

Analysis of Dose-Response Uncertainty using Benchmark Dose Modeling

Jeff Swartout
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment
October 11, 2007

I. Introduction and Methods

I have assessed the uncertainty in modeling the test data sets using the U.S. EPA BMDS (Ver. 1.4.1) software (U.S. EPA, 2007), with a secondary verification using a parametric bootstrap algorithm in S-PLUS[®] (Ver. 6.2 for Windows[®]). I present only the analyses of the (mythical) frambozadrine, nectorine and persimonate data sets, as they offer the most interesting issues for combining results.

The issue of combining response data is first addressed by determining if the individual data sets are similar conceptually. That is, can they be considered to belong to the same population? The statistical aspect of pooling the data is addressed by assessing the net deviance ($-2 \times$ optimized log-likelihood ratio) between the dose-response model fits to the pooled data set and the individual data sets, assuming a specific dose-response family (Stiteler et al., 1993). The test statistic is $D_{\text{pool}} - ED_i$, where D_{pool} is the deviance of the pooled data model fit and ED_i is the sum of the individual model fit deviances from the same fitted model. Data are pooled by combining all dose groups into one data set, keeping each group intact, including controls (i.e., not by summing animals and responders into single groups for the same dose). The test statistic is compared to a Chi-squared distribution with degrees of freedom equal to the difference between the sum of the number of individual model parameters minus the number of parameters in the pooled-data model. As a convention, the probability is reported as $1 - \text{pchisq}$, where pchisq is the cumulative probability of the Chi-squared distribution, with values greater than 0.05 accepted as adequate for pooling the data.

Uncertainty analysis for the Benchmark Dose procedure was accomplished by obtaining lower 95% confidence bounds for selected Benchmark Response (BMR) levels, corresponding to the BMDL_X , where X is the percent response in the tested animals. BMRs ranging from 1% to

99% were evaluated. Confidence bounds in BMDS are determined by the profile likelihood method, assuming an asymptotic relationship of the log-likelihood ratio to the Chi-square distribution.

The parametric bootstrap procedure consists of fitting the selected dose-response model to the raw data by maximum likelihood and treating the fitted response at each dose as the “true” response in the tested animals. Then, with an assumption that the observed response in each dose group can be represented by simple binomial uncertainty¹, a new set of responses is generated based on the fitted probability and sample size (simulating a repeat of the experiment). The selected dose-response model is fit to 1,000 bootstrap samples and the fitted parameters are saved. From the resulting parameter distribution, point-wise confidence bounds can be calculated for the entire dose-response curve. Upper and lower bootstrap confidence bounds can then be compared with the corresponding profile likelihood confidence bounds computed in BMDS. Constraints on the parameter space are an issue for use of either of these methods. Standard likelihood-based methods are sometimes found to perform poorly when there are parameter constraints. The most general results assume no constraints (i.e., an "open parameter space").

II. Frambozadrine Data Analysis

The frambozadrine data are reproduced in Table II-1.

Table II-1. Frambozadrine dose-response data

Dose (mg/kg-day)	Total no rats	Incidence (hyperkeratosis)	Percent response
Male			
0	47	2	4.3
1.2	45	6	13
15	44	4	9.1
82	47	24	51
Female			
0	48	3	6.3
1.8	49	5	10
21	47	3	6.4
109	48	33	69

¹ The parametric bootstrap procedure does not account for over-dispersion (i.e., extra-binomial variance) in the response data.

The questions to be answered are:

- 1) Do we need separate dose-response relations for males and females?
- 2) Does combination alter the uncertainty in response?

Frambozadrine BMDS model-fitting results

Table II-2 shows the results of the model fitting from BMDS. Any model fit with a p-value greater than 0.1 was considered adequate. Goodness-of-fit was evaluated on the basis of the Akaike Information Criterion (AIC) value, with lower values indicating better fit. The deviance statistic is given for determination of the statistical propriety of combining the data. The BMDS output files for selected model fits are provided in Appendix A. BMD and BMDL values are in units of mg/kg-day.

Table II-2. BMDS results for frambozadrine

Model	AIC	p-value	Deviance	BMD₁₀	BMDL₁₀
Male					
Multistage (2 ⁿ)	150.4	0.28	2.551	33.7	12.0
gamma	152.3	0.12	2.478	40.8	10.3
log-logistic	152.3	0.12	2.478	42.3	12.0
log-probit	152.3	0.12	2.478	36.9	13.6
Weibull	152.3	0.12	2.478	44.5	10.8
Female					
Multistage (2 ⁿ)	142.4	0.43	1.742	34.4	21.2
gamma	143.3	0.41	0.667	71.2	25.0
log-logistic	143.3	0.41	0.667	82.4	24.7
log-probit	143.3	0.41	0.667	66.7	24.3
Weibull	143.3	0.41	0.667	86.4	25.2
Pooled data					
Multistage (2 ⁿ)	289.0	0.60	4.498	34.1	22.9
gamma	290.1	0.60	3.603	43.5	25.7
log-logistic	290.1	0.48	3.380	44.3	26.0
log-probit	289.9	0.51	3.192	44.3	25.9
Weibull	290.5	0.54	4.036	41.2	24.6

On the basis of the AIC values, the best fitting model for all three data sets is the 2nd order multistage. The multistage model fits are shown in Figure II-1. Given the small differences in the AIC values, however, all of the models are virtually equivalent for goodness-of-fit. BMDL₁₀ values are similar for all models. The multistage (best-fitting), and Weibull (most flexible)

models were chosen for uncertainty analysis. Pooling the data across gender for the same species and strain is reasonable conceptually, and is done commonly. The pooled-data test statistics are 0.205 and 0.891 for the multistage and Weibull models, respectively. The Chi-squared p-values for the test statistics, given 3 degrees of freedom, are 0.98 and 0.83 for the multistage and Weibull models, respectively, indicating that pooling the data is reasonable. Note, in particular that the BMDL₁₀ values for the pooled data are higher than those for either of the individual-gender response data sets – almost twice the male response BMDL₁₀. This behavior is of particular significance in defining the point-of-departure (POD) for deriving a Reference Dose (RfD).

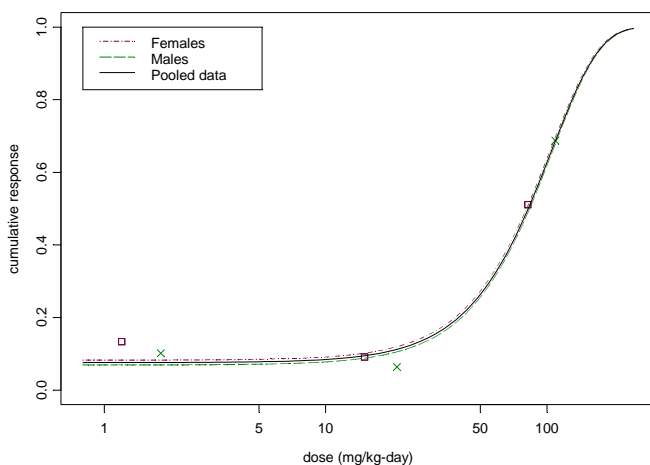


Figure II-1. Multistage (2nd order) model fits to the frambozadrine data

Frambozadrine uncertainty analysis results

The results of the Benchmark Dose and bootstrap uncertainty analyses of the frambozadrine data for the 2nd order multistage model is shown in Table II-3. The BMD, BMDL and ratio (BMDLr) of the BMD to the BMDL are given for selected BMRs. The analogous values (ED, EDL, EDLr) for the bootstrap simulation are shown for comparison. The BMDLr and EDLr values provide a quick measure of the relative statistical uncertainty for any given combination of model and response level (BMR). The 5% and 10% BMRs are generally applied in the derivation of RfDs. The 1% BMR results are included to show the behavior of the models

at response levels farther away from the data. BMD, BMDL, ED and EDL values are in units of mg/kg-day.

Table II-3. BMD and bootstrap results for frambozadrine (2nd order multistage)

BMR¹	BMDL²	BMD³	BMDLr⁴	EDL⁵	ED⁶	EDLr⁷
Male						
0.01	1.15	10.4	9.0	1.35	8.90	6.6
0.05	5.86	23.5	4.0	6.85	20.7	3.0
0.10	12.0	33.7	2.8	13.8	30.3	2.2
Female						
0.01	2.55	10.6	4.2	1.59	9.05	5.7
0.05	11.6	24.0	2.1	7.77	21.2	2.7
0.10	21.2	34.4	1.6	15.6	30.9	2.0
Pooled data						
0.01	2.92	10.5	3.6	1.75	9.37	5.4
0.05	12.8	23.8	1.9	8.67	21.9	2.5
0.10	22.9	34.1	1.5	16.9	32.0	1.9

¹ Benchmark Response level (fraction responding)

² lower 95% confidence bound on the BMD (from BMDS)

³ maximum likelihood estimate (from BMDS)

⁴ ratio of BMD to BMDL

⁵ lower 95% bootstrap confidence bound (bootstrap analog of BMDL)

⁶ median bootstrap ED estimate (bootstrap analog of BMD)

⁷ ratio of ED to EDL

As would be expected, Table II-3 shows that the uncertainty increases with decreasing BMR, that is, as the prediction moves farther away from the data. The BMDLr values for the male response data are about twice as large as for the female and pooled response data. The bootstrap EDLr values are much closer together than the corresponding BMDLr values. The EDLr values are somewhat larger than the corresponding ones for the BMDLr, except for the male response data. In particular for the BMDLr_{0.01}. is Pooling the data reduces the uncertainty at all BMR levels, particularly with respect to the male response data. The 90% bootstrap confidence bounds for the multistage model fit to the pooled frambozadrine data are shown in Figure II-2.

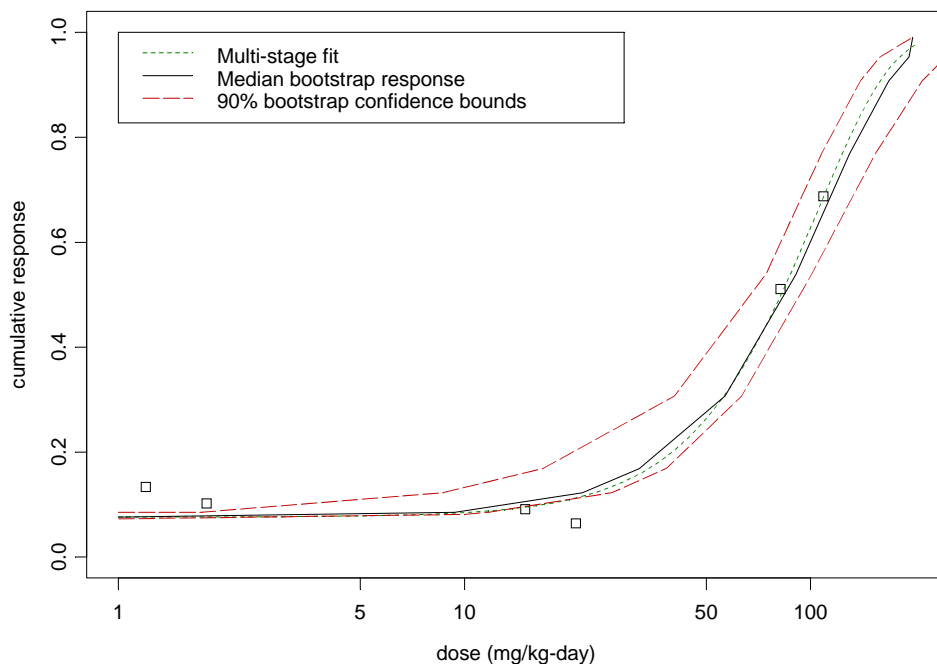


Figure II-2. 90% bootstrap confidence bounds on the multistage model fit to the pooled frambozadrine data

The results of the Benchmark Dose and bootstrap uncertainty analyses of the frambozadrine data for the Weibull model are shown in Table II-4.

Table II-4. BMD and bootstrap results for frambozadrine (Weibull model)

BMR	BMDL	BMD	BMDLr	EDL	ED	EDLr
Male						
0.01	0.678	19.8	29.2	1.13	20.7	18.3
0.05	4.61	34.7	7.5	6.19	36.0	5.8
0.10	10.7	44.5	4.2	13.4	45.3	3.4
Female						
0.01	5.76	68.3	11.9	4.34	27.8	6.4
0.05	16.1	80.4	5.5	12.2	43.7	3.6
0.10	25.2	86.4	3.4	19.8	53.2	2.7
Pooled data						
0.01	5.34	15.8	3.0	4.85	15.9	3.3
0.05	15.4	30.7	2.0	14.2	31.1	2.2
0.10	24.6	41.2	1.7	22.9	41.7	1.8

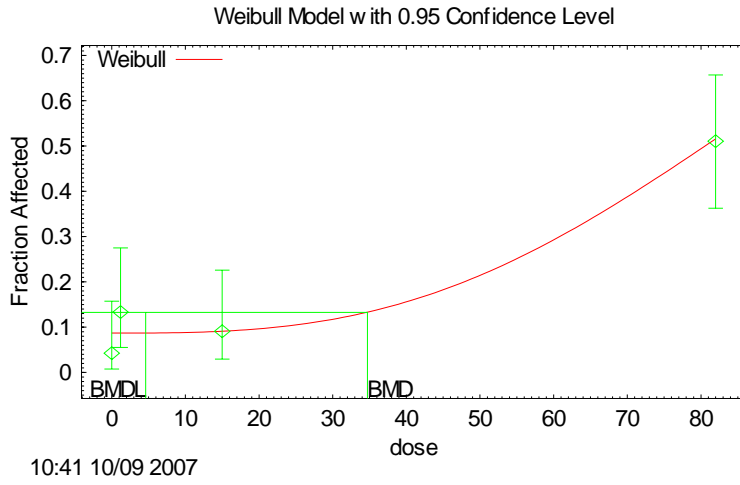
Except for the pooled data, the BMDLr values for the Weibull model are larger and more variable than those for the 2nd order multistage model (compare to Table II-3). In addition, the

BMD uncertainty is greater than the bootstrap uncertainty for the individual data sets, with the greatest difference at the lowest BMRs. Another significant difference between the two models is in the central-tendency estimate (BMD or ED) for any given BMR, with Weibull BMD values 1.5 to 3 times greater than multistage BMDs. As for the multistage model, the bootstrap uncertainty is somewhat less than the BMD uncertainty in most cases. BMDL and EDL values are similar, but there is a large disparity between the BMDs and EDs for the female response data not seen for either the male response data (Weibull fit) or the multistage model fits to the female data. Figure II-3 shows the BMDS plots (BMR = 0.05) for these three model fits. The Weibull fit for the female data (Fig. 3.c) is much shallower at the low end than either of the other two fits. The corresponding likelihood profile will also be flat in that dose region, resulting in a large difference between the BMD and BMDL. The primary problem here is the wide gap in female response between the last two doses, rising from essentially background level to over 50%. Large model-specific differences can arise, depending on the relative shape flexibility of the model. Ideally, observed responses in the vicinity of the BMR are needed to “anchor” the model fit appropriately (U.S. EPA, 2000). The pooled-data Weibull fit reduces the uncertainty significantly when compared to either of the individual data set fits.

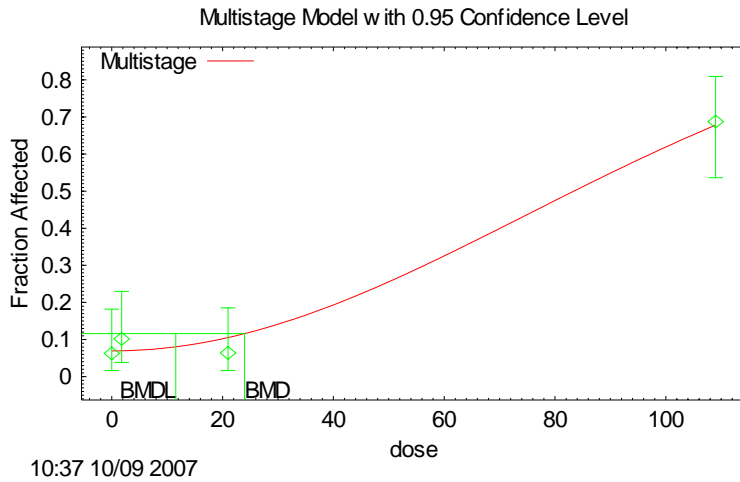
Conclusions for frambozadrine uncertainty analysis

1. We do not need separate dose-response relations for males and females.
2. Combining the response data decreases dose-response uncertainty somewhat for the multistage model and significantly for the Weibull model. Given the lack of response near the BMR for the female response data, however, a BMD analysis might not be appropriate. At the least, highly flexible models, such as the Weibull, probably should not be used.

a. Male response data, Weibull model



b. Female response data, 2nd-order multistage model



c. Female response data, Weibull model

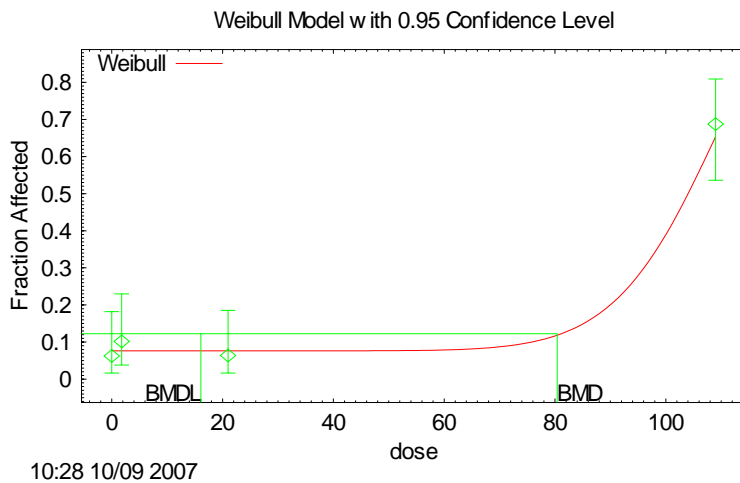


Figure II-3. Frambozadrine BMD₀₅ / BMDL₀₅ plots

III. Nectorine Data Analysis

The nectorine data are reproduced in Table III-1. The animals in each study were different and only the indicated endpoint was evaluated in each study (i.e., the tumors are independent).

Table III-1. Nectorine inhalation dose-response data

	Concentration (ppm)			
	0	10	30	60
Lesion	# response / # in trial			
Respiratory epithelial adenoma in rats	0/49	6/49	8/48	15/48
Olfactory epithelial neuroblastoma in rats	0/49	0/49	4/48	3/48

The question to be answered is:

What is the uncertainty in response as function of dose for either respiratory epithelial adenoma OR olfactory epithelial neuroblastoma?

Nectorine BMDS model-fitting results

Table III-2 shows the results of the model fitting from BMDS. Human-equivalent exposures are not evaluated. As uncertainty in animal-to-human scaling is not evaluated, the uncertainty in the human-equivalent inhalation unit risks (IUR) will be identical to that for the rats. Model acceptability and goodness-of-fit criteria are the same as for the frambozadrine analysis. The BMDS output files for selected model fits are provided in Appendix B. The multistage and Weibull models were fit to the data. The multistage model is the standard default cancer model (U.S. EPA, 2005). As for frambozadrine, the Weibull model was used to show the effect of a more shape-flexible model.

Table III-2. BMDS results for nectorine

Model	AIC	p-value	BMD ₁₀	BMDL ₁₀	Multistage IUR ¹
Respiratory epithelial adenoma in rats					
Multistage (1 ^N)	143.6	0.44	15.2	11.3	0.00883
Weibull	143.9	0.75	8.71	0.259	
Olfactory epithelial neuroblastoma in rats					
Multistage (1 ^N)	55.2	0.42	70.1	39.9	0.00251
Weibull	57.2	0.24	70.4	39.9	

¹ 0.1/BMDL₁₀ for the multistage model fit in units of (mg/kg-day)⁻¹

The 1st order multistage model provides an adequate fit ($p > 0.10$) to both data sets and is selected as the primary model. Under the EPA guidelines, if a p-value of at least 0.10 is obtained by the 1st order fit, higher order models are not used even if the p-values are higher. The Weibull model provides a much better fit for the adenoma endpoint. While the Weibull would not be used for derivation of an IUR, it is useful for uncertainty analysis. The Weibull fit to the neuroblastoma data is essentially identical to the 1st order multistage, with a Weibull exponent parameter value of 0.99 (virtually exponential). Assuming independence of the two tumor types, the combined risk of either tumor can be calculated exactly from the binomial risk formula: $IUR_1 + [1 - IUR_1] \times IUR_2$, where IUR_1 and IUR_2 are the two independent IURs. A close approximation can be obtained by simply adding the IURs if the sum is less than 0.1. In this case, the combined risk of either tumor is 0.011 by either method. BMDS, however, does not estimate confidence limits for the IUR, as it is only an approximation of a 95% upper confidence bound. In addition, the probability of the sum of multiple IURs cannot be estimated from the BMDS output. The U.S. EPA cancer guidelines (U.S. EPA, 2005) do not address this issue, which is an active area of investigation. An attempt to address the uncertainty distribution of the combined IURs is attempted in the following uncertainty analysis.

Nectorine uncertainty analysis results

The results of the Benchmark Dose and bootstrap uncertainty analyses of the nectorine data for the multistage model is shown in Table III-3. A 2nd order parameter was fit in the bootstrap simulation to account for bootstrapped responses that could not be fit adequately by the 1st order

model. The BMD results are strictly 1st order. The 2nd order parameter was zero when the 2nd order model was fit to the raw data.

Table III-3. BMD and bootstrap results for nectorine (multistage model)

BMR	BMDL	BMD	BMDLr	EDL	ED	EDLr
Respiratory epithelial adenoma						
0.01	1.08	1.45	1.3	1.15	1.73	1.5
0.05	5.51	7.39	1.3	5.86	8.72	1.5
0.10	11.3	15.2	1.3	12.0	17.3	1.4
Olfactory epithelial neuroblastoma						
0.01	3.81	6.69	1.8	4.60	9.76	2.1
0.05	19.4	34.1	1.8	22.9	39.3	1.7
0.10	39.9	70.1	1.8	43.8	63.0	1.4

Because the 1st order multistage model has only one parameter, the BMDLr values are constant for any given data set and relatively small, indicating low uncertainty. The 2nd order multistage bootstrap results represent a slight increase in uncertainty at the lower BMR values. Otherwise, The BMDS and bootstrap results are similar.

Table III-4 shows the results of the Weibull model uncertainty analysis.

Table III-4. BMD and bootstrap results for nectorine (Weibull model)

BMR	BMDL	BMD	BMDLr	EDL	ED	EDLr
Respiratory epithelial adenoma						
0.01	2.3×10^{-7}	0.191	8.4×10^5	4.8×10^{-5}	0.222	4600
0.05	0.0037	2.70	730	0.043	2.92	68
0.10	0.259	8.71	34	1.05	44.9	8.9
Olfactory epithelial neuroblastoma						
0.01	3.1×10^{-28}	6.60	2.1×10^{28}	0.094	8.43	89
0.05	2.79	34.1	12	16.3	39.9	2.4
0.10	39.9	70.4	1.8	43.2	64.4	1.5

The BMDS results show extreme uncertainty, particularly for low BMRs and the adenoma endpoint. The bootstrap results are characterized by large variability also, but are much more stable. The Weibull exponent for the fit to the adenoma response data is 0.67, indicating a supralinear response. High BMDLr values are usually associated with supralinear fits. Most 2-parameter models that have real support only above zero will exhibit this behavior, as the variance increases without bound as supralinearity increases (i.e., as the exponent, or shape parameter, approaches zero). For this reason, the U.S. EPA recommends constraining the models

to avoid sublinearity (U.S. EPA, 2000). However, a large number of dose-response data sets exhibit supralinear behavior when unconstrained models are fit. Thus, although the extreme uncertainty shown for the Weibull analysis of these data is not to be considered too seriously, the results serve as a caution when interpreting confidence bounds based on constrained model fits.

IUR distributions for each individual tumor type are approximated from the bootstrapped-response model fits by dividing the BMR (0.1) by the ED₁₀ at each iteration. The combined risk is determined by summing the (MLE) IURs for randomly paired iterations for each tumor type. The results are shown in Table III-5, which lists the empirical quantiles for each distribution. The values in Table III-5 are approximate, as only 500 paired bootstrap samples were used in the calculation. Probability density plots of the distributions are shown in Figure III-1.

Table III-5. Bootstrap Inhalation Unit Risk distributions for nectorine ([mg/kg-day]⁻¹)

									Fractile
0.01	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.99	
Respiratory epithelial adenoma									
0.0033	0.0039	0.0042	0.0049	0.0059	0.0069	0.0079	0.0084	0.095	
Olfactory epithelial neuroblastoma									
0.00027	0.00059	0.00080	0.0012	0.0016	0.0019	0.0021	0.0023	0.0027	
Combined tumors									
0.0045	0.0052	0.0056	0.0064	0.0074	0.0085	0.0095	0.0010	0.011	

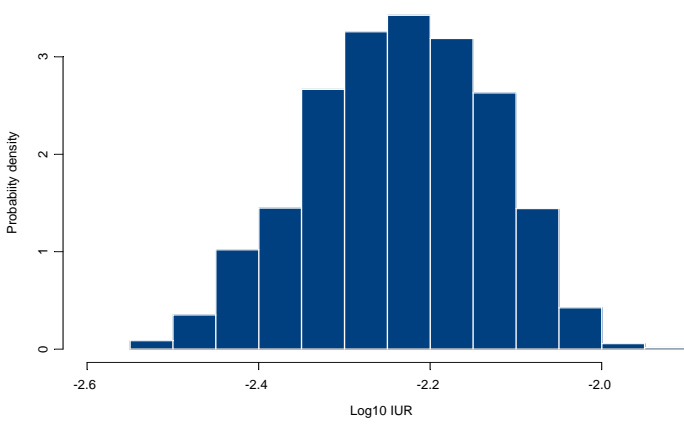
Note that summing the quantiles directly does not yield the same answer, such that summing the 95% upper-bound multistage cancer IURs as calculated by BMDS will not give a 95% upper bound, but something farther out in the tail of the distribution. In this case, the direct sum of the BMDS IURs of 0.011 is equivalent to a 99th percentile in the bootstrap distribution. The 90% bootstrap confidence interval for the adenoma IUR is 0.0039 to 0.0084 (mg/kg-day)⁻¹, spanning a 2.2-fold risk interval. The 90% bootstrap confidence interval for the neuroblastoma IUR is 0.00059 to 0.0023 (mg/kg-day)⁻¹, spanning a 3.9-fold range. The 90% bootstrap confidence interval for the combined IUR is 0.0052 to 0.0010 (mg/kg-day)⁻¹, spanning a 1.9-fold interval.

Conclusions for nectorine uncertainty analysis

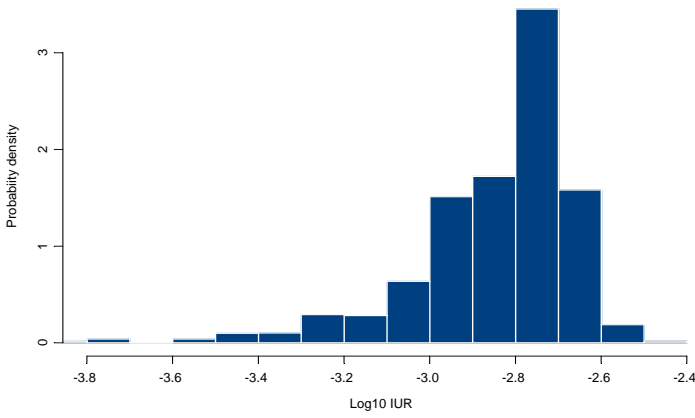
1. The constant-shape 1st order multistage model indicates low uncertainty compared to the shape-flexible Weibull model. Assuming, however, that a mutagenic (i.e., one-hit) mode of action would apply, the Weibull model would not be considered for deriving an inhalation unit risk.

1. Based on the 90% bootstrap confidence intervals, the uncertainty in the combined response is less than that for either of the individual tumor responses.

a. Respiratory epithelial adenoma



b. Olfactory epithelial neuroblastoma



c. Combined tumors (either or)

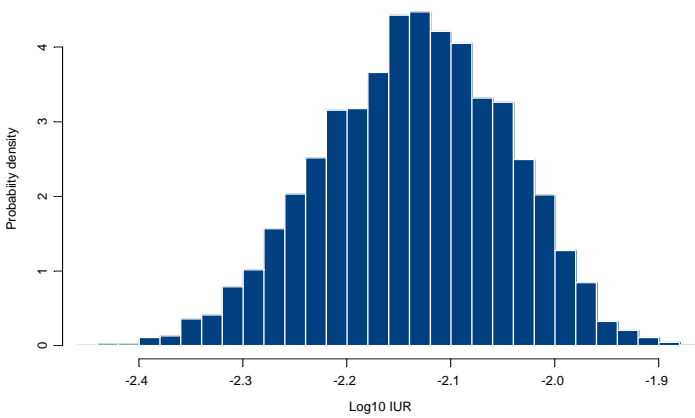


Figure III-1. Bootstrap Inhalation Unit Risk distributions for nectorine

IV. Persimonate Data Analysis

The persimonate data are reproduced in Table IV-1.

Table IV-1. Persimonate inhalation dose-response data

	Exposure level (ppm)	Metabolized dose (mg/kg-day)	Survival-adjusted tumor incidence	Percent response
B6C3F1 male mice	0	0	17/49	34.7
	18	27	31/47	66.0
	36	41	41/50	82
Crj:BDF1 male mice	0	0	13/46	28.3
	1.8	3.4	21/49	42.9
	9.0	14	19/48	39.6
	45	36	40/49	81.6

The questions to be answered are:

- 1) Can we combine these studies?
- 2) Does it affect our uncertainty?

Persimonate BMDS model-fitting results

Table IV-2 shows the results of the model fitting from BMDS. The analysis is based on the metabolized dose rather than the external exposure levels. Human-equivalent exposures are not evaluated. Model acceptability and goodness-of-fit criteria are the same as for the frambozadrine analysis. The BMDS output files for selected model fits are provided in Appendix C.

Table IV-2. BMDS results for persimonate

Model	AIC	p-value	Deviance	BMD ₁₀	BMDL ₁₀
BC63F1 male mice					
Multistage (1 ^N)	175.2	0.49	0.4714	3.70	2.71
Crj:BDF1 male mice					
Multistage (1 ^N)	242.5	0.06	5.616	3.54	2.53
Multistage (2 ^N)	241.6	0.10	2.673	12.2	3.45
gamma	241.1	0.14	2.215	16.1	5.06
log-logistic	241.1	0.14	2.211	16.1	6.82
log-probit	241.1	0.14	2.216	15.9	7.93
Weibull	241.1	0.14	2.201	16.2	4.60

Pooled data					
Multistage (1 st)	414.1	0.26	6.499	3.60	2.86
Multistage (2 nd)	412.7	0.54	3.110	10.7	3.82
gamma	412.5	0.57	2.915	13.3	4.80
log-logistic	412.3	0.61	2.706	13.8	6.80
log-probit	412.3	0.61	2.707	14.3	7.79
Weibull	412.7	0.54	3.103	11.5	4.27

Only the 1st order (one-parameter) multistage model can be fit to the BC63F1 data, as there are not enough degrees of freedom to fit any of the other models and still allow for statistical comparison. This model is essentially a quantal-linear, or exponential, model. On the basis of the AIC values, the best fitting models for the Crj:BDF1 data set are the gamma, log-logistic, log-probit and Weibull, all with identical fits. The log-logistic and log-probit fit the pooled data best. The multistage model fits are shown in Figure IV-1. Given the small differences in the AIC values, however, all of the 2-parameter models are virtually equivalent for goodness-of-fit. BMDL₁₀ values vary, however, by about a factor of two among the better-fitting models for each data set. Pooling the data across strains within a species is reasonable conceptually. The only common model across the data sets is the 1st order multistage, so the pooling test is based on the deviance for that model fit even though the fit is marginal for the Crj:BDF1 data ($p = 0.06$; the U.S. EPA BMD guidance [U.S. EPA, 2000] suggests rejecting fits with a p-value less than 0.1). The pooled-data test statistic is 0.4116. The Chi-squared p-value for the test statistic, given 2 degrees of freedom, is 0.81, indicating that pooling the data is reasonable. For conducting the uncertainty analysis on the Crj:BDF1 and pooled data sets, the 1st order multistage and Weibull models are chosen. The Weibull, as for the frambozadrine and nectorine analyses, is selected for shape flexibility. The 1st order multistage, although not the best fitting model, would be the preferred model under the U.S. EPA cancer risk assessment guidelines (U.S. EPA, 2005) for modeling the pooled response data (if not for the Crj:BDF1 data).

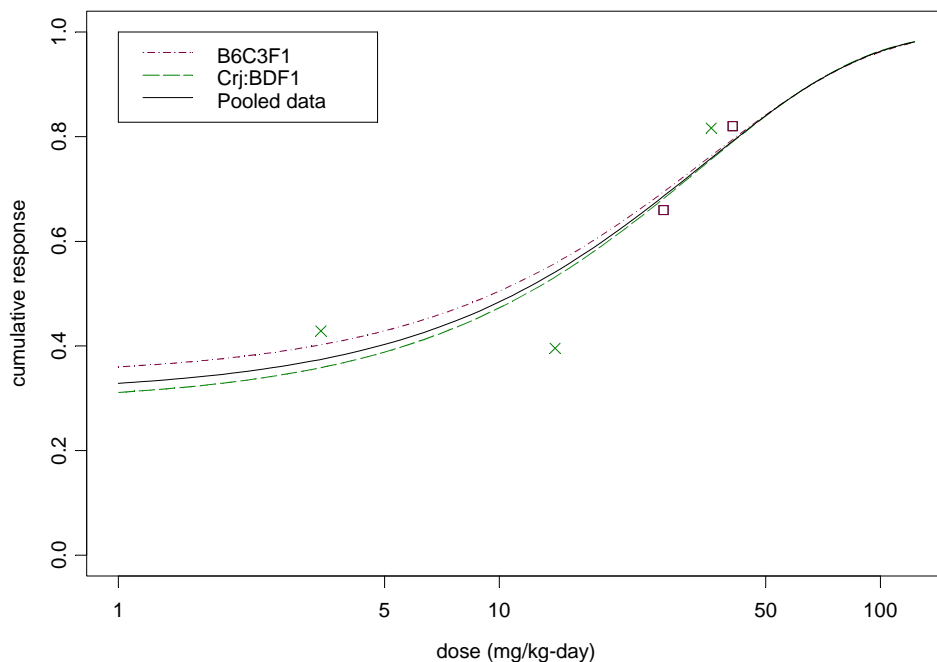


Figure IV-1. Multistage (1st order) model fits to the persimonate data

Persimonate uncertainty analysis results

The results of the Benchmark Dose and bootstrap uncertainty analyses of the persimonate data for the 1st order multistage model is shown in Table IV-3 and Figure IV-2 (pooled data).

Table IV-3. BMD and bootstrap results for persimonate (1st order multistage)

BMR	BMDL	BMD	BMDLr	EDL	ED	EDLr
BC63F1 male mice						
0.01	0.259	0.353	1.4	0.249	0.348	1.4
0.05	1.32	1.80	1.4	1.27	1.78	1.4
0.10	2.71	3.70	1.4	2.61	3.65	1.4
Crj:BDF1 male mice						
0.01	0.242	0.338	1.4	0.237	0.336	1.4
0.05	1.23	1.73	1.4	1.21	1.72	1.4
0.10	2.53	3.54	1.4	2.49	3.52	1.4
Pooled data						
0.01	0.273	0.343	1.3	0.270	0.344	1.3
0.05	1.39	1.75	1.3	1.38	1.76	1.3
0.10	2.86	3.60	1.3	2.83	3.61	1.3

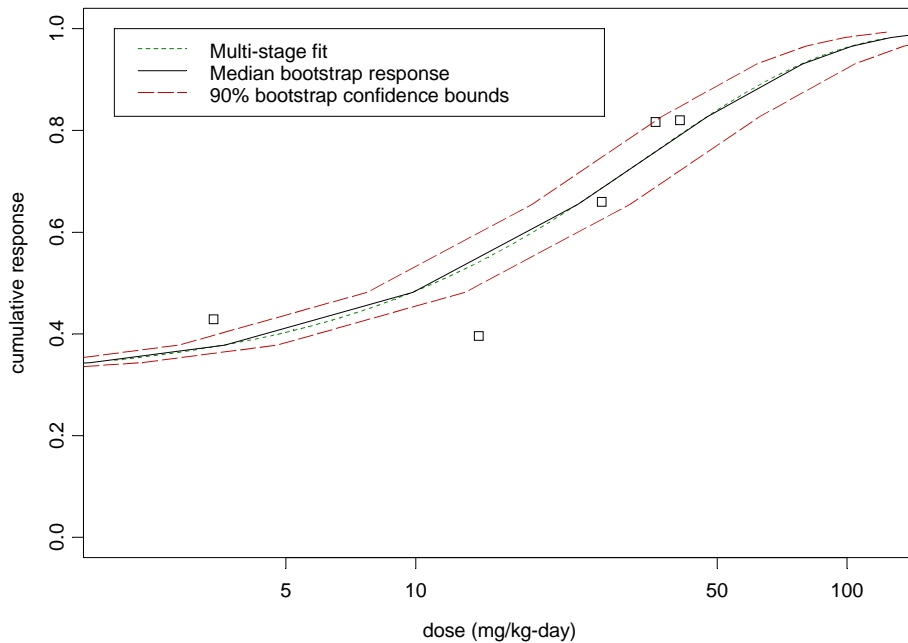


Figure IV-2. 90% bootstrap confidence bounds on the multistage model fit to the pooled persimonate data

The BMD and BMDL values are similar across the data sets for any given BMR. The same holds for the ED and EDL values. Because the dose-response model has only one parameter, the BMDLr and EDLr values are constant for any given data set and relatively small, indicating low uncertainty.

The results of the Benchmark Dose and bootstrap uncertainty analysis of the persimonate data for the Weibull model are shown in Table IV-4.

Table IV-4. BMD and bootstrap results for persimonate (Weibull model)

BMR	BMDL	BMD	BMDLr	EDL	ED	EDLr
Crj:BDF1 male mice¹						
0.01	0.609	7.57	12.4	1.25	7.02	5.6
0.05	2.48	12.8	5.2	3.84	12.2	3.2
0.10	4.60	16.2	3.5	6.37	15.7	2.5
Pooled data²						
0.01	0.503	3.70	7.4	0.600	3.73	6.2
0.05	2.22	8.15	3.7	2.51	8.17	3.3
0.10	4.27	11.5	2.7	4.68	11.5	2.5

The most notable difference from the multistage model results is the increase in BMDLr and EDLr values, substantially increasing the uncertainty estimate. This result is expected, given the greater flexibility of the Weibull model. All of the BMDs and EDs are increased compared to the multistage model, as well, but consistent with each other (i.e., within the Weibull model results). The BMDLr values are larger than the EDLr values, particularly for the $BMR_{0.01}$ results for the Crj:BDF1 response data, similar to the results for the frambozadrine Weibull analysis.

Conclusions for parsimonate uncertainty analysis

1. The two studies can be combined.
2. Combining the response data does not affect the uncertainty for the 1st order multistage model, but reduces uncertainty slightly for the Weibull model based on the BMD analysis. The uncertainty is not affected for the Weibull model, however, if based on the bootstrap analysis.

IV. General Conclusions

The BMD approach tends to overestimate, with respect to the bootstrap method, the dose-response uncertainty at low BMR values for these particular data sets, primarily for responses relatively far from the lowest observed non-background response and for smaller sample sizes. Overall, however, there is not too much of a difference between BMD and bootstrap results for these data. Most of these sample data sets, however, are relatively well behaved, in that the most of the model fits are very good and all are linear (exponential) or sublinear. For supralinear data sets (e.g., Weibull exponents < 1), such as the nectorine respiratory epithelial adenoma response data, BMDLr values tend to increase without bound as the supralinearity increases (results not shown). A useful follow-on exercise would be to evaluate different uncertainty analysis techniques on more supralinear response data.

References

Stiteler, W.M., L.A. Knauf, R.C. Hertzberg and R.S. Schoeny. 1993. A statistical test of compatibility of data sets to a common dose-response model. *Reg. Tox. Pharm.* **18**:392-402

U.S. EPA. 2000. Benchmark Dose Technical Guidance Document. [External Review Draft]. EPA/630/R-00/001.

U.S. EPA. 2005. Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum, Washington, DC. EPA/630/P-03/001B. Available at: www.epa.gov/cancerguidelines

U.S. EPA. 2007. Benchmark dose software (BMDS) version 1.4.1 (last modified 2006). Available at: <http://www.epa.gov/ncea/bmds/>