of the intercept on three species defines three subgroups (strata) of data, one for each species. The same parameter can be stratified on more than one variable. For example, the intercept might be stratified on both Species and Target. If there are two target organs for each of three species, then there are six strata, each corresponding to a distinct combination of species and target. CatReg will provide six intercept estimates, one for each species-target combination. To stratify the intercept on Species and Target, enter those two variable names in response to CatReg's query on whether to stratify the intercept (enter both on the same line, separated by a space, or on separate lines). In the same way, CatReg queries for variables on which to stratify the coefficients of dose and time.

Stratification is often conducted to test if the value of a regression parameter (e.g., the intercept) is the same for two or more values of a variable (e.g., Species). The user typically wants to produce an exposure-response curve that is suitably "accurate" by taking account of different parameter values that may occur between species, endpoints, etc., but that achieves that objective as simply as possible (i.e., with the minimal number of parameters). Stratification can be a way of implementing toxicological considerations, as the following example illustrates.

Suppose the data contain mortality results from experiments using rats and mice. The basic explanatory variables are atmospheric toxicant dose and time of exposure. The response score equals 0 for surviving animals and 1 for animals that died. In this case, the maximum severity score is S = 1. Assuming C and T enter the model logarithmically, the basic model has the form

$$L[\Pr(Y = 1|C,T)] \; = \; \alpha_1 + \beta_{1^*} \log_{10}(C) + \beta_{2^*} \log_{10}(T),$$

where L is the link function (the inverse function of H in Appendix C, Eq. 1a). Logarithmic scaling is typical when the explanatory variables range over two or more factors of 10. Because of different rates of respiration and metabolism among rats and mice, it may be reasonable to assume the internal dose for rats should be rescaled compared to that of mice. One possibility is to assume that a dose of C for rats is equivalent to a dose of kC for mice, where k is common to all mice in the study. Then, for the mice,

$$L[\Pr(Y = 1|C,T)] \; = \alpha_1 + \beta_{1^*} \log_{10}(kC) + \beta_{2^*} \log_{10}(T).$$
$$= [\alpha_1 + \beta_{1^*} \log_{10}(k)] + \beta_{1^*} \log_{10}(C) + \beta_{2^*} \log_{10}(T).$$

This shows that the mice, in effect, have a different intercept than do the rats, namely $\alpha_1^{MU} = \alpha_1^{RT} + \beta_{1^*} \log_{10}(k)$, where MU and RT refer to mouse and rat parameters, respectively. By stratifying the intercept parameter, the data are allowed to determine the estimate of the conversion factor k. Whether k is significant would be determined by testing $\alpha_1^{MU} = \alpha_1^{RT}$. Refer to Section 7.1 for an explanation of how to perform hypothesis testing in CatReg.[1]

6.7   Clustering

On the Model and BMD tab, select the variable(s) from the list boxes that are part of a cluster. The list boxes display the variables for selection.

---

1 Note in particular the discussion at the end of that section on the testing of the intercept parameter estimate, which is done slightly differently from other parameters because the intercept parameter estimates are reported in CatReg as incremental changes from a reference intercept parameter estimate.

An input file may consist of subsamples of data from common sources that causes them to be more similar to each other than to observations from another source.

Example 1: Suppose there are reports from three different "identical" experiments conducted at three different laboratories. The data from each laboratory may be considered a cluster because of the following reasons:

- There are likely some differences among laboratories in the way subjects were fed, their animal suppliers, the age of the subjects, the conditions under which the subjects have been taken off the study, the protocol for histopathology (or just different histopathologists), etc.
- The differences can be viewed as random effects by thinking of the specific laboratories as a random sample among a population of laboratories.
- The magnitude of the differences among the specific laboratories may vary.
- The laboratories are reasonably homogeneous (i.e., there is not one or more of them that is unreliable or consistently different in some way from the others).

Example 2: Suppose that an experiment is conducted wherein pregnant female rats are exposed to a toxic substance. Each rat gives birth to a litter and the pups are examined for specific health effects. Each litter could be considered a cluster sample.

Clustering is necessary whenever there is reason to suspect that batches of data are correlated (i.e., when the design of the study involves cluster sampling). The cluster variable should uniquely identify each batch of correlated data uniquely.

Cluster labels might be text identifiers, identification numbers, combinations of variables, etc.

The only requirements are that observations from the same cluster have the same cluster label, and those from different clusters have different cluster labels. If no cluster variables are specified, the program treats all data as being independent.

CatReg assumes that responses from the same cluster are correlated, whereas observations from different clusters are independent. It adjusts for the cluster sampling effect using the method of generalized estimating equations (GEE). The cluster adjustment affects standard errors, confidence limits and hypothesis tests (p-values), but it does not affect parameter estimates or the deviance (a statistic used to measure the fit of exposure-response curve to the data). For technical background on GEE, see Simpson et al. (1996b) and Diggle et al. (1994). It also may be noted that cluster sampling invalidates the large sample F distribution of the generalized F-statistic. However, it is common practice to compute F as a rough guideline (see Venables and Ripley, 1994, p. 187). In any case, the R^2 statistic gives an idea of how much variation in the response is accounted for by the explanatory variables (see Section 7 for information on how F and R^2 are computed). Ignoring clusters of observations typically leads to underestimation of variability in estimates and confidence bounds that are inappropriately narrow.

6.8   Link Function

A link function is a function applied to the exposure-response curve to transform it to a simple linear relationship in dose and time. By also transforming the observed responses, the link function reduces the mathematical complexity of estimating the parameters. The parameter estimates then are substituted into the (untransformed) exposure-response curve.

CatReg provides three different link functions for the exposure-response curves.

Table 10: *Link Functions*

The Link Function...	Corresponds to Probability Function...
Logit	Logistic
Probit	Normal

The Link Function...	Corresponds to Probability Function...
Cloglog (complementary log-log)	Gumbel

A comparison of how well the different link functions fit the data may be assessed using the AIC (Akaike Information Criteria, Akaike (1974)). A link with a smaller AIC provides a closer fit to the data. The AIC for different link functions may be compared only if there are no changes in the data or in use of the data filtering and stratification options. Otherwise, the differences in AIC may result from the changes.

The parameter estimates and their statistical characteristics, including standard errors and significance levels, are routinely output by CatReg, along with an analysis of deviance table to assess model fit and a table of estimates of extra risk concentrations (ERC) (concentrations at which extra risk is a user-specified value) for 1-, 4-, 8-, and 24-hour exposure times.

6.9   Model Form

There are two choices of models, Model 1 (cumulative odds model) and Model 2 (unrestricted cumulative model).

CatReg provides a choice of two models, the cumulative odds model and the unrestricted cumulative model.

The Cumulative Odds Model is described by the following equation:

$$\Pr(Y \geq s | C,T) = H[\alpha_s + \beta 1 * f1(C) + \beta 2 * f2(T)] \qquad \textbf{Model 1}$$

The Unrestricted Cumulative Model is described by the following equation:

$$\Pr(Y \geq s | C,T) = H[\alpha_s + \beta 1_s * f1(C) + \beta 2_s * f2(T)] \qquad \textbf{Model 2}$$

CatReg refers to any model of the form of Model 1 as a cumulative odds model because the model is expressed in terms of the cumulative probabilities, or odds, for $Y \geq s$.

Note that Model 1 is a special case of Model 2 wherein parameters $\beta 1_s$ and $\beta 2_s$ do not depend on s (which denotes severity level). That is, Model 1 is a simplification of Model 2 in which only the intercept term can vary across severity levels, not the coefficients of dose or time (a restriction called parallelism).

In other words, the cumulative odds model (Model 1) states that:

- the probability that a severity level *s* or higher will occur at a given dose or concentration (*C*) and time (*T*) is given by the exposure-response curve (logistic, normal, or Gumbel, determined by the choice of link function), and

- the intercept parameters may differ by severity level (i.e., a different intercept for each severity level), but the coefficients for *C* and *T* do not differ by severity.

A primary use of fitting Model 2 is to test whether the simpler Model 1 is adequate. Model 2, although more general than Model 1, has the undesirable feature that the regression lines for different severity levels may cross. Often the crossing is well outside the range of values of interest, so the model can be used to make empirical risk estimates. The user has the option to add an additional parameter, γ, which represents a hypothetical background concentration or dose, in some circumstances.

6.9.1   Understanding the Model Equations

The left-hand side of the Model 1 and Model 2 equations above is read as follows: the probability that a response of severity level *s* or greater occurs, given that concentration or dose is *C* and

time is *T* (time refers to exposure duration). No expression for *s* = 0 is included because this is the minimal category, and *Y* is always greater than or equal to 0 (i.e., Pr(*Y* ≥ 0|*C,T*) = 1).

The right-hand side is described as follows:

- H is a probability function taking values between 0 and 1, for which the user has three choices: logistic, normal, and Gumbel.

- The parameter α_s is the intercept for severity level s, *s* = 2,...,*S* (to be called the intercept or severity parameters). The severity parameters are ordered as $\alpha_1 \ni \alpha_2 \ni ... \ni \alpha_s$. This constraint is a consequence of the requirement that the probability of exceeding a lower score is larger than the probability of exceeding a higher score for any fixed levels of *C* and *T*.

- In Model 1, the parameter β_1 determines the dependence of the response on concentration or dose (I), whereas β_2 determines the dependence on time (*T*). In Model 2, the parameters are also indexed by *s* because they may change values with severity level *s*.

- Current choices for f_1 and f_2 are "untransformed" and "base-10 logarithm." Other transformations of *C* and *T* may be obtained by transforming the input data.

- Parameters are α_s, β_{1s} (to be called the coefficient of concentration or dose), β_{2s} (to be called the coefficient of time or duration), for severity levels *s* running from 1 to *S*. All parameters may be stratified on variables, as discussed in Section 6.6.

6.10   Assume Zero Background Response

The option to assume zero background response is unavailable and permanently set to "No" unless the option to Log10 transform the concentration (Dose) parameter is chosen. In all other cases, the background response is not assumed to be zero and is estimated. When the concentration (Dose) term is Log10 transformed, then you can choose Yes or No for the Assume Zero Background Response option.

- If you choose Yes, then the implied probability of an adverse response at zero concentration or dose is zero and observations at zero concentration are uninformative (treated the same as if they were filtered out). **Warning:** This option should only be chosen if it is not reasonable to expect severity > 1 responses at zero concentration or dose (e.g., when evaluating mortality from acute exposures) and when the observed data do not have severity > 1 responses at zero concentration or dose.

- If you choose No, then CatReg adds a hypothetical background concentration or dose to the administered doses, denoted as the parameter γ (gamma) that is estimated by maximum likelihood and displayed in the summary table of parameter estimates.

For example, If Assume Zero Background Response is set to No, an experimental concentration or dose of 50 mg/m3 is treated as an observation at dose (50 + γ) mg/m3, where γ is estimated from the data simultaneously with the other model parameters. If the logarithmic option is chosen for time (Hours), the implied probability of an adverse response at zero time is zero and observations at zero time are uninformative (treated the same as if they were filtered out).

6.11   Worst Case Analysis

CatReg provides an option to do a worst-case analysis when there is at least one record in the input file that contains censored data:

In a worst case analysis, censored responses are treated as occurring at their highest possible (worst) severities.

Although the graphical presentation of censored points will not change in a worst-case analysis, higher estimates of risk will be produced than those for the corresponding censored analysis. Comparison of risk estimates from the two methods provides an indication of the sensitivity of the results to the severity scoring.
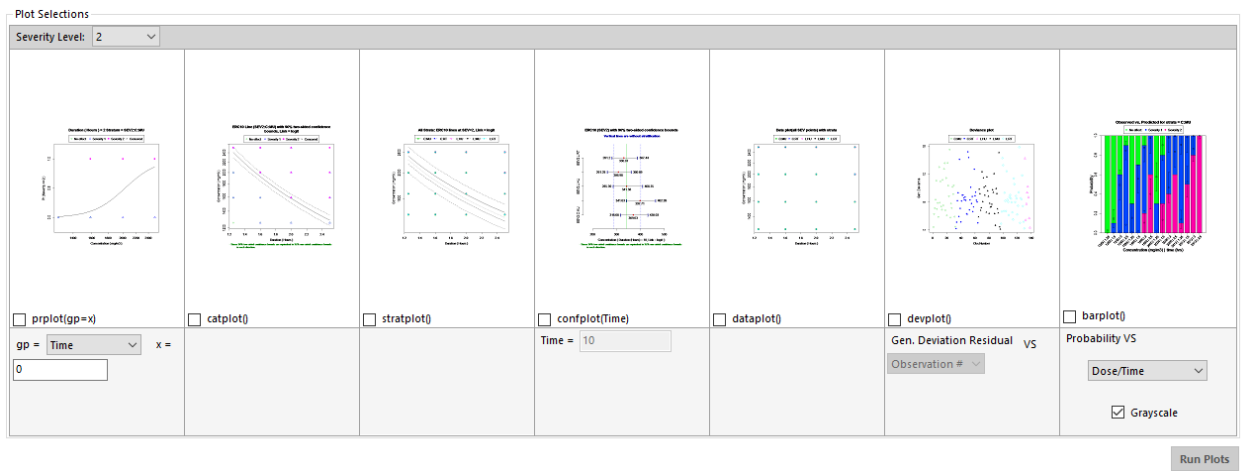
When the worst case option is selected, the output file contains the line: Type of analysis: Worst-case.

When the worst-case option is not selected, the output line is: Type of analysis: Censored.

7.0   PLOTS TAB

CatReg includes functions for making plots after an exposure-response curve has been fit to the data.

Figure 11.—Plots tab.

To specify a plot:

1. Select the Severity Level from the drop down list. The choices are 0, 1, 2 (default), and All.

2. Select one or more plot functions. The thumbnail illustrations indicate the type of graph that CatReg will generate and display. Note that some plots (such as prplot and barplot) require additional inputs.

3. Select **Run Plots**. The plots will be displayed in their own window (see Figure 12).

7.1   About the CatReg Plot Files: Locations, Formats, Names

CatReg saves any plots you've selected to the directory that contains the currently loaded input file.

CatReg names the files based on the input file's name and the select plot type, plus any options/parameters for filtering, stratification, etc. that were selected. Examples: ChemxHu_barplot-SEV2-MU.png, Chemx_catplot-SEV2.emf.

Note	If you re-run an analysis without changing any parameters and generate a plot, CatReg will overwrite the previous plot file.

The plots are saved in either .PNG format (barplot) or .EMF format (all other plots). You can import these files into PowerPoint, GnuPlot, or other graphics application for printing.

For the EMF-formatted plots only, you can use PowerPoint, GnuPlot, or another graphics program to manipulate or edit the plot and its components. For example, import an EMF plot into PowerPoint, right-click on the graphic, and select the Group>Ungroup command to edit each element of the plot (text, points, lines, etc.).

7.2   Copying Plots from the Graphs Window

Right-click on any plot in the window and select Copy to Clipboard. You can then paste the plot into PowerPoint, GnuPlot, or other graphics application for printing.

Figure 12.—Plots generated by CatReg.
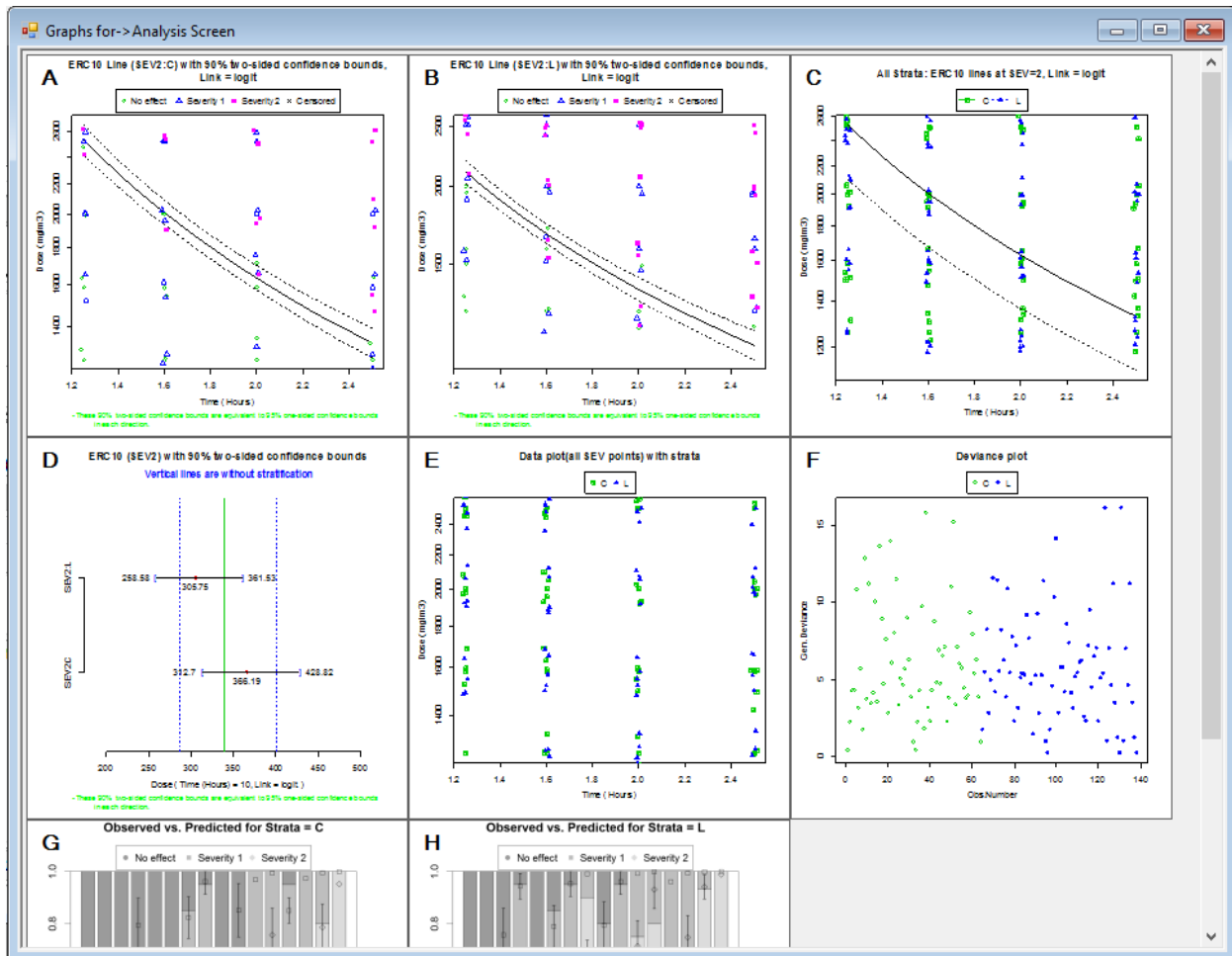
7.3   Plot Functions and Options

The following tables describe the plot functions and specific options available within CatReg.

Table 11: *Plot Functions and Options*

Function	Description
prplot(gp=x)	Probability of exceeding a specified severity level on y-axis, dose or time on x-axis. Plots the probability curve as a function of dose, keeping time fixed, or as a function of time with dose fixed. Use the dropdown list to set the time or dose level to be used.
catplot()	Dose on y-axis, time on x-axis. Plot of ERC line with confidence interval for a single severity level and stratum, and response data for all severity levels.
stratplot()	Dose on y-axis, time on x-axis. Plot of ERC lines and response data for all strata for a single severity level.
confplot(Time)	Strata for a single severity level displayed on y-axis, dose on x-axis. Plot of ERC with confidence interval for the unstratified model, and for individual strata of the model, for a single exposure time.
dataplot()	Dose on y-axis, time on x-axis. Plots response data for all severity levels combined by stratum.

| Function	Description
devplot()	Generalized deviance residuals on y-axis and data observation number, log-dose, or log-time on x-axis.
barplot()	Bar plot of probability of actual incidences of data, with calculated probability as a point, and error bars for each.

7.4   Jitter

CatReg's Jitter setting (which can be toggled from **Preferences>Jitter**) can be useful to get a sense of data density in cases where observations pile on top of one another.

The Jitter function adds "noise" – small random offsets to coincident observations. Although CatReg bounds the amount of offset, the randomness of the offset causes each jitter of the data to produce a slightly different plot.

You can see the effects in the following figure. In plot a., it is impossible to discern the density of data as it appears to be a grid of points. But with jitter enabled in plot b, we can see the data distribution more easily. While plot b would not be useful for analyzing the data (CatReg provides text results for that), the jittered graph provides a quick visual look at the underlying data.

Figure 13.—Plots without jitter (a.) and with jitter (b.).
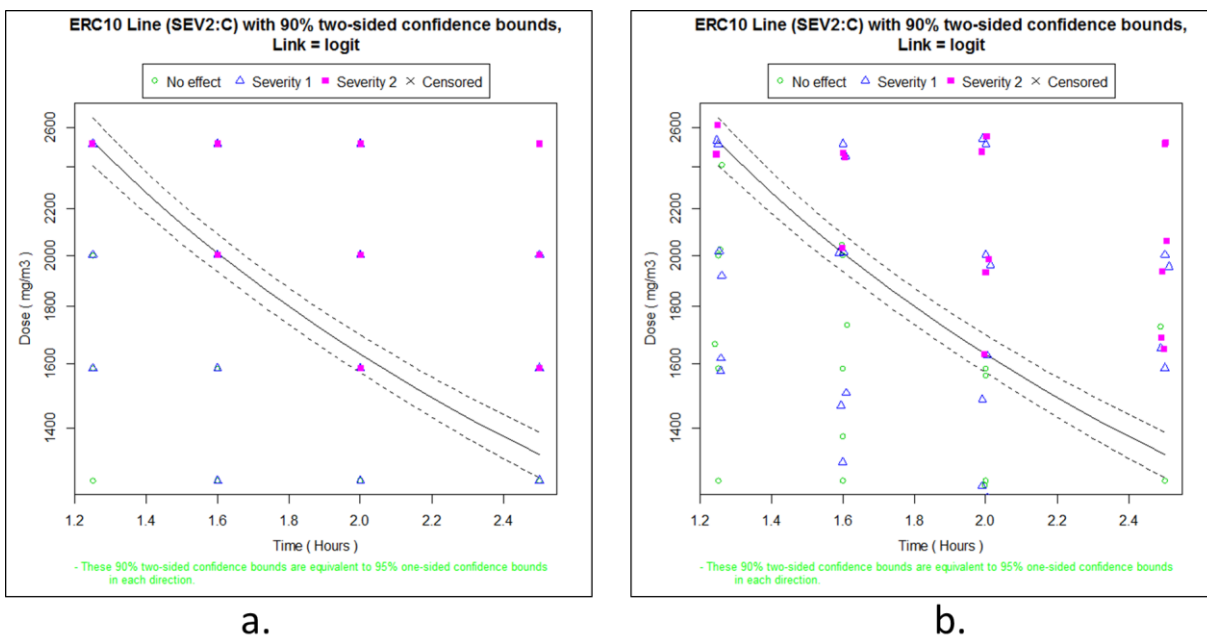
After you toggle the Jitter preferences setting, you will need to select the **Run Plots** button to display plots with the new setting.

7.5   Doses and Times for Designated ERC and Severity (catplot)

The catplot() plots the response data and graphs the extra risk concentration (ERC) with confidence interval, for a single severity level and stratum.

Exposure concentration or dose is on the y-axis and exposure duration (time) on the x-axis.

The current ERC settings are used for ERC percentile and severity level.

The confidence interval percentile is 90% (two-sided) and cannot be overridden.

By default, dose and time are graphed on a log-linear scale. This type of graph is useful for showing how extra risk changes with dose or time. Be aware that the weighting of individual points is not shown.

Each call to catplot generates a new graphics window. Thus, repeated use of catplot allows comparison of results across strata.

Figure 14.—Illustration of catplot().
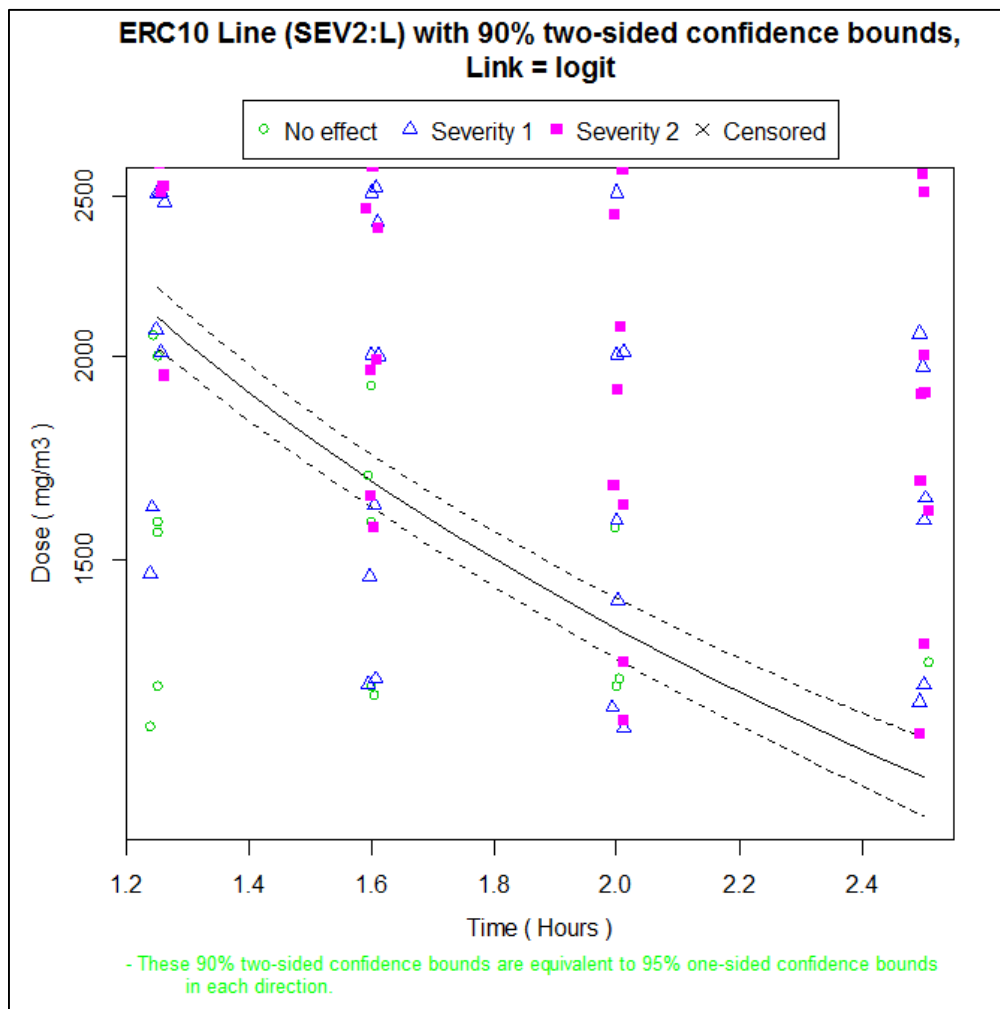
The x-y locations of symbols on the graph indicate the exposure doses and times of observations on mice with the central nervous system as the target organ. The symbol itself indicates the severity category, as shown in the legend. The lines on the graph are the estimated ERC10 (solid line) and upper and lower 95% one-sided confidence bounds (dashed lines) (equivalently, two sided 90% confidence bounds). Slicing the graph vertically at a specific time gives the confidence interval that confplot would graph for that time. As expected, longer times require lower doses to achieve the estimated 10% level of extra risk.

7.6 Doses and Times for Designated Probability and Severity of Strata (stratplot)

The command stratplot() plots the response data by stratum for all severity levels and graphs the extra risk concentration (ERC) (without confidence interval) by stratum for a single severity level.

Exposure concentration or Dose is on the y-axis and exposure duration or Time on the x-axis.

The current ERC settings are used (for the ERC percentile and ERC severity level).

Response to a pop-up menu determines whether to include a legend.

By default, Dose and Time are graphed on a log-linear scale.

Be aware that the weighting of individual points is not shown. The number of points in each severity category will be displayed in the R command window, along with the number of hidden points.

This graph provides comparison of the ERC curves for different strata by plotting them on the same graph.

Figure 15.— Illustration of stratplot().
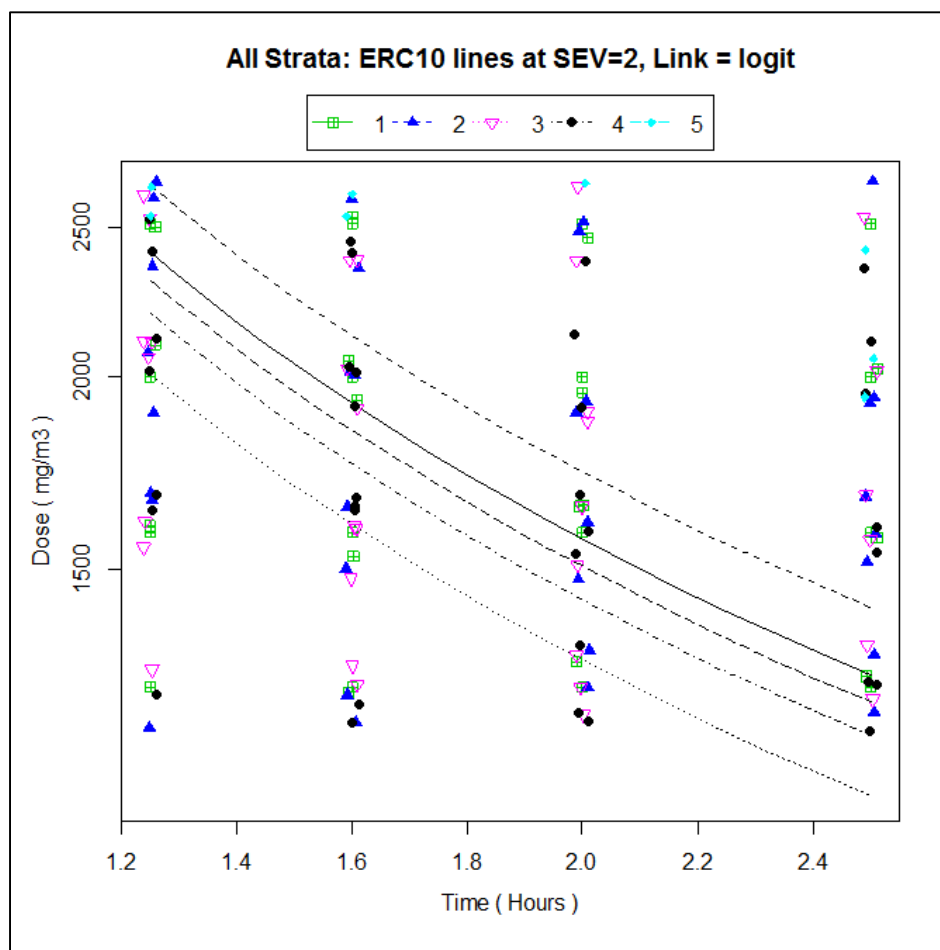
7.7   Extra Risk Concentration at Desired Time (confplot)

The command confplot(Time) displays the extra risk concentration (ERC) with confidence interval for the unstratified model and for individual strata of the stratified model, for a single severity level at the exposure duration set in the "Time =" field. The default value is 10, for 10 hours exposure.

The time argument is required unless time is not included in the data as an explanatory variable.

The current ERC settings are used for the ERC percentile and severity level.

The confidence interval percentile is 90% (two-sided).
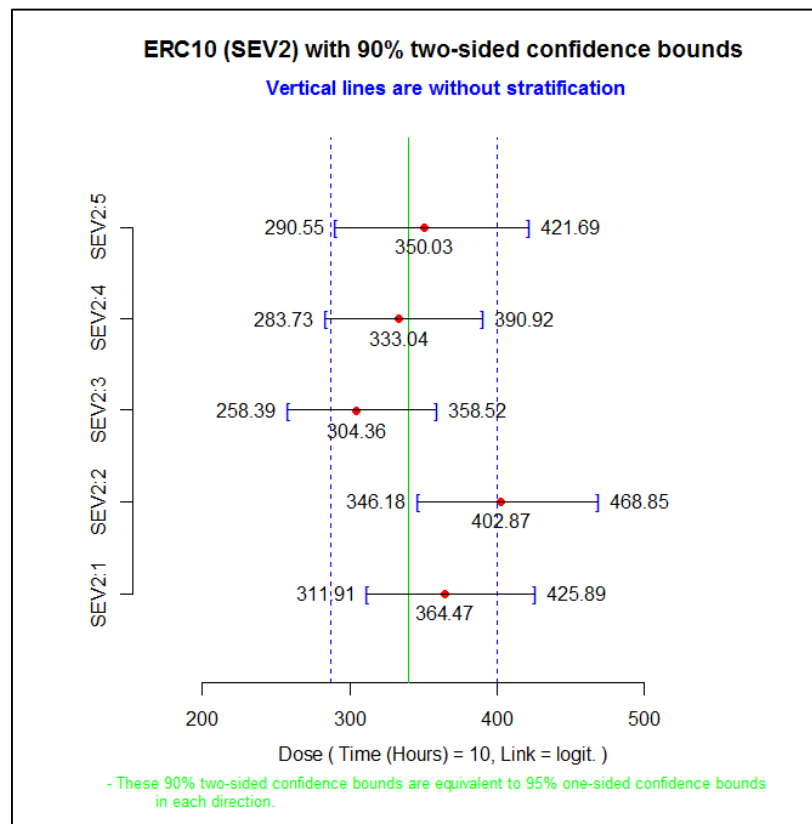
This graph is useful for comparing ERC estimates and confidence intervals among strata, and comparing individual strata with the unstratified model.

Figure 16.—Illustration of confplot().

The above figure was produced with Time=10 hours. Dose is on the x-axis. The central dot for each stratum is the estimated ERC10 for severity level 1.

The vertical solid line is the estimated ERC10 for the unstratified exposure-response curve, and the vertical dashed lines are the associated confidence intervals (90% for two-sided bounds, determined from the ERC settings).

To display the confidence interval for another duration, repeat the confplot command and reset the Time field.

7.8   Probability Versus One Explanatory Variable (prplot)

The command prplot(gp=x) displays the probability curve (i.e., the exposure-response curve) as a function of concentration or dose, keeping time fixed, or as a function of time with dose fixed.

If either of these variables is constant in the data, then prplot graphs the probability against the non-constant variable.

If both dose and time vary in the data , then you need to tell prplot which variable to hold constant. You can specify either Time or Dose from the "gp=" dropdown list:

* To plot the response probability versus time at a fixed dose, select gp=Dose.
* To plot the response probability versus dose at a fixed time, select gp=Time.

The severity level is determined by the current ERC settings. The choice of stratum and whether to include a legend are determined by response to pop-up menus.

The function prplot can help assess whether the fitted probability curve is consistent with the data, and for representing the risk over a range of exposure levels.

Figure 17 was produced with gp=Time and x=1.25. The curve plots the probability of the occurrence of a severity level 1 response or greater for the liver target organ in mice (species MU) as a function of concentration or dose, with time fixed at 1.25 h. The value used for time need not be an exposure time in the data. When that occurs, the probability curve is displayed but there are no data to display.

Figure 17.—Illustration of prplot().
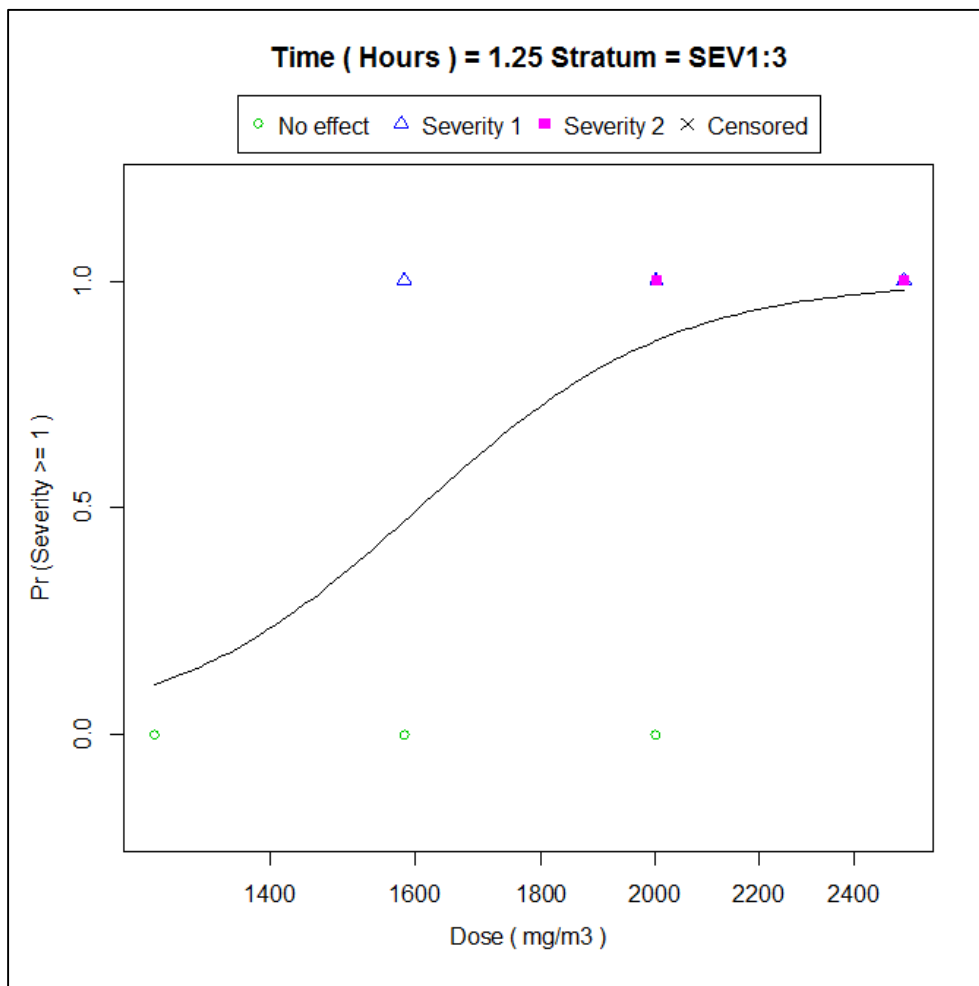
In these graphs, the vertical location of a symbol represents whether the response at a particular dose was equal to or greater than the severity of interest (adverse, severity 1, for the above figure). No effect responses are plotted at Pr = 0. Adverse effects (severity 1) and severe effects (severity 2) are both plotted at Pr = 1. There are no censored observations. If the data had contained some observations censored as [0,1], they would have been graphed as "×" at Y = 0.5 on the severity ≥1 graph, and as "×" at Y = 0 on the severity ≥2 graph. This is because, in the first case, it is not known whether the censored observation meets the severity threshold (i.e., whether severity ≥1), whereas in the second case the severity is known to be less than 2.

7.9 Data Plotted by Stratum (dataplot)

The command dataplot() plots the response data for all severity levels by stratum, without a response probability curve, as shown in the following figure.

Exposure concentration or dose is on the y-axis and exposure time on the x-axis. Notice that this figure is very similar to the stratplot() figure, except that the dataplot() figure does not show the ERC10 lines.

Figure 18.—Illustration of dataplot().
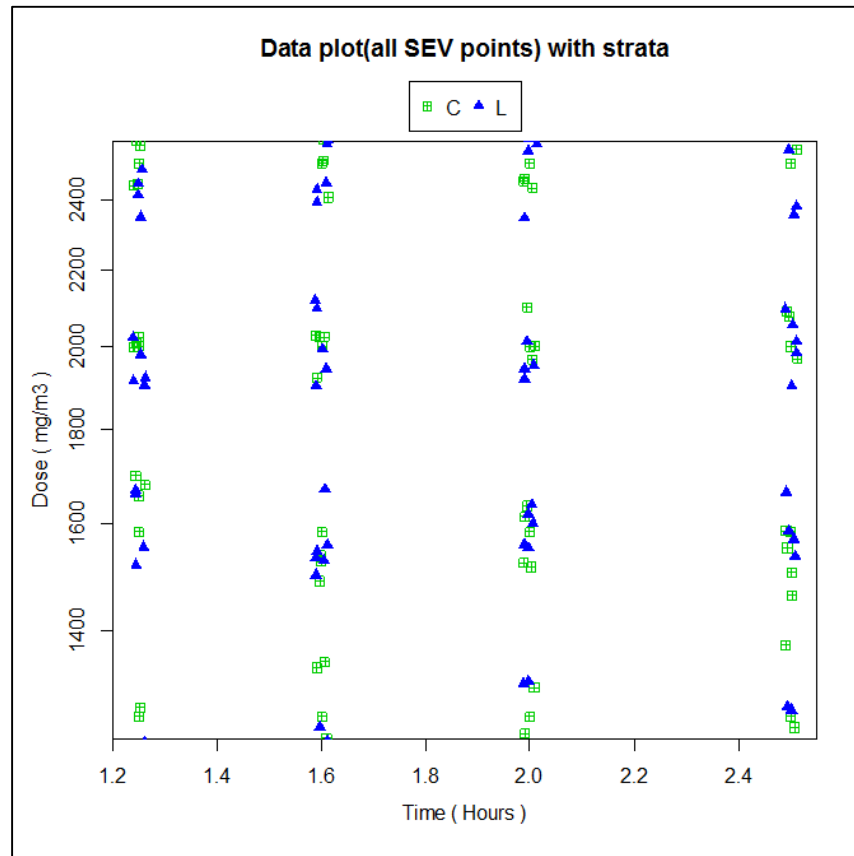
7.10 Contribution to Deviance for Individual Datum (devplot)

How well are individual observations explained by the exposure-response curve?

In linear regression, residual plots help to address that question, that is, plots of the differences between observed and predicted responses.

Here the response is binary or ordinal, so there is a scaling problem in defining a "residual." Various possibilities for resolving this problem exist.

CatReg uses the individual components of the deviance statistic to measure how well individual observations are explained by the exposure-response curve. Using deviance residuals is common practice in generalized linear modeling (see Venables and Ripley [1994, p. 188-189]).

For interval-censored ordinal data, other types of residuals are undefined, whereas the deviance residuals are defined in the usual way. The observations that contribute to any lack of fit of the exposure-response curve can be identified by examination of the individual contributions to the deviance.

The devplot() function produces a diagnostic plot of generalized deviance residuals versus observation number (excluding filtered data), dose (or log-dose), or time (or log-time), depending on the user selection from a pop-up menu.

If the model is stratified, the points are labeled by strata.

Generally, plotting deviances versus observation number (Obs.Number) is a good choice. It provides a representation of the relative effectiveness of the model in fitting the different observations or strata. Devplot does not require jittering because each observation is represented uniquely under the Obs.Number option. The fit is suspect if one or a few observations have much larger deviance residuals than the remaining observations because the fit may be unduly influenced by these observations. If one stratum has large deviances, the model may be inadequate for this stratum. Rerunning the model without this stratum would allow one to determine whether the results for the other groups are heavily influenced by the poorly fit subset.

Plots of deviance versus dose (or log-dose) or time (or log-time) are useful in studying the adequacy of the functional form of the regression relationship. Trends in the deviances would suggest a problem with the functional form. One should be aware that differences in the density of the data at different dosages or times will affect the perception. Regions of the plot with more data will tend to have more spread in the deviances because of random variation, even if the model is adequate for the data.

Figure 19 was produced by selecting Observation # from the pop-up menu.

Figure 19.—Illustration of devplot() — Example 1.
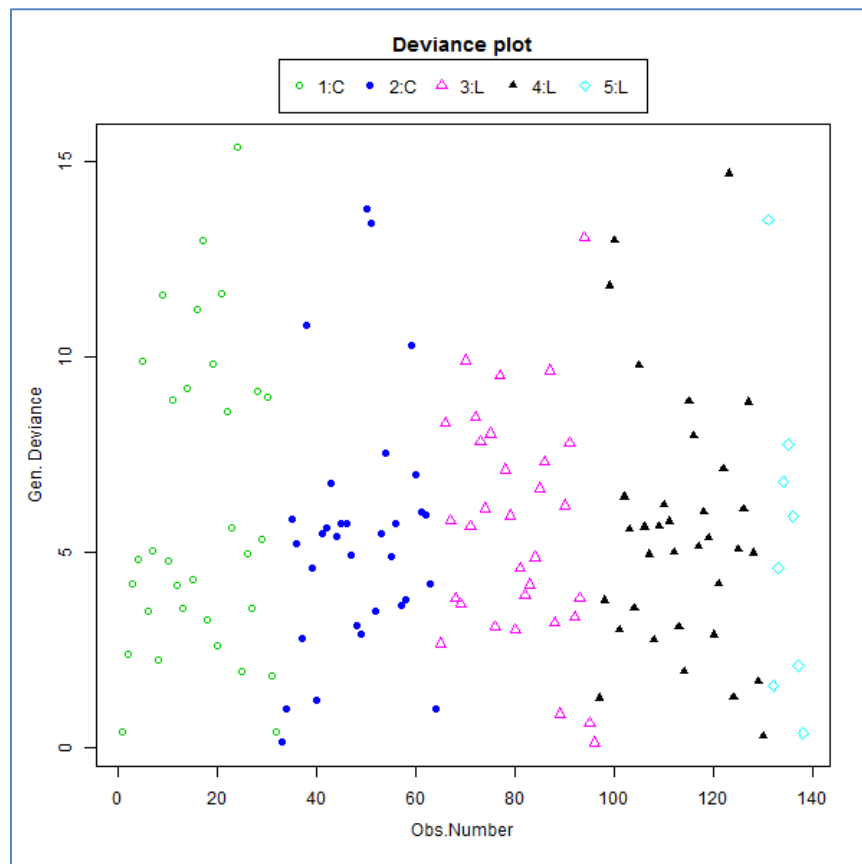
Figure 20 is a devplot() example for a different set of dose-response group data. The input and output files are not shown. The targets are C= central nervous system and L=liver. The species are HU (human), MU (mice), and RT (rats). Intercepts were stratified on "Species" and "Target". The notable feature of the figure is that the liver data for mice are relatively poorly described by

the curve (i.e., residual deviances are relatively high) in comparison to most of the other data. These data may or may not be influential on the fitted exposure-response curve. To examine whether they are, the curve could be refit after filtering out liver data for mice.

Figure 20.—Illustration of devplot() — Example 2.
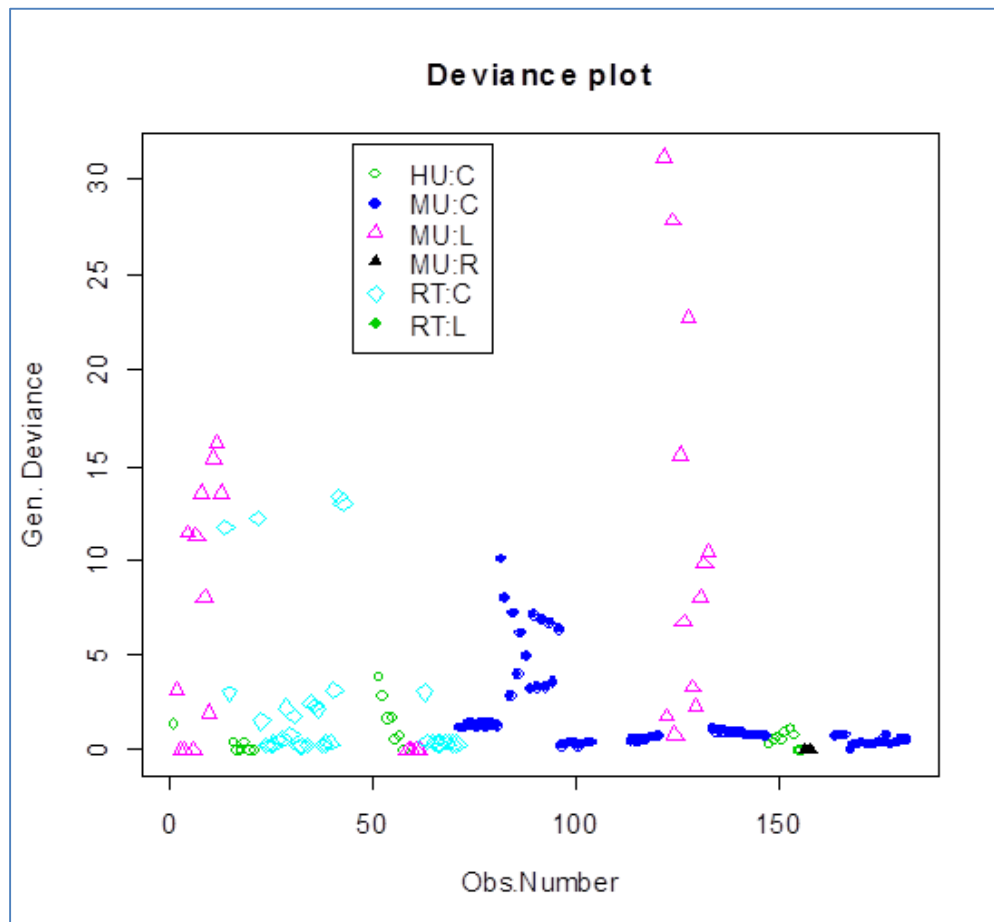
7.11 Observed and Predicted Probability Versus Dose/Time combination (barplot)

The barplot() displays a stacked bar plot showing the observed probability of each severity level for a given dose/time combination.

The plot overlays each bar with predicted probability points for each dose/time value at each severity level (except the "no effect" level) along with error bars to represent the prediction's standard error.

The barplot() offers the following X-axis options for Dose/Time combination:

- Dose/Time – sorts by dose first, and then time for each dose value
- Time/Dose – sorts by time first, and then dose for each time value
- Dose*Time – sorts by the product of dose*time

The plot also features a Grayscale option (checked by default) to output the plots in grayscale rather than color.

Figure 21 plots the Observed and Predicted Probability vs. Dose/Time for the central nervous system target organ in mice (species MU)

Figure 21.—Illustration of barplot()
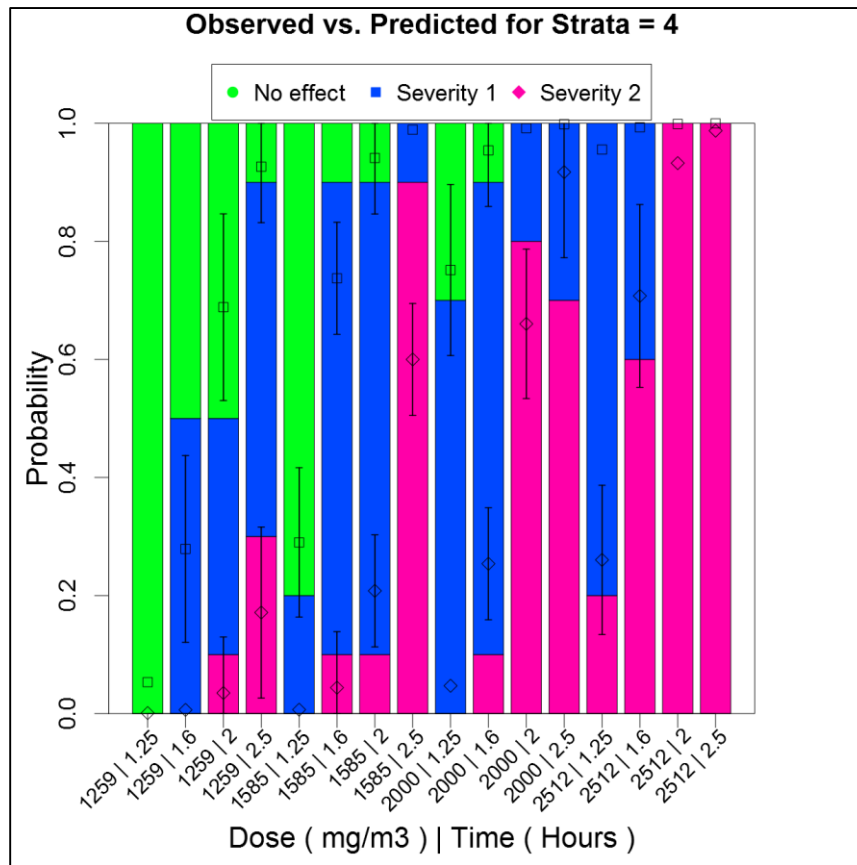
8.0   HYPOTHESES TAB

Figure 22.—Hypotheses tab.

From the Hypotheses tab, you can test exposure-response hypotheses. The Hypotheses tab is disabled until you have run an analysis. After running an analysis, the controls are enabled and eligible parameters from the analysis are loaded into the Intercept, Dose, and Time parameter lists.

- CatReg automatically runs parameter equality tests, based on the data. Those results are displayed in the **Automatic Tests** section.

- Simply select and drag parameters from the Available Parameters box to the appropriate Test for Equality boxes. For example, under Intercept, you can drag SEV1 to the left Test for Equality box and SEV2 to the right box.

- Select **Clear Tests** to clear the Test for Equality boxes and start over.

- Select **Run Tests** to run the specified hypotheses tests.

- After running the tests, CatReg displays a new window showing the text-based results. The window's title bar indicates the file name and location (in CatReg's Data\OptionFiles directory).

- The hypotheses results are saved as a text file with the extension *.ANX.

Figure 23.—Hypotheses results.

8.1   Testing Parameters

The output file includes tabular results for a simple test of the hypothesis that a parameter is zero. Dividing the estimated coefficient by the standard error provides a Z statistic for the corresponding parameter. This statistic provides a one-degree-of-freedom test of the null hypothesis that the parameter equals zero.

Under the null hypothesis, the Z statistic has an approximate standard normal distribution. The larger the sample, the better the approximation is.

The p-value in the table gives the significance level of the test. A parameter that is not significantly different from zero may be considered a candidate for removal and simplification of the exposure-response curve. The p-values apply to individual parameters considered singly, however, and further testing is needed to test the joint hypothesis that more than one parameter is zero, that two or more parameters are equal, or a combination of the two.

CatReg lists all the parameters in the Parameters list box. You can select a parameter and then select the appropriate buttons to test the parameter(s) for removal or to select a group of parameters to test for equality (the default is none in both cases).

The idea is to express the hypothesis to be tested as a set of constraints on the model coefficients.

A test is then conducted of the hypothesis or joint hypotheses, if more than one was entered. The test is a (generalized Wald-type) chi-square test of the null hypothesis that all of the specified constraints hold. The distribution of the test statistic is derived from the sampling distribution of the estimated model coefficients, and it takes into account any cluster sampling.

One may want to test for a sex difference, or whether there are interspecies differences in the exposure-response.

For such a test, the null hypothesis is that the specified parameters are equal. A p-value less than 0.05 is usually taken as evidence that the hypothesis should be rejected (i.e., the specified parameters are not equal).

The parameters to be tested as equal must, of course, be included in the exposure-response curve; this is accomplished by stratifying. For example, to test that there is no difference between species in the coefficients of dose, the user would stratify the Dose parameter on species on the Dataset and Variables tab.

When CatReg queries which coefficients to test for equality, the parameters to be tested (if any) are entered from the Parameters list on the Hypotheses tab.

To test for no difference between species in the coefficients of dose, you would enter the coefficient for each species. Some care needs to be used when tests involve stratified intercept parameters because stratified intercept parameters are represented in CatReg as increments relative to a reference. CatReg assigns the lowest alphabetical term to be the reference.

For illustration, consider the example provided in Section 6.6, "Stratifying." The intercept is stratified on species (rat and mouse) in this example. In this case, the MU:INTERCEPT parameter estimate will be the reference as it alphabetically precedes RT:INTERCEPT. Thus, the SEV1 intercept value reported by CatReg will apply to mice. The reported rat intercept value will be an incremental difference from the mouse reference value. To test whether the mouse and rat SEV1 intercept estimates are equal, one would perform a Z-test of the null hypothesis that RT:INTERCEPT is zero. To perform this test in the CatReg Hypothesis tab, simply select and drag the RT:INTERCEPT parameter to one side and "Value of 0" to the other side of the Intercept "Test for Equality of Parameters" boxes, then select "Run Tests."

9.0   ASSESSING MODEL FIT

This section focuses on the fit of the exposure-response curve rather than on its parameters.

In linear regression analysis, it is common practice to consider the proportion of variation accounted for by the model, the so-called R^2 statistic, as a measure of the model's explanatory power. This statistic, which ranges in value between 0 and 1, is the ratio of the model and total sums of squares. These sums of squares, along with degrees of freedom, F-tests, and so on usually are reported in the form of an analysis of variance table. Standard texts such as Weisberg (1985) describe the use and interpretation of these statistics.

CatReg provides generalized analysis of variance and R^2 statistics for assessing the explanatory capacity of the exposure-response curve. These are derived from deviance statistics for hierarchical models. Following McCullagh and Nelder (1989) and Venables and Ripley (1994), this type of analysis is called the analysis of deviance. The analysis of deviance statistics and R^2 statistic are in a table in the output file.

After running CatReg, the fitted curve is available for further analysis. As an example of the analysis of deviance, consider the output file in Figure 24.

The summary output, which shows the coefficient estimates, indicates that seven parameters have been estimated. There are 64 treatment groups (2 species x 2 targets x 4 doses x 4 times), so total degrees of freedom (df) is 64 x (3 severity levels - 1) = 128 (unadjusted for SEV1 and SEV2). The estimates of the SEV1 and SEV2 intercepts provide no information about how well the curve fits the data, so it is customary to adjust the total degrees of freedom for them in the analysis of deviance table, leaving 128 - 2 = 126 degrees of freedom. The model has 5 parameters, aside from SEV1 and SEV2, which are MU:L:INTERCEPT, RT:C:INTERCEPT, RT:L:INTERCEPT, LG10C0NC, and LG10TIME. This leaves 126 – 5 = 121 degrees of freedom for the residual deviance.

The analysis of deviance table for the example shown in Figure 24 partitions the total deviance into the sum of two components: the "model" deviance and the "residual" deviance. The total deviance is the deviance when the only parameters in the exposure-response curve (or "model" in the terminology being used here to show the comparison with the analysis of variance) are the intercepts, SEV1 and SEV2. This is referred to as the null model because it contains no explanatory variables. The residual deviance is simply the deviance of the fitted model, which includes the 5 parameters, aside from SEV1 and SEV2, i.e., MU:L:INTERCEPT, RT:C:INTERCEPT, RT:L:INTERCEPT, LG10C0NC, and LG10TIME, as explanatory variables. The model deviance is that part of the total deviance that is explained by the model (the proportion is the generalized R^2). In this example, 46.4% of the total deviance is explained by the model. R^2 is a general measure of the proportion of the variation in the response that is accounted for by the explanatory variables.

The mean deviance entries, labeled as "Mean.Dev," are computed as deviance divided by degrees of freedom. An approximate F-test of the model is obtained as the ratio of model to residual mean deviations. This is an approximate F-statistic in large samples under the ordinal regression model with independent responses. This statistic tests the null hypothesis that all explanatory variables can be dropped from the model.

Generally, the F-test will reject the null hypothesis unless the sample size is very small or the model fits poorly. It merely verifies that there is some relationship between the response and the explanatory variables. The generalized R^2 statistic often will be of more direct interest as a measure of the explanatory value of the variables in the model.

Cluster sampling invalidates the large sample F distribution of the generalized F-statistic. However, it is common practice to compute F as a rough guideline; see Venables and Ripley (1994, p. 187). In any case, the R^2 statistic gives an idea of how much variation in the response is accounted for by the explanatory variables.

Figure 24.—Output file from CatReg. Input file:chemx.csv. Model 1. logit. Scales: log10.
Stratification: intercept on species and target. Assumed Zero Background Risk.

```
CatReg Version 3.1
Run Date:  Thursday, 15 Jun 2017 at 12:45:6

Source data file: c:/users/mbrown06/documents/catreg/catreg_31beta3c/data/chemx.csv

Type of analysis: Censored

Input file   : c:/users/mbrown06/documents/catreg/catreg_31beta3c/data/chemx.csv

Model Variable Mapping:
    Dose:  mg/m3
    Time:  Hours
    Incidence:  Incid
    N:  Nsub
    Sev Lo:  SevLo
    Sev Hi: [Not Used]

Model Specifications:
    Model:  cumulative odds model
    Link:  logit
    Assume Zero Background Response:  Y
    Worst Case Analysis:  n
    Scale for Dose(Concentration): log10( mg/m3 )
    Scale for Time(Duration): log10( Hours )

BMD Specifications:
    Risk:  extra risk
    BMR:  0.1
    Confidence Level(%):  95
    Time:  0.5 1 2 4 8

Filtered data: none
Clustering   : none
Stratification:
     Intercept    : Species Target
     Dose(Concentration):
     Time(Duration)     :

Message      :
Iterations   : 39 14
Deviance     : 741.0266
Residual DF  : 121
AIC          : 755.0266


Coefficients:
              Estimate Std. Error   Z-Test=0    p-value
SEV1         -66.2375249  4.2947663 -15.422847   0.00001
SEV2         -70.3128027  4.4629521 -15.754774   0.00001
MU:C:INTERCEPT   0.0000000  0.0000000         NA        NA
MU:L:INTERCEPT   1.5443003  0.2803179   5.509103   0.00001
RT:C:INTERCEPT  -0.8580588  0.2701988  -3.175658   0.00149
RT:L:INTERCEPT   0.7735915  0.2703681   2.861253   0.00422
LG10CONC        19.6220725  1.2858838  15.259600   0.00001
LG10TIME        18.0738765  1.2265874  14.735091   0.00001
```

```
Variance:
                    SEV1         SEV2 MU:L:INTERCEPT RT:C:INTERCEPT
SEV1            18.4450174 19.1496806    -0.32566226     0.13047885
SEV2            19.1496806 19.9179410    -0.34292950     0.14029845
MU:L:INTERCEPT -0.3256623 -0.3429295     0.07857815     0.03192034
RT:C:INTERCEPT  0.1304789  0.1402984     0.03192034     0.07300738
RT:L:INTERCEPT -0.1730734 -0.1814300     0.03822081     0.03365877
LG10CONC       -5.5095904 -5.7215951     0.08581225    -0.04903025
LG10TIME       -3.4384286 -3.6292401     0.08062509    -0.04554744
               RT:L:INTERCEPT    LG10CONC     LG10TIME
SEV1              -0.17307335 -5.50959042 -3.43842860
SEV2              -0.18143003 -5.72159505 -3.62924010
MU:L:INTERCEPT     0.03822081  0.08581225  0.08062509
RT:C:INTERCEPT     0.03365877 -0.04903025 -0.04554744
RT:L:INTERCEPT     0.07309891  0.04053521  0.04091319
LG10CONC           0.04053521  1.65349721  0.97163480
LG10TIME           0.04091319  0.97163480  1.50451662


Analysis of Deviance Statistics:

Generalized R-squared: 0.464


          DF  Deviance Mean.Dev  Gen.F pvalue
Model      5  642.5451  128.509 20.984      0
Residual 121  741.0266    6.124
Total    126 1383.5717


#########################################################################
Note: About  46.4 % of the variation in the response is accounted
for by the explanatory variables in the current model fit.

The p-value of the model fit is <= 0.05. This is generally considered
significant, indicating that the current model fit is acceptable.
#########################################################################


**************************** ERC summary ****************************



< ERC10 at specific time point(s) >
 : One-sided 95% lower bound and one-sided 95% upper bound

* Risk Type : extra risk

*  ERC10 at SEV2:MU:C
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 5606.0094          4972.9013          6319.7196        408.4255
1.0 Hours 2960.5699          2769.3104          3165.0386        120.2035
2.0 Hours 1563.4961          1493.3890          1636.8944         43.6072
4.0 Hours  825.6924           758.0997           899.3118         42.8733
8.0 Hours  436.0535           378.5090           502.3466         37.5186


*  ERC10 at SEV2:MU:L
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 4676.8304          4178.5047          5234.5862        320.3478
1.0 Hours 2469.8644          2322.1674          2626.9555         92.5903
2.0 Hours 1304.3514          1241.2888          1370.6178         39.2971
4.0 Hours  688.8364           627.6448           755.9938         38.9591
8.0 Hours  363.7790           313.2137           422.5075         33.0993
```

```
*  ERC10 at SEV2:RT:C
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 6199.8772          5475.8851           7019.5918     468.0497
1.0 Hours 3274.1953          3051.3295           3513.3390     140.3247
2.0 Hours 1729.1238          1652.3140           1809.5041      47.7660
4.0 Hours  913.1615           841.3680            991.0809      45.4587
8.0 Hours  482.2465           420.3186            553.2985      40.2960


*  ERC10 at SEV2:RT:L
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 5119.5203          4559.8071           5747.9379     360.3605
1.0 Hours 2703.6518          2536.5080           2881.8097     104.8931
2.0 Hours 1427.8160          1361.1725           1497.7224      41.4924
4.0 Hours  754.0389           689.3606            824.7856      41.1111
8.0 Hours  398.2128           344.0843            460.8564      35.3702
```

Note:
1. These One-sided 95% lower bound and one-sided 95% upper bound
confidence intervals are equivalent to
the lower bound and upper bound of two-sided 90% confidence intervals.

****

**************************** ERC summary ****************************

< ERC10 at specific time point(s) >
 : One-sided 95% lower bound and one-sided 95% upper bound

* Risk Type : extra risk

```
*  ERC10 at SEV1:MU:C
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 3475.0846          3129.8727           3858.3718     221.0445
1.0 Hours 1835.2147          1729.5769           1947.3045      66.1458
2.0 Hours  969.1887           912.5364           1029.3582      35.4899
4.0 Hours  511.8349           460.2481            569.2037      33.0579
8.0 Hours  270.3033           229.7032            318.0794      26.7462


*  ERC10 at SEV1:MU:L
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 2899.0999          2623.0928           3204.1490     176.3342
1.0 Hours 1531.0334          1442.0397           1625.5193      55.7405
2.0 Hours  808.5487           754.9296            865.9760      33.7292
4.0 Hours  426.9998           380.3650            479.3523      30.0230
8.0 Hours  225.5014           189.8653            267.8260      23.5819


*  ERC10 at SEV1:RT:C
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 3843.2146          3450.7259           4280.3455     251.6993
1.0 Hours 2029.6265          1911.0637           2155.5451      74.2722
2.0 Hours 1071.8589          1013.0276           1134.1068      36.7859
4.0 Hours  566.0556           511.4375            626.5066      34.9186
8.0 Hours  298.9376           255.2595            350.0897      28.7068
```

```
*  ERC10 at SEV1:RT:L
              ERC10 Lower Bound of ERC10 Upper Bound of ERC10 Standard Error
0.5 Hours 3173.5169             2866.2209             3513.7591       196.4977
1.0 Hours 1675.9548             1579.5678             1778.2236        60.3518
2.0 Hours  885.0826              829.6099              944.2646        34.8282
4.0 Hours  467.4179              418.1326              522.5124        31.6636
8.0 Hours  246.8464              208.6956              291.9714        25.1955



Note:
1. These One-sided 95% lower bound and one-sided 95% upper bound
confidence intervals are equivalent to
the lower bound and upper bound of two-sided 90% confidence intervals.

***********************************************************************


***********************Automated Test(s)***********************

**********************Test Equality of Intercepts*********************


Hypothesis Test Results
-----------------------

Coefficients tested for removal:  None

Coefficients tested for equality:

   Group 1 :   SEV1 SEV2

Test statistics:

   Chisquare    df    p-value
    261.1419     1      1e-05

##############################################################################

The P Value of the equality test is <= 0.05. This is generally considered
significant, indicating that all the tested parameters should be retained
in the model.

##############################################################################


***********************Test Removal of Time********************


Hypothesis Test Results
-----------------------

Coefficients tested for removal:  LG10TIME

Coefficients tested for equality:  None

Test statistics:

   Chisquare    df    p-value
    217.1229     1      1e-05
```

##

The P Value of the removal test is <= 0.05. This is generally considered
significant, indicating that the tested parameters should not be removed
from the model.

##

************************Test Removal of Dose********************

Hypothesis Test Results

Coefficients tested for removal: LG10CONC

Coefficients tested for equality: None

Test statistics:

    Chisquare   df   p-value
     232.8554   1   1e-05

##

The P Value of the removal test is <= 0.05. This is generally considered
significant, indicating that the tested parameters should not be removed
from the model.

##

10.0 MENUS, TOOL BARS, & STATUS BARS

At all times, CatReg displays a menu bar at the top of the CatReg application window, and a status bar at the bottom of the window.

This section describes the different options provided by the CatReg menus and status bars.

10.1 File Menu

Command	Function
New Analysis	Opens a new Analysis Screen window.
Open Analysis	Displays an Open dialog box, where you can navigate to and select a previously saved analysis file (*.anx).
Exit	Closes the CatReg application. CatReg will prompt you to save any unsaved changes.

10.2 Preferences Menu

Command	Function
Jitter	(Plots tab) Adds small random offsets to coincident observations to produce a plot that better shows density of the data.

10.3 Help Menu

Command	Function
About	Window describing the CatReg Sponsors and Credits, CatReg program version.
User Guide	Displays the CatReg User Guide PDF, located in the CatReg program directory
Check for Updates…	Displays Update Preferences dialog for CatReg to periodically check for new versions or for the user to manually check for updates.

10.4 Text Window (Results) Menus

10.4.1 File

Command	Function
Save	Saves changes to the selected file.
Save As...	Displays the Windows Save As dialog box so the file can be saved under a new filename.
Print	Prints the contents of the selected window.
Print Setup	Displays the Windows Print Setup dialog box, where you can select such options as orientation, margins, paper size, and so on.
Print Preview	Displays a window showing how the file will look when printed.
Close	Closes the selected window. CatReg will prompt you to save any unsaved changes.

10.4.2 Edit

Command	Function
Undo (Ctrl+Z)	Undo the most recent change.
Cut (Ctrl+X)	Selected data is cut from an active output file.
Copy (Ctrl+C)	Selected data is copied to the clipboard.
Paste (Ctrl+V)	Cut/copied data is pasted into output file at cursor location.

Command	Function
Select All (Ctrl+A)	Selects all text in current active window.

10.4.3 Preferences

Command	Function
Word Wrap	Toggle word wrap.
Font	Display font selection dialog.

10.5   Dataset Context Menu

With data loaded in the Analysis Screen, right-click to display the following options.

Command	Function
Load Dataset…	Displays an Open dialog box, where you can navigate to and select an input file (*.csv).
Regroup Dataset…	Displays the Grouping dialog box.
Save Dataset As…	Displays a Save dialog box, where you can choose to overwrite an existing input file or save the displayed data with a new name.

10.6   Status Bar

Each window in CatReg has its own status bar and communicates different information.

Status Bar Location	Functions
CatReg application window	Displays the results of actions executed within session and data windows, such as when rows or columns are inserted, a session is saved, or when a parameter options file is opened for editing.
Analysis Screen	Displays such messages as "Right-click on Dataset control for additional option(s). or "Done processing!"

11.0   KEEPING CATREG UP TO DATE

After you start CatReg, and assuming your computer has an Internet connection, the software will check the CatReg server for any updates. You can change the frequency of this check by selecting **Help>Check for Updates**.

Figure 25.—Update Preferences dialog box.
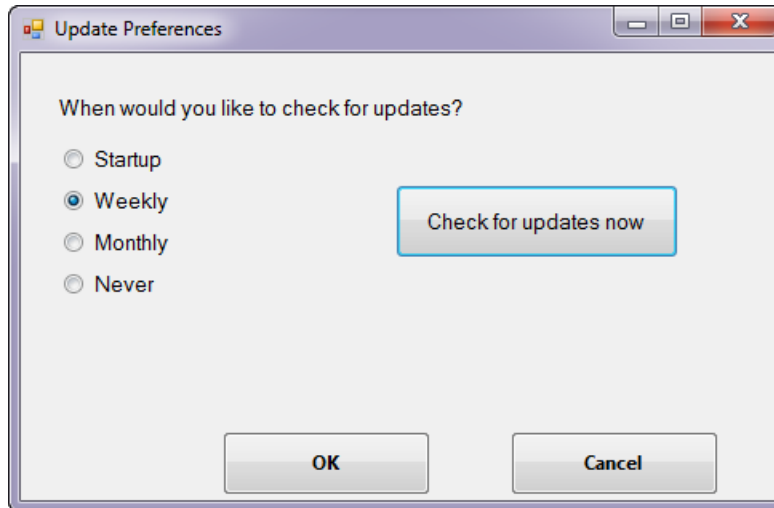
From this dialog box, you can select whether CatReg checks for a new update:

- On startup
- Weekly (the default)
- Monthly

Selecting Never means you will have to manually check for updates by selecting the **Check for updates now** button.

Note that CatReg *must be running* so it can check for and notify you of an update.
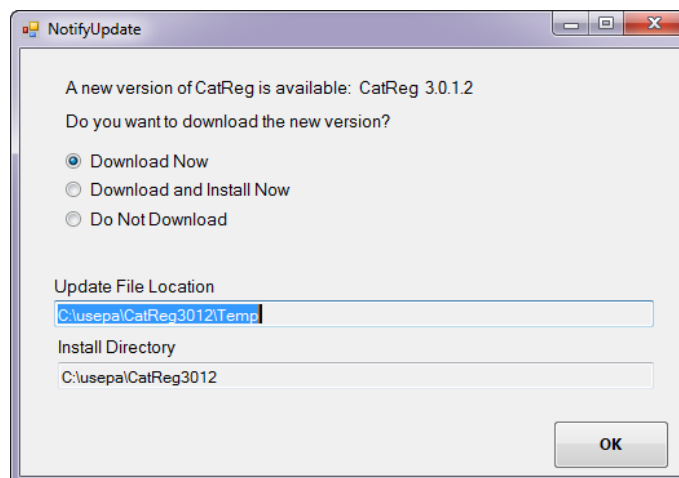
11.1   When an Update is Ready for Download

When CatReg detects that an update is ready to download, it will display an update notification dialog.

Figure 26.—Update Notification dialog box.

You can choose to:

- Download the update and install it yourself later (the default)

- Have CatReg download and automatically install the update. The dialog shows the Update File Location where CatReg will download the new update file and the Install Directory where the new CatReg directory will be created if automatically installed. These fields are read-only and cannot be edited; CatReg enforces this information to ensure a stable installation process.

- Not download the update

If you choose to download and install the software yourself, you can follow the instructions described in the CatReg readme file included with the install package.

If you choose to have CatReg automatically download and install the update, then CatReg displays an information box after installation confirming the update.

Note that this automatic install process is done under your user account and does not require a system administrator.

11.2   To Use the Newly Installed Software

Following an automatic installation, the older CatReg version does not shut down; it remains open.

To use the newly installed CatReg software:

1. Close the old CatReg version.

2. Review the readme file to learn about new features, bug fixes, or other need-to-know items regarding the new software.

3. Follow the instructions in the readme file to:

    - Create a shortcut for the new CatReg version on your desktop

    - Determine that the new CatReg version works properly on your computer

    - Fix possible issues with displaying the help file

11.3   Managing Your Existing Files After an Automatic Update

The automatic install does not replace or displace your existing data and results files. Your old version of CatReg and related input filesw remain in their current locations. Input files do not have to be moved from their current folder locations to work with the new version of CatReg. Follow your organization's guidelines for locating and maintaining data and files.

12.0   REFERENCES

Akaike, M. (1974) A new look at statistical model identification. IEEE Transactions on Automatic Control AU-19: 716-722.

Diggle, P. J.; Liang, K.-Y.; Zeger, S. L. (1994) Analysis of longitudinal data. New York, NY: Clarendon Press.

Huber, P. J. (1967) The behavior of maximum likelihood estimates under nonstandard conditions. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Volume 1. Berkeley, CA: University of California Press.

Liang, K.-Y.; Zeger, S. L. (1986) Longitudinal data analysis using generalized linear models. Biometrika 73: 13-22.

McCullagh, P.; Nelder, J. A. (1989) Generalized linear models. 2nd ed. London, United Kingdom: Chapman and Hall.

Morgan, B. J. T. (1992) Analysis of quantal response data. London, United Kingdom: Chapman & Hall.

Simpson, D. G.; Carroll, R. J.; Xie, M.; Guth, D. J. (1996a) Weighted logistic regression and robust analysis of diverse toxicology data. Commun. Stat. 25: 2615-2632.

Simpson, D. G.; Carroll, R. J.; Zhou, H.; Guth, D. J. (1996b) Interval censoring and marginal analysis in ordinal regression. J. Agric. Biol. Environ. Stat. 1: 354-376.

U.S. Environmental Protection Agency. (2000) CatReg software documentation. Research Triangle Park, NC: Office of Research and Development, National Center for Environmental Assessment; report no. EPA/600/R-98/053.

U.S. Environmental Protection Agency. (2006) CatReg Software User Manual (R-Version). Research Triangle Park, NC: Office of Research and Development, National Center for Environmental Assessment; report no. EPA/600/R-04/006. Downloadable from the EPA Web page, "CATREG SOFTWARE FOR CATEGORICAL REGRESSION ANALYSIS."

Venables, W. N.; Ripley, B. D. (1994) Modern applied statistics with S-Plus®. New York, NY: Springer-Verlag.

Weisberg, S. (1985) Applied linear regression. 2nd ed. New York, NY: Wiley.

13.0   GLOSSARY

Akaike Information Criteria (AIC) - The deviance (-2 times the log of the maximized value of the likelihood function) + 2 times the number of parameters in the model.

Categorical regression - A model expressing the probabilities of different response categories as functions of explanatory variables.

Cluster sample - A data set comprised of subsamples of data from common sources. For example, a data set may contain several data records per laboratory from several different laboratories. The subgroup of data records from each individual laboratory would represent a cluster sample.

Cumulative odds regression - Ordinal regression model for directly modeling the probabilities of exceeding different severity levels.

Deviance - The minimized value of twice the negative logarithm of the likelihood function.

Extra risk concentration (ERC) – Concentration or dose at which extra risk is a user-specified value.

ERC settings - Three numbers: ERC percentile, ERC severity level, and ERC percentile for confidence intervals.

Filtering - Exclusion of selected data records from the analysis. This capability "filters out" selected data without altering the data input file.

Generalized estimating equation - An equation depending on the data and the parameter values, such that solving for the parameter values yields consistent estimates.

Hierarchical models - An ordered series of models, such that each model is a special case of the next one in the series.

Inf - Infinite value

Interval censored data - Data for which the response is known only to lie in an interval of values.

Likelihood function - For categorical response data, a model for the joint probability of the observed data values, expressed as a function of the model parameters.

Link function - A function applied to the categorical response probability to transform the categorical regression model to linear units.

Meta-analysis - The analysis of data from multiple studies to determine overall trends and increase power.

NaN - Not a number

Ordinal data - Data reported as ordered categories. The order is meaningful, but the numerical difference between ordered categories is not.

Parallelism - The coefficients of dose and time in the exposure-response model (i.e., the probability function) do not change with severity level (parallelism applies to Model 1, the cumulative odds model, but not to Model 2, the unrestricted cumulative model).

Probability function - The function of the explanatory variables that gives the probability of exceeding a given severity level.

Proportional odds regression - Ordinal regression in which the log-odds of exceeding different severity levels are parallel across severity categories. It is a special case of cumulative odds regression with the logistic link function.

Stratification - To create subsets of data by allowing the model parameters to vary by subset. Covariate information such as species, sex, and target organ may be used as a basis for creating the sub sets.

APPENDIX A: DISTRIBUTION OF CONTINUOUS RESPONSE DATA OVER SEVERITY LEVELS

Response data are sometimes measured on a continuous scale, with the mean and standard error reported. Section 5.11.3 contains a hypothetical example for lung weight in rats in which a mean and standard error were reported for each treatment group. The lung weight data were distributed over the severity levels shown for "Ref.id" = 2 in Table 7 for analysis by CatReg. The following description explains how that was done and provides an example calculation for reference. A technical explanation is provided at the end of the appendix.

The distribution of lung weights in healthy, unexposed rats is needed, either estimated from control animals in the experiment or "known" from other sources. It is assumed here that the distribution of lung weights is normally distributed with mean 1.0 g and standard deviation 0.05 g. (Note: The normal distribution is assumed simply for illustration. The same idea could be applied to other distributions). The user then determines weight intervals for severity levels to be used. For this purpose, it may be helpful to estimate first the highest weight that might be considered in the "normal" range for unexposed animals. The weight 1.15 g, which is three standard deviations above the mean, is an upper bound on virtually all lung weights in unexposed animals (i.e., a weight above 1.15 g is above the normal range of lung weights). The following correspondence was made between severity levels and lung weights for the example: Sev 0 (<1.15 g), Sev1 (1.15 to 1.50 g), and Sev2 (>1.50 g).

Suppose that the average lung weights and standard errors shown in Table A-1 were reported for treatment groups in "Ref.id" = 2, Table 7.

> **Note:** If SE is the standard error from a treatment group of size n, the estimate of the standard deviation is SE.

There is a separate mean and standard error reported for each treatment group, denoted by , and SEi, respectively for index i, with i = 1,...,13. The ith treatment group is assumed to be a sample from a normal distribution with unknown mean and standard deviation, denoted by μi and σi, respectively. The treatment groups are rather small (ni = 10 for all i), so it was assumed that the standard deviation was the same for treatment groups with similar estimates of the standard deviation (i.e., $SE_i\sqrt{n_i}$). The standard deviation was assumed equal for indices 1, 2, 6, and 10 (to be called Group A), 3, 7, and 11 (Group B), 4, 8, and 12 (Group C), and 5, 9, and 13 (Group D) in Table A-1. An estimate of the common standard deviation in Group A, σA, is calculated from SE_1, SE_2, SE_6, and SE_{10} as follows (estimation of the standard deviations for the other three groups is similar).

Let σ_A^2 be the common variance for Group A. The variance estimate from Sample 1 is $S_1^2 = n_1 SE_1^2 = 10\ (0.03)^2$. The estimate of σ_A^2, denoted by S_A^2 is the sum of the estimates of σ_A^2 from samples 1, 2, 6, and 10, weighted by their degrees of freedom (df) (9 for each sample). This gives $S_A^2 = 1/36\ [9(10)(0.03)^2 + 9(10)(0.025)^2 + 9(10)(0.025)^2 + 9(10)(0.025)^2] = 0.00694$. The proportion of Sample 1 with lung weights less than 1.15 g is estimated as follows. If X is a new observation for Sample 1, then

$$T = (X - \bar{X}_1)/\left(\frac{S_A^2(n_1 + 1)}{n_1}\right)^{0.5}$$

has a t-distribution with nA – 4 df, where n1 is the number of observations in Sample 1 (i.e., n1 = 10), and nA is the number of observations in Group A (i.e., nA = 40). Pr (X < 1.15 g) is estimated by Pr (T < (1.15 − 1.1)/[0.00694 (11/10)0.5] = 0.715, shown in Table A-2 under Sev0 for "Index" = 1. To estimate Pr (1.15 ≤ X < 1.5), first estimate Pr (X < 1.5) as above, except with 1.15 g replaced by 1.5 g, and then subtract the estimate of Pr(X < 1.15). Similarly, the relationship Pr(X ≥ 1.5 = 1 – Pr(X < 1.5) is used to estimate Pr (X ≥1.5). The estimated proportions of each sample with lung weights in the intervals <1.15, 1.15 to 1.5, and ≥1.5 are displayed in Table A-2.

Table A-1: *Summary Data for "Ref.Id" = 2 in Table 7*

Index	Exp.	Group	mg/m3	Hours	Average Lung Weight ($\overline{X}$)	Standard Error
1	1	1	330	2	1.1	0.03
2	1	2	360	2	1.2	0.025
3	1	3	390	2	1.5	0.04
4	1	4	410	2	1.8	0.08
5	1	5	460	2	1.7	0.1
6	2	1	460	1	1.2	0.025
7	2	2	510	1	1.3	0.04
8	2	3	560	1	1.6	0.08
9	2	4	610	1	1.8	0.1
10	3	1	560	0.5	1.3	0.025
11	3	2	610	0.5	1.5	0.04
12	3	3	660	0.5	1.6	0.08
13	3	4	710	0.5	1.8	0.1

Table A-2: *Estimated Proportions in Severity Levels*

Group	Index	$S\sqrt{\dfrac{n+1}{n}}$	$\overline{X}$	$\dfrac{1.15-\overline{X}}{S\sqrt{\dfrac{n+1}{n}}}$	df	Sev0 Pr (X < 1.15)	$\dfrac{1.5-\overline{X}}{S\sqrt{\dfrac{n+1}{n}}}$	Sev1 Pr(1.15 < X < 1.5)	Sev2 Pr(X > 1.5)
A	1	0.0874	1.1	0.572	36	0.715	4.577	0.285	0
A	2	0.0874	1.2	−0.572	36	0.285	3.432	0.714	0
B	3	0.1327	1.5	−2.638	27	0.007	0	0.493	0.500
C	4	0.2653	1.8	−2.450	27	0.011	−1.131	0.123	0.866
D	5	0.3317	1.7	−1.1658	27	0.127	−0.603	0.149	0.724
A	6	0.0874	1.2	−0.572	36	0.285	3.432	0.714	0.001
B	7	0.1327	1.3	−1.130	27	0.134	1.507	0.794	0.072
C	8	0.2653	1.6	−4.146	27	0	−0.377	0.354	0.645
D	9	0.3317	1.8	−1.960	27	0.030	−0.904	0.157	0.813
A	10	0.0874	1.3	−1.716	36	0.047	2.288	0.939	0.014
B	11	0.1327	1.5	−2.638	27	0.007	0	0.493	0.5
C	12	0.2653	1.6	−4.146	27	0	−0.377	0.354	0.645
D	13	0.3317	1.8	−1.960	27	0.030	−0.904	0.157	0.813

APPENDIX B: TECHNICAL DISCUSSION

B.1   Link Functions

Without loss of generality, link functions will be discussed in the context of Model 1, described by Eq. 1a in , specifically,

$$\Pr(Y \geq s|C,T) = H[\alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$$

CatReg currently supports three forms for H:

Logistic	$H(x) = e^x / (1 + e^x)$,

Normal $H(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-z^2}{2}}, dz$ and

Gumbel $H(x) = 1 - \exp(-e^x)$.

The inverse of H, which is denoted by L, is called the *link* function in the statistical literature. The link functions corresponding to the probability functions given above are

Logit $L(p) = \log[p / (1-p)]$,

Probit $L(p) = 100\ p$th percentile of normal (0,1), and

Cloglog $L(p) = \log[-\log(1 - p)]$,

where *p* is any number between 0 and 1. The link function and probability function are inverse to each other in the sense that $H[L(p)] = p$ and $L[H(x)] = x$. Applying the link function to both sides of Model 1 gives

$$L[\Pr(Y \geq s|C,T)] = \alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T),$$

$$s = 1,2,...S.$$

This expression shows that the link-transformed probability follows a linear model.

The use of a link function is essential here. Without it, the linear model becomes unbounded, and one is led to absurd estimates of probabilities for extreme values of *C* and *T*, namely negative probabilities or probabilities greater than 1. Moreover, link functions may be derived from a basic assumption that the ordinal severity score corresponds to exceeding an underlying toxic response threshold (see discussion below). Under this approach, the ordinal response score is called a quantal response because it is a quantization of the underlying response (see, e.g., Morgan, 1992). Any ordinal regression model of the type given in Model 1 may be interpreted as a quantal response model. When β_1 and β_2 are constant across severity categories data from one severity category adds information to the modeling of another category.

A toxicological interpretation of the link function follows from quantal response analysis. In particular, let *Z* denote a particular measure of health for a randomly selected subject given exposure to concentration or dose *C* for time *T*. Larger values of *Z* correspond to a healthier individual. Suppose that the health variable *Z* is distributed in the population as

$$\Pr(Z \leq x) = H(\frac{x - \mu}{\sigma})$$

where *x* is a possible value for an individual's health, μ determines the median health for the given level of exposure, and σ is a measure of the population variation in health. In particular, if *H* is Gaussian, then the health levels in the population are distributed normally with mean μ and standard deviation σ. The larger μ is, the healthier the population is at the given level of exposure.

To model the situation where health deteriorates with increasing exposure, suppose that

$$\frac{\mu}{\sigma} = -\beta_1 * f_1(C) - \beta_2 * f_2(T).$$

Further, suppose that a toxic reaction of severity $Y \geq s$ occurs if the health measure Z is below a threshold $\sigma\alpha_s$, where σ is specific to the health measure. Then, under exposure (C, T), the probability of toxic severity of category s or higher is

$$\Pr(Y \leq s) = \Pr(Z \leq \sigma\alpha_s) = H[\alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$$

This is precisely the ordinal regression model given in Model 1. It is apparent that the link function is a reflection of the underlying distribution of Z, which cannot be measured directly, but its distribution, in particular the dependence on C and T, can be estimated indirectly from the toxicological response data.

B.2   Interval Censoring

Although censored responses do not provide as much information as fully scored responses, they do provide some information about the model. This information is used in the maximum likelihood estimation and the generalized maximum likelihood estimation described in Section B.3. In fitting the model by maximum likelihood, it is necessary to compute the probability of the observed response as a function of the model parameters. Table B-1 shows how these probabilities are computed in a three-category scoring system with interval censoring, using Model 1.

Table B-1: *Interval Probabilities for Model 1 with Three Severity Categories*

Interval	Probability
(0,0)	$1 - H[\alpha_1 + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$
(0,1)	$1 - H[\alpha_2 + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$
(1,1)	$H[\alpha_1 + \beta_1 * f_1(C) + \beta_2 * f_2(T)] - H[\alpha_2 + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$
(1,2)	$H[\alpha_1 + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$
(2,2)	$H[\alpha_2 + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$

B.3   Parameter Estimation

B.3.1   Maximum Likelihood Estimation

The likelihood function is defined to be the joint probability density of the data, viewed as a function of the parameters. In categorical regression, the response variables are discrete, so the likelihood may be interpreted as the probability that an investigation would result in the particular values that were observed. This probability depends on the unknown parameters. Maximum likelihood estimates the unknown parameters by the values that maximize the likelihood of the observed data. It is often more convenient to work with the logarithm of the likelihood. For this purpose, it is common to define the deviance function:

Deviance = −2 * log(likelihood).

The deviance is a nonnegative measure of model fit. Maximizing the likelihood is mathematically equivalent to minimizing the deviance. The factor 2 is included because it is the correct multiplier for certain likelihood-based, goodness-of-fit tests. Smaller deviances (larger likelihoods) correspond to a closer fit of the model to the data. A deviance of 0 would indicate a perfect fit, that is, a "saturated" model. Generally, a deviance of zero would indicate a model that is too complicated. The deviance value shown in CatReg summary statistics is the "residual deviance" discussed in Section 9.0: "ASSESSING MODEL FIT."

If all data are independent, the likelihood function for interval-censored ordinal regression has a simple form. For $i = 1,2,..., n$, let Y_i denote the ordinal response, and let C_i and T_i denote the dose and time of exposure for the ith experimental subject, respectively. Y_i may be known only to lie in an interval. To account for this, let L_i and U_i denote the lower and upper endpoints of the known interval for Y_i, respectively. If it is known that $Y_i = k$, then $L_i = U_i = k$. For convenient reference, denote the probability of severity s or greater by

$$P_i(s) = \begin{cases} 1, \\ H[\alpha_s + \beta_1 f_1(C) + \beta_2 f_2(T_i)], \\ 0, \end{cases}$$

$$\begin{array}{l} if\ s = 0; \\ if\ s - 1, ..., S; \\ if\ s > S. \end{array}$$

Then the deviance for interval-censored ordinal regression is given by

$$\text{Deviance} = -2 \sum_{i=1}^{N} \log[\Pr(L_i \leq Y_i \leq U_i | C_i, T_i)]$$

$$-2 \sum_{i=1}^{N} \log[P_i(L_i) - P_i(U_i + 1)]$$

Parameter estimates are computed by iteratively minimizing the deviance. CatReg uses the R function `optim()` to perform this optimization.

B.3.2   Generalized Likelihood Estimation

Weighted ordinal regression analysis corresponds to a modified likelihood in which the probability associated with the ith observation is raised to a positive power w_i. This results in a modified likelihood with a weighted deviance:

$$\text{Deviance} = -2 \sum_{i=1}^{N} w_i \log[P_i(L_i) - P_i(U_i + 1)]$$

If the weights do not correspond to incidences, then this likelihood corresponds to a nonstandard ordinal regression model. In this situation, it is more common to interpret the deviance as a generalized criterion and to assume that the usual ordinal regression model holds. Under this assumption, the generalized deviance still leads to consistent estimates of the parameters, but it does not correspond to the likelihood of the data. Instead, the estimator is defined by a generalized estimating equation, which provides the basis for computing valid large-sample confidence intervals and test statistics.

A further modification arises when the data are cluster sampled. The likelihood for cluster-sampled data does not have the simple form given above. Rather, it involves a product of multiple integrals of conditional likelihoods. Such likelihoods are computationally challenging and the results may be sensitive to the specification of the correlation structure. An alternative approach is to assume the ordinal regression model holds in a population-average sense. Consistent estimates then may be obtained quite generally, without making extensive distributional assumptions about the correlation structure. To achieve this, CatReg takes the expression derived above as a "working deviance" criterion. Minimizing it leads to consistent estimates under the population-average model. The main impact on the analysis compared to a standard likelihood analysis is the use of generalized estimating equation methods for making statistical inferences. In particular, rather than reporting the inverse information matrix as the estimated parameter covariance, the well-known sandwich formula is used.

Most applications of CatReg will involve cluster-sampled data and possible weighting of observations. As noted above, CatReg uses the weighted independence criterion as an estimating criterion, but computes confidence intervals and hypothesis tests without assuming the criterion is the likelihood. This approach has a long history in the statistical literature. Huber (1967) derived the large-sample theory of "maximum likelihood" estimators when the working likelihood is different from the actual likelihood of the data. In the literature on robust statistics, this type of estimator is called an "M-estimator" because it generalizes maximum likelihood. Liang and Zeger (1986) extended the method to the analysis of correlated data, based on a "working" correlation structure, without assuming the working correlation structure was correct. This general approach is widely used in biometry, econometrics, and survey sampling.

B.4   Confidence Limit Calculations

CatReg uses the method of generalized estimating equations, which is well accepted in the literature (see Diggle et al., 1994), for the calculation of confidence limits for cluster-sampled data. The classical likelihood ratio inferences do not apply to cluster-correlated data because the likelihood ratio test assumes independent responses. The application of generalized likelihood ratio tests for correlated data is, however, an area of active research that may produce usable results in the future.

Confidence intervals and hypothesis tests about the parameters rely on a large-sample normal approximation to the joint distribution of the parameter estimates. The main steps in the derivation of this approximation are as follows. First, assuming the data are cluster sampled, write the generalized deviance as

$$GD = -2 \sum_{i=1}^{N} \sum_{j=1}^{n_i} w_{ij} \log[P_{ij}(L_{ij}) - P_{ij}(U_{ij} + 1)],$$

where $-2\, w_{ij}\log[P_{ij}(L_{ij}) - P_{ij}(U_{ij} + 1)]$ is the contribution of the jth individual from cluster i, $j = 1, \dots, n_i, i = 1, \dots, N$. Let B denote the vector of all model parameters. The GD estimate of B solves a generalized estimating equation

$$0 = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \Psi_{ij},$$

where the summand Ψ_{ij} is given by

$$\Psi_{ij} = w_{ij} \frac{dP_{ij}(L_{ij}) - dP_{ij}(U_{ij} + 1)}{P_{ij}(L_{ij}) - P_{ij}(U_{ij} + 1)},$$

and $dP_{ij}(t)$ denotes the vector of derivatives of $P_{ij}(t)$, with respect to the components of B, evaluated at the estimated parameters. Expanding the estimating equation in a Taylor series leads to a large-sample normal approximation. The estimated parameter vector, $\hat{B}$, is approximately multivariate normal. The mean of the approximating normal distribution equals the true value of B, and the covariance matrix is given by the sandwich formula

$$\text{Est.Cov}(\hat{B}) = J^{-1}CJ^{-1}$$

where J is given by

$$J = -\sum_{i=1}^{N} \sum_{j=1}^{n_i} d\Psi_{ij},$$

and C is a covariance estimate for the total score, given by

$$C = \sum_{i=1}^{N} \left(\sum_{j=1}^{n_i} \Psi_{ij} \right) \left(\sum_{k-1}^{n_i} \Psi_{ik} \right)'$$

If the working likelihood were the actual likelihood of the data, then J and C would estimate the same matrices, and the usual inverse information, J^1, would estimate the covariance of $\hat{B}$. Further details are given in Simpson et al. (1996a).

Standard errors of individual parameter estimates are obtained as the square roots of the diagonal elements of the estimated covariance. Confidence intervals and hypothesis tests derive from the normal approximation for $\hat{B}$.

APPENDIX C: TECHNICAL BACKGROUND: MODELS AND EXTRA RISK

C.1   Exposure-response Models

Let Y denote a dependent variable that represents the severity or intensity of the response. Assume Y is an ordinal score taking one of the values (0,1,...,S). A score of 0 corresponds to the lowest severity (e.g., no adverse effect), and a score of S corresponds to the highest severity (e.g., lethal, in a toxicological context). Categorical regression is a method for modeling the probability distribution of Y as a function of the explanatory variables, concentration or dose (C) and time (T). It employs a generalized linear model (McCullagh and Nelder, 1989) for the dependence of the probabilities of different severity categories on the explanatory variables.

CatReg provides a choice of two models (with variations on both):

- Model 1 (the cumulative odds model) described by Eq. 1a, and
- Model 2 (the unrestricted cumulative model) described by Eq. 1b.

CatReg refers to any model of the form of Model 1 as a cumulative odds model because the model is expressed in terms of the cumulative probabilities, or odds, for $Y \geq s$. Note that Model 1 is a special case of Model 2 wherein parameters β_{1s} and β_{2s} do not depend on s (which denotes severity level, as discussed below). A primary use of fitting Model 2 is to test whether the simpler Model 1 is adequate.

Model 2, although more general than Model 1, has the undesirable feature that the regression lines for different severity levels may cross. Often the crossing is well outside the range of values of interest, so the model can be used to make empirical risk estimates. The user has the option to add an additional parameter, γ, which represents a hypothetical background concentration or dose, in some circumstances (see Section 6.10: "Assume Zero Background Response").

For s = 1,2,…,S,

Model 1

$$\Pr(Y \geq s | C, T) = H[\alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$$ **Eq. 1a**

Model 2

$$\Pr(Y \geq s | C, T) = H[\alpha_s + \beta_{1s} * f_1(C) + \beta_{2s} * f_2(T)]$$ **Eq. 1b**

The left-hand side is read as follows: the probability that a response of severity level s or greater occurs, given that concentration or dose is C and time is T (time refers to exposure duration). No expression for s = 0 is included because this is the minimal category, and Y is always greater than or equal to 0 (i.e., Pr($Y \geq 0 | C, T$) = 1).

The right-hand side is described as follows:

- H is a probability function taking values between 0 and 1, for which the user has three choices: (1) logistic, (2) normal, and (3) Gumbel (described further in Appendix B).
- The parameter α_s is the intercept for severity level s, s = 2,...,S (to be called the intercept or severity parameters). The severity parameters are ordered as $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_s$. This constraint is a consequence of the requirement that the probability of exceeding a lower score is larger than the probability of exceeding a higher score for any fixed levels of C and T.
- In Model 1, the parameter β_1 determines the dependence of the response on concentration or soe(C), whereas β_2 determines the dependence on time (T). In Model 2, the parameters are also indexed by s because they may change values with severity level s.
- Current choices for f_1 and f_2 are "untransformed" and "base-10 logarithm." Other transformations of C and T may be obtained by transforming the input data.

- Parameters are α_s, β_{1s} (to be called the coefficient of concentration or dose), β_{2s} (to be called the coefficient of time or duration), for severity levels s running from 1 to S. All parameters may be stratified on variables, as discussed under "Stratifying" in Section 6.6.

The normal and logistic distributions are symmetric, each having median equal to zero. The Gumbel distribution is skewed, with a lower tail similar to that of the logistic distribution and a lighter upper tail. Figure 27 displays these three probability distributions. To compare the shapes effectively, the distributions have been rescaled to have medians = 0 and equal 25th percentiles (labeled as "EC25" on the horizontal axis). The scaled logistic and normal distributions are very close for much of the range and differ substantively only in the extreme tails. The Gumbel distribution is skewed, with a heavier tail on the left and a lighter tail on the right.

For each of the three choices of the probability function H there is an inverse function of H, called the link function, that transforms it to a simple linear function in dose and time. CatReg requests the name of the link function instead of the name of the probability function. The corresponding link functions (in parentheses) are: logistic (logit), normal (probit), and Gumbel (cloglog). There is further discussion of linking in Section 6.8 and Appendix B; the latter includes an example of how link functions may be derived from a basic assumption that the ordinal severity score corresponds to exceeding an underlying toxic response threshold.

Figure 27.—Normal, logistic, and Gumbel probability functions shifted and scaled to have equal medians and 25th percentiles.
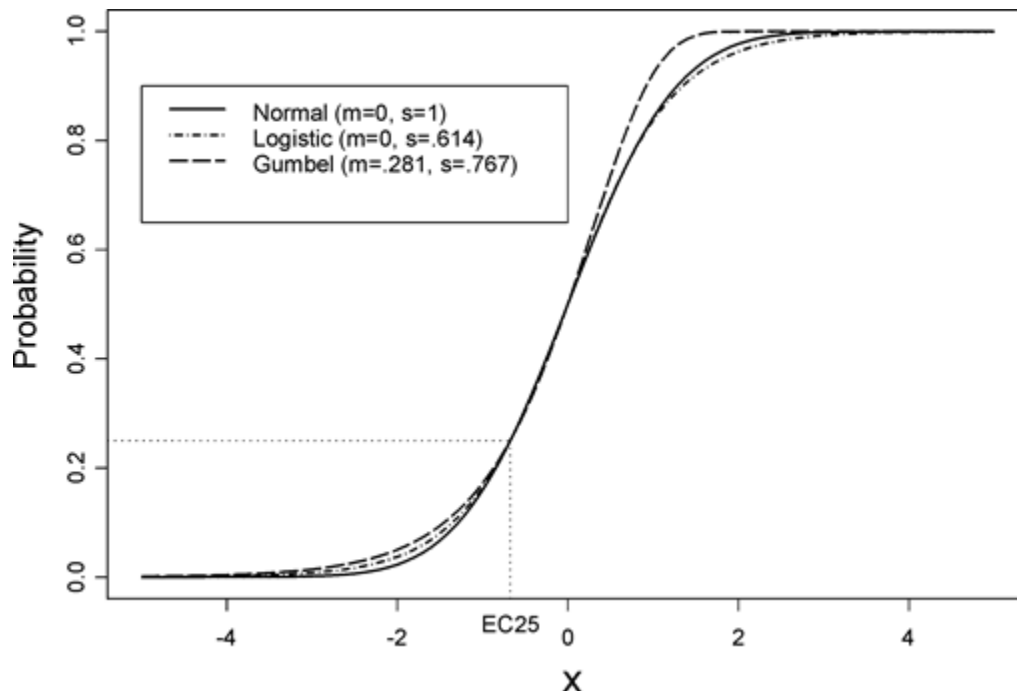
C.2   Extra Risk Concentration (ERC)

Extra risk at concentration or dose C=c and time T=t, at severity level s, is defined as

$$\frac{\Pr(Y \geq s | C = c, T = t) - \Pr(Y \geq s | C=0, T=t)}{1 - \Pr(Y \geq s | C=0, T=t)} \qquad \textbf{Eq. 2}$$

For q between 1 and 100, inclusive, ERCq, at time T = t, is the concentration c for which equation (4-2) equals q/100. For example, ERC10 at T = 2 (exposure time of 2 hours) for severity level 1 is the value of c that satisfies

$$\frac{\Pr(Y \geq 1 | C = c, T = 2) - \Pr(Y \geq 1 | C=0, T=2)}{1 - \Pr(Y \geq 1 | C=0, T=2)} \qquad \textbf{Eq. 3}$$

In words, ERC10 at T = 2 for severity level s is the exposure concentration or dose at which the probability is 0.10 of an adverse effect of level s or higher *due to exposure of two hours*, i.e., given the adverse effect would not have occurred from other causes ("background causes") during that time.