



Benchmark Dose Modeling of *In Vitro* Genotoxicity Data: a Reanalysis

Xiaoqing Guo and Nan Mei

Division of Genetic and Molecular Toxicology, National Center for Toxicological Research, Jefferson, AR, USA

Abstract

The methods of applied genetic toxicology are changing from qualitative hazard identification to quantitative risk assessment. Recently, quantitative analysis with point of departure (PoD) metrics and benchmark dose (BMD) modeling have been applied to *in vitro* genotoxicity data. Two software packages are commonly used for BMD analysis. In previous studies, we performed quantitative dose-response analysis by using the PROAST software to quantitatively evaluate the mutagenicity of four piperidine nitroxides with various substituent groups on the 4-position of the piperidine ring and six cigarette whole smoke solutions (WSSs) prepared by bubbling machine-generated whole smoke. In the present study, we reanalyzed the obtained genotoxicity data by using the EPA's BMD software (BMDS) to evaluate the inter-platform quantitative agreement of the estimates of genotoxic potency. We calculated the BMDs for 10%, 50%, and 100% (i.e., a two-fold increase), and 200% increases over the concurrent vehicle controls to achieve better discrimination of the dose-responses, along with their BMDLs (the lower 95% confidence interval of the BMD) and BMDUs (the upper 95% confidence interval of the BMD). The BMD values and rankings estimated in this study by using the EPA's BMDS were reasonably similar to those calculated in our previous studies by using PROAST. These results indicated that both software packages were suitable for dose-response analysis using the mouse lymphoma assay and that the BMD modeling results from these software packages produced comparable rank orders of the mutagenic potency.

Key words: Benchmark dose, Quantitative analysis, *In vitro* genotoxicity, Mouse lymphoma assay

INTRODUCTION

The benchmark dose (BMD) is calculated by using a mathematical dose-response curve estimated from all data and reflects the shape of the dose-response. The BMD concept was first used in 1984 as an alternative to the “no observed adverse effect level” (NOAEL) to derive a point of departure (PoD) for toxicity data (1). For the risk assessment of a particular chemical, different methods may predict different NOAEL values at different endpoints

(2). The NOAEL and its related exposure level, the “lowest observed adverse effect level” (LOAEL), are usually selected from an actual experimental dose levels and are therefore affected considerably by the experimental design, such as dose range, and each dose selected (3). The benchmark procedure can be applied to various types of data, such as “dichotomous” and “continuous” data, to determine the dose that causes a prescribed adverse response. The BMDL is defined as the lower confidence limit of the BMD that corresponds to a given increase in response over the background response. The BMD approach has been endorsed by the U.S. Environmental Protection Agency (EPA) as an acceptable replacement for the NOAEL (4). In 1995, the EPA's National Center for Environmental Assessment initiated a project to develop benchmark dose software (BMDS) and in 1999, the BMDS was released and has since been improved and enhanced (5).

In 1993, a workshop on benchmark dose methodology was held to assess the feasibility and implications of the use of BMD and the workshop participants supported the

Correspondence to: Xiaoqing Guo, 3900 NCTR Road, Jefferson, AR 72079, USA

E-mail: xiaoqing.guo@fda.hhs.gov

Nan Mei, 3900 NCTR Road, Jefferson, AR 72079, USA

E-mail: nan.mei@fda.hhs.gov

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

use of a BMD for all noncancer risk assessment (6). For example, the BMD approach was applied to developmental toxicity (7) and neurotoxicity (8). In 2001, Crump and colleagues compared the T25 (a 25% increase in the incidence of a specified tumor type) estimation method used in the European Union (i.e., the E.U. approach) and BMD modeling data from 276 carcinogenesis bioassays conducted as part of the National Toxicology Program (NTP), and concluded that the BMD estimate had less bias and less variation for estimation of the potency of carcinogens (9). In the past decade, dose-response modeling and BMD methods have also been applied to analyze microarray data to identify novel biomarkers of exposure and to examine potential toxicity (10). When applied for the evaluation of gene transcription data in quantitative cancer and noncancer risk assessment, the transcriptional BMD values can provide a better understanding of how chemical exposures disturb the cellular and molecular networks (11).

Genotoxicity tests measure the type of damage to genetic material that might lead to cancer. Therefore, genotoxicity assays are often used for the determination of cancer potency parameters and establishing the carcinogenic mode of action (MoA) for risk assessment (12). Both *in vivo* and *in vitro* genotoxicity studies are considered important for the evaluation of the mechanisms of carcinogenicity; and genotoxicity assessments offer the advantage of measuring a property related to human disease. Historically, data obtained from these studies are subject to a qualitative determination, i.e., either a “positive” or “negative” outcome, although a positive dose-response is also required for a positive determination (13). The BMD approach has been applied to *in vivo* genotoxicity data, including data from the *in vivo* micronucleus assay, the *in vivo* comet assay, *in vivo* DNA adduct formation, and transgenic rodent gene mutation assays (14-16).

Recently, quantitative assessment with BMD modeling and PoD metrics has been employed for *in vitro* genotoxicity data (17,18). During the 6th International Workshop on Genotoxicity Testing (IWGT) in 2013, the Working Group on Quantitative Approaches to Genetic Toxicology Risk Assessment (QWG) indicated the need for quantitative dose-response analysis of genotoxicity data, the development and evaluation of the methods used to derive PoDs, and the definition of the methodologies used to assess exposure-induced risks (19,20). Therefore, the methods used in applied genetic toxicology are in a transition period from qualitative hazard identification to quantitative dose-response analyses (21), as evidenced in a special issue on these topics published in 2016 in the journal *Mutagenesis*. In our laboratory, we have applied PoD and other quantitative metrics to data obtained from the mouse lymphoma assay (MLA) and performed quantitative analyses of the relative mutagenicity of five chemical compo-

nents representative of different classes of chemicals found in tobacco smoke (13). More recently, we used BMD modeling to provide a quantitative evaluation of the mutagenicity of four piperidine nitroxides [i.e., 2,2,6,6-Tetramethylpiperidine-1-oxyl (TEMPO), 4-hydroxy-TEMPO, 4-oxo-TEMPO, and 4-methoxy-TEMPO] with different substituent groups on the 4-position of the piperidine ring (22), and six cigarette whole smoke solutions (WSSs) prepared by bubbling machine-generated whole smoke through dimethyl sulfoxide (DMSO) (23). The findings demonstrated the suitability of BMD modeling for potency ranking of *in vitro* studies on structurally different chemical agents (13), structurally similar compounds (i.e., different derivatives of one chemical) (22), and different complex chemical mixtures (23).

BMD modeling is primarily conducted using two software packages, PROAST software, developed by the Netherlands National Institute for Public Health and the Environment (RIVM) (24), and BMDS, developed by the U.S. EPA (25). In previous studies, we used the PROAST software to perform quantitative dose-response analyses of the MLA data produced by testing four piperidine nitroxides (22) and six WSSs (23). In the present study, we reanalyzed these genotoxicity data from the MLA using the EPA's BMD software to evaluate the quantitative agreement of the estimates of genotoxic potency.

BENCHMARK DOSE MODELING

A reanalysis of the *in vitro* genotoxicity data for BMD modeling was performed using the EPA's Benchmark Dose Software (version 2.7, released on August 18, 2017) in accordance with the BMD technical guidance (25). Briefly, a continuous model was used in the reanalysis as all response variables can take a continuous range of values (22,23). The BMDS analysis included exponential, Hill, linear, polynomial, and power models. The models with absolute values of scaled residuals > 2 (indicating a disagreement between the predicted and observed means) were excluded, and the BMD_{10} [a 10% increase over the background mutant frequency (MF)] was calculated for each data set using all five models. Among the models (out of five models) with a reasonable fit to the dataset, the model that produced the lowest Akaike's Information Criterion was chosen as the best data fitting model (26), and was used for the generation of the BMDs, BMDLs, and BMDUs for 10%, 50%, 100% (i.e., two-fold increase), and 200% increases in MF above the control. The BMDU and BMDL refer to the two-sided (upper and lower bounds) 95% confidence intervals (CIs) of the BMD. The BMDL is the most widely preferred PoD. It should be noted that when performing a BMD modeling analysis, there may be a need for collaboration between toxicologists and statisticians to determine the most appropriate model.

BENCHMARK DOSE MODELING OF MOUSE LYMPHOMA ASSAY DATA FROM FOUR NITROXIDES

The MLA determines genetic changes affecting the expression of the thymidine kinase (*Tk*) gene located on murine chromosome 11, and is widely recommended as a genotoxicity assay by many regulatory authorities. The qualitative positive responses for the MLA are defined as those where in which there was a dose-related increase in MF and the induced MF in one or more chemical-treated samples in excess of the global evaluation factor (27). A quantitative analysis of the dose-response MLA data and BMD modeling using the EPA's BMDs for the determination of genetic toxicity PoD metrics was conducted previously (13) and is also discussed in this study.

The widespread use of piperidine nitroxides has raised safety concerns, and TEMPO is one of the frequently used piperidine nitroxide catalysts (28). In our previous studies, we reported that TEMPO exerted a mutagenic effect in mouse lymphoma cells, induced micronuclei in human lymphoblastoid TK6 cells (29), and resulted in intracellular reactive oxygen species (ROS) generation and DNA damage in mammalian cells (28). The addition of ring a different group to the 4-position of the piperidine ring is a sim-

ple method to form derivatives of TEMPO. Recently, we investigated the mutagenicity of four structurally similar piperidine nitroxides (TEMPO, 4-hydroxy-TEMPO, 4-methoxy-TEMPO, and 4-oxo-TEMPO) using the MLA and quantified their mutagenic potencies by using PROAST (22). The BMDL₁₀ was used to rank the potency of the nitroxides and when L5178Y mouse lymphoma cells were metabolically activated, 4-oxo-TEMPO was the nitroxide displaying the most potent mutagenicity in the MLA, followed by TEMPO > 4-methoxy-TEMPO > 4-hydroxy-TEMPO. The MLA dose-response data were also analyzed by using different BMD₅₀₋₂₀₀ values, and the four nitroxides showed clear differences in the mutagenic potency in the presence of S9, with 4-oxo-TEMPO exhibiting the lowest BMD values for BMD₁₀₋₂₀₀ (22).

These MLA mutagenicity data for the four TEMPO derivatives were reanalyzed by BMD modeling using the EPA's BMD software. First, the BMD₁₀ with its BMDL₁₀ and BMDU₁₀ were calculated by using the EPA's BMD software with five data fitting models (26). The data fitting model that produced the lowest Akaike's Information Criterion value was considered to be the most suitable model for fitting the data. The results indicated that the exponential model appeared the best fit for the calculation of BMD₁₀ for two nitroxides, 4-methoxy-TEMPO and 4-

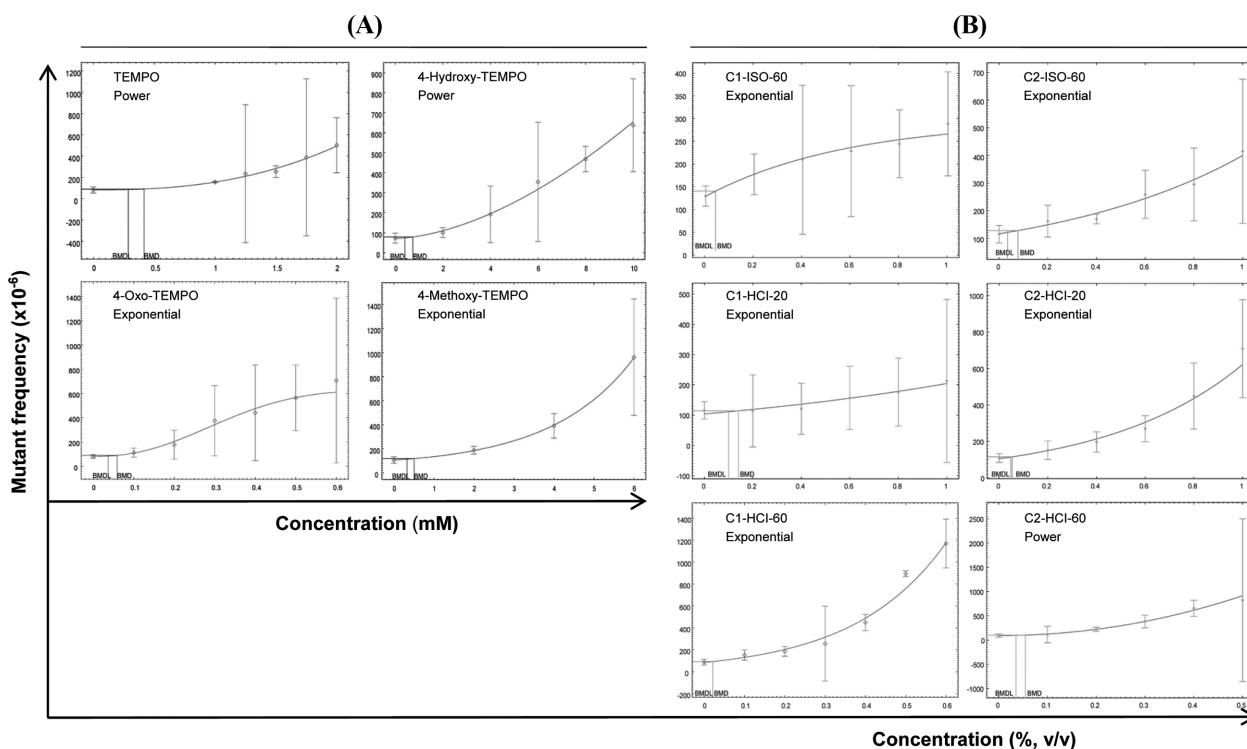


Fig. 1. Quantitative analysis of the mutagenic dose-responses induced by four piperidine nitroxides (A) and six WSSs (B). The original dose-response data were published previously (22,23), and the BMD₁₀, BMDL₁₀, and BMDU₁₀ were calculated by using the EPA's BMD software. The indication of exponential or power shows the model used for the calculation. The bars represent the calculated 95% confidence interval of each MF.

Table 1. BMDL₁₀ (with ranking) of four piperidine nitroxides and six whole smoke solutions (WSSs) in the mouse lymphoma assay

Group	Sample ID	BMDL ₁₀	
		BMDS [#]	PROAST [§]
Nitroxides (mM)	TEMPO	0.29 [2]*	0.26 [2]
	4-Hydroxy-TEMPO	0.41 [4]	1.02 [4]
	4-Oxo-TEMPO	0.04 [1]	0.06 [1]
	4-Methoxy-TEMPO	0.33 [3]	0.77 [3]
WSSs (%, v/v)	C1-ISO-60	0.09 [5]	0.24 [5]
	C2-ISO-60	0.07 [4]	0.17 [4]
	C1-HCI-20	0.09 [6]	0.25 [6]
	C1-HCI-60	0.02 [1]	0.04 [2]
	C2-HCI-20	0.05 [3]	0.12 [3]
	C2-HCI-60	0.04 [2]	0.04 [1]

[#]Calculated using the EPA's Benchmark Dose Software (BMDS).

[§]Previously published data (22,23) calculated using the PROAST Software.

*The number in square brackets is the rank of the potency.

oxo-TEMPO, whereas the power model provided the best fit for the TEMPO and 4-hydroxy-TEMPO data (Fig. 1A). The rank order of BMDL₁₀ for piperidine nitroxides generated by the EPA's BMDS was the same as that calculated by using PROAST (Table 1), although the numerical concentrations generated by the two software programmes differed slightly.

When the MLA data were analyzed using BMD₅₀, BMD₁₀₀, and BMD₂₀₀, similar BMDL potency rankings were also observed (Table 2). 4-Oxo-TEMPO produced the lowest BMDL values for BMDL₁₀₋₂₀₀, whereas 4-hydroxy-TEMPO had the highest BMDL values for all BMDLs. Generally, the lower the BMDL value, the higher the potential mutagenicity for a given agent. When comparing the BMDL/BMDU range for each chemical, the BMD₁₀ value for 4-oxo-TEMPO was classified separately from the other three chemicals (i.e., two groups), and the BMD₅₀, BMD₁₀₀, and BMD₂₀₀ values were separated into three groups. Although 4-methoxy-TEMPO had a relatively lower value for BMDL₁₀₋₂₀₀ than 4-hydroxy-TEMPO, their BMDL/BMDU ranges overlapped. Previously, using the covariate approach in PROAST, the 4-methoxy-TEMPO

BMDL/BMDU range was completely separated from the 4-hydroxy-TEMPO BMDL/BMDU range (22). This difference is not surprising because these two approaches were calculated using different models; the values for 4-hydroxy-TEMPO were calculated by using the power model as the accepted model to provide the PoD in the EPA's BMDS and the PROAST software used the exponential model (Fig. 1A). Frequently, different models compatible with the same dataset may result in different BMDLs with somewhat higher or lower risk estimates. With the exponential models, users can produce outcomes for each model of the exponential family using the EPA's BMDS and select the best model according to the BMD technical guidance (25), whereas PROAST automatically selects the best model. It should also be noted that PROAST and BMDS make different assumptions for statistical distributions (30).

BENCHMARK DOSE MODELING OF MOUSE LYMPHOMA ASSAY DATA FROM SIX WHOLE SMOKE SOLUTIONS

The methods to calculate the NOAEL, breakpoint-dose (BPD; formerly named Td), and BMD₁₀ for genotoxicity endpoints have been previously described (18,31). The NOAEL, LOAEL, BMDL₁₀, and BMDL_{1SD} (i.e., one standard deviation increase over the background) PoD metrics can be easily determined for most gene mutation and chromosomal damage studies; whereas BPDs cannot always be calculated owing to data limitations and constraints of the statistical methods (18). In the MLA data, the values of BMDL₁₀ were lower than the other PoDs, and therefore, the distribution of BMDL₁₀ values produced the lowest PoD (13). Among the methods utilized in the studies, including the lowest positive response by the global evaluation factor and mutagenic potency, the BMD approach appeared to be the most reasonable for the quantitative description of genotoxicity data generated from the MLA (13,23).

In a previous study, six WSSs were generated from two commercial cigarette products [Cigarette type #1 (C1) and Cigarette type #2 (C2)] with two smoke conditions [the International Organization for Standardization (ISO) and

Table 2. The lower and upper 95% confidence intervals derived from the benchmark dose (BMD) estimates for four piperidine nitroxides by using the BMDS

Piperidine nitroxide	BMD ₁₀ (mM)			BMD ₅₀ (mM)			BMD ₁₀₀ (mM)			BMD ₂₀₀ (mM)		
	BMDL	BMDU	BMDU/BMDL	BMDL	BMDU	BMDU/BMDL	BMDL	BMDU	BMDU/BMDL	BMDL	BMDU	BMDU/BMDL
4-Oxo-TEMPO	0.04	0.09	2.38	0.09	0.15	1.62	0.14	0.20	1.42	0.20	0.27	1.33
TEMPO	0.29	0.56	1.94	0.64	0.95	1.50	0.89	1.21	1.36	1.23	1.54	1.25
4-Methoxy-TEMPO	0.33	0.74	2.25	1.25	1.96	1.56	2.05	2.81	1.37	3.12	3.85	1.24
4-Hydroxy-TEMPO	0.41	1.17	2.89	1.28	2.60	2.03	2.09	3.68	1.76	3.39	5.24	1.55

Health Canada Intense (HCI) methods] and different numbers of cigarettes (i.e., 20 and 60) (23). The MFs from these six WSSs in the MLA, in the presence of S9 metabolic activation, were reanalyzed by using the EPA's BMDS. The results showed that the exponential model provided the best fit to calculate BMD₁₀ for all the WSSs, except C2-HCI-60, for which the power model was used. That the same model was chosen for five of six dose-responses may be a result of the similarity of the test articles, which were prepared from the mainstream smoke of cigarettes in the same brand family. The dose-response BMD₁₀ modeling

plots using the *Tk* gene MF data are presented in Fig. 1B.

The evaluation of BMDL_{10s} and BMD_{10s} by using BMDS generated relatively lower numerical values than NOAEL, LOAEL, and other metrics (13). The rank order of BMDL₁₀ for six WSSs generated by using the EPA's BMDS was similar to the rank order calculated by using PROAST, except the order of top two samples (i.e., C1-HCI-60 and C2-HCI-60) was reversed (Table 1). In addition, the numerical values of BMDL₁₀ calculated by the BMDS were smaller than those generated by PROAST, although the exponential model was used for the calcula-

Table 3. Comparison and rankings of the benchmark doses (BMDs) producing a 10%, 50%, 100%, and 200% increase over the background frequency (BMD₁₀₋₂₀₀) for six WSSs, calculated by using the BMDS

WSS ID	BMD ₁₀ (% <i>, v/v</i>)		BMD ₅₀ (% <i>, v/v</i>)		BMD ₁₀₀ (% <i>, v/v</i>)		BMD ₂₀₀ (% <i>, v/v</i>)	
	BMDS [#]	PROAST [§]	BMDS	PROAST	BMDS	PROAST	BMDS	PROAST
C1-ISO-60	0.12 [5]*	0.31 [5]	0.51 [5]	0.69 [5]	0.88 [5]	0.98 [5]	1.39 [5]	1.38 [5]
C2-ISO-60	0.08 [4]	0.21 [4]	0.33 [4]	0.47 [4]	0.56 [4]	0.67 [4]	0.89 [4]	0.94 [4]
C1-HCI-20	0.13 [6]	0.33 [6]	0.56 [6]	0.73 [6]	0.95 [6]	1.03 [6]	1.51 [6]	1.45 [6]
C1-HCI-60	0.04 [1]	0.06 [2]	0.13 [2]	0.13 [2]	0.20 [2]	0.19 [2]	0.30 [2]	0.26 [2]
C2-HCI-20	0.05 [2]	0.15 [3]	0.23 [3]	0.33 [3]	0.39 [3]	0.47 [3]	0.62 [3]	0.66 [3]
C2-HCI-60	0.05 [2]	0.05 [1]	0.12 [1]	0.12 [1]	0.17 [1]	0.16 [1]	0.24 [1]	0.23 [1]

[#]Calculated using the EPA's Benchmark Dose Software (BMDS).

[§]Previously published data (23) calculated using the PROAST Software.

*The number in square brackets is the potency rank of the WSSs.

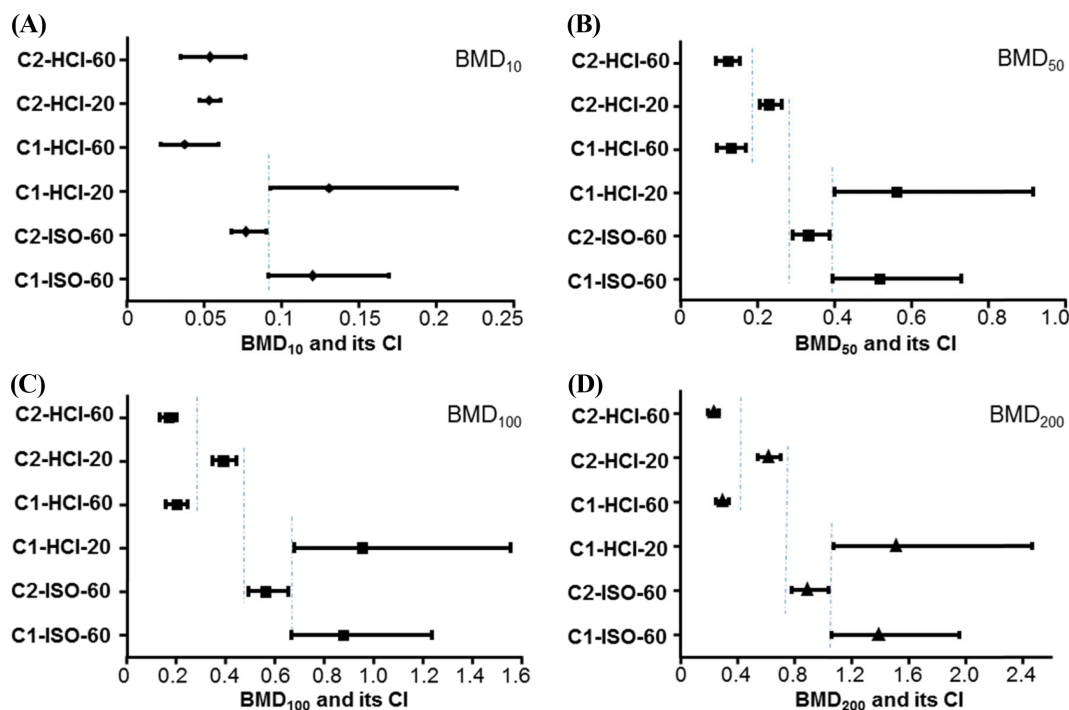


Fig. 2. A comparison of BMD values for WSSs in the mouse lymphoma assay. The BMDs (BMD₁₀, BMD₅₀, BMD₁₀₀, and BMD₂₀₀) estimates producing a 10%, 50%, 100%, or 200% increase over the background frequency were calculated using the selected models in the BMDS (A, B, C, and D). The bars represent the calculated 95% confidence interval (CI) for each value. The lower and upper limits derived from the BMD estimates were used to differentiate between the responses based on non-overlapping confidence intervals.

tion in both software programmes. Given that the two approaches were implemented with different model assumptions, this is not unexpected, and should be considered when developing a PoD during chemical risk assessments.

We also calculated the BMDs, BMDLs, and BMDUs for 50%, 100%, and 200% increases above the background response in order to achieve better discrimination between the dose responses (Table 3, Fig. 2). For example, the BMDS analysis of the six dose-responses divided the BMD_{10s} into two groups for which the upper and lower 95% CIs did not overlap (Fig. 2A). Similarly, it was possible to divide the BMD_{50s} , BMD_{100s} , and BMD_{200s} into four groups without overlapping 95% CIs (Figs. 2B, 2C, 2D). The BMD potency ranking analysis indicated that when 60 cigarettes of type #1 or #2 were analyzed by using the HCI smoking condition, the CIs of all $BMD_{10-200s}$ overlapped (Fig. 2), which suggested that C1-HCI-60 and C2-HCI-60 were not distinguishable from each other based on their mutagenic response (i.e., BMDs). In contrast, when 60 cigarettes of type #1 or #2 were analyzed by using the ISO smoking conditions, the CIs of all BMDs were separated from each other, which indicated that the mutagenic responses induced by these two samples were significantly different. In addition, BMD potency ranking of BMD_{10-200} also distinguished the mutagenicity of samples induced by smoking 20 cigarettes of type #1 or #2 under the HCI regimen (C1-HCI-20 vs. C2-HCI-20) and the mutagenicity of samples induced by smoking different numbers of the same cigarette, i.e., C1-HCI-20 vs. C1-HCI-60 and C2-HCI-20 vs. C2-HCI-60. Furthermore, when comparing the same cigarette with same cigarette number, different smoking conditions resulted in different mutagenicity responses, as indicated by BMD potency rankings of BMD_{50} , BMD_{100} , and BMD_{200} for C2-HCI-60 vs. C2-ISO-60 and C1-HCI-60 vs. C1-ISO-60 (Figs. 2B, 2C, 2D).

The BMD values and rankings estimated in this study using the EPA's BMDS were comparable with those calculated in our previous study using PROAST. There were only small differences in the ranking of the top two samples between the two estimates by the PROAST and BMDS for BMD_{10} ; in addition, the numerical values of the BMD_{50} , BMD_{100} , and BMD_{200} generated by both software programmes were very similar and followed the same order for all WSSs (Table 3). Given that different software is used, practical differences between the two approaches are expected. However, in the context of the data modeled in this study, it was observed that the modeling results were comparable in terms of the order of potency. It should be noted that the results from PROAST (Table 3) were calculated by conducting covariate analyses [i.e., considering a variable (e.g., different WSS samples) as a covariate], to provide additional BMD precision (31). Our previous study indicated that the weak muta-

genic potential and the imprecision in determining dose-responses in different studies may contribute to the variability of CI values (23).

CONCLUSIONS

Covariate analysis is not an option in the EPA's BMDS, yet the use of both BMDS and PROAST with a covariate approach produced identical grouping patterns for WSSs when the analyses were conducted at BMD_{50} , BMD_{100} , or BMD_{200} . Our data indicate that, despite small differences in the data output for the two software packages, both were functionally equivalent for conducting dose-response analysis and generating mutagenicity potency rankings. Currently, the U.S. EPA and the Netherlands' RIVM are collaborating to achieve consistency between the BMDS and PROAST software packages (32).

BMD modeling has been used for more than 20 years in the U.S. for risk assessment (4), but received little attention in Europe before 2009 (33). Recently, BMD modeling has been more widely accepted, especially in applied genetic toxicology (21). To the best of our knowledge, this study is the first comparison of mutagenicity potency ranking on the same *in vitro* genotoxicity data from the MLA conducted with BMD values generated from two software packages, BMDS and PROAST. Both software programmes are regularly used for the estimation of BMDs. Although there are some limitations in our previous studies (22,23) and this study, our approach to making this comparison was to identify the analysis routine that produced the greatest differentiation between the mutagenic effects produced by different derivatives of one chemical, such as nitroxides, or different complex chemical mixtures, such as WSSs.

ACKNOWLEDGMENTS

We thank Drs. Dayton Petibone and Dong Wang for their critical review of this manuscript. The information in this manuscript is not a formal dissemination of information by the U.S. Food and Drug Administration (FDA). This publication represents the views of the authors and does not represent FDA position or policy.

CONFLICT OF INTEREST

The authors have no conflict of interest.

Received August 10, 2018; Revised August 16, 2018; Accepted August 30, 2018

REFERENCES

1. Crump, K.S. (1984) A new method for determining allow-

- able daily intakes. *Fundam. Appl. Toxicol.*, **4**, 854-871.
2. Cronin, M.T.D., Enoch, S.J., Mellor, C.L., Przybylak, K.R., Richarz, A.N. and Madden, J.C. (2017) *In silico* prediction of organ level toxicity: Linking chemistry to adverse effects. *Toxicol. Res.*, **33**, 173-182.
 3. Brandon, E.F., Bulder, A.S., van Engelen, J.G., Mahieu, C.M., Mennes, W.C., Pronk, M.E., Rietveld, A.G., van de Ven, B.M., Ten Voorde, S.E., Wolterink, G., Slob, W., Zeil-maker, M.J. and Bessems, J.G. (2013) Does EU legislation allow the use of the Benchmark Dose (BMD) approach for risk assessment? *Regul. Toxicol. Pharmacol.*, **67**, 182-188.
 4. EPA (1994) Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry. U.S. Environmental Protection Agency [Accessed on 2018 Aug 7]. Available from: <https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=30001K30004C.PDF/>.
 5. EPA (1999) What is benchmark dose software (BMDS)? U.S. Environmental Protection Agency [Accessed on 2018 Aug 7]. Available from: <https://www.epa.gov/bmds/what-benchmark-dose-software-bmds/>.
 6. Barnes, D.G., Daston, G.P., Evans, J.S., Jarabek, A.M., Kavlock, R.J., Kimmel, C.A., Park, C. and Spitzer, H.L. (1995) Benchmark Dose Workshop: criteria for use of a benchmark dose to estimate a reference dose. *Regul. Toxicol. Pharmacol.*, **21**, 296-306.
 7. Mattison, D.R. and Sandler, J.D. (1994) Summary of the workshop on issues in risk assessment: quantitative methods for developmental toxicology. *Risk Anal.*, **14**, 595-604.
 8. Slikker, W., Jr., Crump, K.S., Andersen, M.E. and Bellinger, D. (1996) Biologically based, quantitative risk assessment of neurotoxicants. *Fundam. Appl. Toxicol.*, **29**, 18-30.
 9. Van Landingham, C.B., Allen, B.C., Shipp, A.M. and Crump, K.S. (2001) Comparison of the EU T25 single point estimate method with benchmark dose response modeling for estimating potency of carcinogens. *Risk Anal.*, **21**, 641-656.
 10. Burgoon, L.D. and Zacharewski, T.R. (2008) Automated quantitative dose-response modeling and point of departure determination for large toxicogenomic and high-throughput screening data sets. *Toxicol. Sci.*, **104**, 412-418.
 11. Thomas, R.S., Clewell, H.J., 3rd, Allen, B.C., Wesselkamper, S.C., Wang, N.C., Lambert, J.C., Hess-Wilson, J.K., Zhao, Q.J. and Andersen, M.E. (2011) Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. *Toxicol. Sci.*, **120**, 194-205.
 12. EPA (2005) Guidelines for carcinogen risk assessment. U.S. Environmental Protection Agency [Accessed on 2018 Aug 7]. Available from: http://www2.epa.gov/sites/production/files/2013-2009/documents/cancer_guidelines_final_2013-2025-2005.pdf/.
 13. Guo, X., Heflich, R.H., Dial, S.L., Richter, P.A., Moore, M.M. and Mei, N. (2016) Quantitative analysis of the relative mutagenicity of five chemical constituents of tobacco smoke in the mouse lymphoma assay. *Mutagenesis*, **31**, 287-296.
 14. Hernandez, L.G., Slob, W., van Steeg, H. and van Benthem, J. (2011) Can carcinogenic potency be predicted from *in vivo* genotoxicity data? a meta-analysis of historical data. *Environ. Mol. Mutagen.*, **52**, 518-528.
 15. Cao, X., Mittelstaedt, R.A., Pearce, M.G., Allen, B.C., Soeteman-Hernandez, L.G., Johnson, G.E., Bigger, C.A. and Heflich, R.H. (2014) Quantitative dose-response analysis of ethyl methanesulfonate genotoxicity in adult gpt-delta transgenic mice. *Environ. Mol. Mutagen.*, **55**, 385-399.
 16. Paini, A., Scholz, G., Marin-Kuan, M., Schilter, B., O'Brien, J., van Bladeren, P.J. and Rietjens, I.M. (2011) Quantitative comparison between *in vivo* DNA adduct formation from exposure to selected DNA-reactive carcinogens, natural background levels of DNA adduct formation and tumour incidence in rodent bioassays. *Mutagenesis*, **26**, 605-618.
 17. Gollapudi, B.B., Johnson, G.E., Hernandez, L.G., Pottenger, L.H., Dearfield, K.L., Jeffrey, A.M., Julien, E., Kim, J.H., Lovell, D.P., Macgregor, J.T., Moore, M.M., van Benthem, J., White, P.A., Zeiger, E. and Thybaud, V. (2013) Quantitative approaches for assessing dose-response relationships in genetic toxicology studies. *Environ. Mol. Mutagen.*, **54**, 8-18.
 18. Johnson, G.E., Soeteman-Hernandez, L.G., Gollapudi, B.B., Bodger, O.G., Dearfield, K.L., Heflich, R.H., Hixon, J.G., Lovell, D.P., MacGregor, J.T., Pottenger, L.H., Thompson, C.M., Abraham, L., Thybaud, V., Tanir, J.Y., Zeiger, E., van Benthem, J. and White, P.A. (2014) Derivation of point of departure (PoD) estimates in genetic toxicology studies and their potential applications in risk assessment. *Environ. Mol. Mutagen.*, **55**, 609-623.
 19. MacGregor, J.T., Frotschl, R., White, P.A., Crump, K.S., Eastmond, D.A., Fukushima, S., Guerard, M., Hayashi, M., Soeteman-Hernandez, L.G., Kasamatsu, T., Levy, D.D., Morita, T., Muller, L., Schoeny, R., Schuler, M.J., Thybaud, V. and Johnson, G.E. (2015) IWGT report on quantitative approaches to genotoxicity risk assessment I. Methods and metrics for defining exposure-response relationships and points of departure (PoDs). *Mutat. Res. Genet. Toxicol. Environ. Mutagen.*, **783**, 55-65.
 20. MacGregor, J.T., Frotschl, R., White, P.A., Crump, K.S., Eastmond, D.A., Fukushima, S., Guerard, M., Hayashi, M., Soeteman-Hernandez, L.G., Johnson, G.E., Kasamatsu, T., Levy, D.D., Morita, T., Muller, L., Schoeny, R., Schuler, M.J. and Thybaud, V. (2015) IWGT report on quantitative approaches to genotoxicity risk assessment II. Use of point-of-departure (PoD) metrics in defining acceptable exposure limits and assessing human risk. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.*, **783**, 66-78.
 21. White, P.A. and Johnson, G.E. (2016) Genetic toxicology at the crossroads—from qualitative hazard evaluation to quantitative risk assessment. *Mutagenesis*, **31**, 233-237.
 22. Guo, X., Seo, J.E., Bryce, S.M., Tan, J.A., Wu, Q., Dial, S.L., Moore, M.M. and Mei, N. (2018) Comparative genotoxicity of TEMPO and 3 of its derivatives in mouse lymphoma cells. *Toxicol. Sci.*, **163**, 214-225.
 23. Guo, X., Heflich, R.H., Dial, S.L., De, M., Richter, P.A. and Mei, N. (2018) Quantitative differentiation of whole smoke solution-induced mutagenicity in the mouse lymphoma assay. *Environ. Mol. Mutagen.*, **59**, 103-113.
 24. RIVM (2013) PROAST. The Dutch National Institute for Public Health and the Environment [Accessed on 2018 Aug 7]. Available from: http://www.rivm.nl/en/Documents_and_publications/Scientific/Models/PROAST/.
 25. EPA (2012) Benchmark dose technical guidance. U.S. Envi-

- ronmental Protection Agency [Accessed on 2018 Aug 7]. Available from: <https://www.epa.gov/risk/benchmark-dose-technical-guidance/>.
26. Davis, J.A., Gift, J.S. and Zhao, Q.J. (2011) Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1. *Toxicol. Appl. Pharmacol.*, **254**, 181-191.
 27. Moore, M.M., Honma, M., Clements, J., Bolcsfoldi, G., Burlinson, B., Cifone, M., Clarke, J., Delongchamp, R., Durward, R., Fellows, M., Gollapudi, B., Hou, S., Jenkinson, P., Lloyd, M., Majeska, J., Myhr, B., O'Donovan, M., Omori, T., Riach, C., San, R., Stankowski, L.F., Jr., Thakur, A.K., Van Goethem, F., Wakuri, S. and Yoshimura, I. (2006) Mouse lymphoma thymidine kinase gene mutation assay: follow-up meeting of the International Workshop on Genotoxicity Testing--Aberdeen, Scotland, 2003--Assay acceptance criteria, positive controls, and data evaluation. *Environ. Mol. Mutagen.*, **47**, 1-5.
 28. Guo, X., Chen, S., Zhang, Z., Dobrovolsky, V.N., Dial, S.L., Guo, L. and Mei, N. (2015) Reactive oxygen species and c-Jun N-terminal kinases contribute to TEMPO-induced apoptosis in L5178Y cells. *Chem. Biol. Interact.*, **235**, 27-36.
 29. Guo, X., Mittelstaedt, R.A., Guo, L., Shaddock, J.G., Heflich, R.H., Bigger, A.H., Moore, M.M. and Mei, N. (2013) Nitroxide TEMPO: a genotoxic and oxidative stress inducer in cultured cells. *Toxicol. In Vitro*, **27**, 1496-1502.
 30. EFSA (2017) Update: use of the benchmark dose approach in risk assessment. *The EFSA Journal*, **15**, 4658.
 31. Wills, J.W., Johnson, G.E., Doak, S.H., Soeteman-Hernandez, L.G., Slob, W. and White, P.A. (2016) Empirical analysis of BMD metrics in genetic toxicology part I: *in vitro* analyses to provide robust potency rankings and support MOA determinations. *Mutagenesis*, **31**, 255-263.
 32. PHE (2014) Carcinogenic dose response: defining a point of departure and potency estimates. Public Health England [Accessed on 2018 Jul 5]. Available from: <https://www.gov.uk/government/publications/carcinogenic-dose-response-defining-a-point-of-departure-and-potency-estimates/>.
 33. EFSA (2009) Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessment *The EFSA Journal*, **1150**, 1-72.