

Comment	EPA Response <sup>1</sup>
<b>Overarching</b>	
<p>1. DoD believes this to be an extraordinarily comprehensive instruction on the derivation of Benchmark dose (BMD) and the use of EPA's BMD software. It provides extensive information on the mathematics, definitions and rationales of application for quantal, discontinuous, and continuous data for cancer and noncancer responses. Although the authors stress that "The document is not intended as a primer on BMD modeling", "the guidance document provides a step-by-step process to be used in evaluating studies and endpoint types that are appropriate for modeling, selecting the BMR level, modeling fitting and BMD computation, judging the fit of the model, and calculating the BMDL." This appears to be a distinction without a difference. DoD feels that the EPA Benchmark Dose Technical Guidance document is intended to provide a useful resource to risk assessors and dose-response modelers across EPA. The document contains significant science and science policy changes since it was last subject to public review (which we believe was in 2000) and external expert peer review. To that end, DoD's comments highlight these changes and recommends either deletions or external peer review (for new scientific practices) or public review (for new science policy) to ensure that EPA is following its own policies with respect to document review.</p> <p>EPA leadership needs to consider the value of external peer and public reviews and the consequences of failing to have science and science policy issues that effect multiple program offices reviewed.</p>	<p>EPA has benefitted from a decade of experience employing BMD methodology in numerous programmatic and ORD assessments and has used that experience to improve the quality of the document. As noted in the comment, the document has been revised incorporating changes. In order to be responsive to the interagency and other reviewers EPA has made changes since the IAR draft document. EPA has revised the document based on comments from two rounds of interagency reviews in addition to peer review. EPA would like to emphasize a few key points:</p> <ul style="list-style-type: none"> <li>• Science in the field of statistics governing basic BMD methodology has not changed significantly over the past decade;</li> <li>• EPA has incorporated experience gained from assessments to improve the quality of the BMD Technical Guidance;</li> <li>• Assessors from other agencies have requested a final release of the document to address the uncertainty associated with and,</li> <li>• In weighing the benefits of an additional round of public comments to a more expeditious release of an important guidance document, EPA considers the time factor to be the more beneficial.</li> </ul>

<sup>1</sup> **Text in bold, red font indicates changes to be made to document in response to comments.**  
**Text in bold, green font indicates comments that are in error.**

Comment	EPA Response <sup>1</sup>
<p>2. The guidance could be strengthened by a more thorough discussion of the sources and types of uncertainty (other than uncertainty due to measurement or sampling error) inherent in calculation of a benchmark dose and discussion of communicating that uncertainty to those risk assessors and risk managers who may use the BMD in assessments or risk management decisions. We recognize that we have made this comment before.</p>	<p>Section 2.3 addresses the modeling of the data including noting sources of uncertainty. For a document of this type EPA feels that the content on uncertainty is appropriate for the level at which the document is developed and for the intended audience.</p>
<p>3. This is one of those cases where a response to agency comments would have been helpful -- what did they agree to change, if they didn't change, why not (what was their rationale)?</p>	<p>Only an internal draft was developed.</p>
<p>4. Bottom line is that we really think this needs a public review given what appears to be much significant new science policy since the last review. We know they are anxious to get this out - but given the years of drafting what harm to put some of these issues to bed?</p>	<p>See Response to #1.</p>
<p><b>Specific and Editorial Comments</b></p>	
<p>5. Page 21. The modification makes the first bullet internally inconsistent. The default UF is 10; a UF of "up to 10" can not be a default, as there are an infinite number of values that the UF can take. Either use the original text or correctly state that the default is 10, and that a value of between 1 and 10 may be selected depending on the particular information.</p>	<p>Comment appears to apply to page 4. <b>EPA agrees that the content will be clearer if the use of “default” is removed.</b></p>
<p>6. Page 22, Section 2.2 states: “The Agency does not currently have guidance to assist in making such judgments for the selection of the appropriate response level, or BMR, to use with BMD modeling for particular applications (e.g., for calculating reference doses or relative potency factors), and such guidance is beyond the scope of this document, This section outlines some general principles to consider along with case-specific issues.” Consistent with the exec summary and guidance provided in the document, EPA should clarify that that guidance on defaults for reporting purposes are provided. We suggest a footnote which states: “However, we note that guidance on defaults for reporting certain information is recommended. For quantal data, an extra risk of 10% is the BMR for standard reporting. For continuous data, it is recommended that the BMD corresponding to one control SD from the control mean response be presented for reporting purposes.”</p>	<p>The commenter recommends adding a footnote that provides the section’s conclusions to an introductory section, before the details leading to the document’s recommendation have been presented. <b>The introduction can be clearer that the section addresses what the commenter is looking for.</b> Otherwise, the specific suggestion leads to odd organization, with important information which is the main purpose of the document being repeated in random footnotes. Footnotes are for further clarification that would otherwise break up the continuity of the document.</p>

Comment	EPA Response <sup>1</sup>
<p>7. Page 22/23 EPA states: “For comparing potencies across chemicals or endpoints (e.g., for chemical rankings) for dichotomous data, a response level of 10% extra risk has been commonly used to define BMDs, also known as effective doses (i.e., ED10s).” Suggest adding a footnote which states: “It is important to note that while this guidance suggests this value as a standard default for reporting purposes, the BMD choice depends on the specific application and on biological considerations (including adversity). These determinations are not the subject of this guidance document.”</p>	<p>Disagree. The commenter misunderstands the point of the example application. The recommended text is redundant and confusing rather than clarifying (see previous response about odd use of footnotes).</p>
<p>8. Page 23 suggested insert: “this is in contrast to using lower <u>or upper</u> bounds for PODs for reference values or cancer potency estimates.”</p>	<p>Reporting recommendations will also vary depending on the purpose of the assessment. Reporting the lower 95% bound has been the standard practice for risk assessment for many years. When appropriate the assessor is encouraged to include the upper bound, as noted in Section 2.4 Reporting Recommendations.</p>
<p>9. Page 24 suggested edit: “Alternatively, in the absence of any cogent basis for selecting a BMR for continuous data, a BMR of one control SD (or lower <u>or higher</u> if warranted by statistical and biological considerations) change from the control mean can be used, as is recommended as the standardized reporting level for comparisons for continuous data.”</p>	<p><b>EPA may consider this clarification.</b></p>
<p>10. Page 26, suggested edit: “In the absence of any other idea of what level of response to consider adverse, a change in the mean equal to one control SD (or lower, e.g., 0.5 SD, for frank effects) from the control mean should be used.” Shouldn’t it be lower or higher based on the goals of the risk assessment? Since this is technical guidance for broad use, not just use, it seems that EPA should not be setting confusing defaults in the absence of information. If there is no information, than the one control SD seems appropriate, the parenthetical then makes it messy.</p>	<p><b>EPA agrees that this qualifier is not as helpful as was intended. It will be revised consistent with the suggestion.</b></p>
<p>11. Page 28, suggested edit to reflect the wording on EPA’s website (<a href="http://www.epa.gov/ncea/bmds/bmds_training/methodology/intro.htm">http://www.epa.gov/ncea/bmds/bmds_training/methodology/intro.htm</a>), the edit is intended to recognize the value in addressing the typical practice of BMR</p>	<p>EPA disagrees with this revision. The website will be revised to coordinate with the guidance document.</p>

Comment	EPA Response <sup>1</sup>
<p>selection – i.e. just because a 5% BMR is easily supported, that doesn't necessarily mean that it is typically used. "From a statistical standpoint, <del>most reproductive and developmental studies with nested study designs easily support a BMR of 5%.</del> <u>reproductive and developmental studies having nested study designs often have greater sensitivity, and for such studies a BMR of 5% has typically been used.</u>"</p>	
<p>12. Page 37, shouldn't this mention that using an approach which drops dose groups is not preferred? Similarly, shouldn't example A2 on page 51 also note that this is not a preferred approach?</p>	<p>EPA disagrees that omitting data that may inform the dose-response curve is not a preferred approach; there are cases in which dropping data would be the preferred approach. EPA acknowledged the complexity of the issue by revising the text and editing the content provided by OMB to be more succinct.</p>
<p>13. Page 39, Kang et al. (2000) is cited in the first paragraph of section 2.3.7. I recommend deleting this cite because the article does not propose Bayesian model averaging which is the point of the sentence. I note the list of References has already deleted Kang et al.</p>	<p>Kang et al. (2000) should be cited in the document because it was one of the first publications to suggest model averaging of any sort; deleting it was an oversight. <b>Will doublecheck whether it was cited in the right place.</b></p>
<p>14. Page 42 suggested edit: "Since there is no clear remaining biological or statistical basis on which to choose among them, <del>the lowest BMDL may be selected as a reasonable conservative estimate</del> <u>a policy decision must be made. This science informed judgment should be appropriate for the context of the analysis and use of this judgment should be clearly communicated to risk managers.</u> Additional analysis and discussion may be appropriate, which might include the use of additional models, the examination of the parameter values for the models used, or an evaluation of the BMDs to determine if the same pattern exists as for the BMDLs. Discussion of the decision procedure should always be provided.</p>	<p>EPA disagrees. The current text reflects EPA's current policy posture.</p>
<p>15. Page 42. This entire section was not in the draft that was externally peer reviewed. The section contains many scientific and science policy determinations on which experts might disagree. It is also in conflict with the second bullet on page 43.</p>	<p>See response to comment #1.  EPA disagrees; added text is largely clarification of existing procedures.</p>

Comment	EPA Response <sup>1</sup>
<p>Adding new science or science policy text after external peer review is not consistent with EPA's standard practice. Public review of new science or science policy would appear appropriate</p>	
<p>16. Page 43. The statement that EPA does not have any guidance regarding selection of the BMD is not correct. EPA's 2005 cancer guidelines state that the POD, and by inference the BMR, should be near the range of the available data. As EPA's 2005 cancer guidelines are cited in this document, the statement should be corrected.</p>	<p>In my copy, this is page 22. Will make certain that the cancer guidelines are not mischaracterized.</p>
<p>17. Page 43. The paragraph starting "For some data sets" was not in the external peer review draft and contains new science policy statements.</p> <p>Adding new science policy text after external peer review is not consistent with EPA's standard practice. Recommend external peer / public review be pursued.</p>	<p>See response to comment #1</p>
<p>18. Page 45. First bullet: now the list of criteria contains a "preferred approach". This was not in the external review draft.</p> <p>Science policy, e.g., preferred approaches, should be publically reviewed.</p>	<p>Not clear what this refers to. If it is BMRs for continuous data, this is just a clarification of what was in the EPR draft. Policy decisions are the purview of EPA. See response to comment #1</p>
<p>19. Page 45. In the added text, only biological reasons should be used to lower from one SD, as the change must have a biological significance, not just be able to be statistically different from. If "statistical" is not removed, this new text becomes a new policy position that has not been externally peer reviewed. Options for addressing this concern would appear to be: delete "statistical" or seek external/public peer review for new policy.</p>	<p>EPA disagrees that this statement is a new policy. It is just a general recommendation that different types of scientific judgment be used to reach a suitable conclusion.</p>
<p>20. Page 46, figure 4, suggest editing box which states "use lowest BMDL". This is a policy choice and its not clear that it belongs in technical guidance that has many varied uses.</p>	<p>EPA's approach to the use of BMD methodology is to identify the model that offers the best fit and only when there are no other alternatives does EPA select the lowest BMDL. From a practical standpoint, it is unreasonable to model all data. It is up to the risk assessor and risk manager to decide what constitutes</p>

Comment	EPA Response <sup>1</sup>
<p>21. Page 47. Paragraph starting "without biological" through discussion of dynamic approach is new to the document and concludes that the dynamic model is not recommended.</p> <p>The recommendation that the dynamic model be used has not subjected to public nor the external peer reviewer. At a minimum, no conclusions should be presented or if keeping this recommendation, then pursue public review.</p>	<p>“appropriate” data.</p> <p><b>The commenter is mistaken that issue has not been subjected to public or external peer review.</b> It was included in the 2000 EPR review draft and discussed at the meeting. The EPR panel concluded that this method of BMR selection should <i>not</i> be used. The document retains the description of the approach to be clear and transparent about which practice is being discouraged.</p>
<p>22. Page 49. The paragraph beginning "the simplest" is not consistent with other EPA documents. Is EPA saying that neither the multistage model nor the Hill model have biological interpretations? DoD suggests deleting this paragraph.</p>	<p>No, EPA is not saying that these models do not have biological interpretations. <b>The text will be clarified to indicate that it is the parameters of a curve-fitting model that do not have a biological interpretation.</b></p>
<p>23. Page 49, suggest deleting the following statement: “Recall that it is the fit in the low-dose range that is of greatest interest (Section 2.3.5).”</p>	<p>This statement is used to clarify the decision making process. <b>EPA will qualify the statement to indicate that the low-dose range is usually of greatest interest in risk assessment applications.</b></p>
<p>24. Page 50. The statement that the IRIS program prefers the multistage model for cancer bioassays requires a reference. While it is accurate that similar statements appear in some, but not all, IRIS documents, DoD is not aware of any guidance or other written reference to this practice.</p> <p>EPA should provide a publically available reference to verify this point or delete this new and potentially very significant policy statement or subject this document to external/public review.</p>	<p>EPA already replaced “policy” with “practice.” (do we want to insert “preferred” before “practice.”)</p>
<p>25. Page 51, suggested edit: “For risk assessment purposes, for example, the range is large enough that the model with the lowest BMDL <del>would</del> <u>could</u> be considered preferable, as a reasonable conservative estimate <u>if the risk assessment and risk management goal was to choose the most conservative value.</u> <u>Another plausible approach could be to choose the BMDL that represented the midpoint of the three</u></p>	<p>The current statement is a policy determination. Nonetheless, this would depend on the purpose of the assessment.</p>

Comment	EPA Response <sup>1</sup>
<p>values; this would again be a policy decision as all three of the models had <u>acceptable fit.</u></p>	
<p>26. Page 52, suggested edit: “When the BMR is near or below the lowest dose, the rationale for eliminating data at the highest dose(s) is that the data at the highest dose may be the least informative of responses in the lower dose region of interest, i.e., near the BMR. This is true for both dichotomous and continuous data. <u>However, it needs to be noted that these data could still impact the dose- response curve and thus consideration of a sensitivity analysis is helpful.”</u></p>	<p>Consistent with this comment, the document already notes many caveats about dropping higher doses, and recommends presentation of the results of fitting the full data set. EPA disagrees that further explanation is needed.</p>
<p>27. Page 53 suggested edit: “The risk assessor has several choices at this point: (1) choose another model or (2) choose another dataset or endpoint, or (3) <u>as a last resort</u> drop a dose group and refit the model of choice.” Edit suggested because other approaches which use all the data and are informed by the full dose-response curve are more appropriate than dropping some data.</p>	<p>EPA disagrees that dropping a dose group is a last resort, because in some cases it is actually the preferred approach. The document already explains these considerations in detail (see response to #12).</p>
<p>28. Page 54. The reasons to consider restraining a power variable to 1 have been deleted and thus weaken the document. The reason given was not supralinearity, which can occur with some models with an exponent of 1, e.g., the Michaelis-Menten model as mentioned in the previous paragraph. EPA should correct its explanation of why power functions should be greater than 1. Note that the example in the previous text is incomplete, as all of the functions using a log of dose are also undefined at zero.</p>	<p>(Cannot find specific deleted text.) <b>The commenter is mistaken that the reasons to restrain a power parameter have been deleted from the document. EPA can add cross-references to Section 2.3.3.3 from Example A.2.</b> It is beyond the scope of the TG to get into numerical computation issues; check particular software packages for how zero dose is handled in these models</p>
<p>29. Page 54. The conclusions reached regarding supralinearity in this section are new science and new science policy. EPA should delete the discussion of supralinearity or submit the document for external peer/public review.</p>	<p>See response to comment #1. (The commenter exaggerates.)</p>
<p>30. Page 60 suggested edit: “The three highest doses, at 30, 100 and 300, are quite far from the BMD; if we drop those doses, we will be eliminating doses with responses that the model cannot account for very well, and, since they are far from</p>	<p>See response to #37.</p>

Comment	EPA Response <sup>1</sup>
<p>the BMD, we should not be eliminating much information about the actual location of the BMD. <u>As noted in Section 2.4, dropping dose groups should be carefully undertaken and conducted and transparently presented. A clear justification for the rational should be provided. Risk assessors may want to instead consider (or present for comparison purposes) the results of a NOAEL/LOAEL approach instead of dropping 50% of the dose groups.</u></p>	
<p>31. Page 62. In the last paragraph, it is very important to distinguish between parameters and estimated parameters. The term "parameters" is used many times in this paragraph, and only qualified once. DoD believes the qualifier "estimated" should be applied more frequently. EPA should ensure that "estimated" is used to appropriately qualify "parameter" each time the qualifier is necessary.</p>	<p>[The text in question is probably at the bottom of p. 38 in the compare draft, or p. 51 in terms of running page count. So there is some uncertainty in what the problem really is.] EPA believes the necessary corrections have already been made.</p>
<p>32. Section 2.3.8. In this section, the discussion jumps between "confidence intervals", "one-sided confidence intervals", and "two-sided confidence intervals" without describing which BMD type is being used and why. The last sentence where the process is described is vague. EPA should clarify which type is being used and provide a justification.</p>	<p>EPA disagrees that this section needs to be more specific, given the general scope of the document. <b>The scope of the document can be repeated here.</b></p>
<p>33. Page 65. Deletion of the caveat about the lowest BMDL being an outlier is not justified. Consideration of outliers is important, especially as both this property and the lowest AIC may be reflective of the previously added section on "improving" the fit of the model.</p>	<p>EPA is not saying outliers cannot be discarded. The text was left out because clear guidance on what to consider an outlier cannot be provided currently.</p>
<p>34. Page 65 states: "The AIC is lower for the 1st-degree model suggesting that this is the preferred model." However we note the 2 values are 157.3 vs 158.7. As these values are within 2 units of each other, shouldn't they be considered similar? Its not clear this makes one preferred over the other. Similarly on page 66, shouldn't the two AIC values be considered similar (450.6 vs 452.5)?</p>	<p>The document has clarified the use of AIC, and discourages simple averaging of BMDLs. Specific applications may call for different interpretations.</p>
<p>35. Section 3.2.6. This section on "improving" the fit of the model is new and contains major science policy issues, e.g., dropping high dose(s), that have been criticized by external peer reviewers in specific cases where that has been done, e.g., EPA's SAB on ethylene oxide. Furthermore, the cautionary note in footnote 16 about dropping data "as a last resort" was deleted in this version.</p>	<p>The external peer review of this document did not criticize dropping high doses; in fact it encouraged more guidance and examples. See response to #37.</p>

Comment	EPA Response <sup>1</sup>
<p>The decision to include recommendations of procedures that have been strongly criticized by EPA's SAB and other external peer reviewers should be carefully considered. At a minimum, the caveat from footnote 16 should be retained and the document should be presented for scientific peer and public review.</p>	
<p>36. A.2 Figure A3.3 Because this example hasn't been externally peer reviews, a critical flaw in the analysis has been missed and needs to be addressed. EPA's analysis without dropping the highest dose is biologically plausible, i.e., as the response approaches 100%, i.e., above 70%, the response becomes asymptotic so that it will not predict responses greater than 100%. When the highest dose is ignored, the fit may be somewhat better, but the predicted response can be seen to be increasing exponentially above a 70% response. As EPA states, graphical analysis can be one of the criteria for goodness of fit, however this example, in which the graphical representation is NOT biologically plausible, should be rejected. Biological plausability should be a critical criteria for BMD modeling. Use a biologically plausible example.</p>	<p>[Commenter must mean Figure A.2.2.]</p> <p>The external peer review requested further guidance and examples on dropping the high dose; EPA added one focused on dichotomous data.</p> <p>EPA states in this draft that any time dose groups are dropped, the resulting fit is only applicable to the reduced range. <b>This can be repeated or cross-referenced in Example A.2.</b></p>
<p>37. Section A.2 The concept of "dropping" high dose(s) to get a better fit -- sometimes when models have adequate fit -- has been repeatedly criticized, including by EPA's SAB.?</p> <p>The new concept and the new example should either be deleted or the document should undergo external peer review and public comment to ascertain its acceptability.</p>	<p>The guidance document clarifies that dose groups should not be dropped to improve an already adequate fit. <b>This point can be repeated or cross-referenced in Example A.2.</b></p> <p>The external peer review supported dropping dose groups:</p> <ol style="list-style-type: none"> <li>1. "Favored dropping the high doses and modeling only the low-dose region if the high doses complicated the dose-response modeling."</li> <li>2. "[M]odels are very sensitive to the high-dose region, so tweaking something in this region can have a big effect at the low end. This is wrong from a biological point of view, since noise in the high-dose region should not be driving changes in the low-</li> </ol>

Comment	EPA Response <sup>1</sup>
<p>38. From A.3. The "solution" to the problem in this example has been substantially changed from that of the external review draft and is now dependent on dropping data points (as discussed in other comments) rather than a procedure that does not require ignoring some of the relevant data. As EPA states that the advantage of using BMD modeling is that ALL of the data are used, if there are alternative procedures that use all of the data, these should be preferred, especially in a guidance document. Dropping only the higher doses is especially significant when the Hill model is used, as this model is symmetrical around the 50% response and therefore dropping only high doses without dropping an equivalent amount of low doses is implicitly weighing the lower doses for this model. Finally, this dropping of only the higher doses for a symmetrical model may be why (Table A.3.9) the <i>nonsymmetrical</i> polynomial model now has the better fit. It is not possible to say for sure, as EPA does not include the polynomial fit before the doses are dropped. Moreover, the change in response to the initial lack of fit has introduced additional errors.</p> <p>EPA should return to the previous resolution of the lack of goodness of fit for this data set, as it has an adequate solution that does not require ignoring some of the data. <b>EPA should emphasize that ignoring data should be the last alternative, not an easy "fix"</b>. Correct errors that have been introduced. EPA should seek external and public review of the document due to these science and science policy changes.</p>	<p>dose region.”</p> <p>EPA disagrees that Example A.3 has been substantially changed. The main revision was to delete the dynamic range BMR, as recommended by the EPR panel.</p> <p><b>EPA will clarify that an advantage of BMD modeling is that much more of the data are used than with the NOAEL/LOAEL approach.</b></p> <p>Not clear what previous resolution is meant.</p>
<p>39. A. 4. The statement that the multistage model has been the long standing model used for cancer bioassays is misleading. Until EPA's 2005 cancer guidelines, the linearized multistage model was used. This constrained version has quite different properties than the unconstrained model. Now, EPA does NOT have a preferred model for cancer bioassays, although the IRIS program appears to have one -- although we can not find any guidance on this subject.</p> <p>This paragraph contains other errors.</p> <ol style="list-style-type: none"> <li>1. The choice of BMR should be independent of the choice of method of extrapolation from the POD. This is emphasized in EPA's cancer guidelines.</li> </ol>	<p>EPA disagrees that these two forms of the multistage model are very different. <b>EPA will include Subramaniam et al. (2007) as substantiation.</b></p> <p><b>EPA will clarify BMR selection for cancer bioassays.</b></p> <p><b>One-sided vs. two-sided BMDLs: some further internal discussion needed.</b></p>

<b>Comment</b>	<b>EPA Response<sup>1</sup></b>
<p>2. The BMDL is a one-sided confidence limit. As the estimation of confidence limits in this case should not be constrained, a two-sided calculation would be expected, and this constraint should also be noted in the discussion.</p> <p>Revise the misleading sentence. EPA should correct all of the errors in this paragraph.</p>	
<p>40. Page 109. The example of a linear dose-response curve dose not fit the definition. Response = b x dose is linear. Response does not vary linearly with dose for Response = a + b x dose. It is also inconsistent with the discussion of thresholds on page 55.</p> <p>Example should be corrected.</p>	<p><b>Commenter is mistaken about the definition of linear dose-response curves.</b> Not clear which discussion of thresholds is meant.</p>