Benchmark Dose Software (BMDS)

VERSION 3.2
USER GUIDE

EPA

Authors and Reviewers

Federal Authors
Jeff Gift, Ph.D.
J Allen Davis, M.S.
Todd Blessinger, Ph.D.
    U.S. EPA
    Office of Research and Development
    Center for Public Health and Environmental Assessment (CPHEA)
    RTP, NC, Cincinnati, OH and Washington, DC

Matthew Wheeler, Ph.D.
    The National Institute for Occupational Safety and Health (NIOSH)

Contract Authors
Louis Olszyk
Cody Simmons, Ph.D.
Yadong Xu, Ph.D.
Bruce Allen, M.S.
Michael E. Brown, M.S.
    General Dynamics Information Technology / ARA
    US EPA, 109 T.W. Alexander Dr.
    Research Triangle Park, NC 27711

Reviewers
Laura Carlson, Ph.D.
    EPA Center for Public Health and Environmental Assessment

Stephen Gilbert, Ph.D.
    The National Institute for Occupational Safety and Health (NIOSH)

Contents

Table of Tables

Table of Figures

Abbreviations

AIC	Akaike's Information Criterion
BIC	Bayesian Information Criterion
BMD	benchmark dose
BMDL	benchmark dose lower confidence limit
BMDS	Benchmark Dose Software
BMDU	benchmark dose upper confidence limit
BMR	benchmark response
CI	confidence interval
CPHEA	Center for Public Health and Environmental Assessment
CV	coefficient of variation
EPA	Environmental Protection Agency
IRIS	Integrated Risk Information System
LOAEL	lowest-observed-adverse-effect level
LPP	Log Posterior Probability
LSC	Litter specific covariate
MLE	maximum-likelihood estimation
NCEA	National Center for Environmental Assessment
NCTR	National Center for Toxicological Research
NIEHS	National Institute of Environmental Health Sciences
NIOSH	National Institute for Occupational Safety and Health
NOAEL	no-observed-adverse-effect level
POD	point of departure
RfC	inhalation reference concentration
RfD	oral reference dose
SD	standard deviation
SE	standard error
SEM	standard error of the mean
US EPA	United States Environmental Protection Agency

1.0   Overview

The U.S. Environmental Protection Agency (EPA) Benchmark Dose Software (BMDS) was developed as a tool to facilitate the application of benchmark dose (BMD) methods to EPA hazardous pollutant risk assessments. This user guide provides instruction on how to use the BMDS but is not intended to address or replace EPA BMD guidance. However, every attempt has been made to make this software consistent with EPA guidance, including the Risk Assessment Forum (RAF) Benchmark Dose Technical Guidance Document. (U.S. EPA, 2012).

1.1   What is BMDS

BMDS is an application that facilitates dose-response modeling. BMDS models are currently accessible via an Excel-based user interface.

1.1.1   Dose-Response Modeling

Dose-response modeling is a technique, often used in toxicology and risk assessment, for quantitatively relating exposure (the dose) to a biological outcome (the response). It may be thought of as an elaborate form of regression, which is the statistical technique used to explore or represent the relationship(s) between two (or more) variables. In the dose-response context, the dose term (e.g., mg of chemical per kg body weight per day) is most often viewed as the "cause" of the response (e.g., presence of a tumor or other manifestation of disease or a measure of the weight of some organ that might be susceptible to the toxic effects of the exposure).

BMDS collects together, and provides easy access to, numerous dose-response models that the user may select from and/or compare, to make predictions about the quantitative relationship between dose and response. One specific focus of this software is the estimation of a benchmark dose (BMD), including bounds (e.g., 95% confidence intervals) on such estimates.

The BMD is a dose estimated to produce a response level of a defined (benchmark) magnitude. The online BMDS Glossary defines BMD as follows:

An exposure due to a dose of a substance associated with a specified low incidence of risk, generally in the range of 1% to 10%, of a health effect; or the dose associated with a specified measure or change of a biological effect.

1.1.2   Types of Responses Modeled

One other key aspect of dose-response modeling is that the models, statistical assumptions, and techniques that it uses depend on the type of response under consideration.

For BMDS, as reflected many times over in this user guide, the distinctions that are made, i.e., for which separate and distinct modeling approaches are applied, can be categorized with respect to the following three types of response: continuous endpoints, dichotomous endpoints, and nested dichotomous endpoints.

The key features of those response (endpoint) types are as follows:

- <u>Continuous Endpoint</u>: the response is measured on a continuous scale, so its valid values are real numbers (often restricted to positive values, but not always). Organ weight, body weight, concentration levels of biological markers — these are all examples of continuous endpoints whose values might be affected by exposure to the compound under consideration.
- <u>Dichotomous Endpoint</u>: the response here is the presence or absence of a disease state or other "counter" of system malfunction. In this case, each experimental unit will either have the response or not. Dichotomous dose-response models predict the proportion of such units that have the response. Tumor responses are a subcategory of dichotomous endpoints with cancer-specific model selection considerations that are automated in the BMDS Multi-tumor model; for more information, see Section 10.0, "Multiple Tumor Analysis," on page 105.
- <u>Nested Dichotomous Endpoint</u>: as for Dichotomous endpoints, the presence or absence of a disease state or counter is the response. However, in this case each experimental unit may have more than one such count; i.e., the counts are nested within an experimental unit. The most common example of such an endpoint is with developmental toxicity experiments in which the experimental unit is the pregnant dam and the fetuses or offspring from each dam's litter are examined for the presence or absence of an effect (e.g., malformation).

Collectively, the application of the methods for fitting mathematical models to data is referred to as BMD modeling or the BMD approach. BMDS facilitates these operations by providing simple data-management tools and an easy-to-use interface to run multiple models on one or more dose-response datasets.

1.1.3 Presentation of Model Results

Model results are presented as <u>textual and graphical outputs</u> that can be printed or saved and incorporated into other documents. Results from all models include:

- Model-run options chosen by the user
- Goodness-of-fit information
- BMD
- Estimates of the bounds (e.g., confidence limits) on the BMD (notated BMDL and BMDU for the lower bound and upper bound, respectively).

1.2   How EPA Uses BMD Methods

EPA uses BMD methods to derive risk estimates such as reference doses (RfDs), reference concentrations (RfCs), and slope factors, which are used along with other scientific information to set standards for human health effects.

Prior to the availability of tools such as BMDS, noncancer risk assessment benchmarks such as RfDs and RfCs were determined from no-observed-adverse-effect levels (NOAELs), which represent the highest experimental dose for which no adverse health effects have been documented.

However, using the NOAEL in determining RfDs and RfCs has long been recognized as having limitations:

- It is limited to one of the doses in the study and is dependent on study design
- It does not account for variability in the estimate of the dose-response
- It does not account for the slope of the dose-response curve

- It cannot be applied when there is no NOAEL, except through the application of an uncertainty factor (Kimmel and Gaylor, 1988; Crump, 1984) .

A goal of the BMD approach is to define a starting point of departure (POD) for the computation of a reference value (RfD or RfC) or slope factor that is more independent of study design. The EPA Risk Assessment Forum has published technical guidance for the application of the BMD approach in cancer and non-cancer dose-response assessments (U.S. EPA, 2012).

1.3   History of BMDS Development

Research into model development for BMDS started in 1995 and the first BMDS prototype was internally reviewed by EPA in 1997. After external and public reviews in 1998-1999, and extensive Quality Assurance testing in 1999-2000, the first public version of BMDS, version 1.2, was released in April 2000.

The BMDS release history is documented in two places on the BMDS website:
- All releases from 1995-2017, up through BMDS 2.7
- All BMDS 3.x releases

1.4   What's New in BMDS 3.2

As of version 3.2, Bayesian models have been added for continuous response endpoints. **Please note, however, that BMDS 3.2 represents a *preview*-version of Bayesian continuous endpoint modeling.**

The preview Bayesian continuous models have not been formally reviewed and approved by the EPA for risk assessment purposes. Such models are not recommended for use in EPA risk assessments at this time. EPA welcomes feedback on these preview models.

Users can run individual continuous Bayesian model runs as Normal, Lognormal, or combined Normal and Lognormal.

Peer review of the Bayesian continuous models is planned for later in 2020.

Note   At this time, EPA does not offer technical guidance on Bayesian modeling.

Check the BMDS website and join the BMDS mailing list to receive updates on the availability of new guidance and finalized models.

Note   These preview models are new. Users acknowledge they have not been extensively tested, and formally reviewed and approved by the EPA for risk assessment purposes.

BMDS 3.2 also includes the following enhancements and fixes:
- BMDS graphs now extend the plots to zero dose for all models. All plots will begin at dose=0, even if the data has a lowest dose greater than 0.
- Corrected the equation used to plot the Dichotomous Hill dose-response curve.
- Upon request through BMDS's e-Ticket support site, users can obtain a digitally signed version of BMDS 3.2 to show that BMDS Excel Macros are from a trusted developer. This should meet most organizations' security requirements for VBA-based applications. Note that this signed version of BMDS is not being distributed broadly because EPA cannot guarantee that it will work on all Excel platforms. Users can sign the unsigned BMDS Excel macros themselves by following Microsoft's instructions for digitally self-signing the macros.

The BMDS Release History page lists all features, enhancements, fixes, and changes for each BMDS 3 release.

1.5   Future of BMDS

EPA plans to continually improve and expand the BMDS system. Current plans include:
- Developing an online version of BMDS that will be integrated with the EPA Health & Environmental Research Online (HERO) database and Health Assessment Workspace Collaborative (HAWC) website,
- Adding a nested model for continuous responses,
- Adding covariate analysis tools,
- Adding trend tests, and
- Improving the Word report feature.

We welcome and encourage your comments on the BMDS software. Please provide comments, recommendations, suggested revisions, or corrections through our eTicket support site.

1.6   BMDS Online Resources

1.6.1 BMDS Website

The BMDS website contains the most up-to-date source of information and of updates pertaining to BMDS.

In addition to the latest downloadable version of BMDS. the site includes links to troubleshooting and usage tips, the BMDS 3 Release History, links to technical guidance, external and peer-review information on models used in BMDS, and opportunities to participate in the development of the next generation of BMDS models.

1.6.2 BMDS Mailing List

The BMDS mailing list is the best way to stay current with software development, training opportunities, and other information relevant to your work with BMDS.

The BMDS mailing list is low-traffic; members receive about 3-5 announcement per year.

To join the mailing list, please sign up on the BMDS website.

The website includes an archive of previous BMDS announcements.

1.6.3 BMDS Glossary

For definitions of terms used in this guide, please refer to the online BMDS Glossary. The glossary items can be exported to other formats, such as PDF or Excel.

Another good source of dose-response terminology is the Integrated Risk Information System (IRIS) Glossary.

1.6.4 BMDS Model Source Code

If you are a developer, you can download the BMDS model source code, which enables you to build the model DLLs used by BMDS to calculate dose-response results.

Note, however, that the model source code does not include the user interface source code (written in Excel's VBA language; the USER INTERFACE source code controls the appearance and functionality of the Analysis Workbook).

The model source code download page includes more information on constraints and third-party libraries needed to build the software.

1.6.5 BMDS eTicket Support

The BMDS eTicket site serves as our online Help Desk. Submit questions, concerns, comments, or suggestions on any aspect of the software or its usage, and someone from the BMDS Development Team will reply.

2.0 Setting Up BMDS

2.1   System Requirements

BMDS requires the desktop version of Microsoft Excel 2010, 2013, or 2016 (32- or 64-bit) for Windows or later with macros enabled (see Section 2.4 on page 15 for details).

Recommended: use the Office 365 or standalone Microsoft Office installation methods rather than installing via the Microsoft Store.

BMDS is compatible with 64-bit versions of Microsoft Windows.

Note   BMDS does *not* run on macOS systems; instead, we recommend installing a Windows virtual machine and running BMDS from there.

2.2   Step 1: Download BMDS 3

The latest version of BMDS is always available from the BMDS Download page. Follow the instructions on that page to download the .zip file containing BMDS.

See the next section for instructions on installation.

2.3   Step 2: Unzip BMDS 3 to a Folder

IMPORTANT!

BMDS <u>MUST BE UNZIPPED</u> TO ITS OWN FOLDER OR IT WILL NOT WORK!

Attempting to run BMDS from within its zip file will cause the application to fail.

BMDS can be unzipped to any folder where the user has read/write privileges. Administrator privileges are not required.

The following instructions are written specifically for Windows 10, which is the EPA standard desktop.

1. Locate the zipped file downloaded in Step 1.
2. Right-click the zipped file and, from the context menu, select **Extract All**. The "Extract Compressed (Zipped) Folders" dialog box displays.

Figure 1. Select "Extract All…" from the context menu.
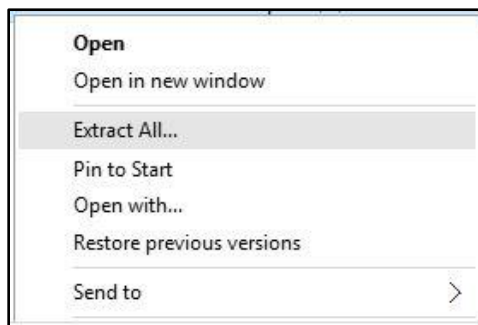
3. In the "Extract Compressed (Zipped) Folders" dialog box, shown in **Figure 2** below, enter or browse to the folder where the BMDS folder should be extracted. (Each version of BMDS should live in its own folder.) Check the "Show extracted files when complete" checkbox to open the BMDS folder after the zip extraction completes.

Keep the file path short: For best results, place the extracted folder (and its subfolders) in the simplest, shortest directory, without special characters or spaces, and for which you have administrative rights (for most users, this will be C:\Users\[user's LAN ID], but sometimes includes C:\).

Figure 2. *Extract Compressed (Zipped) Folders dialog box.*
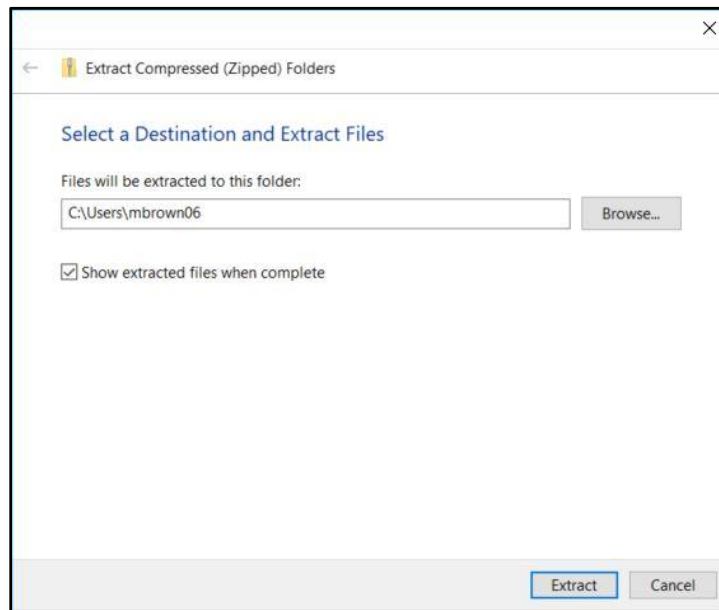
4. Select the **Extract** button to start the extraction.
5. To launch BMDS, select "bmds3.xslm" (or "bmds3" if your Windows operating system does not show file extensions).

2.4   Step 3: Accept EULA on Startup; Enable & Digitally Sign Excel Macros

On first startup, BMDS displays the End-User License Agreement (EULA) panel. Review the text and select either **Accept** or **Decline**. A small "Options saved" dialog appears; select **OK**.

2.4.1 REQUIRED: Enable Macros on First Startup

You will need to enable Excel macros on first startup for BMDS to work.

Follow the instructions for enabling Excel macros from the Microsoft Office web site.

2.4.2 Optional: Digitally Sign Macros

In the version of BMDS that is distributed via the BMDS website, the BMDS Excel Macros are not digitally signed.

If a signed version is required, follow Microsoft's instructions for digitally self-signing the macros in your copy of BMDS.

Users whose organizations require a digitally signed version of BMDS (to show that BMDS Excel Macros are from a trusted developer) can request one through the BMDS e-Ticket system. This should meet most organizations' security requirements for VBA-based applications.

Warning	The digitally signed version of BMDS 3.2 may not work consistently across different versions of Excel, including 32- and 64-bit versions.

2.5 Step 4: Create a BMDS Desktop Icon (Optional)

Some users may find it more convenient to run BMDS from a desktop shortcut icon. To do so:
1. Delete any older BMDS shortcut icons on the desktop.
2. In Windows Explorer, navigate to the newly installed BMDS application folder.
3. Right-click the BMDS3.xlsm file (or "BMDS3" if the system does not display file extensions). A context menu appears.
4. Click Send To. A submenu appears.
5. Click "Desktop (Create Shortcut)." Windows creates a shortcut to the file on the desktop.

2.6 Previous BMDS Installations

It is not necessary to uninstall previous 2.x or 3.x versions of BMDS.

However, ensure each BMDS installation is run from its own directory. Installed in this way, there should be no problems.

- For more information on backwards-compatibility with previous BMDS versions, see Section 3.6, "Backwards-Compatibility," on page 25.
- For more information on models not included in BMDS 3 from previous BMDS versions, see Section 3.7, "Models Not Included in BMDS 3," on page 26.

2.7 Troubleshooting

2.7.1 Extract BMDS from its Zip File Before Running

One of the most common BMDS support issues arises when attempting to run BMDS 3 from within its zip file.

When run from within the zip file, BMDS cannot exchange data and results between the application and its supporting dynamic link libraries (DLLs). This leads to interrupted file operations and frequent error messages.

To operate properly, BMDS 3 *must* be extracted from the downloaded zip file into its own directory, as explained in Section 2.3, "Step 2: Unzip BMDS 3 to a Folder," on page 14.

2.7.2 Resolving Some User Interface Display Issues

There may be situations where Excel improperly displays some onscreen controls, such as buttons or check boxes (for example, button text is hidden or truncated, or check boxes are misaligned).

Figure 3. Excel controls, such as check boxes or buttons, may sometimes appear misaligned.
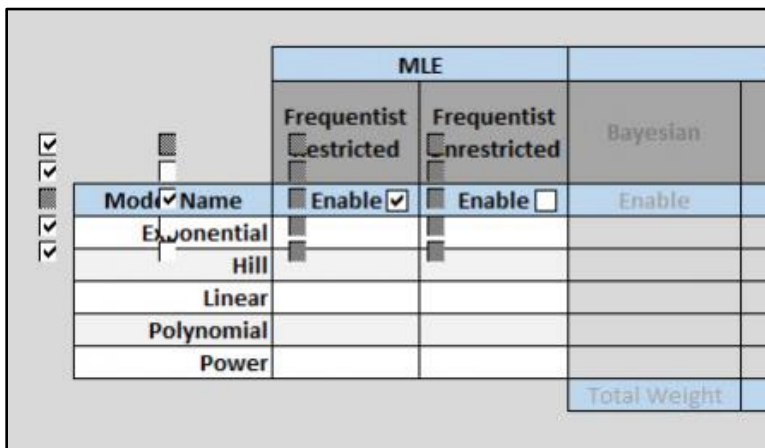
To resolve, select the "Display Settings" button in the status bar at the bottom of the Excel window and then select "Optimize for Compatibility." The workbook may need to be closed and re-opened for this change to take effect.

Figure 4. The Display Settings button.
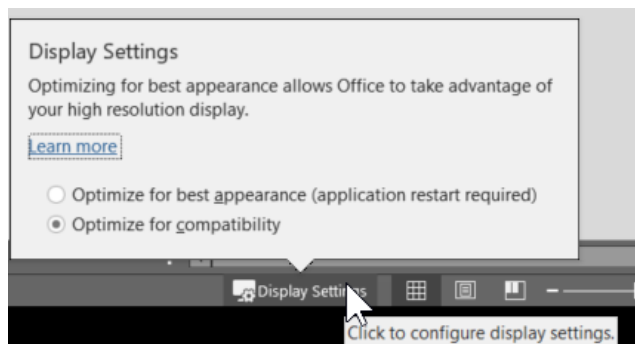
2.7.3 Avoid Using Windows Reserved Characters in File and Path Names and Datasets

BMDS allows any character, except for Windows reserved characters, to be used when naming files, directories, or datasets that BMDS will access.

The following Windows reserved characters are disallowed and cannot be used for naming files or datasets:

< (less than)

> (greater than)

: (colon)

" (double quote)

/ (forward slash)

\ (backslash) (use this character for specifying network drive paths)

| (vertical bar or pipe)

? (question mark)

* (asterisk)

2.7.4 Slow Performance

Some activities, such as generating the Results workbook, may slow BMDS performance
or the performance of other applications.

Using the 64-bit version of Excel may aid performance in some cases, as Excel memory
management can be a contributing factor.

Limiting the number of batch processes may help; try limiting the number of combinations
of Option Sets, models, and datasets that BMDS is analyzing.

An all-purpose fix is to reboot the computer to clear the RAM and to run as few
applications as possible while BMDS is processing.

2.7.5 Request Support with eTicket

For any technical problem related to running BMDS, please submit a problem report at
the BMDS eTicket site. With eTicket, users can request help, ask a question, or check on
the status of an existing issue.

3.0   BMDS 3 Basics

3.1   Excel-based User Interface

BMDS 2.7 and its predecessors were primarily standalone Windows applications. All aspects of the program were contained within the BMDS 2.x user interface window.

1. Users specified their datasets and analysis options in the Windows user interface.
2. The user interface sent the data to individual model executables for analysis.
3. The model executables returned the results to the user interface.

BMDS 3 works quite differently. BMDS 3 uses highly customized Microsoft Excel workbooks for analysis and results display.

1. A read-only Analysis workbook holds the datasets and analysis options.
2. The Analysis workbook sends the data to custom programs called dynamic link libraries (DLLs) for model analysis.
3. The DLLs' results are displayed in a separate Results workbook.
4. Each Results workbook holds the results for an analysis; a single analysis can be made up of multiple datasets.
5. From the Analysis workbook, a user can select a specific Results workbook to completely re-load the datasets and model options for re-running or for further configuration.

3.2   Analysis Workbook (bmds3.xlsm)

To start BMDS 3, double-click the bmds3.xlsm file in the BMDS program folder. The bmds3.xlsm file is the BMDS Analysis Workbook.

Figure 5. BMDS3 folder contents, with bmds3.xlsm workbook file.
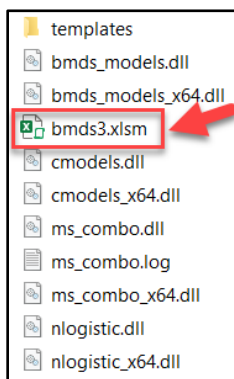
Basic information about the bmds3.xlsm/Analysis Workbook file:

- It contains the macros, user forms, and other data needed to render analyses, create the Results Workbooks, and so on.
- The file is write-protected. No user can save any changes they make to this file.
- The BMDS macros are also protected and cannot be viewed or edited.

Users enter datasets, modeling, and reporting options for an analysis in the BMDS 3 Analysis Workbook. Users can specify modeling options from intuitive forms and picklists. All calculations are performed within the Analysis Workbook.

Figure 6. The Analysis Workbook on first opening, with the Main tab displayed.

The Analysis Workbook is designed to facilitate performing and tracking dose-response analyses of multiple datasets having continuous, dichotomous, or nested dichotomous responses. There is an additional capability to model multiple cancer (dichotomous) endpoints and derive BMDs related to their combination.

Depending on the needs of the risk assessment, users can focus a BMDS 3 analysis on datasets associated by:

- study (e.g., for chemicals with a large database of studies)
- chemical (e.g., for chemicals that are not well-studied)
- health outcome (e.g., for chemicals with health outcomes that have been assessed in multiple studies and/or by multiple response measures)

For more details, refer to Section 4.0, "Defining and Running an Analysis," on page 27.

3.2.1 Help Tab

The Help tab contains links to BMDS's online support tools, documentation download, glossary, and technical guidance. Many of these links are described in more detail in Section 1.6, "BMDS Online Resources," on page 12.

3.2.2 Main Tab

The primary workspace of BMDS, where models, datasets, and option sets are queued for analysis. Full instructions on how to use the controls on this tab are described in Section 4.0, "Defining and Running an Analysis," on page 27

3.2.3 Data Tab

Figure 7. Data tab, on first opening.
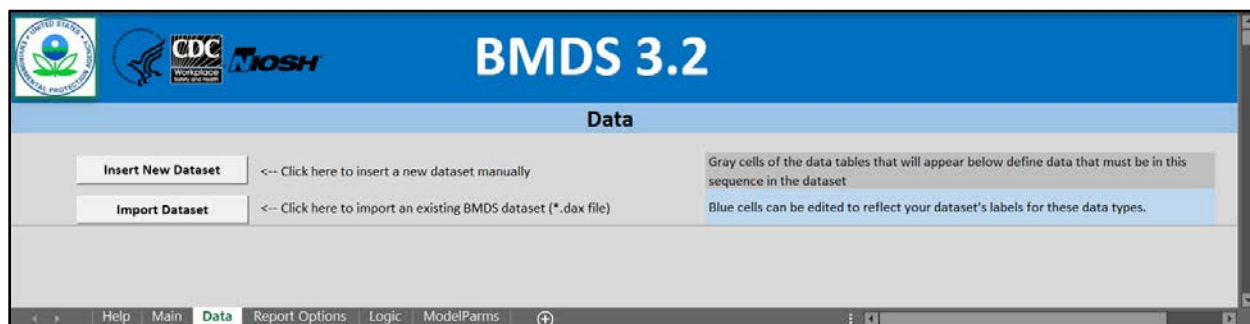
Users can specify one or more datasets from the Data tab. All datasets specified here will be listed on the Main tab, where specific options can be set for each dataset.

Datasets can be imported from earlier BMDS versions or can be manually entered.

Note the help text describing the utility of the gray- and blue-colored cells.

For more information on the Data tab, refer to Section 4.2, "Step 2: Add Datasets," on page 29.

3.2.4 Report Options Tab

Figure 8. Report Options tab, on first opening.
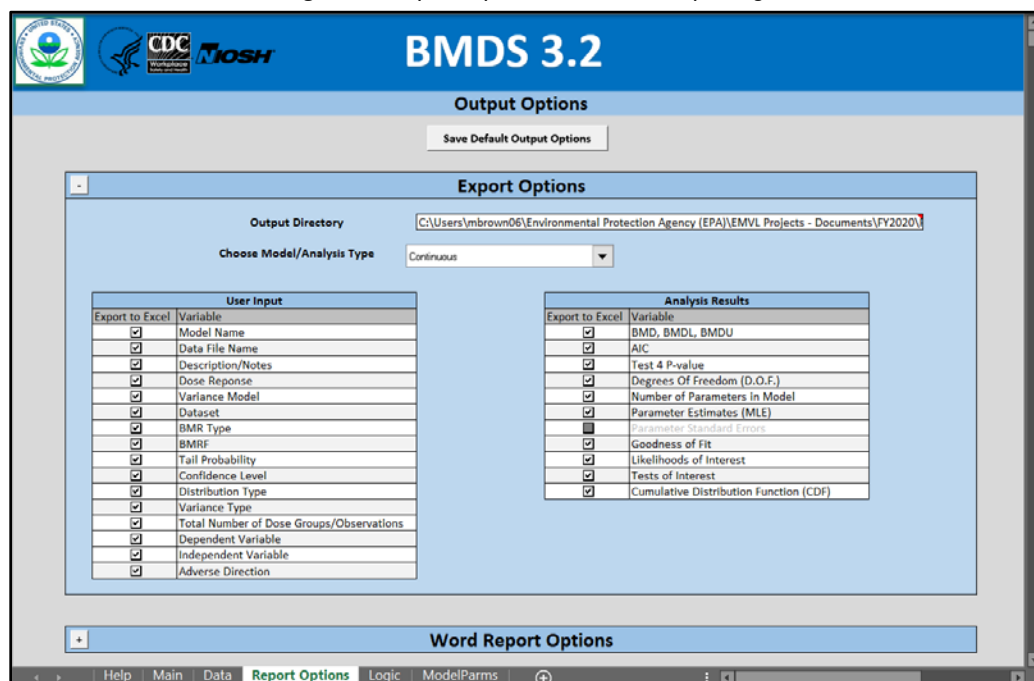
From the Report Options tab, the user can specify the data values to be included in the BMDS analysis results. The user can also specify elements to include in a Word-based report.

For more details, refer to Section 4.7, "Step 5: Prepare Summary Word Report(s)," on page 40.

3.2.5 Logic Tab

Figure 9. Logic tab, with EPA default recommendation decision logic.

BMDS 3's results include automatic recommendations regarding model selection that are consistent with the 2012 EPA Benchmark Dose Technical Guidance (U.S. EPA, 2012).

These criteria can be altered in the Logic tab of the BMDS 3 Analysis Workbook. Decision logic can be turned on or off, and specific criteria can be enabled or disabled for different dataset types. Notice that the logic depends on what type of data is being analyzed (continuous, dichotomous, nested).

For more details, refer to Section 12.0, "BMDS Recommendations and Decision Logic," on page 114.

3.2.6 ModelParms Tab (Model Parameters)

Figure 10. *Model Parameters tab.*

The ModelParms tab displays all of the models that BMDS runs, along with each model's parameters and their "specifications."

The values listed on the ModelParms tab are actual inputs to the program; they are read and are used by the model executables. The values are password-protected and cannot be changed/edited by the user.

For more details, refer to Section 5.2, "Model Parameters," on page 43.

3.3 Settings Workbook (.xlsx)

A Settings Workbook is created when the "Save Analysis" feature of BMDS is used (see Section 4.3).

The Settings Workbook contains all of the modeling options (Main tab), dataset information (Data tab), reporting options (Report tab), and model selection logic options (Logic tab) entered/selected/accepted by the user.

The Settings Workbook contains only settings for an analysis but not the results; this is because the analysis options were saved before an analysis was run.

The naming convention for the Settings Workbook is

[AnalysisName].xlsx

where AnalysisName is the content of the "Analysis Name" field from the Main tab.

The Settings Workbook file is stored in the directory location identified by the user in the "Select Output Directory" field of the Main tab. It can be loaded into BMDS later via the "Load Analysis" feature (see Section 4.4).

3.4   Results Workbook (.xlsx)

A Result Workbook is created when the **Run Analysis** button is selected on the Analysis Workbook's Main tab.

A bar showing the modeling status appears; as Excel compiles the Results Workbook, the display will not be updated.

For individual dataset analyses, BMDS records all model results in a separate Results Workbook for each dataset analyzed. For example, if a user analyzes five datasets, BMDS will create five separate Results Workbook files.

The Results Workbook filename for a given dataset is constructed as follows:

[DatasetName]_analysis.xlsx

- DatasetName = the user-specified name of the dataset from the Data tab
- _analysis = appended by BMDS to the end of the filename

The above naming convention is used for continuous, dichotomous, and nested modeling results Workbooks.

The naming convention is slightly different for multi-tumor analyses, which generally involve multiple datasets. For multi-tumor analyses, the Results Workbook naming convention is:

[AnalysisName]-multitumor.xlsx

- AnalysisName = the Analysis Name field contents from the Main tab
- -multitumor = appended by BMDS to the end of the filename

Figure 11. *Results Workbook files (.xlsx).*
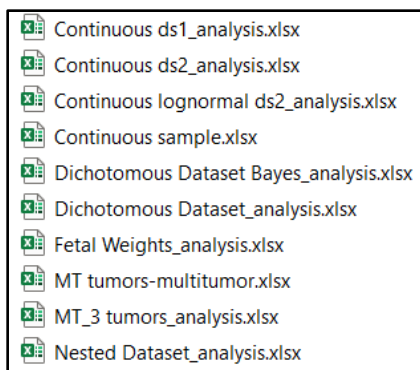
Note   A best practice with BMDS is to save BMDS Settings and Results Workbooks to their own folder outside the BMDS 3 program folder.

All the options used in the analysis are saved in the Results Workbook so they can be re-run the analysis later (see Section 4.4).

The Results Workbook contains a copy of the dataset, the dataset description entered by the user in the Analysis Workbook, and individual results tabs for each set of options specified.

Figure 12. A Results Workbook's Summary tab.
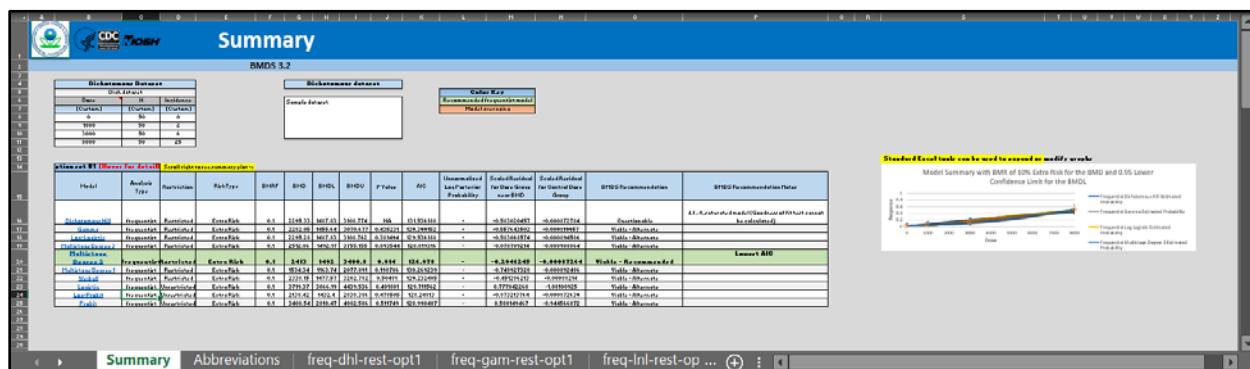
On the Summary tab, scroll to the right to view the summary results graph (see *Figure 13*). All of Excel's tools can be used for editing images and graphs.

Figure 13. Summary model results graph, on the Summary tab.
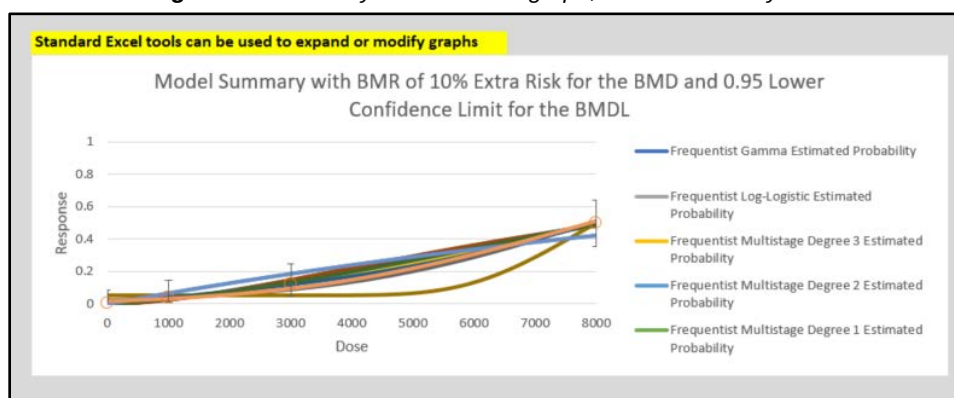
For more details on BMDS results, refer to Section 4.5, "Step 4: Run Analysis," on page 36.

3.5 Upgrades to Pre-BMDS 3 Models

BMDS continuous models have been upgraded to include a "Hybrid" modeling capability and Lognormal response options. For more details, see Table 5 on page 54.

Also, all pre-existing models have been re-coded to facilitate their maintenance and improve their performance in terms of stability, accuracy, reliability, and speed.

Note BMDS 3 handles Akaike Information Criterion (AIC) calculations somewhat differently from BMDS 2.x to facilitate comparing models with different likelihoods (i.e., Normal vs. Lognormal). For more details, refer to Section 11.1, "AIC for Continuous Models," on page 112.

3.6 Backwards-Compatibility

BMDS 3.2 retains backwards compatibility with BMDS 2.7 and BMDS Wizard 1.11.

The intent of all upgrades to BMDS is to improve optimization and estimation of parameters, including benchmark doses (BMDs).

Note	BMDS 3.x and 2.x will return different values for the log-likelihood and AIC when run on the same data. For more details, refer to Section 7.4.4, "AIC and Model Comparisons," on page 60.

For experienced users, BMDS 3 resembles the pre-existing BMDS Wizard in these ways:
- Excel-based
- Enables users to see and specify modeling options in a single tab
- Includes auto-selection features for identifying the "best" results in accordance with EPA recommendations or user-defined logic
- Documents all inputs and outputs in a single results workbook
- Provides flexible print options for displaying results in Microsoft Word tables formatted in a manner suitable for presentation in a risk assessment

Please file a ticket on the BMDS eTicket site for any questions or concerns regarding BMDS results.

3.7   Models Not Included in BMDS 3

BMDS 3 contains all the models and features that were available in BMDS 2.7 and BMDS Wizard 1.11 except for:
- Dichotomous background-dose models
- Rai and Van Ryzin nested dichotomous model
- ToxicoDiffusion model
- ten Berge model, which has been superseded by EPA's categorical regression software CatReg, which has the same functionality but with added features and options
- NCTR (National Center for Toxicological Research) nested dichotomous model (slated for inclusion in a future BMDS release)

These models can be accessed in BMDS 2.7, which is available from the BMDS website as an archive version of BMDS.

4.0 Defining and Running an Analysis

To initiate a new analysis (a "session" of dose-response model runs), open the file bmds3.xlsm from the BMDS3 program folder. Open the file from the directory in which it is saved or select it on the appropriate shortcut icon.

The user needs to enable Excel macros on first startup for BMDS to work.

Note: Throughout this User Guide, the bmds3.xlsm file will be referred to as the Analysis Workbook.

Figure 14. BMDS3 folder contents, with bmds3.xlsm file highlighted.
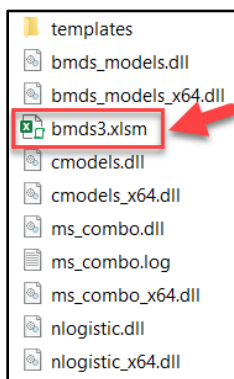
When bmds3.xslm opens, the Analysis Workbook's Main tab is displayed.

Figure 15. The Analysis Workbook on first opening.
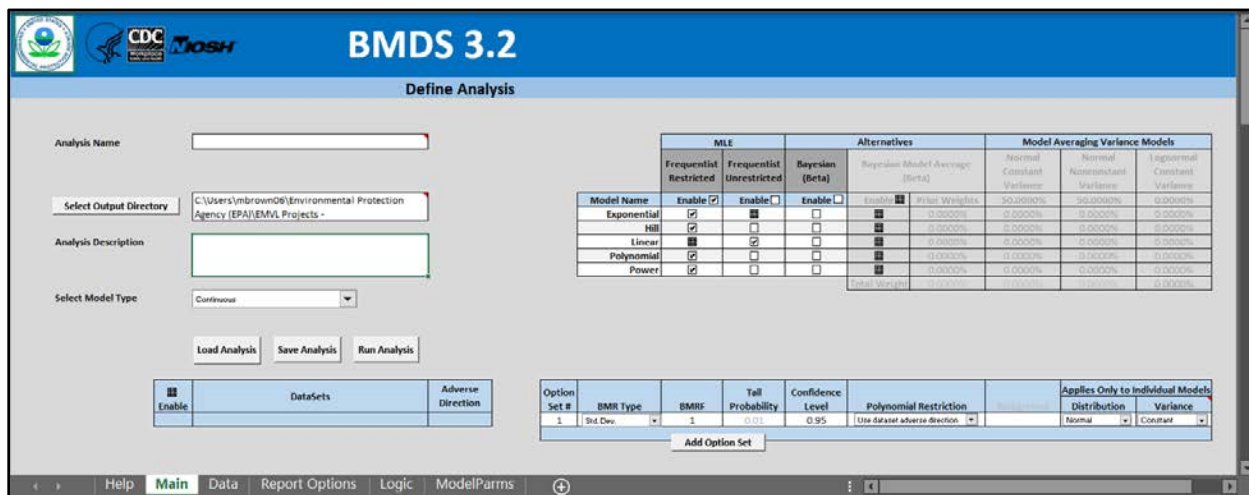
The Analysis Workbook contains five tabs:

- *Help*—contains links to the BMDS web site support materials, including this user guide, readme, release notes, Benchmark Dose Technical Guidance, and eTicket for user-support questions.
- *Main*—define an analysis to be run, save analysis configuration, load saved analysis configurations, and run an analysis. For more details on these functions, refer to Section 4.1, "Step 1: Analysis Documentation," on page 28.and Section 4.3, "Step 3a: Select and Save Modeling Options," on page 34.

- *Data*—define the dataset to be analyzed, either by entering the data manually or loading a BMDS .dax dataset file (compatible with BMDS 2.7 .dax files). For more details, refer to Section 4.2, "Step 2: Add Datasets," on page 29.
- *Report Options*—define the types of input and analysis results to be included in the results files, and also define what should be included in the Microsoft Word-based report file. For more details, refer to Section 4.7, "Step 5: Prepare Summary Word Report(s)," on page 40.
- *Logic*—the BMDS Model Recommendation Decision Logic rules can be reviewed and, if necessary, tweaked by the user. For more details, refer to Section 12.0, "BMDS Recommendations and Decision Logic" on page 114.

4.1   Step 1: Analysis Documentation

4.1.1 Enter an Analysis Name and Description

Figure 16. Fields for Analysis metadata.

Provide a name for the analysis. BMDS will use this to create the Results Workbook filename.

While not a required step, providing a fuller, free-text Analysis Description is useful for analyses to be saved for future use or consideration. These can include more detailed notes to describe the dataset; these notes will be displayed in the Settings Workbook, Results Workbook, and Word Report file.

4.1.2 User Settings and the Output Directory

Default settings are automatically loaded when BMDS first opens. For the most part, the initial Main, Data, Report Options, and Logic tab settings will always be the same. The only exception is the "Selected Output Directory" field, which is initially set to its value when BMDS was last closed.

By default, BMDS saves results to its install directory. The **Select Output Directory** button (as shown in *Figure 16*) displays a File Manager dialog for selecting a different directory. The output directory specified by the user is displayed in the "Select Output Directory" field

Note   Hover the mouse cursor over the field to display the complete output directory path; this is helpful in cases when the address is too long for the box.

The Output Directory will store a Settings Workbook (per analysis, using Analysis Name), Results Workbooks (one per dataset, using Dataset Name), and Word Report files (one per dataset, using Dataset Name).

The output directory specified on the Main tab is shown at the top of the Report Options tab also, as shown in *Figure 17*.

Figure 17. Report Options tab, with the same Output Directory as specified on the Main tab.
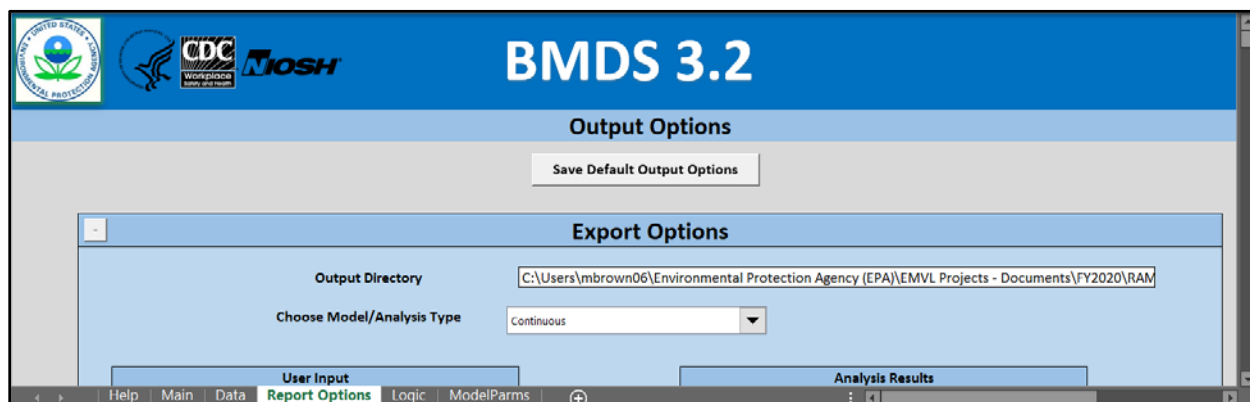
If the output directory is changed using the **Select Output Directory** button on the Main tab, BMDS will display the "Save New Settings?" dialog box. Select **Yes** to set the new directory as the default.

Figure 18. Save New Settings dialog box.
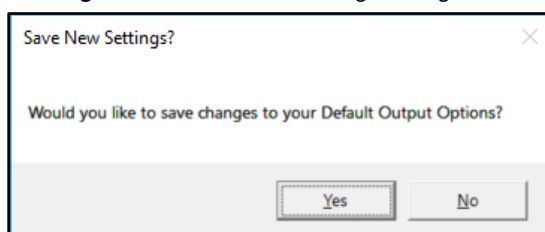
BMDS then displays the following confirmation. Select OK to continue.

Figure 19. Confirmation dialog box for changing default output options.
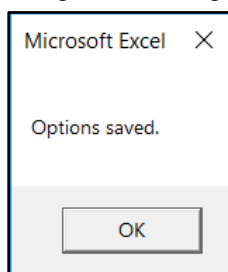
4.2   Step 2: Add Datasets

After entering the analysis documentation information in the Main tab, the dose-response data should be entered in the Data tab.

The user can add multiple datasets associated with four response types:
- Summarized continuous (e.g., mean and SD)
- Individual continuous (e.g., dose and response for each test subject)
- Dichotomous (e.g., lesion incidence)
- Nested dichotomous (e.g., developmental study) responses.

Note   The user can enter multiple datasets of different model types — continuous, dichotomous, and/or nested — on the Data tab. However, in the Main tab,

datasets that cannot be modeled with the selected model type will be greyed out
and un-selectable (see Section 4.3 below). For example, continuous and
dichotomous datasets must be analyzed in separate analysis runs.

4.2.1 Importing .dax Datasets

Previous versions of BMDS saved datasets in .dax format; BMDS 3 also saves datasets
to the .dax format.

To import existing .dax datasets
1. On the Data tab, select the **Import Dataset** button.

Figure 20. *Data tab buttons for specifying datasets.*
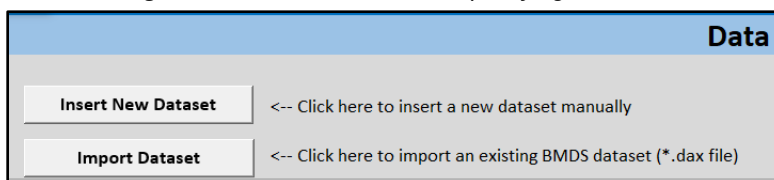
2. The Import Dax File dialog box displays.

Figure 21. *Import Dataset dialog.*
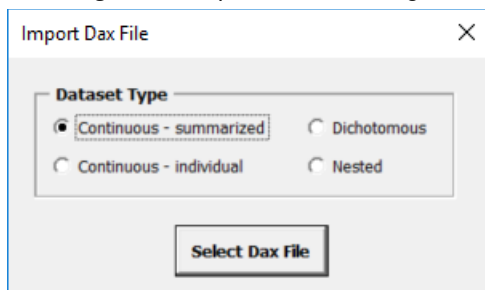
3. Select the dataset type, then select the **Select Dax File** button. The Select a Dax File
dialog box displays.

4. Navigate to the dataset's location, select the .dax file, and press **Open**.

5. The Map Data Columns dialog box will be displayed; from the picklists, select the
appropriate dataset header that corresponds to the variable type that BMDS is
expecting.

Figure 22. *Continuous-summarized dataset import dialog.*

Figure 23. *Continuous-individual dataset import dialog.*

Figure 24. *Dichotomous dataset import dialog.*

Figure 25. *Nested dataset import dialog.*

6. Select the **Import Data** button. BMDS imports the dataset into the Data tab, as shown in *Figure 26*.

Figure 26. *A .dax dataset imported into BMDS 3.*
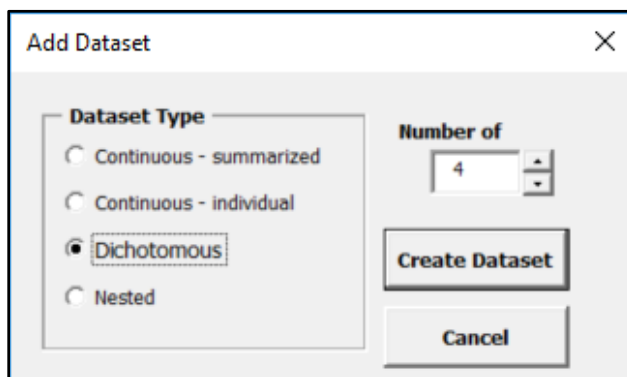
Note   Follow the guidelines for entering a unique dataset name. Also, refer to Section 4.2.7, "The Difference between the Gray and Blue Cells," on page 34.

4.2.2 Inserting a New Dataset

1. On the Data tab, select **Insert New Dataset**. The Add Dataset dialog box displays. Specify the Dataset Type and the number of rows.

Figure 27. Add Dataset dialog.
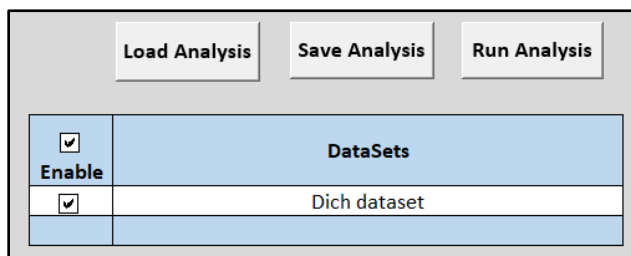
2. Select **Create Dataset**. BMDS inserts a table with the specified number of empty rows.

Figure 28. Example of empty dataset table created for dichotomous data.

3. Double-click in the blue cell and change "DataSet Name1" to a more meaningful description.

4. Edit the "[Add user notes here]" cell or leave blank.

5. The grayed cells cannot be edited. They indicate acceptable data types in the sequence that BMDS requires for a proper model analysis. The blue cells marked "[Custom]" can be edited by the user if the dataset employs different variable names for these data types. For more details, refer to Section 4.2.7, "The Difference between the Gray and Blue Cells," on page 34.

6. Enter the data into the remaining cells. Or, copy and paste data that comes from another table, spreadsheet, or program (such as a prior BMDS version).

7. Click on the Main tab to display it. The Datasets table displays the dataset name that was entered. The checked Enable column tells BMDS to run an analysis on the selected dataset.

Figure 29. The Main tab after a dataset has been specified.

4.2.3 Inserting Multiple Datasets

To add more datasets to the Data tab, select either the **Import** or **Insert Dataset** buttons and specify the dataset as described above. BMDS automatically adds the dataset to the end of the list.

When entering a lot of datasets, note that by default BMDS "freezes" the top portion of the tab so that the **Import** and **Insert Dataset** buttons are always visible.

4.2.4 Editing Datasets

After a dataset has been entered, a small **Edit** button appears beside the dataset.

Figure 30. Example dataset with Edit button.

Edit	Dich dataset		
	[Add user notes here]		
	Dose	N	Incidence
	[Custom]	[Custom]	[Custom]
	0	50	0
	1000	50	2
	3000	50	6
	8000	50	25

Click the **Edit** button to display a palette of Edit Dataset commands. These commands work only on the selected dataset.

Figure 31. Edit Dataset commands
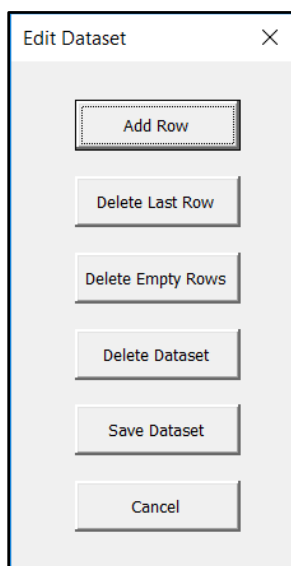
- **Add Row**—Inserts a duplicate of the last row at the end of the table.
- **Delete Last Row**—Deletes the last row in the table.
- **Delete Empty Rows**—Deletes all empty rows. (Note that BMDS will skip rows if the dose value is missing.)
- **Delete Dataset**—Deletes the entire dataset.
- **Save Dataset**—Saves the dataset in .dax format. Click the button, navigate to where the dataset should be saved, and click **Save**.
- **Cancel**—Closes the palette window.

4.2.5 Enter Unique Names for Each Dataset

Enter a unique name for each dataset.

BMDS uses this name to reference the dataset on the Main tab (where users can select datasets to include in a modeling analysis) and to name all Result Workbook and Word Report files generated from modeling the dataset.

Note   Avoid naming datasets using Windows reserved characters, such as the less than sign (<), vertical pipe (|), and so on. For a list of reserved characters to avoid, refer to Section 2.7.3, "Avoid Using Windows Reserved Characters in File and Path Names and Datasets" on page 17.

4.2.6 Datasets Can Have Empty Rows

The dataset can have empty rows. A BMDS analysis run will skip any rows for which the dose value is missing.

4.2.7 The Difference between the Gray and Blue Cells

Figure 32. Explanatory text on the Data tab for the gray and blue cells.
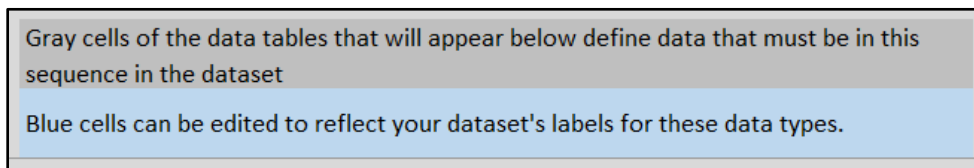
Grayed cells for datasets cannot be edited. They indicate acceptable data types in the sequence that BMDS requires for a proper model analysis.

The blue cells underneath can be edited if the dataset employs different variable names for these data types.

4.3   Step 3a: Select and Save Modeling Options

All models and modeling options available for use in an analysis can be selected on the Main tab.

Options can be saved and reloaded at any time before or after running an analysis.

An analysis can involve the use of any one of four Model Types:
- Continuous
- Dichotomous
- Dichotomous – Multi-tumor
- Dichotomous – Nested

As noted previously, an Analysis Workbook can include a mix of any of the data types on the Data tab. However, only one type of data can be run in a single analysis. So, for example, an Analysis Workbook can include Continuous, Dichotomous, and Nested datasets defined on the Data tab, but when it comes time to run an analysis, the user will need to run separate analyses for each data type.

Of course, the user is free to run two (or more) analyses from a single Analysis Workbook, each potentially implementing an analysis of a different data type.

Users can run the selected models against multiple, user-defined modeling "Option Sets" and multiple datasets.

The BMDS Main tab lists the datasets entered in the Data tab, enabling the user to choose datasets of the appropriate Model Type to analyze using the selected Models and Option Sets.

Each model type offers a different set of models and/or modeling options:

- For details on continuous endpoint model options, refer to Section 7.3, "Options," on page 52
- For details on dichotomous endpoint model options, refer to Section 8.2, "Options," on page 74.
- For details on nested dichotomous endpoint model options, refer to Section 9.3, "Options," on page 93.

4.4 Step 3b: Load, Save, or Run an Analysis

REMEMBER: No Data are Stored in the Analysis Workbook

A side-effect of the write-protection on bmds3.xslm is that the user cannot save data in the workbook itself. However, all datasets, modeling, reporting options, logic settings, etc. can be saved and reloaded for later use.

Figure 33. Analysis function buttons.
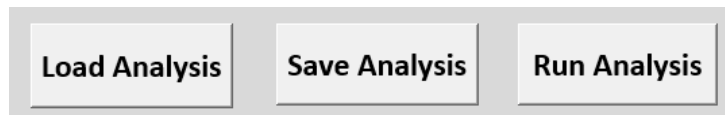
Save Analysis (without running)**.** Select this button to save a workbook with no results created.

To save analysis options without running an analysis

1. Specify the options, datasets, etc. for the analysis. The Analysis Name field will serve as the filename.
2. Select the **Save Analysis** button. BMDS creates a Settings Workbook file in the Output Directory, opens the file, and displays the following message on the Settings Workbook file's Summary tab.

Figure 34. BMDS displays the following message on the Settings Workbook's Summary tab.

This analysis was saved without running. Please open the BMDS analysis workbook (BMDS3.xlsm) then click the 'Load Analysis' button to load this analysis.

Run Analysis. Select this button after specifying an analysis. BMDS will automatically create a Results Workbook and (if selected) Word Report files in the user-specified Output Directory.

All Analysis Workbook specifications are saved in a Settings Workbook or Results Workbook file so the corresponding Analysis Workbook can be re-generated. When the **Load Analysis** button is selected, BMDS will re-generate the Analysis Workbook. All datasets, as well as modeling, model selection logic and reporting options from the original Analysis Workbook are reloaded into the Analysis Workbook. The user can then re-run the analysis using different options.

To load and re-run an analysis

To re-run an analysis later, or run it with different parameters or additional data:

1. Open the bmds3.xslm file.

2. Select **Load Analysis** from the Main tab.

3. An Open file manager dialog displays. Navigate to a saved BMDS Results Workbook file or previously saved analysis file, select it, and select **Open**.

4. BMDS loads the selected workbook's saved dataset and parameters into bmds3.xlsm. From there, specify new parameters, add additional data or datasets, create additional Model-Option Sets, etc.

Note   When an Analysis Workbook is re-generated, all options and datasets in the open Analysis Workbook will be overwritten. BMDS will display a warning so the user can save the open Analysis Workbook file or overwrite it.

4.5   Step 4: Run Analysis

On the Main tab, in the Datasets table, select one or many datasets to run in the analysis.

Click the Enable checkbox in the header to toggle selection/deselection of all loaded datasets. Only datasets that correspond to the selected model type will be selectable.

Figure 35. With Continuous as the selected Model Type, selecting the Enable checkbox selects the continuous dataset(s) that can be run in the current analysis.

Select the **Run Analysis** button on the Analysis Workbook's Main tab to begin the modeling run. A bar showing the modeling status appears; display updating is suspended as Excel compiles the Results Workbook.

The user can rename the files as desired. The BMDS Results Workbook file naming conventions are described in Section 3.3, "Settings Workbook ," on page 23.

Each dataset will be saved to its own Results Workbook. All model options specified as part of the analysis are also saved to the new Results Workbook.

4.6   Step 5: Review Results

After the user selects the **Run Analysis** button, BMDS creates and opens a Results Workbook for the specified analysis.

At the bottom of the Results Workbook window are the tabs that contain the Summary and model/option-specific results for the completed analysis. Right-click on Excel's tab-selection triangle to view a list of all the tabs in the file.

Figure 36. *Viewing the list of Results Workbook tabs in Excel by right-clicking the tab-selection triangle.*
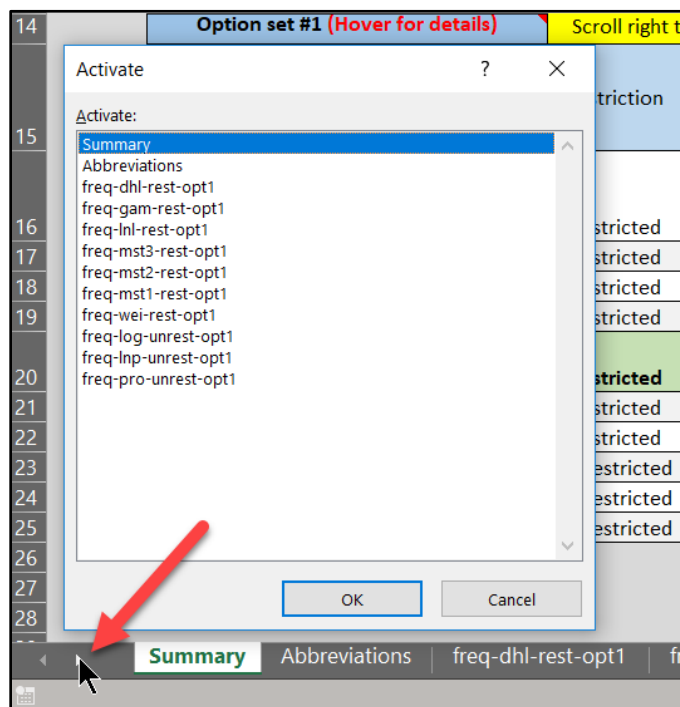
Each model's results tab contains its own BMD graph along with more detailed reports related to the Benchmark Dose, model parameters, goodness of fit, and analysis of deviance, among other results.

4.6.1 Model Abbreviations Tables

Each tab in the Results Workbook follows a naming convention to uniquely identify each result.

The Abbreviations tab summarizes all of the three-letter abbreviations for the model type used in the analysis. For example, the Abbreviations tab for an analysis with Dichotomous models will contain only abbreviations for dichotomous models.

The following tables list all the models available in BMDS 3, their abbreviations (used in results tab name and for other labeling purposes within the program), and their version numbers.

Model version numbers were re-set from the version numbers used for BMDS 2.7. Prior to BMDS 3, models were coded as separate executable files with their own version numbering.

For BMDS 3, models are now coded into a single DLL and so now carry (unless otherwise noted) the same version number.

Some users may wish to cite the version numbers in publications to establish the source of their results.

Table 1. *Frequentist Continuous Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Exponential	exp	1.1 (07/19/2019)
Hill	hil	1.1 (07/19/2019)
Linear	lin	1.1 (07/19/2019)
Polynomial	ply	1.1 (07/19/2019)
Power	pow	1.1 (07/19/2019)

Table 2. *Frequentist Dichotomous Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Gamma	gam	1.1 (07/19/2019)
Logistic	log	1.1 (07/19/2019)
Log-Logistic	lnl	1.1 (07/19/2019)
Log-Probit	lnp	1.1 (07/19/2019)
Multistage	mst	1.1 (07/19/2019)
Probit	pro	1.1 (07/19/2019)
Weibull	wei	1.1 (07/19/2019)
Quantal Linear	qln	1.1 (07/19/2019)
Dichotomous Hill	dhl	1.1 (07/19/2019)

Table 3. *Frequentist Nested Dichotomous Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Nested Logistic	nln	2.20 (04/27/2015)

Table 4. *Other Models: Abbreviations & Versions.*

Name	Abbreviation	Version (Date)
Bayesian model averaging (dichotomous)	bma	1.1 (07/19/2019)
Multi-tumor (MS_Combo)	msc	1.8 (04/27/2015)

4.6.2 Other Common Abbreviations on Tabs

The other common abbreviations used in the Results Workbook tab names are:

- Frequentist ("freq") or Bayesian ("bayes").
- Dichotomous Model Averaging ("DichoMA"), if Model Averaging was selected.
- Restricted ("rest"), Unrestricted ("unrest"). The user can choose to run both restricted and unrestricted from the Analysis Workbook's Main tab.
- "dset*n*" = dataset1, dataset2, and so on.
- "opt*n*" = identifying the Option Set on the Analysis Workbook's Main tab that was used to generate the results, in case multiple Option Sets were specified. If only one Option Set was identified, then this will always be "opt1."

> **Example**   A results tab named "freq-gam-rest-opt1" contains results for the frequentist gamma model, running restricted, and using the first Option Set defined in the Analysis Workbook.

4.6.2.1 Dichotomous - Multi-tumor (MS_Combo) Abbreviations

Figure 37. *Multi-tumor Results Workbook tabs.*

Abbreviations	**frequentist-msc-option1**	freq-mst3-rest-dset1-opt1	freq-mst2-rest-dset1-opt1	freq-mst1-rest-dset1-opt1

The following abbreviations are unique to the Multi-tumor results:
- "msc" = MS_Combo.
- "mst*n*" = Each individual multistage ("mst") results tab considered as part of the MS_Combo result.

4.6.2.2 Dichotomous – Nested Abbreviations

Figure 38. *Nested Results Workbook tabs.*

freq-nln-rest_lsc+ilc+_opt1	freq-nln-rest_lsc+ilc-_opt1	freq-nln-rest_lsc-ilc+_opt1	freq-nln-rest_lsc-ilc-_opt1

The following abbreviations are unique to the Nested results:
- "nln" = Nested Logistic model. At this time, there is only a single nested model.
- "lsc" = Litter Specific Covariate. The "+" indicates the LSC was included; the "-" indicates LSC was not included.
- "ilc" = Intralitter Correlation. The "+" indicates that ILC estimates were included; the "-" assumes ILC is zero.

4.7   Step 5: Prepare Summary Word Report(s)

On the BMDS 3 Analysis Workbook's Report Options tab, users can select modeling inputs and results to report for each model type, "Export Options," and "Word Report Options."

Figure 39. Analysis Workbook's Report Options tab.

- The selected Export Options affect both the Result Workbook and Word Report files that are generated from an analysis.
- Export Options ("User Input" and "Analysis Results") are set separately for each of the different model/analysis types. So "Continuous" will have its own set of enabled export options, Dichotomous-Nested its own set, and so on. Select the model type from the Choose Model/Analysis Type dropdown menu.
- Word Report Options are applied for creating tabular documentation of modeling results in Microsoft Word.

Note   Because the Word report may take a few minutes to compile, we recommend running the Word Report Options *after* the analysis results have been verified. If the analysis results are satisfactory, then re-run with the "Create Word Report" option checked.

Note   When using the Word Report option, it is recommended that the user specify on the Main tab only a single Model-Option Set combination. The more datasets and options specified, the longer BMDS takes to generate the report. Word Report generation improvements will be addressed in a future BMDS version.

5.0   Modeling in BMDS

Recall that the three endpoint types that BMDS can model are continuous, dichotomous, and nested dichotomous. Chapters 6, 7 and 8 discuss the particulars for each type of endpoint/modeling. There is one additional situation considered separately, i.e., the "MS_Combo" option. Note however that it is merely a special case of dichotomous modeling of multiple endpoints. Details for the MS_Combo modeling options are discussed in Section 10.1, "Dichotomous—Multi-tumor Models and Options," on page 105.

Presented here, however, are general modeling considerations common to all model types.

5.1   Frequentist and Bayesian

BMDS now allows the user to perform either (or both) Bayesian or non-Bayesian analyses.

BMDS refers to the non-Bayesian approach as "frequentist" or "maximum-likelihood estimation (MLE)." That approach is based on likelihood calculations. Models fit by these methods report maximum likelihood estimates and associated bounds determined by profile likelihood approaches. Presentation of p-values and the like (e.g., goodness-of-fit evaluations) are consistent with the frequentist tradition.

Bayesian analyses, in contrast, update parameter estimates. Distributions describing the a priori uncertainty in the parameter values (the so-called prior distributions) are updated using the data under consideration to yield *a posteriori* distributions. From those, the BMD estimate maximizing the a posteriori likelihood — the so-called maximum *a posteriori probability (*MAP*)* estimate — is reported, as are credible intervals for the BMD.

Note   At this time, EPA does not offer technical guidance on Bayesian modeling or Bayesian model averaging.

5.2 Model Parameters

Figure 40. Model Parameters tab.

The ModelParms tab displays all of the models that BMDS runs, along with each model's parameters and their "specifications."

The values listed on the ModelParms tab are actual inputs to the program; they are read and are used by the model executables. The values are password-protected and cannot be changed/edited by the user.

Specifications are separated by modeling approach: frequentist or Bayesian (see the previous section, 5.1, for a brief description of these approaches).

For the frequentist approach, the specifications consist of initial values (the values used to initiate the optimization of the likelihood) and the constraints on the values of the parameters (a minimum and a maximum). When there is a restricted and an unrestricted form of a model, there are separate frequentist specifications for those two forms, which differ with respect to a bound ("Min" or "Max").

For the Bayesian approach, the specifications define the prior distribution for each parameter. A distribution type (e.g., "Normal") is given, and the parameters that define that prior (e.g., "Mean" and "StdDev"). The priors are also bounded (for numerical stability purposes); the bounds are listed under the columns "Min" and "Max."

For additional information, in relation to the formal mathematical/statistical details, the frequentist and Bayesian model equations are presented in the following sections:

- Section 7.4, "Mathematical Details for Models for Continuous Endpoints in Simple Designs," on page 57
- Section 8.3, "Mathematical Details for Models for Dichotomous Endpoints in Simple Designs," on page 76
- Section 9.4, "Mathematical Details for Models for Nested Dichotomous Endpoints," on page 97

Also see, in particular, the following tables where the priors and model constraints are presented:

- Table 7: The individual Bayesian continuous models (in preview) and their parameter priors, on page 64
- Table 12: Bayesian dichotomous models and their respective parameter priors on page 84

To see how model parameter estimates are reported in the BMDS results, refer to Section 6.3, "Model Parameters ," on page 49.

5.3 Optimization Algorithms Used in BMDS

For frequentist analyses and some Bayesian computations, the NLopt optimization library is used for BMDS 3.0.

Several optimization algorithms available in the library are used to ensure reliability of the estimation:

- For global optimization involving the maximum likelihood or maximum *a posteriori* estimation, the L-BFGS[1] method is attempted first. If it fails to converge, gradient free algorithms "subplex" and "BOBYQA"[2] algorithms are then attempted.
- For profiling, when only non-linear inequality constraints are needed, the COBYLA[3] and MMA[4] approaches are used and compared. In the case the methods return different optimum, the values producing the larger of the two is used.
- For equality-constrained optimization, the augmented Lagrangian algorithm is used and either the L-BFGS, BOBYQA, or the "subplex" algorithm is used in the local optimization step. When two approaches produce different results, the values producing the larger optimum are used.

NLopt 2.4.1 was used when developing the BMDS 3 code. This version is available for download from the NLopt GitHub site.

For more information regarding the algorithms, refer to the NLopt documentation site.

5.4 Bayesian Analyses, including Model Averaging

BMDS model averaging proceeds from the basis of Bayesian analyses, for which the parameters of the models under consideration are updated using the dataset of interest.

Priors for the parameters are defined in sections 7.4.6 and 8.3.6 for continuous (in preview) and dichotomous dose-response models, respectively. Only dichotomous model averaging is available in BMDS 3.2.

For each model, M, there is a likelihood for the data, $\ell(D|M)$, based on the data generating mechanism (binomial sampling in the case of the dichotomous endpoints; Normal or Lognormal distributions for continuous data).

---

[1] Limited-memory BFGS *(*L-BFGS *or LM-*BFGS*)* is an optimization algorithm in the family of quasi-Newton methods that approximates the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS).
[2] Bound Optimization by Quadratic Approximation (BOBYQA) is a numerical optimization algorithm.
[3] Constrained optimization by linear approximation (COBYLA) is a numerical optimization method.
[4] Method of Moving Asymptotes (MMA) is a method for structural optimization.

When one is interested in more than one model, model averaging is an approach that should be seriously considered in lieu of model selection (e.g., basing inferences on one model deemed to be the "best").

Suppose for this development that we are considering K models ($M_k, k = 1, ..., K$).

For each model, BMDS approximates the posterior density for the BMD using a Laplacian approximation; call that density $g_k(BMD|M_k, D)$ for model k. If the parameter vector for model k is denoted θk, let $\hat{\theta}_k$ designate the value of that vector that maximizes the posterior likelihood (the maximum *a posteriori*, or MAP, estimate).

The posterior density of the model averaged BMD is

$$g_{ma}(BMD|D) = \sum_{k=1}^{K} \pi_k(M_k|D)g_k(BMD|M_k, D),$$

where π_k is the posterior probability of model M_k given the data.

Clearly, this approach requires estimation of the posterior probabilities for each model considered. These are the weights for the averaging process. Unlike approaches that have been used elsewhere, we eschew the use of information-criteria-based weights (e.g., those based on Bayesian information criteria or Akaike Information criteria). Rather, BMDS generates weights using the Laplace approximation to the marginal density of the data. That is, for model M_k, 1 ≤ k ≤ K, with parameter vector θ_k of length s, one approximates the marginal density as

$$I_k = (2\pi)^{s/2}\left|\hat{\Sigma}_k\right|^{1/2}\ell\left(D\middle|M_k, \hat{\theta}_k\right)g\left(\hat{\theta}_k\middle|M_k\right)$$

where

$\hat{\theta}_k$ is the MAP estimate,

$\hat{\Sigma}_k$ is the negative inverse Hessian matrix evaluated at $\hat{\theta}_k$,

$\ell\left(D\middle|M_k, \hat{\theta}_k\right)$ is the likelihood of the data, for model k evaluated at the MAP, and

$g\left(\hat{\theta}_k\middle|M_k\right)$ is the value of the prior density for M_k evaluated at the MAP parameter estimates.

To compute the posterior model probabilities for theM_k, one calculates the MAP and then calculates I_k using the preceding equation. The posterior probability of the model is

$$\pi_k(M_k|D) = \frac{w_k I_k}{\sum_{i=1}^{K} w_k I_k},$$

where w_k is the prior probability of model M_k. In BMDS, the user can specify those weights; the default is equal weight for each model (all models being considered are equally probable *a priori*).

This approximation is similar to the Model Averaged Profile Likelihood (MAPL) approach of Fletcher and Turek ([2012](#)). However, while MAPL relies only on the likelihood, our approach incorporates prior information in calculating the marginal profile density of the BMD. In other words, both the likelihood and prior are used. The model-specific density is defined by treating profile density bounds as quantiles of a marginal posterior density for the parameter of interest, and the relation to the present approach and the MAPL approach is justified asymptotically.

This approach can be related to the MAPL framework by substituting the posterior density for the likelihood in each of the steps. This method approximates the marginal likelihood using the posterior MAP estimate and Hessian of the log-posterior.

> **Note** For the model average approach, the dichotomous Bayesian models (described in Section 8.1) are available in the model average. For dichotomous model averaging, the Multistage model is capped to a maximum degree of 2. The reasoning for this follows upon the work of Nitcheva, et al. who show that higher-order polynomials are not necessary given the fact that other models of the model averaging suite (e.g., dichotomous Hill) can provide increased curvature.

The BMDS model-averaged BMD point estimate is the weighted average of BMD MAP estimates from individual models, weighted by posterior weights $\pi_k(M_k|D)$. This is equivalent to the median of the approximate posterior density of θ. For the BMDL or BMDU estimates, the equation defining g_{ma} is integrated. A 100(α)% BMDU estimate or 100(1 - α)% BMDL estimate is the value BMD$_\alpha$ such that:

$$\alpha = \int_{-\infty}^{BMD_\alpha} g_{ma}(\text{BMD}|D) \, d\text{BMD},$$

$$= \sum_{k=1}^{K} \pi_k(M_k|D) \int_{-\infty}^{BMD_\alpha} g_k(\text{BMD}|M_k, D) \, d\text{BMD}.$$

The quantity $\int_{-\infty}^{BMD_\alpha} g_k(\text{BMD}|M_k, D) \, d\text{BMD}$ is approximated by,

$$\int_{-\infty}^{BMD_\alpha} g_k(\text{BMD}|M_k, D) \, d\text{BMD}$$

$$\approx \frac{1}{2}\Pr\left(-2\log[\hat{g}_k(\widehat{\text{BMD}}|M_k, D)] - 2\log[\hat{g}_k(\text{BMD}_\alpha|M_k, D)] < \chi^2_{1,\alpha}\right),$$

where $\hat{g}_k(x|M_k, D)$ is the maximum value of the posterior evaluated at *x*, $\widehat{\text{BMD}}$ is the MAP estimate of the BMD, and $\chi^2_{1,\alpha}$ is the α quantile of a chi-squared random variable with one degree of freedom. The above approximation assumes $BMD_\alpha < \widehat{BMD}$. When $\widehat{BMD} < BMD_\alpha$ the right-hand side of this equation is replaced by

$$\approx 1 - \frac{1}{2}\Pr\left(-2\log[\hat{g}_k(\widehat{\text{BMD}}|M_k, D)] - 2\log[\hat{g}_k(\text{BMD}_\gamma|M_k, D)] < \chi^2_{1,\gamma}\right).$$

This approximation is like the profile-likelihood used when estimating the BMDL and BMDU using the method of maximum likelihood, but in this case $\hat{g}_k(x|M_k, D)$ is the posterior density, which incorporates both the likelihood and the prior.

6.0   Output Common to All Model Types

Dataset-specific Results Workbooks generated by the BMDS 3 Analysis Workbook contain results Model-Option Set combination in separate tabs. Each tab consists of tabular and graphical summaries of the modeling inputs and results.

The purpose of these results is to provide the user with goodness-of-fit criteria and model results to aid in determining the appropriateness of the Model and Option Set to the benchmark dose derivation.

This section describes BMDS model outputs that are **common** to all model types. For details on outputs **specific to each model type**, refer to:

- Section 7.5, "Outputs Specific to Frequentist Continuous Models," on page 65.
- Section 7.6, "Outputs Specific to Bayesian Continuous Models," on page 71.
- Section 8.4, "Outputs Specific to Frequentist Dichotomous Models," on page 86.
- Section 8.5, "Outputs Specific to Bayesian Dichotomous Models," on page 89.
- Section 9.5, "Outputs Specific to Frequentist Nested Dichotomous Models," on page 103.

6.1   Model Run Documentation (User Input Table)

User Input	
Info	
Model	frequentist Dichotomous Hill v1.1
Dataset Name	Inflammation, Chronic Active
User notes	[Add user notes here]
Dose-Response Model	P[dose] = g +(v-v*g)/[1+exp(-a-b*Log(dose))]
Model Options	
Risk Type	Extra Risk
BMR	0.1
Confidence Level	0.95
Background	Estimated
Model Data	
Dependent Variable	Dose
Independent Variable	Effect
Total # of Observations	5

The Results Workbook tabs that are generated for each Model-Options Set contain a User Input table; the User Input table lists the options selected for that Model-Options Set.

For instance, when two users may be comparing results and they obtained different answers, they can consult their respective User Input tables to make sure the settings were the same or if they had used the same (or most current) version of the models.

The User Input tables within the Results Workbook tabs for each Model-Options Set contain the model name and version number, the dataset name, dataset user notes, and modeling options entered and specified by the user on the Analysis Workbook's Main and Data tabs.

6.2 Benchmark Dose Estimates and Key Fit Statistics (Benchmark Dose Table)

Benchmark Dose	
BMD	2295.279512
BMDL	1487.825333
BMDU	3108.76178
AIC	129.5361884
P-value	0.38149439
D.O.F.	1
Chi^2	0.765880782

A results tab's Benchmark Dose table contains the BMD, BMDL, and BMDU estimates, AIC, and the overall goodness-of-fit test p-value, degrees of freedom (D.O.F), and chi-square for the Model-Option set.

6.2.1 AIC

The Akaike Information Criterion (AIC) (Akaike, 1973) value given on the BMDS Results Workbook tabs is calculated as follows:

$$AIC = -2*LL + 2p$$

where LL is the log-likelihood at the maximum likelihood estimates for the parameters, and p is the number of model parameters estimated (and not on a restriction boundary).[5]

The AIC can be used to compare different models fit (using the same fitting method, e.g., least squares or maximum likelihood) to the same data set. Smaller values of the AIC indicate better fit. Although AIC comparisons are not exact (they rely on rules of thumb for interpreting AIC differences), they can provide useful guidance in model selection.

Model-type specific details on the AIC are discussed in the following sections:

- For continuous endpoints, refer to Section 7.4.4 on page 60 and to Section 11.1, "AIC for Continuous Models," on page 112.
- For dichotomous endpoints, refer to Section 8.3.2 on page 78.

6.2.2 P-value

The p-value is computed based on the D.O.F and the Chi^2 value (Chi^2 is assumed to be distributed as a chi-squared distribution having degrees of freedom equal to D.O.F). The p-value measures the "closeness" of the model predictions to the observed data. If the overall p-value is larger than some predetermined critical p-value, then the user might infer that the model appropriately describes the observed dose-response pattern. The critical p-value used by EPA is generally 0.1 but is sometimes relaxed to 0.05 for Multistage model when it is applied to cancer data (U.S. EPA, 2012).

---

[5] For the dichotomous and nested dichotomous models, an additivity constant is not included in the LL calculations.

6.3   Model Parameters Table

Figure 41. *Model Parameters Table, with hover tip explaining the Bounded estimate.*

Model Parameters	
# of Parameters	4
Variable	Estimate
g	1.65207E-08
b1	3.70531E-05
b2	Bounded
b3	7.721E-13

The value of this parameter, 0, is within the tolerance of the bound (see user guide for tolerance limits)

The Model Parameters table includes the estimates for the parameter values that "optimize" the model fit.

Parameter estimates are checked to see if they fall within a given tolerance (1.0e^{-6}) of parameter boundaries. If so, they are marked as "Bounded." This value applies to all parameters.

6.4   Cumulative Distributive Function (CDF) Table

This block is new to BMDS 3.x. CDF stands for "cumulative distribution function," in this case for the BMD estimate. It lists the percentiles associated with the CDF for the BMD being estimated.

Note that the BMD value associated with the CDF value of 0.5 is the MLE of the BMD (and matches the value reported for the BMD in the Benchmark Dose table discussed above).

The CDF block may also correspond to the Benchmark Dose table in terms of the BMDL and BMDU values reported in the latter. Recall that the confidence level specified by the user in the options is a one-sided confidence level. So, if that confidence level is related to one of the cumulative percentiles in the CDF block, the BMD values will match. As an example, if the confidence level specified by the user is 0.95 (95% one-sided confidence limits requested), then the BMDU from the Benchmark Dose table will match the BMD value listed for 0.95 in the CDF block. And, the BMDL will match the BMD value listed for 0.05 in the CDF block.

6.5   Graphical Output

Graphical outputs (plots) are displayed on the Summary tab and on the specific Model-Option tabs of the Results Workbook. The Summary tab shows the plots for *all* Model-

Options Sets run for a given analysis. Individual model results are shown on the tabs corresponding to that Model-Option Set.

Figure 42. Results plot for an individual model.
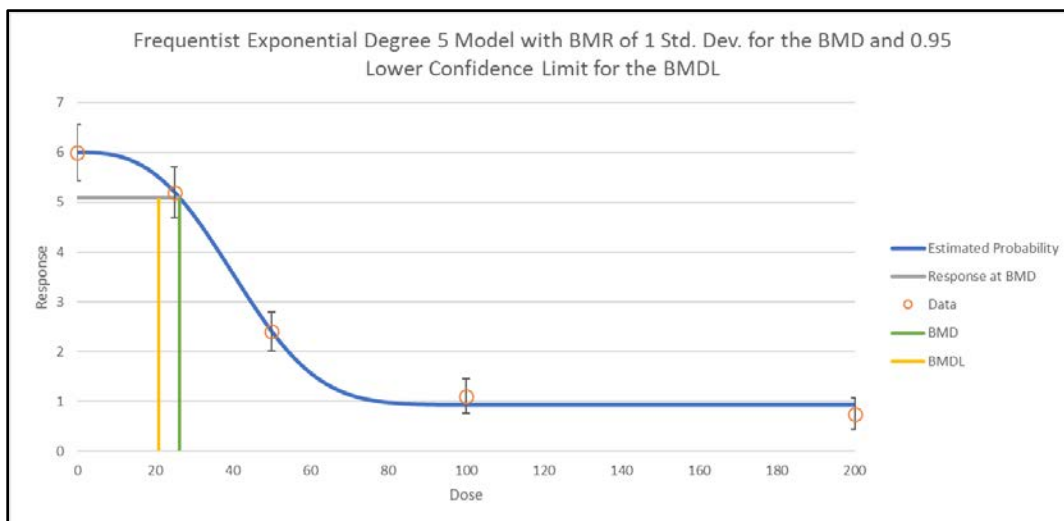
- The BMD and BMDL are indicated by the green and yellow vertical lines, respectively, and are associated with the user-selected benchmark response (BMR), the horizontal grey line.
- The dose-response curve estimated by the model is represented by a blue line.
- The graphical display features can be modified using Excel edit features.
- Data points are shown as orange circles with their individual group confidence intervals.
- Error bar calculations differ slightly based on the endpoint:
  - For continuous endpoints, refer to Section 7.5.4, "Plot and Error Bar Calculation," on page 70.
  - For dichotomous endpoints, refer to Section 8.3.3, "Plot and Error Bar Calculation," on page 79.
  - For nested endpoints, refer to Section 9.4.3, "Plot and Error Bar Calculation," on page 102.

7.0 Continuous Endpoints

Continuous endpoints take on values that are real numbers (as opposed to integers, for example), measuring things that can vary continuously (weights, concentrations, etc.).

The three key features of such measures that need to be specified to estimate a BMD are:

1. What direction of change indicates a toxic response (adverse direction),
2. How should the BMD be defined relative to the change in the response, and
3. How the responses are distributed.

   With respect to the distribution, one needs to consider the type of distribution and the nature of the variability around the center of the distribution. The options available to the user, discussed in Section 7.3, relate to all of those choices.

This section provides details on the following topics:
- Implementation of continuous models in BMDS 3
- Entering continuous model data
- Continuous model options
- Continuous model-specific outputs
- Options for restricting values of certain model parameters
- The Bayesian approach to continuous response modeling, specifying how the priors are defined

7.1 Continuous Response Models

All the traditional frequentist models and options that were available for analyzing continuous response data in previous versions of BMDS are available in BMDS 3.

Figure 43. Default selection of BMDS 3 continuous models, as they appear in the Analysis Workbook. Note that model averaging is disabled for continuous models in BMDS 3.2.

	MLE		Alternatives			Model Averaging Variance Models		
	Frequentist Restricted	Frequentist Unrestricted	Bayesian (Beta)	Bayesian Model Average (Beta)		Normal Constant Variance	Normal Nonconstant Variance	Lognormal Constant Variance
Model Name	Enable ☑	Enable ☐	Enable ☐	Enable ▦	Prior Weights	50.0000%	50.0000%	0.0000%
Exponential	☑	▦	☐	▦	0.0000%	0.0000%	0.0000%	0.0000%
Hill	☑	☐	☐	▦	0.0000%	0.0000%	0.0000%	0.0000%
Linear	▦	☑	☐	▦	0.0000%	0.0000%	0.0000%	0.0000%
Polynomial	☑	☐	☐	▦	0.0000%	0.0000%	0.0000%	0.0000%
Power	☑	☐	☐	▦	0.0000%	0.0000%	0.0000%	0.0000%
				Total Weight	0.0000%	0.0000%	0.0000%	0.0000%

Also, users are now able to use the Hybrid continuous modeling method and the lognormal response distribution assumption (previously only available for Exponential models) for all continuous models.

As in previous versions of BMDS, the user can choose to run the Hill, Polynomial, and Power models either restricted or unrestricted; the Linear model is not restricted and the Exponential models can only be run restricted.

New in BMDS 3.2 are *preview* versions of the Bayesian implementation of continuous endpoint modeling. Peer review of the Bayesian continuous models is planned for later in 2020.

Note   At this time, EPA does not offer technical guidance on Bayesian modeling.

Note   The preview models are new. Users acknowledge they have not been extensively tested, and formally reviewed and approved by the EPA for risk assessment purposes.

7.2   Entering Continuous Response Data

For details on inserting or importing datasets, see Section 4.2, "Step 2: Add Datasets," on page 29.

For summarized continuous response data, the default column headers are "Dose," "N," "Mean" and "Std. Dev."

For individual continuous response data, the default column headers are "Dose" and "Response."

7.2.1 Adverse Direction

Figure 44. Adverse Direction picklist for the selected dataset.

Choices for the Adverse Direction option are "automatic" (default), "up," or "down."

This option refers to whether adversity increases as the dose-response curve rises "up" or falls "down." Manually choose the adverse direction if the direction of adversity is known for the endpoint being studied.

If "automatic" is chosen, BMDS chooses the adverse direction based on the shape of the observed dose-response relationship.

This selection only impacts how the user-designated benchmark response (BMR) is used in conjunction with model results to obtain the BMD.

7.3   Options

On the BMDS 3 Analysis Workbook's Main tab, the user can define multiple Option Sets to apply to multiple user-selected models and multiple user-selected datasets in a single "batch" process. Select the **Add Option Set** button to define a new Option Set configuration.

Figure 45. *Continuous Model options.*

Option Set #	BMR Type	BMRF	Tail Probability	Confidence Level	Polynomial Restriction	Background	Applies Only to Individual Models	
							Distribution	Variance
1	Std. Dev. ▼	1	0.01	0.95	Use dataset adverse direction ▼		Normal ▼	Constant ▼
		Add Option Set						

7.3.1 Defining the BMD

Option Set #	BMR Type	BMRF	Tail Probability	Confidence Level
1	Std. Dev. ▼	1	0.01	0.95
	Std. Dev.		Add Option Set	
	Rel. Dev.			
	Abs. Dev.			
	Point			
	Hybrid - extra risk			

The following options are related to the definition of the BMD and its bounds:

- **Benchmark Response (BMR) Type**, which defines the method of choice for determining the response level used to derive the BMD. For details on these methods, refer to Table 5 on page 54.
- The **BMRF (Benchmark Response factor)** is specific to the BMR Type. Table 5 summarizes the options related to BMR Type and BMRF.
- **Tail Probability** marks the cut-off for defining adversity and applies only to "Hybrid extra risk" BMR Type. If the default setting of 0.01, for example, is used, this indicates that the user has specified that, in the absence of exposure, the probability of a response that is considered adverse is 0.01. This is a "tail probability" in the sense that it specifies how much of the tail of the distribution of responses (upper or lower) is in the adverse range. It implicitly defines the cut-off between normal and adverse responses.
- **Confidence Level** is set to 0.95 by default. This confidence level corresponds to a **one-sided** confidence bound, in either direction. In other words, if the confidence level is set to 0.95, the BMDL is the one-sided 95% lower bound on the BMD; the BMDU is the one-sided 95% upper bound on the BMD. The interval from the BMDL to the BMDU would, in that case, be a 90% confidence interval.

Table 5. *Options related to Continuous BMR Type and BMRF.*

Analysis File, Main Tab Option Name	Verbal Definition: The BMD is the dose yielding …	Mathematical Definition	BMRF Notes
Rel. Dev.	*Relative Deviation*: … the specified change in median response relative to the background median	$$\frac{|m(BMD) - m(0)|}{m(0)} = BMRF$$	BMRF is the "specified change" Default value = 0.1 [10% change in median]
Abs. Dev.	*Absolute Deviation*: … the specified change in median response	$$|m(BMD) - m(0)| = BMRF$$	BMRF is the "specified change" There is no default because it is very endpoint specific
Point	*Fixed Value*: … a median equal to the specified point value	$$m(BMD) = BMRF$$	BMRF is the "specified point value" There is no default because it is very endpoint specific
Std. Dev.	*Standard Deviation*: … the specified change in median relative to the control standard deviation	Normal Responses: $$\frac{|m(BMD) - m(0)|}{\sigma(0)} = BMRF$$ Lognormal Responses: $$\frac{|\ln(m(BMD)) - \ln(m(0))|}{\sigma_L(0)} = BMRF$$	BMRF is the multiple of the standard deviation Default value = 1 [change in median (or log-median) equal to 1 standard deviation (or log-scale standard deviation)]
Hybrid	*Increased Extra Risk*: … the specified extra risk, defined by the estimated distribution and background rate	If high responses are adverse: $$BMRF = \frac{\Pr(X > X_0|BMD) - \Pr(X > X_0|0)}{1 - \Pr(X > X_0|0)}$$ If low responses are adverse: $$BMRF = \frac{\Pr(X < X_0|BMD) - \Pr(X < X_0|0)}{1 - \Pr(X < X_0|0)}$$ where X_0 is a response value and $\Pr(X < X_0|d)$ is the probability that the response, X, is less than X_0 at dose d. For d=0, the latter equals the user-specified "tail probability" and X_0 is then a function of that tail probability and the estimated control-group response-distribution.	BMRF is the extra risk. (Default 0.5) This option also requires specifying a "tail probability" which is the probability of extreme ("adverse") response at dose=0.

Notes: m(x) is the median at dose x. Specifically, m(BMD) is the median at dose=BMD, so BMD is the solution to the equations shown. $\sigma(0)$ is the standard deviation for the control group (d=0) and $\sigma_L(0)$ is the log-scale standard deviation for the control group, used when the responses are assumed to be Lognormally distributed.

7.3.2 Polynomial Restriction

Restrictions on coefficients of the dose terms can be "Use dataset adverse direction" (default; depends on detected or specified adverse direction), "Non-negative" (>0), or "Non-positive" (<0).

7.3.3 Distribution and Variance

The underlying data distribution and the variability around the "center" of that distribution are linked options.

In total, three combinations are allowed:
1. Normal distribution, constant variance (default): each dose group has the same variance, which is estimated by BMDS along with the dose-response model parameters
2. Normal distribution, non-constant (modeled) variance: each dose group may have a different variance, described by a variance model (see Section 7.5.2) with two parameters (α and ρ) relating the dose group's estimated mean value (see below) to the variance. Those two parameters are estimated simultaneously with the parameters of the dose-response model.[6]
3. Lognormal distribution, constant coefficient of variation (CV): for lognormally distributed responses, each dose group has the same CV, which entails that the log-scale variance is constant over dose groups (though the natural-scale variance will differ from group to group[7]).

With respect to the response distribution (Normal or Lognormal), please note the following:

*   The lognormal distribution can only be assumed when the responses are strictly positive. The lognormal distribution is only applicable to positive real values.

*   Regardless of the distribution assumed, the dose-response model under consideration is the representation of the change in the median of the distribution of responses as a function of dose. If we denote the median at dose d by m(d), then it is always true for BMDS that m(d) = f(d), where f(d) is the dose-response

---

[6] The α parameter is returned for all models except for the exponential models, which return $ln(\alpha)$.
[7] CV = standard deviation divided by mean. Log-scale refers to the values of the logarithms of the responses. Natural-scale refers to the values of the responses, untransformed.

function under consideration (see list of possible functions in Table 6 on page 58).

- If the assumed data distribution is Normal, then it is also true that the mean at dose d, μ(d), is equal to the median. Thus, it is common under the Normal assumption to describe the dose-response function as a model of the mean response, and to write μ(d) = f(d), where f(d) is again one of the dose-response functions described in Table 6 on page 58.

When modeling continuous response data, the standard assumption for the BMDS continuous models is that the underlying distributions (one for each dose group) are Normal, with a mean given by the dose-response model and a variance as specified by the user (constant or a function of the mean response). An alternative assumption is that the responses are lognormally distributed.

In BMDS 3 all continuous models allow the user to choose between Normal and Lognormal response distribution assumptions; prior versions of BMDS only allowed this choice for Exponential models.

If the user has access to the individual response data, those data can be log-transformed prior to analysis but, as discussed below, this is not a recommended approach. If the user suspects that the responses are lognormally distributed, the recommended approach is to model the untransformed data assuming the underlying distribution is Lognormal with median values defined by the dose-response function and a constant log-scale variance, corresponding to an assumption of a constant CV.

7.3.3.1 Exact and Approximate MLE Solutions

The exact MLE solution cannot be obtained when the data are assumed to be Lognormally distributed and the data are presented in terms of group-specific means and standard deviations. In that case, the results are "Approximate" MLE solutions. The means (m_L) and standard deviations (s_L) of the log-transformed data are estimated as follows:

estimated log-scale sample mean (m_L): $$m_L = ln(m) - \frac{s_L^2}{2}$$

estimated log-scale sample standard deviation (s_L): $$s_L = \sqrt{\left(ln\left[1 + \frac{s^2}{m^2}\right]\right)}$$

where m and s are the sample mean and sample standard deviation, respectively.

When data are assumed to be lognormally distributed and individual response data are available, BMDS 3 provides an exact maximum-likelihood estimation (MLE) solution. As of BMDS 3.0, the exact solution is the ***only*** solution option implemented when individual observations are input, and the lognormal assumption is chosen. BMDS 3 does not provide an option for computing the approximate solution.

If the user wants to compute an approximate solution from individual observations (e.g., for research purposes), then they should use the following procedure:

1. Compute the group-specific sample means and sample standard deviations.
2. Input those values as would be done for an analysis based on those summary statistics (but still selecting lognormal as the distribution type).

7.3.3.2 Log-transformed Responses are NOT Recommended

Using log-transformed responses in the analysis is not recommended, for the following reasons:

- If the user chooses to log-transform the data prior to analysis, then the interpretation of the BMD and BMDL estimates would have to be considered carefully (and perhaps in consultation with a statistician). Data interpretation when using log-transformed responses will not be the same as when using the natural-scale response values. Indeed, the models—when "transformed back" to the natural scale—will not correspond to any of the standard BMDS models.

For example, if using the power model on log-transformed responses, the user is actually implicitly modeling the medians (on the natural-scale) with the function $e^{(background+slope\times dose)^{power}}$ which is not a standard BMDS model and whose characteristics (e.g., exponential increases in response) may not be those desired by the user.

- Similarly, the interpretation of the BMD will not correspond to simple expressions (e.g., if the BMR is set equal to a relative deviation of 10%, that relative deviation will be assessed on the log-scale and so will not yield BMD or BMDL estimates that correspond to a 10% change in the original mean responses).

For these reasons, log-transforming the response values is not considered a "best practice" and, as stated, should only be applied and interpreted with supporting statistical expertise.

Therefore, in most cases, the user should use non-transformed values and select the lognormal distribution if the data are assumed to be lognormally distributed.

7.4   Mathematical Details for Models for Continuous Endpoints in Simple Designs

Models in this section are for continuous endpoints, such as weight or enzyme activity measures, in simple experimental designs that do not involve nesting or other complications. The models predict the median value of the response, m(dose), expected for a given dose and the variation around that median.

As evidenced by the previous discussion of the options available for continuous models, modeling of continuous endpoints require consideration of more details than do those for dichotomous endpoints in similar designs. This section presents the mathematical and statistical details that determine how estimation is accomplished in BMDS.

7.4.1 Continuous Dose-Response Model Functions

The definitions of the continuous models are fully specified in the following table. Note that $m(\mathrm{dose})$ is the median response for the dose level specified.

Table 6. *The individual continuous models and their respective parameters.*

Model	Parameters	Notes
Linear and Polynomial models $m(\text{dose}) = g + \beta_1 \times \text{dose} + \beta_2 \times \text{dose}^2 + \cdots + \beta_n \times \text{dose}^n$ n is the degree of the polynomial, specified by user and must be a positive integer (maximum value = 21)	g = control response (intercept) $\beta_0 \ldots \beta_n$: polynomial coefficients	Parameter Constraints: none User parameter restriction options: can restrict the value of the polynomial coefficients. Restricting them to be either "non-positive" or "non-negative" guarantees that the resulting function will be strictly decreasing, strictly increasing, or perfectly flat (when all the coefficients are zero). If the coefficients are unrestricted (i.e., an unrestricted form of the model is run), more complicated shapes are possible, and, particularly as the degree of the polynomial approaches the number of dose groups minus one, the polynomial will often be quite ''wavy'.'
Linear $m(\text{dose}) = g + \beta \times \text{dose}$	g = control (intercept) β = slope	Parameter Constraints: none User parameter restriction options: none
Power $m(\text{dose}) = g + \beta \times (\text{dose})^\delta$	g = control response (intercept) β = slope δ= power	Parameter Constraints: $0 \le \delta \le 18$. User parameter restriction options: δ may be further restricted to values ≥ 1. Note: If $\delta < 1$, then the slope of the dose-response curve becomes infinite at the control dose. This is biologically unrealistic and can lead to numerical problems when computing confidence limits, so several authors have recommended restricting $\delta \ge 1$.
Hill[1] $m(\text{dose}) = g + \frac{v \times dose^n}{k^n + dose^n}$	g = control response (intercept) k = dose with half-maximal change n= power v= maximum change	Parameter Constraints: $k > 0$. $$0 \le n \le 18.$$ User parameter restriction options: n may be further restricted to values ≥ 1.
Exponential[1,2] $Exp2: m(\text{dose}) = a \times e^{\pm b \times \text{dose}}$ $Exp3: m(\text{dose}) = a \times e^{\pm (b \times \text{dose})^d}$ $Exp4: m(\text{dose}) = a \times (c - (c - 1) \times e^{-b \times \text{dose}})$ $Exp5: m(\text{dose}) = a \times (c - (c - 1) \times e^{-(b \times \text{dose})^d})$	a = control response (intercept) b = slope c= asymptote term d= power	Parameter Constraints: $a > 0$ $b > 0$ $c > 1$ for responses increasing with dose $0 < c < 1$ for responses decreasing with dose $1 \le n \le 18$. Note: The sign in "" $\pm b$" (Exp2 and Exp3 models) will change depending on the user-designated or auto-detected direction of change: + for responses increasing with dose, - for responses decreasing with dose.

[1] BMDL estimates from models that have an asymptote parameter (including the Hill model) can be unstable when a wide range of parameter values can give nearly identical likelihoods. One indicator of that problem is that the estimated asymptotic response is far outside the range of the observed responses. The user should consult a statistician if this behavior is seen or suspected.

[2] RIVM (National Institute for Public Health and the Environment (Netherlands)). ([RIVM, 2018](#)). PROAST.

Note that the upper bounds for the power parameters in the Power, Hill and Exponential models have been set to 18. That value was selected because it represents a very high degree of curvature that should accommodate almost every dataset, even ones with very (or absolutely) flat dose-response at low doses followed by a very steep dose-response at higher doses.

7.4.2 Variance Model

In addition to the model for the median response as a function of dose, the model for the variance also needs to be defined.

For responses assumed to vary Normally around the median, the variance model is:

$$\sigma_i^2 = exp\{ln(\alpha) + \rho * ln[m(\text{dose}_i)]\},$$

where α (> 0) and ρ are parameters estimated simultaneously with the parameters of the dose-response function (see Table 6 above). As in that table, $m(\text{dose}_i)$ is the predicted median (from the dose-response model under consideration) for the *ith* dose group.

Note that when a constant variance model is specified by the user, the parameter ρ is set to 0 and only α will be estimated. In that case, $\sigma_i^2 = \alpha$

When the responses are assumed to be Lognormally distributed, then the variance modeled is the log-scale variance:

$$\sigma_{Li}^2 = \alpha.$$

Because, for Lognormal data, BMDS is restricted to a constant log-scale variance model (equivalent to a constant coefficient of variation), ρ does not appear in that equation (in essence, it is once again set to 0 under the assumption of Lognormally distributed responses).

The formulation of the variance model shown above allows for several commonly encountered situations. If $\rho = 1$, then the variance is proportional to the median. If $\rho = 2$, then the coefficient of variation is constant, a common assumption especially for biochemical measures and one which mimics the constant coefficient of variation assumption of the Lognormally distributed responses (but without having to assume that the response are in fact Lognormally distributed).

7.4.3 Likelihood Function

For the "frequentist" modeling option, parameter estimates are derived by the method of maximizing the likelihood (i.e., they are maximum likelihood estimates, MLEs). The likelihood functions for the continuous responses are defined here.

Suppose there are G dose groups, having doses

$$\text{dose}_1, \dots, \text{dose}_G$$

with N_i subjects per dose group. Suppose also that y_{ij} is the measurement for the jth subject in the *ith* dose group. The form of the log-likelihood function depends upon whether the responses are assumed to be Normally or Lognormally distributed.

7.4.3.1 Assuming Normally Distributed Responses

For the assumption of **Normally distributed** responses, the log-likelihood function is:

$$LL = -\frac{N}{2}\ln(2\pi) - \sum_{i=1}^{G}\left[\frac{N_i}{2}ln(\sigma_i{}^2) + \frac{(N_i-1)s_i{}^2}{2\sigma_i{}^2} + \frac{N_i\left(\bar{y}_i - m(\text{dose}_i)\right)^2}{2\sigma_i{}^2}\right]$$

where

$\bar{y}_i = \frac{\sum_{j=1}^{N_i} y_{ij}}{N_i}$ (the sample mean for the *ith* dose group),

$s_i^2 = \frac{\sum_{j=1}^{N_i}(y_{ij}-\bar{y}_i)^2}{N_i-1}$ (the sample variance for the *ith* dose group),

$N = \sum_{i=1}^{G} N_i$.

The parameters defining $m(\text{dose}_i)$ and $\sigma_i{}^2$ (see previous two subsections) are optimized to maximize the LL equation value.

7.4.3.2 Assuming Lognormally Distributed Responses

For the assumption of **Lognormally distributed** responses, the log-likelihood function is:

$$LL = -\frac{N}{2}\ln(2\pi) - \sum_{i=1}^{G}\left[N_i\bar{z}_{Li} + \frac{N_i}{2}ln(\sigma_{Li}{}^2) + \frac{(N_i-1)s_{Li}{}^2}{2\sigma_{Li}{}^2} + \frac{N_i(\bar{z}_{Li} - ln(m(\text{dose}_i)))^2}{2\sigma_{Li}{}^2}\right]$$

where

$\bar{z}_{Li}$ = log-scale sample mean for ith dose group, and

$s_{Li}{}^2$ = log-scale sample variance for ith dose group.

As in the case of Normally distributed responses, the parameters defining $m(\text{dose}_i)$ and $\sigma_{Li}{}^2$ (see previous two subsections) are optimized to maximize the LL equation value.

7.4.4 AIC and Model Comparisons

The Akaike Information Criterion (AIC) ([Akaike, 1973](#)) can be used to compare different models fit (by the same fitting method, e.g., by maximizing the likelihood) to the same data set. The AIC is a statistic that depends on the value of LL (see previous section) and the number of estimated parameters, p:

$$AIC = -2 \times LL + 2 \times p$$

Note that the AIC balances the goals of getting the highest LL value possible while being parsimonious with respect to the number of parameters needed to achieve a high LL value. Since the equation for AIC has a negative multiplier for LL (which one wants to be greater) and positive multiplier for p (which one wants to be as small as possible and still get "good fit"), a model with a **smaller value of AIC** than other models is presumed to be the better model on the basis of AIC. Although such methods are not exact, they can provide useful guidance in model selection.

In the current version of BMDS, the number of "estimated parameters" includes only those that have not been estimated to equal a bounding value (either from the model-imposed constraints or user-imposed restrictions (see Table 6).

Note   This counting process may or may not be reasonable, depending on the boundary value that a parameter in question hits.

For example, if the power parameter in a model hits (i.e., is estimated to be equal to) the upper bound of 18, it would usually be the case that one would want to count that parameter as one that is estimated, but BMDS does **not** do that.

For this reason, the user is apprised to consider carefully the cases where parameter bounds have been hit and to consider the implications for issues such as model comparison and model selection.

Note that if a parameter hits a bound for any model, the parameter estimates are maximum likelihood estimates only in the restricted sense that the bounded parameter has been assigned a value and the other parameters are MLEs conditional on that assigned value. Such model results are not strictly comparable with others in terms of AIC. In such a case, the BMD and BMDL could depend on the choice of power parameter; thus, sensitivity analysis is indicated if one intends to rely on the reported BMD or BMDL. This is especially important when considering power parameters that have hit the upper bound of 18.

Note   To facilitate comparing models with different likelihoods (i.e., Normal vs. Lognormal), the log-likelihood is calculated using all of the terms shown in the LL equations. BMDS 2.x did not include the parameter-independent terms in its reported LL values nor for calculating the AIC. As a consequence, BMDS 3.x and 2.x will return different values for the LL and AIC when run on the same data.

Even though the BMDS 3 AIC values for continuous models differ from those in BMDS 2.x versions, if the models have the same underlying distribution, then *the difference of the AICs* will be the same as in previous versions of BMDS. This assumes that the BMDS 3 and BMDS 2.x model fits are the same for the two models being compared. The AIC difference may not be the same if one or more of the model fits differ between the two versions (e.g., if one or more of the BMDS 3 models provide an improved fit to the data over the corresponding BMDS 2.x model).

However, when comparing models having different parametric distributions, the AIC differences will not be the same as previous BMDS versions. For these comparisons, the AIC calculated using the BMDS 3 software is correct and will result in the proper comparison between any two models regardless of underlying distribution.

Caution

A note of caution is required for situations where only the sample mean and sample standard deviation are available (summarized data) and for which the log-scale sample mean and sample standard deviation are only approximated when assuming lognormally distributed responses.

In such cases, the same approximations are made across all the dose-response models. It is therefore strictly valid to compare AIC results across any runs that assumed that the responses were Lognormally distributed.

However, **comparisons of results where one set of results was obtained assuming Normality and one set was obtained assuming Lognormality should be made with caution.**

In the latter case, if the AICs are "similar" (using that term loosely, because no specific guidance can be offered here), then the user ought not to base model selection on AIC differences. Selection when the AIC differences are "larger" may not be problematic, since the approximation used should not be too bad.

A conservative position would be that comparisons of model runs assuming Normally distributed responses to those assuming Lognormal responses should not be made

using the AIC, if the underlying data are presented in summarized form (i.e., only sample means and sample standard deviations are available).

7.4.5 BMDL and BMDU Computation

The estimation of the BMDs, depending on the definition of the BMR type, is specified in Table 5. The derivation of the confidence bounds for the BMD, i.e., the BMDL and BMDU, is defined here.

The general approach to computing the confidence limits for the BMD is the same for all the models in BMDS, and is based on the asymptotic distribution of the likelihood ratio (Crump and Howe, 1985).

Two different specific approaches are followed for the continuous models.

For the Power Model
For the power model, the equations that define the benchmark response in terms of the benchmark dose and the dose-response model are solved for one of the model parameters. The resulting expression is substituted back into the model equations, with the effect of re-parameterizing the model so that BMD appears explicitly as a parameter. A value for BMD is then found such that, when the remaining parameters are varied to maximize the likelihood conditional on that BMD value, the resulting log-likelihood is less than that at the maximum likelihood estimates by exactly

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

For the Polynomial, Hill, and Exponential Models

For the polynomial, Hill, and exponential models, it is impractical or impossible to explicitly reparameterize the dose-response model function to allow BMD to appear as an explicit parameter. For these models, the BMR equation is used as a non-linear constraint, and the minimum value of BMD is determined such that the log-likelihood is equal to the log-likelihood at the maximum likelihood estimates less

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

Occasionally, the following error message may appear for a model: "BMDL computation is at best imprecise for these data." This is a flag that convergence for the BMDL was not "successful" in the sense that the required level of convergence (< 1e-3 relative change in the target function by the time the optimizer terminates) has not been achieved.

7.4.6 Bayesian Continuous Model (Preview) Descriptions

Note   At this time, EPA does not offer technical guidance on Bayesian modeling.

Note   The preview models are new. Users acknowledge they have not been extensively tested, and formally reviewed and approved by the EPA for risk assessment purposes.

From a Bayesian perspective, inference proceeds by defining a data generating mechanism, given a model, M, and its parameters. For our purposes, M would be one of the models listed in Table 6 which determines the probability of response. The data generating mechanism would be the assumption that the observations were obtained from a Normal or Lognormal distribution having the dose-dependent median response defined by one of those models and the variance term defined as described above (according to the user's choice of options).

We can then relate that to the likelihood, here denoted $\ell(D|M)$, which shows explicitly that it is the likelihood of the data, D, conditional on the model. The functional form of the log of the likelihood is presented in Section 7.4.3, "Likelihood Function," on page 59.

The set of preview Bayesian continuous models used in BMDS 3 is identical to the set of models used for frequentist (MLE) approaches (Table 6). In the following, let θ be the vector of parameters that are required to define the any one of those models. So, for example, for the Power model θ = (g, β, δ). The additional consideration incorporated into the Bayesian approach is the specification of a prior distribution for θ. The Bayesian approach takes the specified prior and updates it using the data under consideration to obtain a "posterior" distribution for θ.

BMDS summarizes the posterior for the BMD as follows. The BMD (one might call it a "central estimate" perhaps) is equated to the value obtained by using the maximum *a posteriori* (MAP) θ estimate. The MAP is the value of θ that maximizes the posterior log-likelihood. The posterior density is itself approximated using a Laplacian approximation. This approximation is also used to estimate BMDL and BMDU values, which are the percentiles from that density that correspond to the confidence level selected by the user.

The priors for the BMDS dichotomous models are defined in Table 7.

Important Note

The priors for the parameters are based on scaled doses and scaled responses; BMDS performs this scaling automatically.

BMDS automatically scales the doses by dividing by the maximum dose in the dataset under consideration, i.e., that the doses under consideration range from 0 to 1 (inclusive). **BMDS automatically scales the responses** by dividing by the mean response in the control (or lowest dose) group.

The user does *not* need to scale anything beforehand. That means that the parameter estimates and BMD values returned by the program have been adjusted back to the original scale of the doses and the original scale of the responses specified in the input data file.

Table 7. *The individual Bayesian continuous models (preview) and their parameter priors.*

Model	Constraints	Priors
Polynomial $p(\text{dose}) = g + \beta_1\text{dose} + \beta_2\text{dose}^2 + \cdots + \beta_n\text{dose}^n$	$0 < g < 1e6$ $-1e-4 < \beta_i < 1e4$	$g \sim Lognormal(0,1)$ $\beta_i \sim Normal(0,2)$
Linear $p(\text{dose}) = g + \beta_1\text{dose}$	$0 < g < 1e6$ $-1e-4 < \beta_1 < 1e4$	$g \sim Lognormal(0,1)$ $\beta_1 \sim Normal(0,2)$
Power $p(\text{dose}) = g + \beta \times (\text{dose})^\delta$	$0 < g < 1e6$ $-1e-4 < \beta < 1e4$ $0 < \delta < 40$	$g \sim Lognormal(0,1)$ $\beta \sim Normal(0,1)$ $\delta \sim Lognormal(0.405465, 0.5)$
Hill $p(\text{dose}) = g + \frac{v \times dose^n}{k^n + dose^n}$	$0 < g < 18$ $-18 < v < 18$ $0 < k < 18$ $0 < n < 18$	$g \sim Lognormal(0,1)$ $v \sim Normal(1,2)$ $k \sim Lognormal(-0.69315, 1)$ $n \sim Lognormal(0.405465, 0.2501)$
Exponential $p(\text{dose}) = a \times e^{\pm b \times \text{dose}}$ $p(\text{dose}) = a \times e^{\pm(b \times \text{dose})^d}$ $p(\text{dose}) = a \times (c - (c-1) \times e^{-b \times \text{dose}})$ $p(\text{dose}) = a \times (c - (c-1) \times e^{-(b \times \text{dose})^d})$	$0 < a < 1e6$ $0 < b < 18$ $-20 < ln(c) < 20$ $0 < d < 18$	$a \sim Lognormal(0,1)$ $b \sim Lognormal(0,1)$ $ln(c) \sim Normal(0,1)$ $d \sim Lognormal(0, 0.2501)$

For all models, ρ and α are the parameters of the Variance model: $\sigma_i^2 = exp\{ln(\alpha) + \rho * ln[m(\text{dose}_i)]\}$. The prior for $ln(\alpha)$ in all models is $Normal(0,1)$, constrained to fall between -18 and 18. The parameter ρ is only assigned a prior as specified above when the variance is defined to be non-constant; in that case the prior for $ln(\rho)$ is $Normal(0, 0.2501)$, constrained so that ρ is between 0 and 18. Normal(x, y) denotes a Normal distribution with mean x and standard deviation y. Lognormal(w, z) denotes a lognormal distribution with log-scale mean w and log-scale standard deviation z.

> ***Note*** The preview Bayesian continuous models are new. Users acknowledge they have not been extensively tested, and formally reviewed and approved by the EPA for risk assessment purposes.

7.5   Outputs Specific to Frequentist Continuous Models

Figure 46. Sample Results Workbook tab for a frequentist continuous model run.
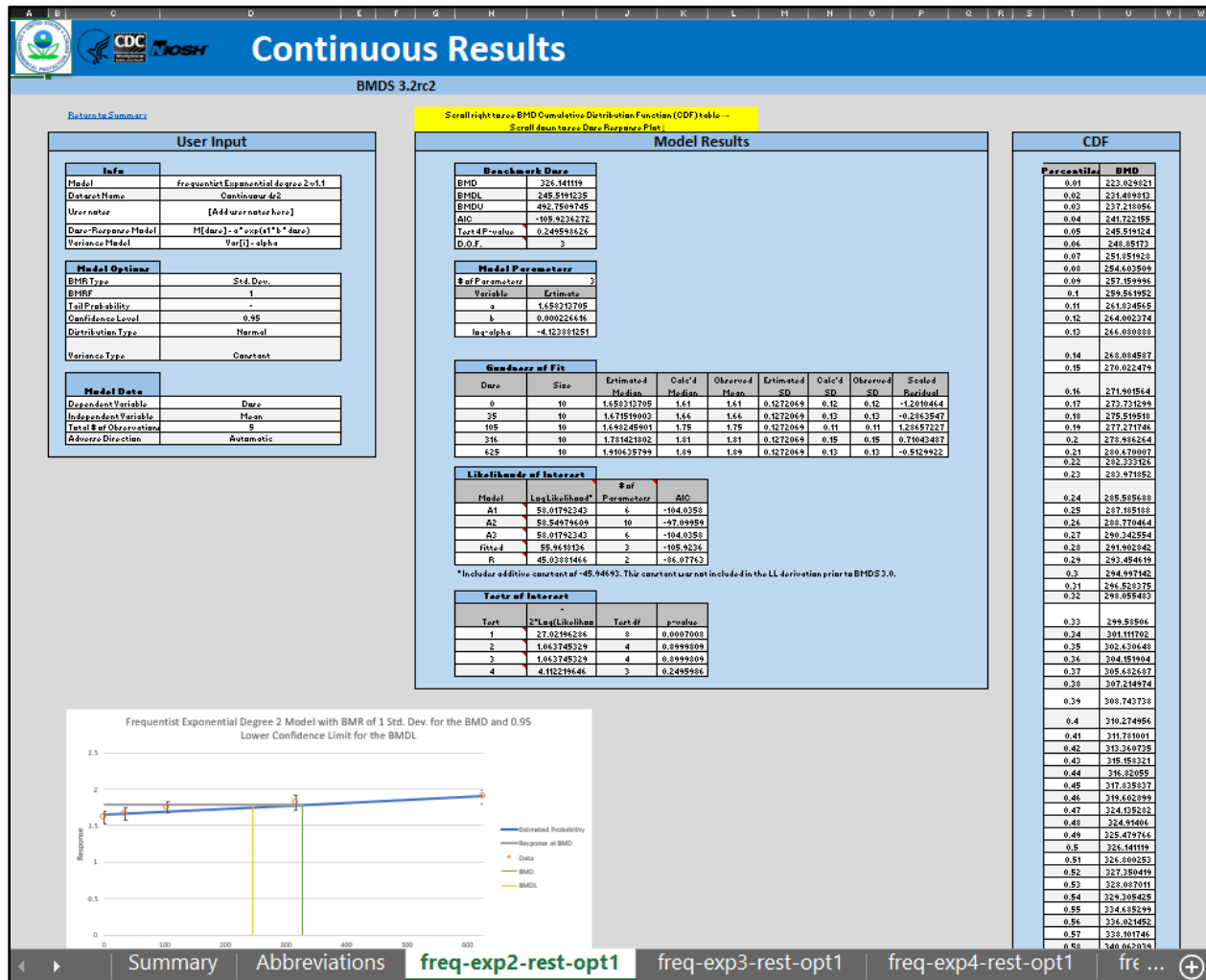
7.5.1 Goodness of Fit Table

Figure 47. Sample Goodness of Fit table, with Normal assumption.

Goodness of Fit		Estimated Median	Calc'd Median	Observed Mean	Estimated SD	Calc'd SD	Observed SD	Scaled Residual
Dose	Size							
0	10	1.658313705	1.61	1.61	0.12720687	0.12	0.12	-1.201046387
35	10	1.671519003	1.66	1.66	0.12720687	0.13	0.13	-0.286354713
105	10	1.698245901	1.75	1.75	0.12720687	0.11	0.11	1.286572266
316	10	1.781421802	1.81	1.81	0.12720687	0.15	0.15	0.710434871
625	10	1.910635799	1.89	1.89	0.12720687	0.13	0.13	-0.512992166

Figure 48. Sample Goodness of Fit table, with Lognormal assumption.

Goodness of Fit		Estimated Median	Calc'd Median	Observed Mean	Estimated GSD	Calc'd GSD	Observed SD	Scaled Residual
Dose	Size							
0	10	1.648797808	1.6055465	1.61	1.12433068	1.077271	0.12	-0.109122205
35	10	1.663398431	1.65493293	1.66	1.12433068	1.081332	0.13	-0.009558383
105	10	1.692599677	1.74655307	1.75	1.12433068	1.064809	0.11	0.161443391
316	10	1.780620575	1.80381636	1.81	1.12433068	1.08625	0.15	0.082632182
625	10	1.909523218	1.8855449	1.89	1.12433068	1.071117	0.13	-0.054910745

The **Goodness of Fit** table displays the model predictions relative to the observed (or "calculated") data that were used as input, one row for each dose group. Generally, one desires to have the model predictions match the input data as well as possible. Note that in this table

- Estimated Median = model predicted median (which equals the mean in the case of Normally distributed data)

- Calc'd Median = the median computed from the observed data. In the case of Normally distributed data, this equals the sample mean. In the case of Lognormally distributed responses, the median is calculated as $exp(z_L)$, where z_L is the log-scale mean, estimated if need be for summarized response data as shown in Section 7.3.3, "Distribution and Variance," on page 55.

- Observed mean = the sample mean.

- Estimated [G]SD = the standard deviation (or geometric standard deviation, in the case of Lognormally distributed data) estimated by the model.

- Calc'd [G]SD = the standard deviation (or geometric standard deviation, in the case of Lognormally distributed data) computed from the observed data. In the case of Normally distributed data, this equals the sample standard deviation of the responses. In the case of Lognormally distributed data, this equals $exp(s_L)$, where s_L is the log-scale standard deviation, estimated if need be for summarized response data as shown in Section 7.3.3, "Distribution and Variance," on page 55.

- Observed SD = sample standard deviation.

- Scaled Residual = For Normal responses, this equals

$$(Calc'd\ Median - Estimated\ Median)/(Estimated\ SD/\sqrt{N_i})$$

whereas for Lognormal responses, the scaled residual equals

$$(ln(Calc'd\ Median) - ln(Estimated\ Median))/(ln(Estimated\ GSD)/\sqrt{N_i})$$

The scaled residual value is a "calibrated" difference between the observations and the model predictions. Regardless of the assumption about the distribution of the responses, it is computed on the scale corresponding to the Normal distribution. Moreover, the denominator for its calculation estimates the degree of uncertainty (standard error of the mean) for the model prediction. Therefore, scaled residual values greater than about 2 in absolute value are indicative of mismatches between predicted and observed values that may indicate poorer fit, at least locally.

7.5.2 Likelihoods of Interest Table

Likelihoods of Interest			
Model	Log Likelihood*	# of Parameters	AIC
A1	58.01792343	6	-104.03585
A2	58.54979609	10	-97.099592
A3	58.01792343	6	-104.03585
fitted	55.9618136	3	-105.92363
R	45.03881466	2	-86.077629

* Includes additive constant of -45.94693. This constant was not included in the LL derivation prior to BMDS 3.0.

The **Likelihoods of Interest** table displays the log-likelihoods, number of parameters, and AIC for five models, including the dose-response model under consideration ("fitted"). Recall that BMDS uses likelihood theory to estimate model parameters and ultimately to make inferences based on risk assessment data. Maximum likelihood is the process of estimating the model parameters; the likelihood function is as large as possible (maximized) given the form of the model under consideration and the data. In other words, parameter values are "chosen" such that the subject model (e.g., polynomial or power) obtains the best possible fit to the data, given the constraints of the model's parameter structure.

As noted previously, the number of parameters for each model excludes parameters that have values on one of the bounds set for their estimation (either bounds specified by the user or those inherent constraints associated with the model; see Table 6 on page 58).

The five log-likelihood models can be used for tests of hypotheses, including tests of fit, that are asymptotically chi-squared**.** that may be of interest to the user. Each of these log-likelihood values corresponds to a model the user may consider in the analysis of the data. The five models are summarized in the following table.

Table 8. *Likelihood values and models for continuous endpoints.*

Model	Description
A1: "Full" Constant Variance Model	$Y_{ij} = \mu_i + e_{ij},$ $Var\{e_{ij}\} = \sigma^2$
A2: "Fullest" Model	$Y_{ij} = \mu_i + e_{ij},$ $Var\{e_{ij}\} = \sigma_i{}^2$
A3: "Full" Model with variance structure specified by the user	$Y_{ij} = \mu_i + e_{ij},$ $Var\{e_{ij}\} = \alpha \times \mu_i{}^\rho$
R: "Reduced" Model	$Y_i = \mu + e_i,$ $Var\{e_i\} = \sigma^2$
Fitted Model	The user-specified model

Model A1 estimates separate and independent means for the observed dose groups (it is "full" or "saturated" in that respect) but posits a constant variance over those groups.

Model A2 is the "fullest" model in that it estimates separate and independent means for the observed dose groups (as in Model A1) and it also estimates separate and independent variances for those groups. There is no assumed functional relationship among the means or among the variances across dose groups. This model is often referred to as the "saturated" model (it has as many mean and variance parameters as

there are dose groups). The log-likelihood obtained for this model is the maximum attainable, for the data under consideration.

Model A3 is similar to model A2 and may only differ with respect to its variance parameters. Model A2 estimates separate and independent means for the observed dose groups (like A1). If the user specifies a constant variance for the fitted model, then model A3 will also assume that and it becomes identical to Model A1. If the user assumes a non-constant variance for the fitted model, then Model A3 will also assume the same functional form for the variance.

The reduced model (R) is the model that implies no difference in mean or variance over the dose levels. In other words, it posits a constant mean response level with the same variance around that mean at every dose level.

The last model, the fitted model, is the user-specified model (e.g., power or polynomial, among others). A user may have reason to believe that a certain model may describe the data well, and thus uses it to calculate the BMD and BMDL.

7.5.3 Tests of Interest Table (Tests of Fit)

Tests of Interest			
Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	27.02196286	8	0.00070084
2	1.063745329	4	0.89998092
3	1.063745329	4	0.89998092
4	4.112219646	3	0.24959863

The **Tests of Interest** table shows the results of four tests based on the log-likelihoods from the Likelihoods of Interest table. The p-values associated with the tests are based on asymptotic properties of the likelihood ratio.

Without getting too technical, the likelihood ratio is just the ratio of two likelihood values, many of which are given in the BMDS output. Statistical theory proves that $-2 *$ $\ln(\text{likelihood ratio})$ converges to a Chi-Square random variable as the sample size gets large and the number of dose levels gets large. These values can in turn be used to obtain approximate probabilities to make inferences about model fit. Chi-Square tables can be found in almost any statistical reference book.

Suppose the user wishes to test two models, A and B, for fit. One assumption that is made for these tests is that model A is "nested within" Model B, i.e., that Model B can be simplified (via restriction of some parameters in Model B) in such a way that the simplified model is Model A. This implies that Model A has fewer varying parameters. As an example, consider that the linear model is a "simpler" or "nested" model relative to the power model because the linear model has the power parameter restricted to be equal to 1.

Note   The model with a higher number of parameters is always in the denominator of this ratio.

Suppose that $L(X)$ represents the likelihood of model X. Now, using the theory, $-2 \times$ $\ln\{\frac{L(A)}{L(B)}\}$ approaches a Chi-Square random variable. This can be simplified by using the fact that the log of a ratio is equal to the difference of the logs, or put,

$$-2 \times \ln\left\{\frac{L(A)}{L(B)}\right\} = -2 \times (\ln\{L(A)\} - \ln\{L(B)\}) = 2 \times \ln\{L(B)\} - 2 \times \ln\{L(A)\}$$

The values in the Likelihoods of Interest table are in fact the log-likelihoods, as discussed above, $\ln\{L(B)\}$ and $\ln\{L(A)\}$, so this likelihood ratio calculation becomes just a subtraction problem. This value can then in turn be compared to a Chi-Square random variable with a specified number of degrees of freedom.

As mentioned in conjunction with the Likelihoods of Interest table, each log-likelihood value has an associated number of parameters. The number of degrees of freedom for the Chi-Square test statistic is merely the difference between the two model parameter counts. In the mini-example above, suppose Model A has 5 fitted parameters, and that Model B has 8. In this case, the Chi-Square value to be compared to would be a Chi-Square with 8 - 5 = 3 degrees of freedom.

In the A vs B example, what is exactly being tested? In terms of hypotheses, it would be:

H_0: A models the data as well as B

H_1: B models the data better than A

Keeping these tests in mind, suppose $2 \times \log\{L(B)\} - 2 \times \log\{L(A)\} = 4.89$ based on 3 degrees of freedom. Also, suppose the rejection criteria is a Chi-Square probability of less than .05. Looking on a Chi-Square table, 4.89 has a p-value somewhere between .10 and .25. In this case, H_0 would not be rejected, and it would seem to be appropriate to model the data using Model A. BMDS automatically does the "table look-up" for the user, and provides the p-value associated with the calculated log-likelihood ratio having degrees of freedom as described above.

The Tests of Interest table provides four default tests. Associated with each of those tests is a "hover box" that can be accessed to show a summarized interpretation of the test results, which includes EPA's interpretation of the test results (i.e., in relation to p-values that have been selected by EPA). However, the computed p-values are presented so that the users are free to use any rejection criteria they want. Each of the four default tests provided for any of the continuous models is discussed in some detail below.

Test 1 (A2 vs R): Tests the null hypothesis that responses and variances don't differ among dose levels. If this test fails to reject the null hypothesis, there may not be a dose-response.

This test compares Model R (the simpler model) to Model A2. Model R is a simpler A2 (or nested within A2) since R can be obtained from A2 by restricting all the mean parameters to be equal to one another and restricting all the variance parameters to be equal to one another. If this test fails to reject the null hypothesis, then there may not be a dose-response, as the inference would be that the simpler model (R) is not much worse than the saturated model. The default p-value for the test (as reported in the Tests of Interest section of the output) is 0.05. A p-value less than 0.05 is an indication that there is a difference between response and/or variances among the dose levels and supports a conclusion to model the data. A p-value greater than 0.05 is an indication that the data may not be suitable for dose-response modeling.

Test 2 (A1 vs A2): Tests the null hypothesis that variances are homogeneous. If this test fails to reject the null hypothesis, the simpler constant variance model may be appropriate.

This test compares A1 (the simpler model) to Model A2. Model A1 is a simpler A2 (or nested within A2) since A1 can be obtained from A2 by restricting all the variance parameters to be equal to one another. If this test rejects the null hypothesis, the inference is that the constant variance assumption is incorrect and a modeled variance is necessary to adequately represent the data. The default p-value for the test (as reported in the Tests of Interest section of the output) is 0.05. A p-value less than 0.05 is an indication that the user should consider running a non-homogeneous variance model. A p-value greater than 0.05 is an indication that a constant variance assumption may be suitable for the dose-response modeling.

Test 3 (A3 vs A2): Tests the null hypothesis that the variances are adequately modeled. If this test fails to reject the null hypothesis, it may be inferred that the variances have been modeled appropriately.

Here, the test is one to see if the user-specified variance model, is appropriate. If the user-specified variance model is "constant variance," then Models A1 and A3 are identical; this test is the same as Test 2, with the same interpretation. If the user-specified variance model is nonconstant ($\sigma_i^2 = \alpha \times \mu_i^\rho$), this test determines if that equation appears adequate to describe the variance across dose groups. Model A3 is the simpler version of Model A2 obtained by constraining the variances to fit the nonconstant variance equation. The default p-value for the test (as reported in the Tests of Interest section of the output) is 0.05. A p-value less than 0.05 is an indication that the user may want to consider a different variance model. A p-value greater than 0.05 supports the use of modeled variance for the dose-response modeling.

Test 4 (Fitted vs A3): Tests the null hypothesis that the model for the mean fits the data. If this test fails to reject the null hypothesis, the user has support for the selected model.

This test compares the Fitted Model to Model A3. The Fitted Model is as simpler Model A3 (or nested within Model A3) because it can be obtained by restricting the means (unrestricted in A3) to be described by the dose-response function under consideration. If this test fails to reject the null hypothesis, the inference is that the fitted model is adequate to describe the dose-related changes in the means (conditional on the form of the variance model; the form of the variance model is the same for the Fitted Model and Model A3). Failure to reject the null hypothesis is associated with the inference that the restriction of the means to the shape of the dose-response function under consideration is adequate. The default p-value for the test (as reported in the Tests of Interest section of the output) is 0.1. A p-value less than 0.1 is an indication that the user may want to try a different model (i.e., the fit of the Fitted Model is not good enough). A p-value greater than 0.1 is an indication that the Fitted Model appears to be suitable for dose-response modeling.

7.5.4 Plot and Error Bar Calculation

Figure 49. Frequentist results plot for continuous data.
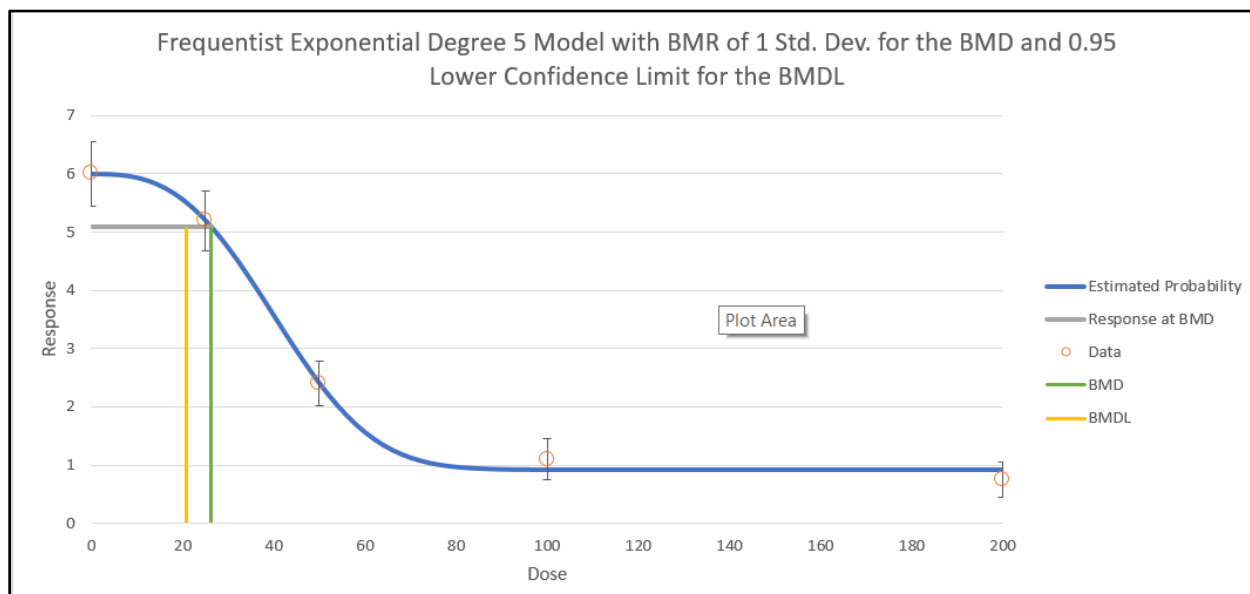
The graphical output, i.e., plot, is a visual depiction of the results of the modeling. Because plots common to all models were discussed in Section 6.5, "Graphical Output," on page 49, here we describe the one additional detail specific to the continuous models, i.e., computation of the error bars:

- The plotting routine calculates the standard error of the mean (SEM) for each group. The routine divides the group-specific observed variance (obs standard deviation squared) by the group-specific sample size.
- The routine then multiplies the SEM by the Student-T percentiles (2.5th percentile or 97.5th percentile for the lower and upper bound, respectively) appropriate for the group-specific sample size (i.e., having degrees of freedom one less than that sample size). The routine adds the products to the observed means to define the lower and upper ends of the error bar.

7.6   Outputs Specific to Bayesian Continuous Models (Preview)

Figure 50. *Bayesian results plot for continuous data.*

To compare the difference between any two Bayesian models, the unnormalized Log Posterior Probability (LPP) is given, which allows the computation of a Bayes factor (BF) to compare any two models. BF equals the exponentiated difference between the two LPP. For example, if one wishes to compare the Log-Logistic model (Model A) (yielding LPP$_A$) to the Multistage 2nd degree model (Model B, with LPP$_B$) one estimates the BF as

$$BF = \exp(LPP_A - LPP_B),$$

This computation assumes that both models have equal probability *a priori.* This value is then interpreted as the posterior odds one model is more correct than the other model and is used in Bayesian hypothesis testing. In the example above, if the Bayes Factor was 2.5, the interpretation would be that the Log-logistic model is *a posteriori* 2.5 times more likely than the multistage model. When these values are normalized into proper probabilities, they are equivalent to the posterior model probabilities given in model averaging (again, assuming equal model probability *a priori*). The table below, adapted from Jeffreys (1998) is a common interpretation of Bayes Factors.

Table 9. *Bayes factors for continuous models.*

Bayes Factor	Strength of Evidence for H$_A$
< 1	negative (supports H$_B$)
1 to 3.2	not worth mentioning
3.2 to 10	substantial
10 to 31.6	strong
31.6 to 100	very strong
100	decisive

For BMDS 3.0, all LPP and corresponding posterior model probabilities are computed using the Laplace approximation. This value is different from the commonly used Bayesian Information Criterion (BIC), and the two should not be confused based upon other model averaging approaches, which use the BIC exclusively. Errors in the posterior probabilities estimated from the BIC are $O(1)$ estimators. Errors in the posterior probabilities estimated using the Laplace approximation are $O(n^{-1})$. This means the latter approximation goes to the true posterior model probability with increasing data and the former, using the BIC, may not go to the true value.

8.0   Dichotomous Endpoints

BMDS includes models for dichotomous endpoints in which the observations are independent of each other. In these models, the dose-response model defines the probability that an experimental unit (e.g., a rat or a mouse in a test of toxicity) will have an adverse response at a given dose. The actual number of animals that have an adverse response is assumed to be binomially distributed.

A specific example of such a dataset is a study in which adult animals are exposed to different concentrations of a toxicant and then evaluated for the presence of liver toxicity.

For models for dichotomous endpoints in which the responses are nested (for example, pups within litters, and litters nested within doses), see Section 9.0, "Nested Dichotomous Endpoints," on page 91.

For dichotomous cancer models, and the combination of model predictions for multiple tumor endpoints, see Section 10.0, "Multiple Tumor Analysis," on page 105.

This section provides details on the following topics:
- Implementation of dichotomous models in BMDS 3
- Entering dichotomous model data
- Dichotomous model options
- Dichotomous model outputs
- Options for restricting values of certain model parameters
- The Bayesian approach to dichotomous response modeling, specifying how the priors are defined and the methods for model averaging

8.1   Dichotomous Response Models

BMDS 3 offers the traditional frequentist dichotomous response models available in previous versions of BMDS plus Bayesian versions of each model, and a Bayesian model averaging feature.

Figure 51. *BMDS 3 dichotomous models, as they appear in the Analysis Workbook.*

	MLE		Alternatives		
	Frequentist Restricted	**Frequentist Unrestricted**	**Bayesian**	**Bayesian Model Average**	
Model Name	**Enable ☐**	**Enable ☐**	**Enable ☐**	**Enable ☐**	**Prior Weights**
Dichotomous Hill	☑	☐	☐	☐	0.0000%
Gamma	☑	☐	☐	☐	0.0000%
Logistic	▨	☑	☐	☐	0.0000%
Log-Logistic	☑	☐	☐	☐	0.0000%
Log-Probit	☐	☑	☐	☐	0.0000%
Multistage	☑	☐	☐	☐	0.0000%
Probit	▨	☑	☐	☐	0.0000%
Quantal Linear	▨	☐	☐	☐	0.0000%
Weibull	☑	☐	☐	☐	0.0000%
				Total Weight	0.000%

Most frequentist models can be run restricted or unrestricted. The EPA default recommendation for initial runs is to restrict the Dichotomous Hill, Gamma, Log-Logistic, Multistage, and Weibull models and un-restrict the Log-Probit model; the Logistic, Probit, and Quantal Linear models have no restricted option (U.S. EPA, 2012).

See Table 10 on page 76 below for the effect of the user selecting the restricted version of the models ("User parameter restriction options"). In general, the restrictions prevent the slope of the dose-response curve from becoming infinite at 0 dose. This is often considered to be biologically unrealistic and can lead to numerical problems when computing confidence limits, so several authors have recommended restricting the appropriate parameter.

8.2   Options

On the BMDS 3 Analysis Workbook's Main tab, the user can define multiple Option Sets to apply to multiple user-selected models and multiple user-selected datasets in a single "batch" process. Select the **Add Option Set** button to define a new Option Set configuration.

Figure 52. Dichotomous Model options.

Option Set #	Risk Type	BMR	Confidence Level	Background
1	Extra Risk ▼	0.1	0.95	Estimated ▼
		Add Option Set		

8.2.1 Risk Type

Option Set #	Risk Type
1	Extra Risk ▼
	Extra Risk
	Added Risk

Choices for the risk type are "Extra" (Default) or "Added."

Added risk is the additional proportion of the total experimental units that respond in the presence of the dose, or the predicted probability of response at dose d, $P(d)$, minus the predicted probability of response in the absence of exposure, $P(0)$:

$$Added\ risk\ at\ dose\ d\ = P(d) - P(0).$$

Extra risk is the additional risk divided by the predicted proportion of animals that will not respond in the absence of exposure $(1 - P(0))$:

$$Extra\ risk\ at\ dose\ d\ = \frac{P(d) - P(0)}{1 - P(0)}.$$

8.2.2 BMR

BMR
0.1

The BMR is the value of risk (extra or added, as specified by the user) for which a BMD is estimated. BMR must be between 0 and 1 (not inclusive). If $P(0) > 0$, then values for BMR greater than $1 - P(0)$ will result in an error when the risk type is added risk. That is

because the maximum added risk that can ever be achieved is $1 - P(0)$. In practice, this should not typically be an issue because one usually is interested in BMR values in the range of 0.01 to around 0.10.

BMR and Graphs
The response associated with the BMR that is displayed in the graphical model output will only be the same as the BMR when $P(0) = 0$.

This is because to obtain the actual response value one must solve for $P(d)$ in the equation for added or extra risk discussed above.

The horizontal bar depicting the response level used to derive the BMD that is displayed in the graphical model output will only be the same as the user-defined BMR (e.g., 10% Extra Risk) when the response at background, P(0), equals zero.

When P(0) does not equal zero, the true response level can be calculated using the Extra Risk equation described above.

8.2.3 Confidence Level

Confidence Level
0.95

The Confidence Level is a fraction between 0 and 1; 0.95 is recommended by EPA (U.S. EPA, 2012).

The value for confidence level must be between 0 and 1 (not inclusive). For a confidence level of x, BMDS will output BMDL and BMDU estimates, each of which is a one-sided confidence bound at level x. For example, if the user sets the confidence level to 0.95 (the default), then the BMDL is a 95% one-sided lower confidence bound for the BMD estimate; the BMDU is a 95% one-sided lower confidence bound for the BMD estimate. In that example, the range from BMDL to BMDU would constitute a 90% confidence interval (5% in each tail outside that interval).

8.2.4 Background

nce	Background
	Zero ▼
	Estimated
	Zero

The user may specify if the background parameter is to be estimated (the default) or is set to zero before fitting and estimation of the other parameters.

Note   It is very important that the user **NOT** set the background to zero if there are more than zero responses in a dose group that has a dose value of zero. This would represent a logically impossible situation, i.e., where the probability of response in that dose group would be zero and yet there were indeed some responses.

8.3   Mathematical Details for Models for Dichotomous Endpoints in Simple Designs

BMDS contains nine models for dichotomous endpoints as defined in the following table.

Table 10. *The individual dichotomous models and their respective parameters*.

Model	Parameters	Notes
Multistage $p(\text{dose}) = g + (1-g)\left(1 - \exp\left[-\sum_{j=1}^{n} \beta_j \text{dose}^j\right]\right)$	g = background β_j = dose coefficients	Parameter Constraints: $0 \leq g < 1$ $$n \leq 23$$ User parameter restriction options: can restrict all β coefficients to $\geq$ 0. Doing so will guarantee that the multistage model will be either perfectly flat or always increasing. Per EPA Guidance, when the Multistage model is used for cancer analyses (e.g., in Multi-tumor analyses) all β coefficients are restricted to be non-negative.
Weibull (and Quantal Linear) $p(\text{dose}) = g + (1-g)(1 - \exp[-\beta \text{dose}^\alpha])$	g = background α = power β = slope	Parameter Constraints: $0 \leq g < 1$ $$0 < \alpha \leq 18$$ $$0 < \beta$$ User parameter restriction options: $1 \leq \alpha$ The **Quantal Linear** model results from setting α = 1.
Gamma $p(\text{dose}) = g + \frac{1-g}{\Gamma(\alpha)} \int_0^{\beta d} t^{\alpha-1} \exp(-t)\, dt$	g = background α = power β = slope	Parameter Constraints: $0 \leq g < 1$ $$0.2 < \alpha \leq 18$$ $$0 < \beta$$ User parameter restriction options: $1 \leq \alpha$ Note that for the unrestricted Gamma model, α > 0.2 for numerical reasons.
Logistic $p(\text{dose}) = \frac{1}{1+\exp[-\alpha-\beta(\text{dose})]}$	α = intercept β = slope	Parameter Constraints: $0 < \beta$ User parameter restriction options: none.

Model	Parameters	Notes
Log-Logistic $p(\text{dose}) = g + \frac{1-g}{1+\exp[-\alpha - \beta \log(\text{dose})]}$	g = background α = power β = slope	Parameter Constraints: $0 \leq g < 1$ $\qquad 0 < \alpha \leq 18$ $\qquad 0 < \beta$ User parameter restriction options: $1 \leq \alpha$
Probit $p(\text{dose}) = \Phi(\alpha + \beta\text{dose})$, where $\Phi(x) = \int_{-\infty}^{x} \phi(t)\mathrm{d}t$ and $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$	α = intercept β = slope	Parameter Constraints: $0 < \beta$ User parameter restriction options: none. Φ is the standard normal cumulative distribution function, ϕ is the standard normal density function
Log-Probit $\qquad p(\text{dose}) = g + (1-g)\Phi[\alpha + \beta \log(\text{dose})]$ $\Phi(x) = \int_{-\infty}^{x} \phi(t)\mathrm{d}t$ and $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}$	g = background α = intercept β = slope	Parameter Constraints: $0 \leq g < 1$ $\qquad 0 < \beta \leq 18$ User parameter restriction options: $1 \leq \beta$ For the log-probit model, the slope of the model will always approach zero as dose approaches zero regardless of the restriction on β. However, depending on the data and parameter estimates, the slope for the log-probit model, for some relatively low doses, perhaps less than those corresponding to the BMR, the slope can be quite steep, which may be manifested in terms of a relatively low value for the BMDL (or perhaps an "NA" result for the BMDL if this causes convergence problems because the steepness entails BMDL estimates that get very small).
Dichotomous Hill $p(\text{dose}) = g + \frac{(v - vg)}{1 + \exp(-\alpha - \beta\log(\text{dose}))}$	g = background v = maximum extra risk α = intercept β = slope	Parameter Constraints: $0 \leq g < 1$ $\qquad 0 < v \leq 1$ $\qquad 0 < \beta \leq 18$ User parameter restriction options: $1 \leq \beta$

8.3.1 Likelihood Function

For the frequentist modeling approach, the dichotomous models are fit using maximum likelihood methods. This section describes the likelihood function used to fit the dichotomous models.

Suppose the dataset has G dose groups with doses:

$$\text{dose}_1, \text{dose}_2, \dots, \text{dose}_G.$$

The total numbers of observations and numbers of responders in each dose group are

$$N_1, N_2, \dots, N_G$$

and, respectively,

$$n_1, n_2, \dots, n_G.$$

The distribution of n_i is assumed to be binomial with probability

$$p_i = p(\text{dose}_i; \ \theta), i = 1, 2, \dots G$$

where θ is a vector of dose-response model parameters (see previous table for lists of parameters for each model). Then the log-likelihood function LL can be written as

$$LL = \sum_{i=1}^{G} LL_i(N_i, n_i, \text{dose}_i; \ \theta)$$

where

$$LL_i(N_i, n_i, \text{dose}_i; \ \theta) = n_i \ln(p_i) + (N_i - n_i) \ln(1 - p_i), i = 1, 2, \dots, G.$$

This expression ignores a constant term that is independent of the parameter vector values and so does not affect estimation of those parameters.

Note from the table above that the upper bound for the power parameter in some of the models, and the slope parameter for some of the other models, has been set to 18. That value was selected because it represents a very high degree of curvature that should accommodate almost every dataset, even ones with very (or absolutely) flat dose-response at low doses followed by a very steep dose-response at higher doses.

If such parameter values are reported to be equal to 18 and/or the estimate in question is reported as "Bounded" (see the description of the output from dichotomous model runs in Section 8.4.2, "Analysis of Deviance Table," on page 87), the parameter estimates are maximum likelihood estimates only in the restricted sense that the parameter in question has been assigned a value and the other parameters are MLEs conditional on that assigned value. Such model results are not strictly comparable with others in terms of AIC. In such a case, the BMD and BMDL could depend on the choice of power parameter; thus, sensitivity analysis is indicated if one intends to rely on the reported BMD or BMDL.

8.3.2 AIC and Model Comparisons

The Akaike Information Criterion (AIC) (Akaike, 1973) can be used to compare different models fit (by the same fitting method, e.g., by maximizing the likelihood) to the same data set. The AIC is a statistic that depends on the value of LL (see previous section) and the number of estimated parameters, p:

$$AIC = -2 \times LL + 2 \times p$$

Note that the AICs for the dichotomous endpoints ignore the parameter-independent term, because LL as defined in the previous section ignores that term. This differs from the case of the continuous endpoints, where the parameter-independent term was not ignored, because its value depended on the assumed underlying data distribution (Normal or Lognormal). For dichotomous endpoints, there is only one assumed distribution for the counts of responders (the Binomial distribution), so the parameter-independent term has no effect and therefore can be ignored.

The AIC balances the goals of getting the highest LL value possible while being parsimonious with respect to the number of parameters needed to achieve a high LL value. Since the equation for AIC has a negative multiplier for LL (which one wants to be greater) and positive multiplier for p (which one wants to be as small as possible and still get "good fit"), a model with a **smaller value of AIC** than other models is presumed to be the better model on the basis of AIC. Although such methods are not exact, they can provide useful guidance in model selection.

In the current version of BMDS, the number of "estimated parameters" includes only those that have not been estimated to equal a bounding value, either from the model-imposed constraints or user-imposed restrictions. For more details, see Table 10 on 76.

Important Note

This counting process may or may not be reasonable, depending on the boundary value that a parameter in question hits.

For example, if the power parameter in a model hits (i.e., is estimated to be equal to) the upper bound of 18, it would usually be the case that one would want to count that parameter as one that is estimated, but BMDS does ***not*** do that.

For this reason, the user is apprised to consider carefully the cases where parameter bounds have been hit and to consider the implications for issues such as model comparison and model selection.

8.3.3 Plot and Error Bar Calculation

Figure 53. *Dichotomous endpoint plot.*
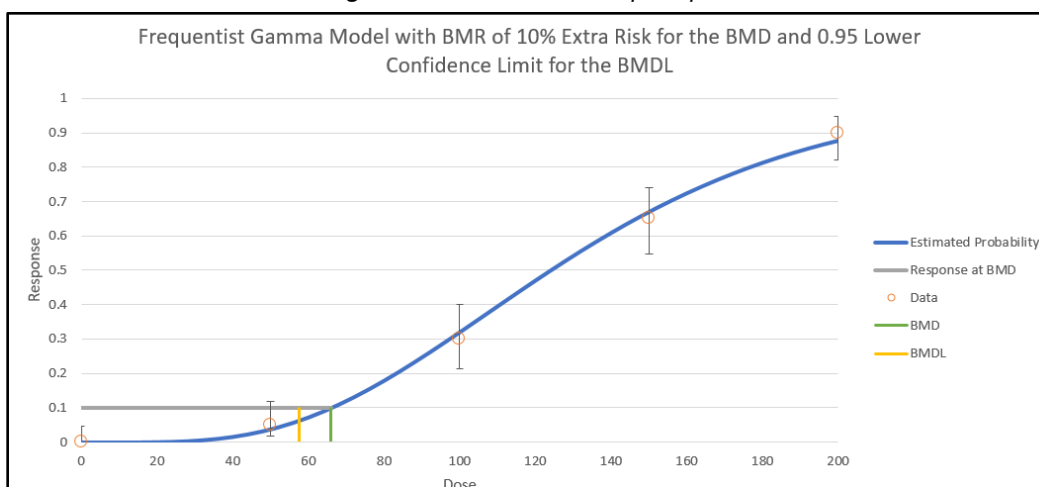
The graphical output, i.e., plot, is a visual depiction of the results of the modeling. Because plots, in general, were discussed in Section 6.5, "Graphical Output," on page

49, here we describe the one additional detail specific to the continuous models, i.e., computation of the error bars:

The error bars shown on the plots of dichotomous data are derived using a modification of the Wilson interval (based on the score statistic) but with a continuity correction method (Fleiss et al., 2003). The calculation finds the proportion, pi, such that

$$\frac{|p - p_i| - \frac{1}{2n}}{\sqrt{\frac{p_i \times (1 - p_i)}{n}}} = z$$

where
- p is the observed proportion
- n is the total number in the group in question
- $z = Z_{1-\frac{\alpha}{2}}$ is the inverse standard normal cumulative distribution function evaluated at $1 - \frac{\alpha}{2}$

This leads to equations for the lower and upper bounds of:

- $LB = \dfrac{(2np+z^2-1)-z\sqrt{z^2-(2+\frac{1}{n})+4p(nq+1)}}{2(n+z^2)}$

- $UB = \dfrac{(2np+z^2+1)+z\sqrt{z^2+(2-\frac{1}{n})+4p(nq-1)}}{2(n+z^2)}$

where $q = 1 - p$.

The error bars shown in BMDS plots use alpha = 0.05 and so represent the 95% confidence intervals on the observed proportions (independent of model).

8.3.4 BMD Computation

The BMD is computed as a function of the parameters of the model under consideration (Table 10). The following table specifies the solutions for the BMD for all of the dichotomous models.

Table 11. *Calculation of the BMD for the individual dichotomous models[8].*

Model	BMD Calculation
Multistage	There is no general analytic form for the BMD in terms of the BMR and the estimated model parameters for the multistage model. Instead, the BMD is the root of the equation $\beta_1 \text{BMD} + \cdots + \beta_n \text{BMD}^n + \ln(1 - A) = 0$, where $$A = \begin{cases} \text{BMR} & \text{extra risk} \\ \dfrac{\text{BMR}}{1 - g} & \text{added risk} \end{cases}$$
Weibull	$$\text{BMD} = \begin{cases} \left[\dfrac{-\ln(1 - \text{BMR})}{\beta}\right]^{\frac{1}{\alpha}} & \text{extra risk} \\ \left[\dfrac{-\ln(1 - \frac{\text{BMR}}{1 - g})}{\beta}\right]^{\frac{1}{\alpha}} & \text{added risk} \end{cases}$$

---

[8] All models represented in Table 11 use the same model forms as presented in Table 10. BMR is the value specified by the user to correspond to the risk level of interest (see 8.3.2).

Model	BMD Calculation
Gamma	Let $G(x; \alpha) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} \mathrm{d}t$ be the incomplete Gamma function and $G^{-1}(\cdot; \alpha)$ be its inverse function. Then $$\mathrm{BMD} = \begin{cases} \dfrac{G^{-1}(\mathrm{BMR}; \alpha)}{\beta} & \text{extra risk} \\ \dfrac{G^{-1}\left(\dfrac{\mathrm{BMR}}{1-g}; \alpha\right)}{\beta} & \text{added risk} \end{cases}$$
Logistic	$\mathrm{BMD} = \frac{\ln\left(\frac{1-Z}{1+Z \times e^{-\alpha}}\right)}{\beta}$ where $Z = \begin{cases} \mathrm{BMR} & \text{extra risk} \\ \mathrm{BMR} \times \frac{1+e^{-\alpha}}{e^{-\alpha}} & \text{added risk} \end{cases}$
Log-Logistic	$$\ln(\mathrm{BMD}) = \begin{cases} \dfrac{\ln\left(\dfrac{\mathrm{BMR}}{1-\mathrm{BMR}}\right) - \alpha}{\beta} & \text{extra risk} \\ \dfrac{\ln\left(\dfrac{\mathrm{BMR}}{1-g-\mathrm{BMR}}\right) - \alpha}{\beta} & \text{added risk} \end{cases}$$
Probit	$$\mathrm{BMD} = \begin{cases} \dfrac{\Phi^{-1}(\mathrm{BMR}[1-\Phi(\alpha)]+\Phi(\alpha))-\alpha}{\beta} & \text{extra risk} \\ \dfrac{\Phi^{-1}(\mathrm{BMR}+\Phi(\alpha))-\alpha}{\beta} & \text{added risk} \end{cases}$$
Log-Probit	$$\ln(\mathrm{BMD}) = \begin{cases} \dfrac{\Phi^{-1}(\mathrm{BMR})-\alpha}{\beta} & \text{extra risk} \\ \dfrac{\Phi^{-1}\dfrac{\mathrm{BMR}}{1-g}-\alpha}{\beta} & \text{added risk} \end{cases}$$
Dichotomous Hill	Added risk: $$\mathrm{BMD} = e^{\dfrac{-\alpha - \log\left(-\dfrac{\mathrm{BMR}-v+gv}{\mathrm{BM}}\right)}{\beta}}$$ Extra risk: $$\mathrm{BMD} = e^{\dfrac{-\alpha - \log\left(-\dfrac{\mathrm{BMR}-v+gv-\mathrm{BMR}gv}{\mathrm{BMR}(-1+gv)}\right)}{\beta}}$$

8.3.5 BMDL and BMDU Computation

BMDS currently calculates one-sided confidence intervals, in accordance with current BMD practice. The general approach to computing the confidence limits for the BMD (called the BMDL and BMDU here) is the same for all the models in BMDS, and is based on the asymptotic distribution of the likelihood ratio (Crump and Howe, 1985). Two different specific approaches are followed in these models.

For the Multistage model, it is impractical to explicitly reparameterize the dose-response model function to allow BMD to appear as an explicit parameter. For these models, the BMR equation is used as a non-linear constraint, and the minimum value of BMD is determined such that the log-likelihood is equal to the log-likelihood at the maximum likelihood estimates less

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

For the remaining models, the equations that define the benchmark response in terms of the benchmark dose and the dose-response model (Table 9) are solved for one of the model parameters. The resulting expression is substituted back into the model equations, with the effect of reparameterizing the model so that BMD appears explicitly as a parameter. A value for BMD is then found such that, when the remaining parameters are varied to maximize the likelihood, the resulting log-likelihood is less than that at the maximum likelihood estimates by exactly

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

In all cases, the additional constraints specify that the BMDL be less than the BMD and the BMDU be greater than the BMD.

8.3.6 Bayesian Dichotomous Model Descriptions

Note   At this time, EPA does not offer technical guidance on Bayesian modeling or Bayesian model averaging.

From a Bayesian perspective, inference proceeds by defining a data generating mechanism, given a model, M, and its parameters. For our purposes, M would be one of the models listed in Table 10 that determines the probability of response. The data generating mechanism would be the assumption that the observations were obtained from binomial sampling, having the dose-dependent probability of response defined by one of those models (with specific values of the parameters in that model).

We can then relate that to the likelihood, here denoted $\ell(D|M)$, which shows explicitly that it is the likelihood of the data, D, conditional on the model. The functional form of the log of the likelihood is presented in Section 8.3.1, "Likelihood Function," on page 78..

The set of Bayesian dichotomous models used in BMDS 3 is identical to the set of models used for frequentist (MLE) approaches (Table 10). In the following, let θ be the vector of parameters that are required to define the any one of those models. So, for example, for the Weibull model θ = (g, α, β). The additional consideration incorporated into the Bayesian approach is the specification of a prior distribution for θ. The Bayesian approach takes the specified prior and updates it using the data under consideration to obtain a "posterior" distribution for θ.

From a Bayesian perspective, functions of θ also have posterior densities. So, using the equations summarized in Table 11, one can derive a posterior distribution for the BMD.

BMDS summarizes the posterior for the BMD as follows. The BMD (one might call it a "central estimate" perhaps) is equated to the value obtained (as in Table 11) by using the maximum *a posteriori* (MAP) θ estimate. The MAP is the value of θ that maximizes the posterior log-likelihood. The posterior density is itself approximated using a Laplacian approximation. This approximation is also used to estimate BMDL and BMDU values, which are the percentiles from that density that correspond to the confidence level selected by the user.

The priors for the BMDS dichotomous models are defined in the following table.

Important Note

The priors for the parameters are based on scaled doses and scaled responses; BMDS performs this scaling automatically.

BMDS automatically scales the doses by dividing by the maximum dose in the dataset under consideration, i.e., that the doses under consideration range from 0 to 1 (inclusive). **BMDS automatically scales the responses** by dividing by the mean response in the control (or lowest dose) group.

The user does *not* need to scale anything beforehand. That means that the parameter estimates and BMD values returned by the program have been adjusted back to the original scale of the doses and the original scale of the responses specified in the input data file.

Table 12. *Bayesian dichotomous models and their respective parameter priors.*

Model	Constraints	Priors	Notes
Multistage $p(\text{dose}) = g + (1-g)\left(1 - \exp\left[-\sum_{i=1}^{N}\beta_i dose^i\right]\right)$	$0 \le g < 1$ $\beta_i > 0$ $N \ge 2$	$\text{logit}(g) \sim \text{Normal}(0,2)$ $\beta_1 \sim \text{Lognormal}(0, 0.25)$ $\beta_i \sim \text{Lognormal}(0,1), i \ge 2)$	The prior over the β₁ parameter reflects the belief that the linear term should be strictly positive if the quadratic term is positive in the two-hit model of carcinogenesis. The difference in priors between Multistage and Quantal Linear models is by design. The objective is to emphasize the higher-order terms in each model. For model averaging purposes, N= 2.
Weibull $p(\text{dose}) = g + (1-g)(1 - \exp[-\beta dose^\alpha])$	$0 \le g < 1$ $\alpha > 0$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0,2)$ $\alpha \sim \text{Lognormal}(ln(2), \sqrt{0.18})$ $\beta \sim \text{Lognormal}(0,1)$	The prior for α entails that there is only a 0.05 prior probability the power parameter will be less than 1. This allows for models that are supra-linear; however, it requires a large amount of data for the α parameter to go much below 1.
Quantal linear $p(\text{dose}) = g + (1-g)(1 - \exp[-\beta dose])$	$0 \le g < 1$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0,2)$ $\beta \sim \text{Lognormal}(0,1)$	The difference in priors between Quantal Linear and the following Multistage model is by design. The objective is to emphasize the higher-order terms in each model. Quantal Linear is *not* the same as Multistage-1. This is important for model averaging purposes.
Gamma $p(\text{dose}) = g + \frac{1-g}{\Gamma(\alpha)}\int_0^{\beta dose} t^{\alpha-1}\exp(-t)\,dt$	$0 \le g < 1$ $\alpha > 0.2$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0,2)$ $\alpha \sim \text{Lognormal}(ln(2), \sqrt{0.18})$ $\beta \sim \text{Lognormal}(0,1)$	The prior for α entails that there is only a 0.05 prior probability the power parameter will be less than 1. This allows for models that are supra linear; however, it requires a large amount of data for the α parameter to go much below 1. The α parameter is also constrained to be greater than 0.2, for numerical reasons.
Logistic $p(\text{dose}) = \frac{1}{1+\exp[-\alpha - \beta dose]}$	$-20 < \alpha < 20$ $\beta > 0$	$\alpha \sim \text{Normal}(0,2)$ $\beta \sim \text{Lognormal}(0.1,1)$	
Log-Logistic $p(d\text{dose}) = g + \frac{1-g}{1+\exp[-\alpha - \beta\ln(\text{dose})]}$	$0 \le g < 1$ $-40 < \alpha < 40$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0,2)$ $\alpha \sim \text{Normal}(0,1)$ $\beta \sim \text{Lognormal}(ln(2), 0.5)$	
Probit $p(\text{dose}) = \Phi(\alpha + \beta\text{dose})$	$-8 < \alpha < 8$ $\beta > 0$	$\alpha \sim \text{Normal}(0,2)$ $\beta \sim \text{Lognormal}(0.1,1)$	$\Phi(\cdot)$ is the standard normal cumulative density function

Model	Constraints	Priors	Notes
Log-Probit $p(\text{dose}) = g + (1-g)\Phi[\alpha + \beta \ln(\text{dose})]$	$0 \le g < 1$ $-8 < \alpha < 8$ $\beta > 0$	$\text{logit}(g) \sim \text{Normal}(0,2)$ $\alpha \sim \text{Normal}(0,1)$ $\beta \sim \text{Lognormal}(ln(2), 0.5)$	$\Phi(\cdot)$ is the standard normal cumulative density function
Dichotomous Hill $p(\text{dose}) = g + \dfrac{v(1-g)}{1+\exp[-a-b\ln(\text{dose})]}$	$0 \le g < 1$ $-40 < a < 40$ $b > 0$ $0 < v \le 1$	$\text{logit}(g) \sim \text{Normal}(-1,2)$ $a \sim \text{Normal}(-3, 3.3)$ $b \sim \text{Lognormal}(ln(2), 0.5)$ $\text{logit}(v) \sim \text{Normal}(0,3)$	

Notes: $\text{logit}(g) = \ln\left(\frac{g}{1-g}\right)$. Normal(x, y) denotes a Normal distribution with mean x and standard deviation y. Lognormal(w, z) denotes a lognormal distribution with log-scale mean w and log-scale standard deviation z.

As the number of observations in a dataset increases, there should be less quantitative difference between the parameters and BMDs obtained from the Bayesian approach and from the frequentist approach.

When there are fewer data points, the priors will affect the Bayesian estimation. The impact may be most noticeable when the data suggest a "hockey-stick" dose-response relationship, or when those data suggest strong supralinear behavior. In these cases, the priors specified above for the Bayesian approach will tend to "shrink back" parameter estimates to obtain smoother dose-response relationships where changes in the slope are more gradual.

8.4　Outputs Specific to Frequentist Dichotomous Models

Figure 54. Sample Results Workbook tab for a dichotomous model run.
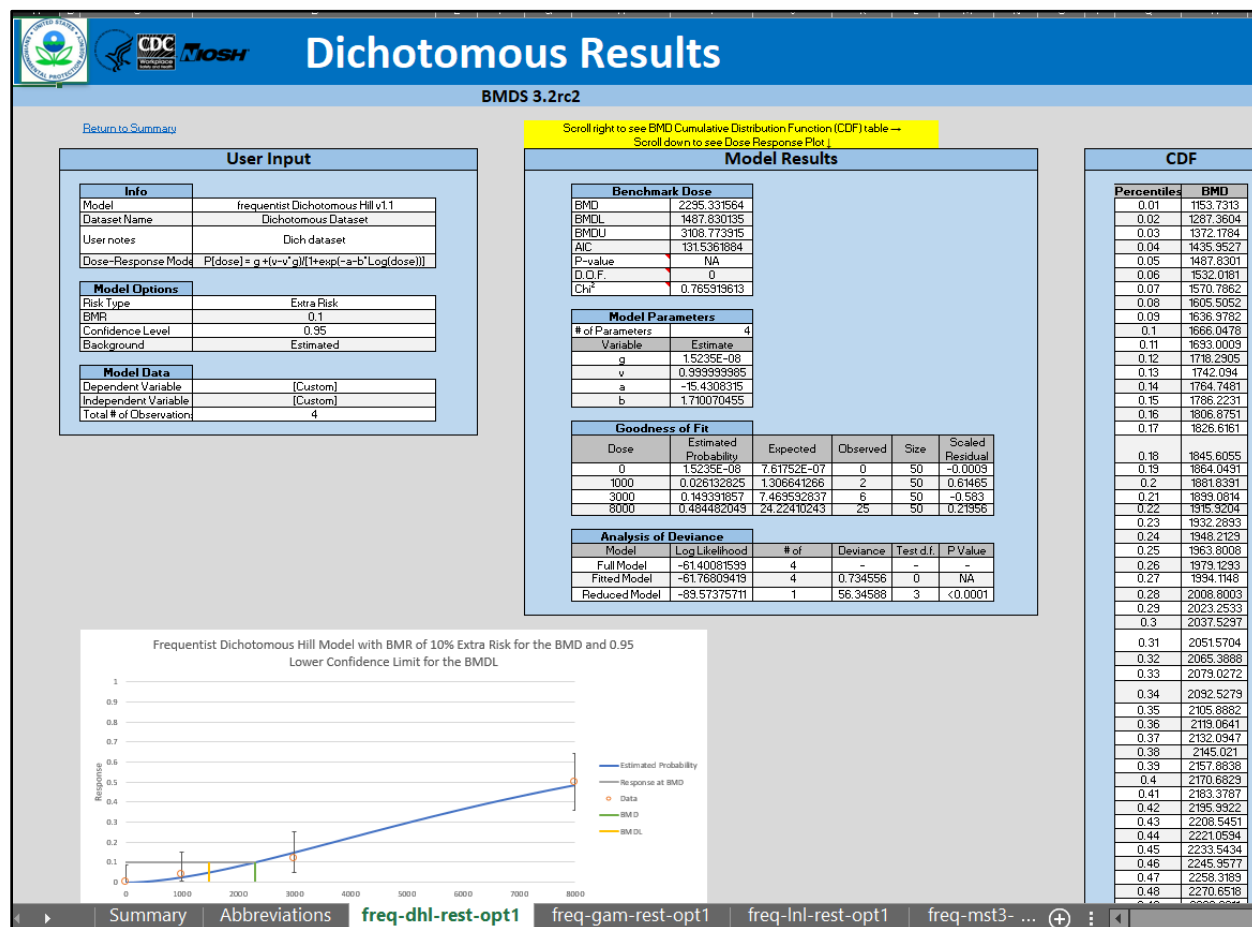
8.4.1　Goodness of Fit Table

Goodness of Fit					
Dose	Estimated Probability	Expected	Observed	Size	Scaled Residual
0	1.5235E-08	7.61752E-07	0	50	-0.000873
1000	0.026132825	1.306641266	2	50	0.6146529
3000	0.149391857	7.469592837	6	50	-0.58302
8000	0.484482049	24.22410243	25	50	0.2195627

This table in the individual model results shows a listing of the data ("Observed" and "Size") as well as the model-estimated probability of response and corresponding expected number of responders. This is a good place for the user to look to judge the appropriateness of the model, in addition to the overall goodness-of-fit statistics reported in the Benchmark Dose Table (see Section 6.2) and the Analysis of Deviance table (Section 8.4.2). If a model fits well, the observed and expected number of responders should be relatively close.

The scaled residual values printed at the end of the table are defined as follows:

$\frac{(\text{Obs} - \text{Expected})}{SE}$,

where "Expected" is the predicted number of responders from the model and SE equals the estimated standard error of that predicted number. For these models, the estimated standard error is equal to $\sqrt{[n \times p \times (1 - p)]}$, where

n is the sample size ("Size" in this table), and

p is the model-predicted probability of response ("Estimated Probability" in this table).

The model's adherence to the data may be called into question if the scaled residual value for any dose group, particularly the control group or the group with dose closest to the BMD, is greater than 2 or less than -2.

8.4.2 Analysis of Deviance Table

Analysis of Deviance					
Model	Log Likelihood	# of Parameters	Deviance	Test d.f.	P Value
Full Model	-61.40081599	4	-	-	-
Fitted Model	-61.76809418	3	0.73455639	1	0.3914103
Reduced Model	-89.57375711	1	56.3458822	3	<0.0001

The analysis of deviance table lists three log-likelihood values.
- The first is for the "full model." The full model posits a separate and independent probability of response for each dose group. There is no dose-response function constraining those probabilities. The log-likelihood displayed is the maximum that could ever be achieved for the given dataset. The number of parameters for the full model is equal to the number of dose groups (each has its own, independent probability parameter).
- The second log-likelihood is for the "fitted model." It is the maximum log-likelihood value obtainable for the model under consideration. It corresponds to the model with the parameters set equal to the values shown in the Parameter Estimates table. The number of parameters equals the number of parameters in that table that are not reported as "Bounded."
- The last log-likelihood value is for the "reduced model." It is the maximum log-likelihood obtainable if one assumed that the same probability of response applied to all the dose groups. There is only 1 parameter for the reduced model, i.e., the assumed constant probability of response.

Associated with each of these three models are three values: Deviance, degrees of freedom (Test d.f.), and P-value. The Deviance is twice the difference between the fitted or reduced model and the full model log-likelihood values. This Deviance is another goodness-of-fit metric: if the Deviance is small, then the "smaller" model (i.e., the fitted or reduced model) describes the data nearly as well as the full model does. Deviance is approximately a chi-squared random variable with degrees of freedom specified by "Test d.f." which is itself the difference in the number of parameters for the two models being compared. The "P-Value" reflects the use of this chi-square approximation to assess significance of the difference in fits. Larger deviances correspond to smaller p-values, so a small p-value indicates that the smaller model does not fit as well as the full model. The user may choose a rejection level (0.05 is common) to test if the model fit is appropriate.

For the fitted model, this is another measure of the fit of the model to the data. For the reduced model, failure to reject that model (p-values greater than the rejection level chosen by the user) might lead the user to infer that there is no dose-related effect on response probabilities.

8.4.3 Additions for the Restricted Multistage Model Only

Figure 55. *Slope Factor (last row in table) appears only on Restricted Multistage Model results.*

Benchmark Dose	
BMD	2512.859821
BMDL	1492.16541
BMDU	3755.154706
AIC	128.8193162
P-value	0.893543819
D.O.F.	1
Chi^2	0.017908162
Slope Factor	6.70167E-05

Some additional assessment tools are presented when the model under consideration is the Multistage model, and it has been run "restricted." This is the EPA standard set-up for modeling cancer data, but these additional results will be shown for any dichotomous endpoint when the Multistage model is run this way.

The Benchmark Dose table for the restricted Multistage model includes an estimate of the slope factor, defined by EPA as the linear slope between the extra risk (0.1) at the BMDL(10) and the extra risk (0) at background (generally 0 dose).

The restricted Multistage model plot also includes a dashed line representing this linear slope.

Figure 56. *The dashed line for the Multistage model plot representing linear slope.*
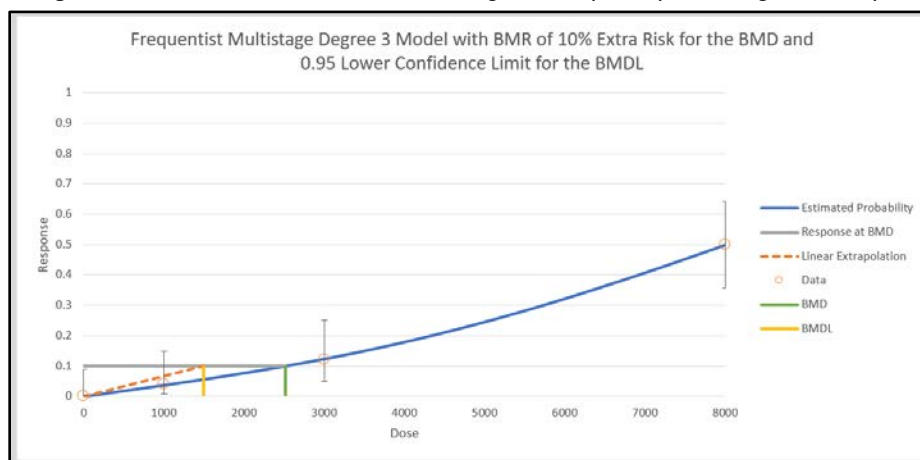
If dose units are in mg/kg-day, this equals the oral slope factor (OSF) as defined by IRIS. If the dose units are µg/m3, this equals the inhalation unit risk (IUR) as defined by IRIS. For more information, see the "IRIS Toxicity Values" section of the Basic Information about the Integrated Risk Information System (IRIS) web page."

8.5 Outputs Specific to Bayesian Dichotomous Models

Figure 57. *Sample Bayesian dichotomous results plot.*

To compare the difference between any two Bayesian models, the unnormalized Log Posterior Probability (LPP) is given, which allows the computation of a Bayes factor (BF) to compare any two models. BF equals the exponentiated difference between the two LPP. For example, if one wishes to compare the Log-Logistic model (Model A) (yielding LPP_A) to the Multistage 2nd degree model (Model B, with LPP_B) one estimates the BF as

$$BF = \exp(LPP_A - LPP_B),$$

This computation assumes that both models have equal probability *a priori.* This value is then interpreted as the posterior odds one model is more correct than the other model and is used in Bayesian hypothesis testing. In the example above, if the Bayes Factor was 2.5, the interpretation would be that the Log-logistic model is *a posteriori* 2.5 times more likely than the multistage model. When these values are normalized into proper probabilities, they are equivalent the posterior model probabilities given in model averaging (again, assuming equal model probability *a priori*). The table below, adapted from Jeffreys (1998) is a common interpretation of Bayes Factors.

Table 13. *Bayes factors for dichotomous models.*

Bayes Factor	Strength of Evidence for H_A
< 1	negative (supports H_B)
1 to 3.2	not worth mentioning
3.2 to 10	substantial
10 to 31.6	strong
31.6 to 100	very strong
100	decisive

For BMDS 3.0, all LPP and corresponding posterior model probabilities are computed using the Laplace approximation. This value is different from the commonly used Bayesian Information Criterion (BIC), and the two should not be confused based upon other model averaging approaches, which use the BIC exclusively. Errors in the posterior

probabilities estimated from the BIC are $O(1)$ estimators. Errors in the posterior probabilities estimated using the Laplace approximation are $O(n^{-1})$. This means the latter approximation goes to the true posterior model probability with increasing data and the former, using the BIC, may not go to the true value.

9.0   Nested Dichotomous Endpoints

In a nested study, for each dose tested, there is a group of experimental units receiving that particular dose of the chemical of interest. For each of those experimental units, a number of dichotomous observations are obtained, i.e., those dichotomous observations are "nested" within the experimental units.

Moreover, because of the nesting, one may suspect that the observations within each experimental unit are more similar to one another than they are to observations from other experimental units. For example, consider a developmental toxicity experiment, in which rodent females ("dams") are exposed to the chemical of interest prior to or during pregnancy. The offspring ("pups") from each litter are examined after birth for the presence or absence of malformations. Because each rodent dam may produce 15-20 pups, the results consist of a set of dichotomous (malformation present or absent) counts for each dam.

Nested dichotomous models are defined so as account for, model, the data structure associated with such experimental designs.

The most common application of the nested models will be to developmental toxicology studies of organisms that have multiple offspring per litter, as do rodents. In these study designs, pregnant dams are given one or several doses of a toxicant, and the fetuses, embryos, or term offspring are examined for signs of abnormal development. In such studies, it is usual for the responses of pups in the same litter to be more similar to each other than to the responses of pups in different litters ("intra-litter correlation," or "litter-effect"). Another way to describe the same phenomenon is that the variance among the proportion of pups affected in litters is greater than would be expected if the pups were responding completely independently of each other.

Observations from such studies might include skeletal structure change, delayed ossification in the bone, or organ structural change to malformation, among many others. Since all those observations are made in pups—but *not* in the mothers—these data are nested data.

Nested models that BMDS has included make available two approaches to this feature of developmental toxicology studies: they use a probability model that provides for extra inter-litter variance of the proportion of pups affected (the beta-binomial probability model: see the "Likelihood Function" section below); and they incorporate a litter-specific covariate that is expected to account for at least some of the extra inter-litter variance. This latter approach was introduced by Rai and Van Ryzin (1985), who reasoned that a covariate that took into account the condition of the dam before dosing might explain much of the observed litter effect. Those authors suggested that litter size would be an appropriate covariate. For the reasoning to apply strictly, the measure of litter size should not be affected by treatment; thus, in a study in which dosing begins after implantation, the number of implantation sites would seem to be an appropriate measure. On the other hand, the number of live fetuses in the litter at term would not be an appropriate measure if there is any dose-induced prenatal death or resorption.

The following screenshot shows how a nested dataset should be formatted for use in BMDS.

Figure 58. *Nested dataset formatted correctly for BMDS analysis.*

Nested ex1			
nested.dax			
Dose	Litter Size	Incidence	Litter Specific Covariate
Dose	N	Resp	Covariate
0	16	1	16
0	9	1	9
0	15	2	15
0	14	3	14
0	13	3	13
0	9	0	9
0	10	2	10
0	14	2	14
0	10	1	10
0	11	2	11
25	14	4	14
25	9	5	9
25	14	6	14
25	9	2	9
25	13	6	13
25	12	3	12
25	10	1	10
25	10	2	10
25	11	4	11
25	14	3	14
50	11	4	11
50	11	5	11
50	14	5	14
50	11	4	11
50	10	5	10
50	11	4	11
50	10	5	10
50	15	6	15
50	7	2	7

As the above figure shows, each litter is on a separate row, showing the dose it received, its sample size ("Litter size"), the number of responders ("Incidence"), and the value of a covariate that will be discussed below ("Litter Specific Covariate").

This section provides details on the following topics:
- Implementation of nested models in BMDS 3
- Entering nested data
- Nested model options
- Nested model outputs
- Options for restricting values of certain model parameters

9.1   Nested Response Model

Figure 59. *BMDS 3 nested model on the Analysis Workbook.*

	MLE	
	Frequentist Restricted	**Frequentist Unrestricted**
Model Name	**Enable** ▨	**Enable** ▨
Nested Logistic	☑	☑
NCTR	▨	▨

BMDS 2.7 contained three nested dichotomous models:

- Nested Logistic
- National Center for Toxicological Research (NCTR)
- Rai and van Ryzin

At this time, the only nested dichotomous model available in BMDS 3 is Nested Logistic. The Nested Logistic Model is the log-logistic model, modified to include a litter-specific covariate.

The NCTR (National Center for Toxicological Research) nested dichotomous model will be included in a future BMDS release.

The NCTR and Rai and Van Ryzin models can be accessed in BMDS 2.7, which is available from the BMDS website as an archive version of BMDS.

9.2   Entering Nested Dichotomous Data

For information on inserting or importing data, see Section 4.2, "Step 2: Add Datasets," on page 29.

The default column headers are "Dose," "Litter Size," "Incidence" and "Litter Specific Covariate" (LSC).

Note   There **must** be data in the LSC row even if the modeling options do not call for the use of LSC.

9.3   Options

On the BMDS 3 Analysis Workbook's Main tab, the user can define multiple Option Sets to apply to multiple user-selected models and multiple user-selected datasets in a single "batch" process. Select the **Add Option Set** button to define a new Option Set configuration.

Figure 60. *Nested Model options.*

Option Set #	Risk Type	BMR	Confidence Level	Litter Specific Covariate	Background	Bootstrap Iterations	Bootstrap Seed
1	Extra Risk ▼	0.1	0.95	Overall Mean ▼	Estimated ▼	1000	Automatic ▼
				Add Option Set			

Unlike previous versions of BMDS, BMDS 3 does not require the user to specify the model form, but rather automatically runs all forms of the available nested models. So, in effect, BMDS 3 runs all four combinations displayed in this table:

Table 14. *All forms of nested models run by BMDS.*

Litter Specific Covariate	Intralitter Correlation- Estimated	Intralitter Correlation- Set to Zero
Included in Model	lsc+ilc+	lsc+ilc-
Not included in Model	lsc-ilc+	lsc-ilc-

The table entries display the abbreviation used by BMDS (e.g., "lsc+ilc+") to name the tabs on the Results Workbook.

The specific definitions of Litter Specific Covariate and Intralitter Correlations are discussed below. Briefly, the covariate is another variable that the model can (optionally) include, one that may help to explain the variation in the response from one experimental unit to another. As noted above, the experimental unit is very often a litter of observations, hence the designation "Litter Specific." The intralitter correlation (again referencing the litter as a common experimental unit), estimates the degree to which observations within the same litter are correlated. If set to zero (one of the options), there is no correlation and the assumption then is that every observation is independent of every other observation (conditional on the model predicted probabilities of response).

9.3.1 Risk Type

Choices are "Extra Risk" (default) or "Added Risk."

Additional risk is the additional proportion of total animals that respond in the presence of the dose, or the probability of response at dose d, $P(d)$, minus the probability of response in the absence of exposure, $P(0)$; $i.e., additional\ risk\ =\ P(d) - P(0)$.

Extra risk is the additional risk divided by the proportion of animals that will not respond in the absence of exposure, $1 - P(0)$; $i.e., extra\ risk\ = \frac{P(d) - P(0)}{1 - P(0)}$. Thus, extra and additional risk are equal when background rate is zero.

9.3.2 BMR

The BMR is the value of risk (extra or added, as specified by the user) for which a BMD is estimated. BMR must be between 0 and 1 (not inclusive). If $P(0) > 0$, then values for BMR greater than $1 - P(0)$ will result in an error when the risk type is added risk. That is because the maximum added risk that can ever be achieved is $1 - P(0)$. In practice, this should not typically be an issue because one usually is interested in BMR values in the range of 0.01 to around 0.10.

BMR and Graphs

The response associated with the BMR that is displayed in the graphical model output will only be the same as the BMR when $P(0) = 0$.

This is because to obtain the actual response value one must solve for $P(d)$ in the equation for added or extra risk discussed in Section 9.3.1.

The horizontal bar depicting the response level used to derive the BMD that is displayed in the graphical model output will only be the same as the user-defined BMR (e.g., 10% Extra Risk) when the response at background, P(0), equals zero.

When P(0) does not equal zero, the true response level can be calculated using the Extra Risk equation described in Section 9.3.1.

9.3.3 Confidence Level

Confidence Level
0.95

The Confidence Level is a fraction between 0 and 1; 0.95 is recommended by EPA (U.S. EPA, 2012).

The value for confidence level must be between 0 and 1 (not inclusive). For a confidence level of x, BMDS will output BMDL and BMDU estimates, each of which is a one-sided confidence bound at level x. For example, if the user sets the confidence level to 0.95 (the default), then the BMDL is a 95% one-sided lower confidence bound for the BMD estimate; the BMDU is a 95% one-sided lower confidence bound for the BMD estimate. In that example, the range from BMDL to BMDU would constitute a 90% confidence interval (5% in each tail outside that interval).

9.3.4 Litter Specific Covariate

nce	Litter Specific Covariate
	Overall Mean ▾
	Overall Mean
	Control Group Mean

Enables the user to account for inter-litter variability by using a litter specific covariate (LSC).

Best practice would dictate that one not use a variable for the LSC if that variable is affected by dose (the other explanatory variable). During the course of the analysis, it is recommended that the models with and without the LSC be compared to determine whether or not the LSC contributes to a better explanation of the observations (e.g., by comparing AIC values).
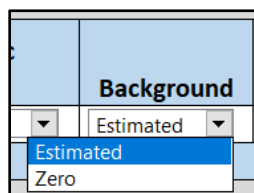
For those runs where the LSC is included in the model, the BMD will depend on the LSC. This Option allows the user to determine if the BMD (and the corresponding plots) will be computed using the "Control Group Mean" value of the LSC or the "Overall Mean" value of the LSC (i.e., averaged across all dose groups; this is the default). See Section 9.4.4,

"BMD Computation," on page 102 for an explanation as to why this option is necessary, and which choice would be preferred for the given dataset.

Basically, the Overall Mean should be used under most circumstances. If the Litter Specific Covariate differs from dose to dose (without any apparent consistent trend with respect to dose), consider using the Control Group Mean.

Note   Carr and Portier (Carr and Porter, 1991), in a simulation study, warn that in situations in which there is no effect of litter size, statistical models that incorporate a litter size parameter, as do the models in BMDS, will often erroneously indicate that there is a litter size effect. Thus, the user should use litter size parameters with caution. Unfortunately, there are currently no good diagnostics for determining whether a litter size effect exists.

9.3.5 Background
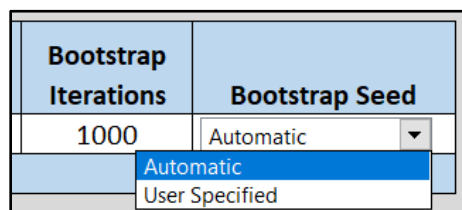
Choices are Estimated (default) or Zero. The user is advised to select Estimated unless, and this will probably be a rare condition, there is very strong evidence (from other studies or ancillary information) that there is absolutely zero probability or response in the absence of exposure.

Note   Do not set Background to Zero when there are responses in a group with dose=0 in the experiment being modeled.

9.3.6 Bootstrapping

Bootstrap Iterations: Specify the number of bootstrap iterations (default is 1000) to run to estimate goodness of fit. It is recommended to keep the value at a minimum of 1000.

Bootstrap Seed: Select this feature to specify a bootstrap seed for the random number generator. Default is that BMDS auto-generates a seed for the random number generator based on the system clock.

For more details, refer to Section 9.5, "Outputs Specific to Frequentist Nested Dichotomous Models," on page 103.

9.4 Mathematical Details for Models for Nested Dichotomous Endpoints

The model that BMDS makes available for nested data is the Logistic Nested Model (Table 15). In the future, the NCTR model will be made available (also presented in Table 15). The user who is interested in the NCTR model (and also the Rai and van Ryzin model) is advised to download BMDS 2.7, which has both of those models for nested data.

Table 15. *Individual nested dichotomous models and their respective parameters .*

Model	Parameters	Notes
Logistic Nested model $p(d) = \alpha + \dfrac{\theta_1 r_{ij} + (1 - \alpha - \theta_1 r_{ij})}{(1 + e^{[-\beta - \theta_2 r_{ij} - \rho \times \ln(\text{dose})]})}$ if dose > 0, and $\alpha + \theta_1 r_{ij}$ if dose = 0	α = Intercept (≥0) ρ = power (≥0, can restrict ≥1) β = slope (≥0) θ_1 = first coefficient for the litter specific covariate θ_2 = second coefficient for the litter specific covariate $\phi_1, \dots, \phi_g$ = intra-litter correlation coefficients	In the model equation, r_{ij} is the litter-specific covariate for the jth litter in the ith dose group. In addition, there are g intra-litter correlation coefficients; g is the number of dose groups. $0 \le \phi_i \le 1$ $(i = 1, \dots, g)$. $1 > \alpha + \rho \ge \theta_1 r_{ij} \ge 0$ for every r_{ij}. If r_m represents either the control mean value for the litter-specific covariate or its overall mean, then the BMD is computed as: $$\text{BMD} = e^{\left\{ \frac{\left[\ln\left(\frac{A}{(1-A)}\right) - \beta - \theta_2 r_m \right]}{\rho} \right\}}$$ where $A = \begin{cases} \text{BMRF} & \text{extra risk} \\ \dfrac{\text{BMRF}}{(1 - a - \theta_1 r_m)} & \text{added risk} \end{cases}$ For the BMDL, the parameter β is replaced with an expression derived from the BMD definition and the BMDL is derived as described in Section 9.4.2.

Model	Parameters	Notes
National Center for Toxicological Research (NCTR) model[9] $p(d) = 1 - e^{\{-(\alpha+\theta_1(r_{ij}-r_m))-(\beta+\theta_2(r_{ij}-r_m))\times \text{dose}^\rho\}}$	α = Intercept (≥0) ρ = power (≥0, can restrict ≥1) β = slope (≥0) θ_1 = first coefficient for the litter specific covariate θ_2 = second coefficient for the litter specific covariate $\phi_1, \dots, \phi_g$ = intra-litter correlation coefficients	In the model equation, r_{ij} is the litter-specific covariate for the jth litter in the ith dose group, r_m is the overall mean for the litter-specific covariate $\theta_1(r_{ij} - r_m) \geq 0$ and $\theta_2(r_{ij} - r_m) \geq 0$ In addition, there are g intra-litter correlation coefficients; g is the number of dose groups. $0 \leq \phi_i \leq 1$ $(i = 1, \dots, g)$; $1 > \alpha + \rho \geq \theta_1 r_{ij} \geq 0$ for every r_{ij}. $\text{BMD} = \left[\dfrac{-(\ln(1-A))}{(\beta + \theta_2 \delta_r)}\right] \times \left(\dfrac{1}{\rho}\right)$ where $A = \begin{cases} \text{BMRF} & \text{extra risk} \\[2mm] \dfrac{\text{BMRF}}{(1 - a - \theta_1 \delta_r)} & \text{added risk} \end{cases}$ For the BMDL, the parameter β is replaced with an expression derived from the BMD definition and the BMDL is derived as described in Section 9.4.2.

---

[9] The NCTR model will be added to a future version of BMDS.

9.4.1 Likelihood Function

Let g represent the number of dose groups. For the *ith* group, there are n_i pregnant females administered dose $dose_i$. In the *jth* litter of the *ith* dose group there are s_{ij} *fetuses*, x_{ij} affected fetuses, and, potentially, a litter-specific covariate r_{ij} which will often be a measure of potential litter size, such as number of implantation sites, though this is not a requirement of the models. In what follows, the dose-response model, which gives the probability that a fetus in the *jth* litter of the *ith* dose group will be affected is represented by

$$p(\mathrm{dose}_i, r_{ij})$$

The beta-binomial distribution can be thought of as resulting from sampling in two stages. First, each litter is assigned a probability, P_{ij} from a beta distribution (beta distributions represent a two-parameter family of probability distributions defined on the interval (0,1)). The parameters of the beta distribution are determined by the administered dose, the litter specific covariate r_{ij} and the degree of intra-litter correlation, v_i. Note that the intra-litter correlation parameter varies among doses. It has been shown (Williams et al., 1988) that when the true intra-litter correlation differs among doses, unbiased estimates of the other parameters in a dose-response model can only be obtained if dose-specific intra-litter correlation parameters are estimated. As a special case, if $v_i = 0$, then this part of the process is completely deterministic, and

$$P_{ij} = p(\mathrm{dose}_i, r_{ij})$$

This allows for the possibility of no litter effect at all.

In the second stage of sampling, s_{ij} *fetuses* are assigned to the litter, and the number of affected fetuses, x_{ij} is sampled from a binomial distribution with parameters P_{ij} and s_{ij}.

The log-likelihood function that results from this process is (Kupper et al., 1986):

$$LL = \sum_{i=1}^{g} \left\{ \sum_{j=1}^{n_i} \left[\sum_{k=1}^{x_{ij}} \ln(p(\mathrm{dose}_i, r_{ij}) + (k-1)\Psi_i) \right. \right.$$
$$\left. \left. + \sum_{k=1}^{s_{ij}-x_{ij}} \ln(1 - p(\mathrm{dose}_i, r_{ij}) + (k-1)\Psi_i) - \sum_{k=1}^{s_{ij}} \ln(1 + (k-1)\Psi_i) \right] \right\}$$

where

$$\Psi_i = \frac{\phi_i}{1 - \phi_i}$$

and

$$\sum_{a}^{b} (\cdot) = 0 \ \ if \ a > b \ (by \ convention).$$

This log-likelihood ignores a term that is independent of the values of the parameters..

9.4.2 Goodness of Fit Information—Litter Data

The "Litter Data" table provides a listing of the data, expected and observed responses and scaled residuals, for each litter.

The scaled residual values printed at the end of the table are defined as follows:

$$\frac{(\text{Obs} - \text{Expected})}{\text{SE}},$$

where "Expected" is the predicted number of responders from the model and SE equals the estimated standard error of that predicted number. For these models, the estimated standard error is equal to $\sqrt{[n \times p \times (1 - p) \times (\theta \times (n - 1) + 1)]}$,

- n is the sample (litter) size,
- p is the model-predicted probability of response, and
- θ is the model-predicted intra-litter correlation coefficient.

The overall model should be called into question if the scaled residual values for several individual dose and litter-specific covariate combinations, particularly for the control group or a dose group near the BMD and for litter-specific covariate values close to the overall mean, are greater than 2 or less than -2.

The goodness-of-fit p-values are calculated using a bootstrap approach.

1. The MLE parameter values are used to generate B pseudo-datasets having the same design features (number of doses and number of litters per dose), litter-sizes, and, if necessary, litter-specific covariate values, as the original dataset. What varies from pseudo-dataset to pseudo-dataset are the number of responding "units" within litters, and those are generated, at random, as dictated by the values of the ML estimates.

2. Once the B bootstrap iterations are generated, a statistic referred to as *chi-square* is calculated for each. The *chi-square statistic* is the sum of the squares of the scaled residuals for each litter, as described above. Higher values of that statistic are indicative of poorer match between the model predictions and the data.

Note In traditional testing situations, the chi-square statistic would be approximated by a chi-squared random variable having a certain degree of freedom, and its "significance" (p-value) would be determined from the appropriate chi-squared distribution function.

3. The chi-square statistic from the original data is computed and compared to the values from the B bootstrap iterations. The p-value is the proportion of chi-square values from the iterations that are greater than the original chi-square value.

   High p-values are indicative of adequate fit (i.e., there was a high proportion of chi-square values associated with pseudo-datasets obtained from data known to be consistent with the model and the ML estimates of the model parameters).

   That calculation is repeated three times, and various percentiles of the generated chi-square statistic are presented. This allows the user to determine if enough bootstrap iterations (B) have been specified. The default iterations for B is 1000 and should probably not be reduced. The user may wish to increase the default if the percentiles for chi-square differ markedly across the three runs (specifically the median and lower percentiles), or if the p-values calculated from the three runs differ markedly. This may only be an issue when the p-value is close to the value (e.g., 0.05 or 0.10) used as a critical value for deciding if the fit of the model to the data is adequate. If there is some variability in the p-values, but they are all greater than 0.20, for example, then one probably need not worry about increasing the value for B.

9.4.3 Plot and Error Bar Calculation

The error bars shown for the plots of nested data are calculated in the same way as those for dichotomous data (and described in Section 8.3.3, "Plot and Error Bar Calculation," on page 79).

However, a Rao-Scott transformation is applied prior to the calculations to express the observations in terms of an "effective" number of affected divided by the total number in each group (the format required for the confidence intervals of simple dichotomous responses).

9.4.4 BMD Computation

BMD computation is like that for dichotomous models with the added wrinkle that a value for a litter-specific covariate (LSC) may be used, in addition to dose, to describe changes in the endpoint. It therefore affects the BMD calculation. If an LSC is included in the model, the user can choose to plot results and compute BMDs for one of two specific values of the LSC, either the overall mean (across all dose groups) or the control group mean. Typically, the overall mean is the preferred choice, but the control group mean might be appropriate in certain situations.

For example, suppose the LSC value varies enough from group to group to be "interesting," but it goes up for some dose groups and down for others in a manner that contra-indicates a dose effect. In this case, the user might decide to use the control group mean LSC when the BMD is close to the background dose (i.e., basically deciding that the LSC of interest in that region is more likely to be the average observed for the control group as opposed to the average across all the groups). If a covariate is found to be affected by dose, i.e., if its value appears to have a consistent trend with respect to dose, its use is discouraged.

Details of the BMD calculation are shown in Table 15 above.

9.4.5 BMDL Computation

BMDS currently only calculates one-sided confidence intervals, in accordance with current BMD practice. The general approach to computing the confidence limit for the BMD (called the BMDL here) is the same for all the models in BMDS, and is based on the asymptotic distribution of the likelihood ratio (Crump and Howe, 1985) .

The approach used for all the nested dichotomous models is the same. The equations that define the benchmark response in terms of the benchmark dose and the dose-response model are solved for one of the model parameters, using either the control group mean or the overall mean of the litter-specific covariate. The resulting expression is substituted back into the model equations, with the effect of re-parameterizing the model so that BMD appears explicitly as a parameter. A value for BMD is then found such that, when the remaining parameters are varied to maximize the likelihood, the resulting log-likelihood is less than that at the maximum likelihood estimates by exactly

$$\frac{\chi^2_{1,1-2\alpha}}{2}$$

9.5   Outputs Specific to Frequentist Nested Dichotomous Models

The nested models use a "bootstrap" approach for evaluating the fit of the model to the data under consideration. That approach takes the model with its MLE parameter values and simulates data sets matching the design (doses, sample sizes, etc.) being modeled. For each simulated data set, the scaled residuals are computed and summed to yield a Chi-square test statistic. The distribution of that test statistic over the iterations is compared to the Chi-square test statistic from the observed data. if the model fits the data well, the observed Chi-square should not be in the upper tail of the Chi-square statistic values from the simulations.

Figure 61. Summary of Bootstrap Fit diagnostics.

Bootstrap Results	
# Iterations	1000
Bootstrap Seed	1588629504
Log-likelihood	-269.4786205
Observed Chi-square	19.6087053
Combined P-value	0.996

The Bootstrap Results table summarizes the result of that test for goodness of fit. It reiterates the user-input number of iterations and displays the seed number used to generate the simulations (which may have been chosen randomly by BMDS). The log-likelihood and the Observed Chi-square test statistic pertain to the observed data. The Combined P-value can be used to infer whether the fit is adequate. Small p-values (e.g., less than 0.05 or 0.10) would indicate poor fit.

Figure 62. Bootstrap Run Details

Bootstrap Runs	Bootstrap Chi-square Percentiles				
Run	P-Value	50th	90th	95th	99th
1	0.994	38.15849454	50.4872245	55.44471	61.746206
2	0.998	38.15084656	50.7198124	55.47846	63.35054
3	0.996	38.15970785	50.458614	53.04604	63.025594
Combined	0.996	38.15970785	50.5109863	54.73189	62.677115

The Bootstrap Runs table gives further details about the bootstrap test of fit. Since it is a random procedure (relying on random generation of datasets with the fitted model as the data-generating process) there is the possibility of noise entering into the computations. Thus, BMDS runs the procedure three times and gets a p-value for each. These can be compared to determine if stability has been achieved. If not, the user may wish to increase the number of iterations. Further details include middle and high-end percentiles for the Chi-square test statistic, that can be further compared to the observed value.

Figure 63. Summarized Scaled Residuals.

Scaled Residuals	
Minimum scaled residual for dose group nearest the BMD	-0.31484
Minimum ABS(scaled residual) for dose group nearest the BMD	0.314837
Average Scaled residual for dose group nearest the BMD	-0.31484
Average ABS(scaled residual) for dose group nearest the BMD	0.314837
Maximum scaled residual for dose group nearest the BMD	-0.31484
Maximum ABS(scaled residual) for dose group nearest the BMD	0.314837

Figure 64. Partial capture of the Litter Data table from Results Workbook.

Litter Data						
Dose	Lit. Spec. Cov.	Est. Prob.	Litter Size	Expected	Observed	Scaled Residua
0	9	0.127585814	9	1.14827	0	-1.147257987
0	9	0.127585814	9	1.14827	1	-0.148141348
0	10	0.13234718	10	1.32347	1	-0.301860366
0	10	0.13234718	10	1.32347	2	0.631328744
0	11	0.137108547	11	1.50819	2	0.431109028
0	13	0.14663128	13	1.90621	3	0.857594396
0	14	0.151392646	14	2.1195	3	0.65653973
0	14	0.151392646	14	2.1195	2	-0.089101984
0	15	0.156154013	15	2.34231	2	-0.243481535
0	16	0.160915379	16	2.57465	1	-1.071325161
25	9	0.301527034	9	2.71374	2	-0.518421334
25	9	0.301527034	9	2.71374	5	1.660602955
25	10	0.297781775	10	2.97782	2	-0.676196332
25	10	0.297781775	10	2.97782	1	-1.367732493
25	11	0.294373053	11	3.2381	4	0.504037656
25	12	0.291294677	12	3.49554	3	-0.314836859
25	13	0.288539894	13	3.75102	6	1.376689417
25	14	0.286101463	14	4.00542	6	1.179530167
25	14	0.286101463	14	4.00542	4	-0.003205497
25	14	0.286101463	14	4.00542	3	-0.594573329
50	7	0.433391608	7	3.03374	2	-0.788462565
50	10	0.410232266	10	4.10232	5	0.577118305
50	10	0.410232266	10	4.10232	5	0.577118305
50	11	0.40315061	11	4.43466	4	-0.267167737
50	11	0.40315061	11	4.43466	4	-0.267167737

In simple dichotomous modeling, there is a single scaled residual for each dose group. For nested designs, the probabilities of response and therefore the scaled residuals will vary across experimental units ("litters"). That variation is shown in the Litter Data table and summarized in the Scaled Residuals table. The summary is an attempt to capture a general impression of the closeness of the observed response rates to those predicted by the model. As is typical, values greater than 2 in absolute value may affect the user's assessment of fit. The Litter Data table shows the model-predicted probability of response and expected number of responders (= Est. Prob * Litter Size).

10.0 Multiple Tumor Analysis

Figure 65. Multi-tumor (MS_Combo) analysis.

10.1 Dichotomous—Multi-tumor Models and Options

The modeling option 'Dichotomous –Multi-tumor (MS_Combo)' is a special application of dichotomous modeling, not previously discussed. It is offered as a convenience to the user who may be interested in evaluating the combined effect of two or more independent, dichotomous responses. It is specialized in the additional sense that it only runs multistage models to fit the dose-response relationship. It returns BMD estimates (and related bounds) for the risk of responding with one or more of the endpoints in question. It is most often limited to analyses of cancer data where the component data sets are for tumors occurring at various sites, hence its name.

As in previous versions of BMDS, BMDS 3 allows users to run the EPA's Multi-tumor (MS_Combo) model to determine the BMD, BMDL and BMDU that is associated with a specified added or extra risk of experiencing at least one of the multiple tumor types. However, unlike previous versions of BMDS, BMDS 3 provides users with the option to manually select, or allow BMDS to "Auto-select," the degree of Multistage model to apply

to each dataset. The auto-selection process follows the most recent EPA technical guidance for selecting the Multistage model degree for the analysis of cancer datasets, which differs from the model selection process described by EPA (2012) for other modeling scenarios.

10.2   Assumptions

The analyses of multiple tumors have the following assumptions and results:
1. The tumors are statistically independent of one another.

> **Note:**   Unless there is substantial biological evidence to indicate that the tumor types are *not* independent—conditional on model parameter values—the approach based on independence is considered appropriate.

2. A multistage model is an appropriate model for each of the tumors separately. The individual multistage models fit to the individual tumors need not have the same polynomial degree, however.
3. The user is interested in estimating the risk of getting one or more of the tumors being analyzed; the results indicate the BMD and BMDL associated with the user-defined benchmark response (BMR) level, where the BMD and BMDL are the maximum likelihood and lower bound estimates of the dose that is estimated to give an extra risk equal to the BMR for the "combination" (getting one or more of the tumors).

In accordance with EPA cancer guidelines (U.S. EPA, 2005), a Multiple Tumor Analysis will always run the restricted form of the Multistage model.

BMDS 3 allows users to have BMDS "Auto-Select" the appropriate polynomial degree of the Multistage model for each tumor dataset. When the "Auto-Select" feature is used, BMDS runs all relevant forms of the Multistage model and selects the polynomial degree to use based on the current EPA Multistage model selection criteria for tumor analyses. This is the default option in BMDS 3.0, but the user can also choose to manually set the polynomial degree for each dataset. In any case, it is ultimately the user's responsibility to ensure that the degree of the polynomial and other selections for modeling parameters are as desired and appropriate for the dataset(s) being analyzed.

10.3   Multi-tumor (MS_Combo) Model Description

> **Note**   Before using MS_Combo, it is strongly recommended that users refer to the Technical Guidance on choosing the appropriate stage of a multistage model for cancer modeling. The Technical Guidance includes background information on the assumptions and application of the BMDS MS_Combo program.

The purpose of the MS_Combo program in BMDS is to allow the user to calculate BMDs and BMDLs for a combination of tumors (corresponding to a defined risk of getting one or more of those tumors) when the individual tumor dose-responses have been modeled using a Multistage-Cancer model.

The output of an MS_Combo run will present the results of fitting each individual tumor (including the BMD and BMDL for that tumor) plus the combined log-likelihood, BMD and BMDL for the combination of specified tumor responses.

In practice, the user should investigate each tumor individually and determine which degree of the Multistage-Cancer model is most appropriate for each individual tumor. That determination will involve all the usual considerations of fit, AIC, etc.

Once a specific form of the Multistage-Cancer model is chosen for each of the tumors of interest (they need not have the same degree across all the tumors in question), the user should specify those choices in the MS_Combo run.

Because of the form of the multistage model, the MLE estimates for the combined risk are a function of the parameter values obtained for the individual tumor multistage model fits. In fact, the combined probability function has a multistage model form:

$$P(d) = 1 - e^{\{-(\beta_0 + \beta_1 d + \beta_2 d^2 + \cdots)\}}$$

and the terms of the combined probability function $(\beta_0, \beta_1, \dots)$ are specified as follows:

$$\beta_0 = \sum \beta_{0i}$$

$$\beta_1 = \sum \beta_{1i}$$

$$\beta_2 = \sum \beta_{2i}$$

etc.

where the sums are over i = 1, …, t, with

t being the number of tumors under consideration, and

β_{xj} being the xth parameter (0, 1, …) for tumor j.

The β_{xj} values are available directly from the Multistage-Cancer runs performed on the individual tumors, but MS_Combo performs the calculations for the user, completing the summations of the individual terms and computing the BMD based on the combined parameter values and the user-specified BMR.

A profile-likelihood approach is used to derive the BMDL.

1. Given the BMD and the log-likelihood associated with the MLE solution, a target likelihood is defined based on the user-specified confidence level (e.g., 95%).
2. That target likelihood is derived by computing the percentile of a chi-square (1 degree of freedom) corresponding to the confidence level specified by the user (actually, the alpha associated with the confidence level, times 2).
3. That percentile is divided by 2 and subtracted from the maximum log-likelihood.
4. That derivation is based on a likelihood ratio test with one degree of freedom; it can be shown that estimating the BMDL corresponds to losing one degree of freedom, regardless of the number of tumors being combined.
5. The BMDL for the combined response (one or more of the tumors of interest) is defined as the smallest dose, D, for which the following two conditions are satisfied:
    i. There is a set of parameters such that the combined log-likelihood using D and those parameters is greater than or equal to the target likelihood), and
    ii. For that set of parameters, the risk at D is equal to the user-specified BMR.

Note that the combined log-likelihood is a function of the fits of the individual tumors (the sum of the individual log-likelihoods), obtained using their tumor-specific β values. Thus, the search for the parameters of the combined Multistage-Cancer model varies the individual-tumor β values in such a way that the individual log-likelihoods add up to a combined likelihood within the range desired (greater than or equal to the target).

However, to satisfy the second constraint, the sums of the individual-tumor parameters (shown above to be the parameters of the combined probability function) are used to evaluate the risk for any proposed BMDL, D.

Note that the individual tumors need not be modeled with the same degree of the Multistage-Cancer model. Any terms not included for an individual tumor are assumed to be zero (and will remain at zero during BMDL optimization) in the summations shown above.

10.4   Entering Multi-tumor Data

Figure 66. *Dataset options for multi-tumor data.*

☑ Enable	DataSets	Degree	Background
☑	tumor 1	auto-select ▾	Estimated ▾
☑	tumor2	auto-select ▾	Zero ▾
☑	tumor3	1 ▾	Estimated ▾

The user can set two options per dataset: (Polynomial) Degree and Background.

10.4.1 Setting Polynomial Degree
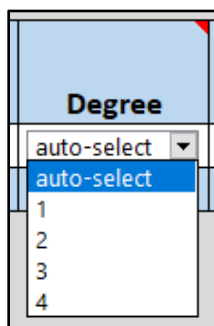
Select whether BMDS auto-selects the appropriate polynomial degree or whether it should be user-specified.

BMDS will recommend a model degree based on the decision logic and settings found on the Logic tab. For more details, refer to Section 12.0, "BMDS Recommendations and Decision Logic," on page 114.

- If a user opts for BMDS' auto-select functionality, the best fitting model is chosen according to the Technical Guidance on choosing the appropriate stage of a multistage model for cancer modeling. If no model can be chosen based on that criteria, then the model is removed from the MS_Combo results.
- The MS_Combo Decision logic uses the user-defined test thresholds from the Logic tab for the following criteria:
  - Goodness of fit p-test (cancer)
  - Ratio of BMD/BMDL (caution)
  - Abs(Residual of interest) too large
  - Abs(Residual at control) too large

- If the results do not meet the "Test Threshold" value set in the "Ratio of BMD/BMDL (Caution)" on the Logic tab, BMDS displays a pop-up message to

the user that "BMD/BMDL ratio > [value]; consider consulting a statistician." However, the model is not removed from the MS_Combo results.

- If a user specifies a degree for a given model, that degree will be used regardless of model fit.

10.4.2 Background
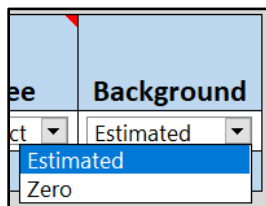
Choices are Estimated (default) or Zero; this should usually be Estimated unless there exists (from other evidence) strong evidence for absolutely zero probability of response in the absence of exposure to the chemical under consideration.

Note   Do not set Background to Zero when there are responses in the control group.

10.5   Options

The Multi-tumor options are the same as for the Dichotomous options.

10.5.1 Risk Type
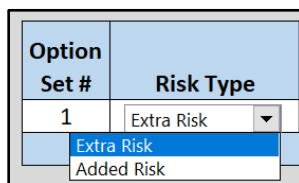
Choices are "Extra Risk" (Default) or "Added Risk."

Added risk is the additional proportion of total animals that respond in the presence of the dose, or the predicted probability of response at dose d, $P(d)$, minus the predicted probability of response in the absence of exposure, $P(0)$. $I.e. added\ risk\ =\ P(d) - P(0)$

Extra risk is the additional risk divided by the predicted proportion of animals that will not respond in the absence of exposure, $1 - P(0)$. $I.e., extra\ risk\ - \frac{P(d) - P(0)}{1 - P(0)}$.

10.5.2 BMR

The BMR is the value of risk (extra or added, as specified by the user) for which a BMD is estimated. BMR must be between 0 and 1 (not inclusive). If $P(0) > 0$, then values for BMR greater than $1 - P(0)$ will result in an error when the risk type is added risk. That is because the maximum added risk that can ever be achieved is $1 - P(0)$. In practice, this should not typically be an issue because one usually is interested in BMR values in the range of 0.01 to around 0.10.

10.5.3 Confidence Level

The Confidence Level is real number between 0 and 1; 0.95 is recommended by EPA (U.S. EPA, 2012).

10.6 Running an Analysis and Viewing Results

When "Run Analysis" is selected a separate Results Workbook of multi-tumor results is created. The workbook will include results for each individual tumor considered separately (using the chosen dataset-specific options), and the corresponding estimate of the BMD and BMDL for the combined tumor probability for the risk type, BMR and confidence levels specified by the user.

Plots for individual multistage model runs will be shown on the individual model results tabs. If the "Auto-Select" feature was used to select the Multistage polynomial degree, the user should verify that the resultant model fits are adequate in the desired dose-response region. If the user wants to try a different Multistage polynomial degree they can re-run the analysis using a specified degree instead of "Auto-Select."

For more information on the Results Workbook tables, refer to Section 8.4, "Outputs Specific to Frequentist Dichotomous Models," on page 86.

10.7 Troubleshooting a Tumor Analysis

If one or more of the tumors is estimated to have a BMD greater than three times the highest dose tested (for that tumor), then the multiple tumor analysis will stop at an intermediate point, i.e., after the fitting has been done for the tumor in question and the magnitude of that BMD has been determined. No tumors listed below that tumor will be analyzed, and no combination will be completed.

It is probably the case that the tumor in question will not add substantially to the estimation of a BMD for the combinations of tumors, assuming other tumors have BMDs less than three times the highest dose; that is because the magnitude of response for the tumor in question has not even reached the benchmark response level for such a high

exposure and so its individual contribution to the risk of getting one or more of the tumors being analyzed will be small in comparison to that for the other tumors. The user might attempt a combination that does not include the tumor in question.

11.0 Special Considerations

11.1 AIC for Continuous Models

To facilitate comparing models with different likelihoods (i.e., Normal vs. Lognormal), the log-likelihood for the Normal and Lognormal distributions are calculated using all normalizing constants. *This results in different numerical AIC values than those given in earlier BMDS versions.*

Even though the BMDS 3 AIC values for continuous models differ from those in BMDS 2.x versions, if the models have the same underlying distribution, then *the difference of the AICs* will be the same as previous versions of BMDS. This assumes that the BMDS 3 and BMDS 2.x model fits are the same for the two models being compared. The AIC difference may not be the same if one or more of the model fits differ between the two versions (e.g., if one or more of the BMDS 3 models provide an improved fit to the data over the corresponding BMDS 2.x model).

However, when comparing models having different parametric distributions, the AIC differences will not be the same as previous BMDS versions. For these comparisons, the AIC calculated using the BMDS 3 software is correct and will result in the proper comparison between any two models regardless of underlying distribution.

Caution

A note of caution is required for situations where only the sample mean and sample standard deviation are available (summarized data) for which the log-scale parameters are only approximated when assuming lognormally distributed responses.

In such cases, the normalizing constant for the lognormal log-likelihood is only approximated. It is the same normalizing constant for any model fit under a lognormal distribution assumption, so comparisons among models using that assumption are valid.

However, **comparisons of results where one set of results was obtained assuming normality and one set was obtained assuming lognormality should be made with caution.**

If the AICs are "similar" (using that term loosely, because no specific guidance can be offered here), then one ought not to base model selection on AIC differences. Selection when the AIC differences are "larger" may not be problematic, since the approximation used should not be too bad.

A conservative position would be that comparisons of models assuming the Normal distribution to those assuming the Lognormal distribution should not be made using the AIC, if the underlying data are presented in summarized form (i.e., only sample means and sample standard deviations are available).

Note AIC values for **dichotomous** models should be the same from BMDS 2.x to BMDS 3.x.

11.2 Continuous Response Data with Negative Means

Data with negative means should only be modeled with a constant variance model.

It may occasionally be the case that, when modeling transformed data, the user will need to model negative data. In this case, the transformation used should be a variance-stabilizing transformation so that a constant-variance model would be appropriate.

If a standard deviation-based BMR is used to define the BMD calculations, then a constant can be added to all the observations (or means) to make the values (means) positive. That will not change the standard deviations of the observations and would allow the user to model the variance.

11.3 Test for Combining Two Datasets for the Same Endpoint

BMDS does not include a formal test for similarity of dose response across covariate values (e.g., across class variables like species or sex). EPA's categorical regression software, CatReg, has that capability.

However, the following procedure can be used in BMDS if there are dose-response data for two experiments that the user is considering combining (e.g., for the two sexes within a species, or two species, etc.).

1. Choose a single model to consider for both datasets.
2. Model the two datasets separately. For each run, record the following:
   - Maximum log-likelihood for each dataset. Add the two log-likelihoods (one from each dataset) to get the **summed** log-likelihood.
   - The number of unconstrained parameters for each dataset. Add those numbers from each run to get the **summed** unconstrained parameters.
3. Combine the data from the two experiments and model them together. Record the following:
   - The maximum log-likelihood for the combined dataset. This will be the **combined** log-likelihood. The fitted model log-likelihoods are reported in the Analysis of Deviance (dichotomous endpoints) or Likelihoods of Interest (continuous endpoints) tables.
   - The number of unconstrained parameters for the combined dataset. This will be the **combined** unconstrained parameters.
4. Subtract the combined log-likelihood from the summed log-likelihood. Then, multiply the difference by 2.
5. Compare the value from Step 4 to a chi-squared distribution. The degrees of freedom for that chi-squared distribution will be the difference between the summed unconstrained parameters (Step 2) and the combined unconstrained parameters (Step 3).

   If the value from Step 4 is in the tail (say, greater than the 95th percentile) of the chi-squared distribution in question, then reject the null hypothesis that the two sets have the same dose-response relationship. If rejection occurs, then infer that it is not proper to combine the two datasets.

12.0 BMDS Recommendations and Decision Logic

Users familiar with the BMDS Wizard application will note that BMDS 3 uses a similar approach to analyzing modeling results and making automatic recommendations regarding model selection that are consistent with the 2012 EPA Benchmark Dose Technical Guidance (U.S. EPA, 2012).

These criteria can be altered in the Logic tab of the BMDS 3 Analysis Workbook, as shown in Figure 67. Decision logic can be turned on or off, and specific criteria can be enabled or disabled for different dataset types. Notice that the logic depends on what type of data is being analyzed (continuous, dichotomous, nested).

Figure 67. BMDS 3 Logic tab with EPA default recommendation decision logic.

Decision-Logic						
Recommend model in Viable Bin?			TRUE			
Recommend model in Questionable Bin?			FALSE		**Reset To Default Logic**	
BMDL range deemed "sufficiently close" to use lowest AIC instead of lowest BMDL in viable models:			3			

Model Recommendation/Bin Placement Logic

Test Description	Test On/Off			Test Threshold (where appropriate)	Bin Placement if Test is Failed	Notes to Show
	Continuous	Dichotomous	Nested			
BMD calculated	On	On	On	N/A	Unusable Bin	BMD not estimated
BMDL calculated	On	On	On	N/A	Unusable Bin	BMDL not estimated
BMDU calculated	Off	Off	Off	N/A	No Bin Change (Warning)	BMDU not estimated
AIC calculated	On	On	On	N/A	Unusable Bin	AIC not estimated
Constant Variance	On	Off	Off	0.05	Questionable Bin	Constant variance test failed (Test 2 p-value < 0.05)
Non-Constant Variance	On	Off	Off	0.05	Questionable Bin	Non-constant variance test failed (Test 3 p-value < 0.05)
Goodness of fit p-test	On	On	On	0.1	Questionable Bin	Goodness of fit p-value < 0.1
Goodness of fit p-test (cancer)	Off	On	Off	0.05	Questionable Bin	Goodness of fit p-value < 0.05
Ratio of BMD/BMDL (serious)	On	On	On	20	Questionable Bin	BMD/BMDL ratio > 20
Ratio of BMD/BMDL (caution)	On	On	On	3	No Bin Change (Warning)	BMD/BMDL ratio > 3
Abs(Residual of interest) too large	On	On	On	2	Questionable Bin	|Residual for Dose Group Near BMD| > 2
BMDS Model Warnings	Off	On	Off	N/A	No Bin Change (Warning)	BMDS output file included warning
BMD higher than highest dose	On	On	On	1	No Bin Change (Warning)	BMD higher than maximum dose
BMDL higher than highest dose	On	On	On	1	No Bin Change (Warning)	BMDL higher than maximum dose
BMD lower than lowest dose (warning)	On	On	On	3	No Bin Change (Warning)	BMD 3x lower than lowest non-zero dose
BMDL lower than lowest dose (warning)	On	On	On	3	No Bin Change (Warning)	BMDL 3x lower than lowest non-zero dose
BMD lower than lowest dose (serious)	On	On	On	10	Questionable Bin	BMD 10x lower than lowest non-zero dose
BMDL lower than lowest dose (serious)	On	On	On	10	Questionable Bin	BMDL 10x lower than lowest non-zero dose
Abs(Residual at control) too large	On	On	On	2	No Bin Change (Warning)	|Residual at control| > 2
Poor control dose std. dev.	On	Off	Off	1.5	No Bin Change (Warning)	Modeled control response std. dev. >|1.5| actual response std. dev.
D.O.F. equals 0	On	On	On	N/A	Questionable Bin	d.f.=0, saturated model (Goodness of fit test cannot be calculated)

Help | Main | Data | Report Options | **Logic** | ModelParms

Based on the decision logic entered by the user as described above, BMDS will attempt to select a "recommended" model. A user must ultimately select a model and may choose to disagree with the BMDS auto-determination.

BMDS 3 automatically generates suggested text for the "BMDS Recommendation" and "BMDS Recommendation Notes" columns of the Results Workbook summary tables and the Word Report File tables.

While some reformatting is allowed in the Results Workbook (e.g., row heights, column widths, and the size, design, and position of plots), the text and numeric results cannot be modified. However, the Word Report files can be modified extensively, and the user is encouraged to take advantage of this flexibility to change and/or expand on the table headers and the justification provided for why a model was selected.

BMDS 3 places each model into one of three different bins:
- Viable—highest quality model, no serious deficiencies found based on user-defined logic but may contain warnings
- Questionable—some serious deficiencies with model based on user-defined decision logic
- Unusable—required outputs such as BMD or BMDL are not estimated

The default settings for factors (tests) that determine bin placement are consistent with EPA Benchmark Dose Technical Guidance (U.S. EPA, 2012) and can be reset by selecting "Reset to Default Logic."

The following default settings that can cause test failure, and thereby affect bin placement, are not explicitly given in the EPA BMD guidance. They have been assigned based on general EPA practice and are, therefore, more open to user discretion.
- BMDL range default fail: > 3-fold
- Constant and non-constant variance p-value[10] default fail: < 0.05
- Ratio of BMD/BMDL (serious) default fail: > 20
- BMD lower than lowest dose (serious) default fail: > 10
- BMDL lower than lowest dose (serious) default fail: > 10

After all models of the same Option Set (i.e., same model run settings such as BMR Type, BMRF, etc.) have been placed into one of three different quality bins, a model is recommended from the "Viable" bin based on BMDL or AIC criteria defined in the 2012 EPA Benchmark Dose Technical Guidance (U.S. EPA, 2012).

12.1   Changing the Decision Logic

BMDS automatically attempts to recommend a best-fitting BMDS model, using the 2012 EPA guidance (U.S. EPA, 2012) and additional criteria as described in the previous section. These criteria can be altered in the Logic worksheet. Decision logic can be turned on or off, and specific criteria can be enabled or disabled for different dataset types.

Based on the decision logic entered by the user, BMDS will attempt to select a model that will be "recommended" as a best-fitting model.

Any changes to the BMDS default logic should be noted in any results or reports.

Some grayed-out cells in the Logic table are not selectable and therefore cannot be edited. Also, values for the columns "Bin Placement if Test is Failed" and "Notes to Show" cannot be edited.

Logic settings are saved with the Settings and Results Workbooks, so loading a previously run analysis will restore any customized logic settings.

---

[10] Examples given in EPA BMD guidance (U.S. EPA, 2012).suggest a criteria of p-value > 0.1 for variance models, but this has since been relaxed in practice. Future EPA guidance will reflect this change.

Enabling/Disabling Model Tests

To turn specific model testing on or off, select the down-pointing triangle for the cell and select "On" or "Off."

Figure 68. Toggling model tests on/off in the Logic table.

Test On/Off		
Continuous	Dichotomous	Nested
On	On	On
On	On	On
Off	Off	Off
On	On	On

To Edit Test Threshold Values

Select the cell and double-click to display the cursor. Edit the cell value as desired.

The user could, for example, change a threshold value for a Dichotomous model analysis and then restore the original value for a Continuous analysis.

Note   When changing the decision logic; an experienced user or statistician should be consulted to ensure the criteria selections are reasonable.

Figure 69. *Flowchart of BMDS 3 model recommendation logic using EPA default logic assumptions.*

Viable bin
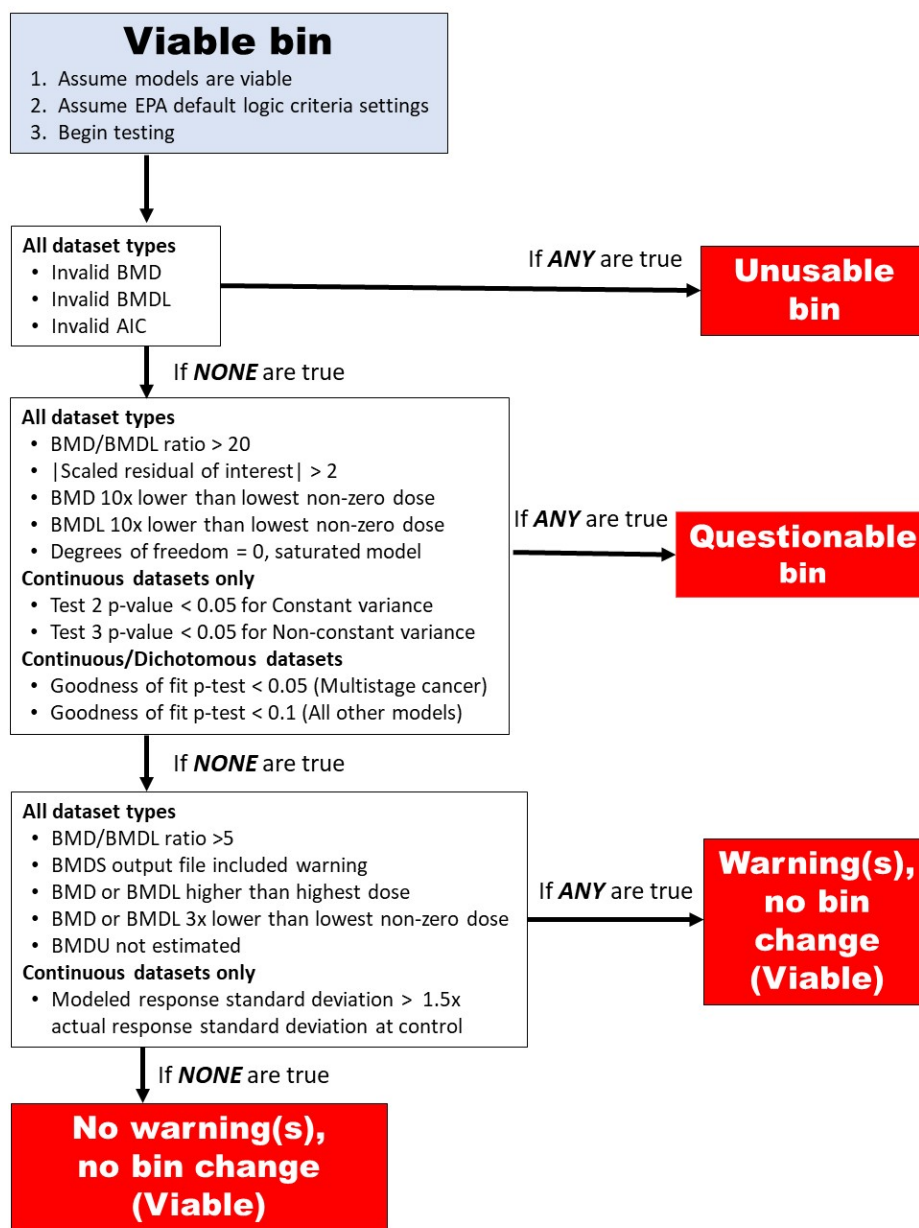1. Assume models are viable
2. Assume EPA default logic criteria settings
3. Begin testing

All dataset types
- Invalid BMD
- Invalid BMDL
- Invalid AIC

If **ANY** are true → **Unusable bin**

If **NONE** are true

All dataset types
- BMD/BMDL ratio > 20
- |Scaled residual of interest| > 2
- BMD 10x lower than lowest non-zero dose
- BMDL 10x lower than lowest non-zero dose
- Degrees of freedom = 0, saturated model

Continuous datasets only
- Test 2 p-value < 0.05 for Constant variance
- Test 3 p-value < 0.05 for Non-constant variance

Continuous/Dichotomous datasets
- Goodness of fit p-test < 0.05 (Multistage cancer)
- Goodness of fit p-test < 0.1 (All other models)

If **ANY** are true → **Questionable bin**

If **NONE** are true

All dataset types
- BMD/BMDL ratio >5
- BMDS output file included warning
- BMD or BMDL higher than highest dose
- BMD or BMDL 3x lower than lowest non-zero dose
- BMDU not estimated

Continuous datasets only
- Modeled response standard deviation > 1.5x actual response standard deviation at control

If **ANY** are true → **Warning(s), no bin change (Viable)**

If **NONE** are true

No warning(s), no bin change (Viable)

Model Recommendation Criteria
- If range of BMDLs from models remaining in "Viable" bin is < 3, recommend BMDL from model with lowest AIC.
- Otherwise, recommend lowest BMDL from models remaining in "Viable" bin.

13.0 References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In BN Petrov; F Csaki (Eds.), 2nd International Symposium on Information Theory (pp. 267-281). Budapest, Hungary: Akadémiai Kiadó.

Carr, GJ; Porter, CJ. (1991). An evaluation of the Rai and Van Ryzin Dose-Response Model in teratology. Risk Anal 11: 111-120. http://dx.doi.org/10.1111/j.1539-6924.1991.tb00581.x

Crump, KS. (1984). A new method for determining allowable daily intakes. Toxicol Sci 4: 854-871. http://dx.doi.org/10.1016/0272-0590(84)90107-6

Crump, KS; Howe, RB. (1985). A review of methods for calculating statistical confidence limits in low dose extrapolation. In DB Clayson; D Krewski; I Munro (Eds.), Toxicological risk assessment: Vol 1 Biological and statistical criteria (pp. 187-203). Boca Raton, FL: CRC Press, Inc.

Fleiss, JL; Levin, B; Paik, MC. (2003). Statistical methods for rates and proportions. Hoboken, NJ: John Wiley & Sons, Inc.

Fletcher, D; Turek, D. (2012). Model-averaged profile likelihood intervals. Journal of Agricultural, Biological, and Environmental Statistics 17: 38-51. http://dx.doi.org/10.1007/s13253-011-0064-8

Jeffreys, H. (1998). The theory of probability. Oxford, United Kingdom: Oxford University Press.

Kimmel, CA; Gaylor, DW. (1988). Issues in qualitative and quantitative risk analysis for developmental toxicology. Risk Anal 8: 15-20. http://dx.doi.org/10.1111/j.1539-6924.1988.tb01149.x

Kupper, LL; Poetier, C; Hogan, MD; Yamamoto, E. (1986). The impact of litter effects on dose-response modeling in teratology. Biometrics 42: 85-98. http://dx.doi.org/10.2307/2531245

Nitcheva, DK; Piegorsch, WW; West, RW. (2007). On use of the multistage dose-response model for assessing laboratory animal carcinogenicity. Regul Toxicol Pharmacol 48: 135-147. http://dx.doi.org/10.1016/j.yrtph.2007.03.002

Rai, K; Van Ryzin, J. (1985). A dose-response model for teratology experiments involving quantal responses. Biometrics 41: 1-9.

RIVM (National Institute for Public Health and the Environment (Netherlands)). (2018). PROAST. Retrieved from https://www.rivm.nl/en/Documents_and_publications/Scientific/Models/PROAST

U.S. EPA (U.S. Environmental Protection Agency). (2012). Benchmark dose technical guidance. (EPA/100/R-12/001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/risk/benchmark-dose-technical-guidance

U.S. EPA (U.S. Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment. (EPA/630/P-03/001B). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/sites/production/files/2013-09/documents/cancer_guidelines_final_3-25-05.pdf