PERFORMANCE WORK STATEMENT
Task Order XXX

Title: External Review (Letter) of *Quantitative risk assessment: Developing a complete Bayesian approach to Dichotomous Dose-Response Model Averaging*

PERIOD OF PERFORMANCE: Task Order award through September 28, 2018

I. Purpose

The Purpose of this purchase order is to provide services to the U.S. Environmental Protection Agency's (EPA) National Center for Environmental Assessment (NCEA), Office of Research and Development (ORD), in organizing and conducting by mail an expert external panel review of the draft NCEA report entitled *Quantitative risk assessment: Developing a complete Bayesian approach to Dichotomous Dose-Response Model Averaging* and associate software source code and executable files.

II. Background

This Section provides an overview of the model averaging methods developed and testing procedures employed. Additional details, model executables and charge questions will be provided to the reviewers as supplemental material (see Task 6).

Model averaging [1-4] is a statistical technique allowing inference over multiple parametric models by producing estimates of the predictor-response relationship as a convex weighted sum of individual models. For quantitative risk assessors, one finds many different model averaging methods dedicated to dose response and benchmark dose estimation [5-9]. Current research shows quantitative risk assessments based upon a single "best model" have poor statistical properties [7,9] and experience with various forms of model averaging has suggested that adoption of a model averaging approach would be superior to the current method of basing inference upon a single model [6,8,9].

To solve these problems, this method proposes a fully Bayesian [10] approach that allows many of these problems to be sidestepped by the inclusion of prior information. In this approach, the parameters of included models are no longer strictly bounded; instead this method opts for "soft bounds" defined by a mildly informative prior distribution. These distributions are specified to put low prior probability on regions often defined outside the boundary of the parameters while placing relatively high prior probability on reasonable parameter values. For example, the US EPA's BMD technical guidance [11] recommends constraining the bounds of the shape parameter of the Weibull model to be greater or equal to one, because values less than 1 lead to an infinite slope of the dose-response curve at dose zero. Now, the shape parameter is allowed to take any number on the positive real line, but the prior gives a small probability, approximately 2.5%, to values less than 1. As such, the model will still describe data in this supralinear region, but such models will only get a high weight if models that are more parsimonious do not describe the data well and the data give large support to the model form. This procedure allows for models that are more extraordinary (e.g. infinite slope or models that give zero response then 100% response in a small dose range) to be fit. Yet, this procedure gives the models high weight when the data

support them and, in the cases where there are limited data, the models are more limited to dose-response shapes that frequently seen in practice.

Models

Consider an animal toxicology experiment with m unique dose groups $d_1,\ldots d_m$ and $n_1,\ldots,n_m$ animals per dose group. For this experiment, let $y_1,\ldots y_m$ be the number of positive responses observed in each dose group. It is frequently assumed that $y_i \sim \text{binomial}(\pi(d_i), n_i)$, where $\pi(d_i)$ is the probability of adverse response. To estimate $\pi(d_i)$ given $y_1,\ldots y_m$, $\pi(d_i)$ is often assumed to be a parametric function of dose. For example, the current US EPA Benchmark Dose Software (US EPA 2016) can be used to estimate the function $\pi(d)$ using one of nine dose-response function. These functions:

$$\pi_1(d) = \gamma + (1-\gamma)[1-\exp(-\beta d)] \quad \text{Quantal linear,} \quad (1)$$

$$\pi_2(d) = \gamma + (1-\gamma)[1-\exp(-\beta_1 d - \beta_2 d^2)] \quad \text{Multistage,} \quad (2)$$

$$\pi_3(d) = \gamma + (1-\gamma)[1-\exp(-\beta d^{\alpha})] \quad \text{Weibull,} \quad (3)$$

$$\pi_4(d) = \gamma + (1-\gamma)\frac{1}{\Gamma(\alpha)}\int_0^{\beta d} t^{\alpha-1}e^{-t}dt \quad \text{Gamma,} \quad (4)$$

$$\pi_5(d) = v\gamma + \frac{(v - v\gamma)}{1+\exp[-a - b\log(d)]} \quad \text{Dichotomous Hill,} \quad (5)$$

$$\pi_6(d) = \frac{1}{1+\exp(\beta_0 + \beta_1 d)} \quad \text{Logistic,} \quad (6)$$

$$\pi_7(d) = \gamma + \frac{1-\gamma}{1+\exp[\beta_0 + \beta_1 \log(d)]} \quad \text{Log-Logistic,} \quad (7)$$

$$\pi_8(d) = \Phi(\beta_0 + \beta_1 d) \quad \text{Probit,} \quad (8)$$

$$\pi_9(d) = \gamma + (1+\gamma)\Phi[\beta_0 + \beta_1 \log(d)] \quad \text{Log-Probit,} \quad (9)$$

As studied previously [7, 12], picking a single model (e.g models (1)-(9)) may result in misrepresentation of the true underlying dose-response relationship and results in significant model uncertainty. Bayesian Model averaging [1, 3] develops a probabilistic framework that incorporates inference from all models considered. One constructs the "model-average model" through combination of the posterior probability distributions that the individual models are correct given data. A large literature applies model averaging to Benchmark dose (BMD) estimation for dichotomous and continuous responses [5-9]; this literature shows model averaging produces BMD estimates closer to the true BMD as well as producing narrower confidence intervals when compared to current practice.

Proposed Model Averaging Approach

The proposed model averaging approach is distinct from previous approaches. It is fully Bayesian and uses maximum a posteriori (MAP) estimation [13] for inference. Additionally, the approach computes the credible intervals bounds using model averaged profile likelihood (MAPL) methods [14]. The approach allows researchers to use prior information in the analysis. Such priors produced regularized estimators of the dose-response curve (i.e., prevents the dose response curve from changing too sharply when there is not enough data), and the maximum *a posteriori* estimation and profile likelihood computation reduces the computational burden necessary compared to computing bootstrap based confidence intervals [15]. The method allows estimation of models having a greater number of parameters than dose groups, which is possible because of the prior information. This allows datasets to be fit using the exact same methodology and the same set of dose-response models regardless of the number of dose groups.

Priors

We use models (1) – (9) in the model averaging procedure and place priors over the parameters for each of the models individually. We chose the priors for each distribution to provide a high prior probability over curves commonly seen in practice while providing lower prior probability on other dose-response curves. This approach allows the data to drive the dose-response estimation but may prevent scientifically unreasonable fits in situations where there are limited data. For example, using the method of maximum likelihood to estimate the parameter α in model (3) one often constrains this value to be less than 18, and values near 18 result in a hockey shaped dose-response curve that implies the probability of an adverse event goes from background to 100% in an extremely small dose range, which may be unrealistic. Assuming high values of this parameter are unlikely, the proposed prior puts exponentially decreasing weight for values of α near 18, and the resultant Bayesian fit may estimate this value to be smaller than its equivalent estimate made using maximum likelihood, which is especially true in cases where there are limited data on the dose response curve. In cases where there are sufficient data, the priors are such that the data overwhelms the prior and one sees minimal differences between the Bayesian estimate and the method of maximum likelihood.

Further, we place priors over the parameters to ensure models do not degenerate into other models. For example, model (2) can degenerate into (1) if $\beta_2 = 0$. We place a prior over this parameter to ensure positivity, which prevents the parameter from getting near zero, and if this value is close to one, the prior is such that model (1) will be preferred over model (2). These priors allow consistent estimation.

For all of the models having a γ parameter, which defines the background probability of the event, we transform this parameter using the logistic CDF, that is, $\gamma = \frac{1}{1+\exp(-\Psi)}$ where Ψ is specified on the entire real line.

Model	**Constraints**	**Priors**	**Notes**
Quantal linear $\pi_1(d) = \gamma + (1-\gamma)(1-\exp[-\beta d])$	$\beta > 0$ $0 \le \gamma \le 1$	$\log(\beta) \sim$ Normal(0,1) $\Psi \sim$ Normal(0,2)	$\gamma = \frac{1}{1+\exp(-\Psi)}$
Multistage $\pi_2(d) = \gamma + (1-\gamma)(1-\exp[-\beta_1 d - \beta_2 d^2])$	$\beta_1 > 0$ $\beta_2 > 0$ $0 \le \gamma \le 1$	$\log(\beta_1) \sim$ Normal(0,0.25) $\log(\beta_2) \sim$ Normal(0,1) $\Psi \sim$ Normal(0,2)	Note the prior over the β_1 parameter expresses the belief that the linear term should be positive if the quadratic term is positive in the two hit model of carcinogenesis.
Weibull $\pi_3(d) = \gamma + (1-\gamma)(1-\exp[-\beta d^{\alpha}])$	$\beta > 0$ $\alpha > 0$ $0 \le \gamma \le 1$	$\log(\beta_1) \sim$ Normal(0,1) $\log(\alpha) \sim$ Normal(log(2),0.18) $\Psi \sim$ Normal(0,2).	Here the prior over α is designed such that there is only a 0.01 prior probability the power parameter will be less than 1. This allows for models that are supra-linear; however, it requires a large amount of data for the α parameter to go much below 1.
Gamma $\pi_4(d) = \gamma + \frac{1-\gamma}{\Gamma(\alpha)} \int_0^{\beta d} t^{\alpha-1} \exp(-t)\, dt$	$\beta > 0$ $\alpha > 0$ $0 \le \gamma \le 1$	$\log(\beta) \sim$ Normal(0, 1) $\log(\alpha) \sim$ Normal(log(2),0.18) $\Psi \sim$ Normal(0,2)	Here the prior over α is designed such that there is only a 0.01 prior probability the power parameter will be less than 1. This allows for models that are supra linear; however, it requires a large amount of data for the parameter to go much below 1.
Dichotomous Hill $\pi_5(d) = \gamma + \frac{\nu(1-\gamma)}{1+\exp[-a - b\log(d)]}$	$0 \le \gamma \le 1$ $0 \le \nu \le 1$ $-\infty < a < \infty$ $b > 0$	a ~ Normal(0, .25) b ~ Normal(log(10),0.0625) $\Psi \sim$ Normal(0,2) $\nu \sim$ Normal(4,2)	$\gamma = \frac{1}{1+\exp(-\Psi)}$
Logistic $\pi_6(d) = \frac{1}{1+\exp[-\beta_0 - \beta_1 d]}$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	$\beta_0 \sim$ Normal(0, 1) $\log(\beta_1) \sim$ Normal(0,2)	
Log-Logistic $\pi_7(d) = \gamma + \frac{1-\gamma}{1+\exp[-\beta_0 - \beta_1 \log(d)]}$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	$\beta_0 \sim$ Normal(0, 1) $\log(\beta_1) \sim$ Normal(log(2),0.25) $\Psi \sim$ Normal(0,2).	$\gamma = \frac{1}{1+\exp(-\Psi)}$
Probit $\pi_8(d) = \Phi(\beta_0 + \beta_1 d)$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	$\beta_0 \sim$ Normal(0,1) $\log(\beta_1) \sim$ Normal(0,1)	
Log-Probit $\pi_9(d) = \gamma + (1-\gamma)\Phi[\beta_0 + \beta_1 \log(d)]$	$-\infty < \beta_0 < \infty$ $\beta_1 > 0$	$\beta_0 \sim$ Normal(0, 1) $\log(\beta_1) \sim$ Normal(log(2),0.25) $\Psi \sim$ Normal(0,2)	$\gamma = \frac{1}{1+\exp(-\Psi)}$

Given data, we fit models (1)-(9) individually, and we compute individual BMDs and BMDLs from the approximate distribution of the BMD given model M_k, i.e., $Pr(\text{BMD} \mid M_k, D)$. Additionally, we calculate an estimate of the BMD as well as the BMDL using model averaging. As the method is Bayesian, the approach for computing the BMD as well as the BMD lower bound is different from past methods.

Weight Calculation

In previous approaches to benchmark dose calculation using model averaging (e.g., see Bailer, Noble and Wheeler [5]), weights were calculated using either the BIC or AIC, where the AIC is used primarily in frequentist model averaging [2]. The proposed approach generates weights using the Laplace approximation to the marginal density of the data [16]. That is for model M_k, $1 \leq k \leq 9$, on approximates the marginal density as:

$$I_k = (2\pi)^{s/2}|\Sigma_k|^{1/2}\ell\left(D\middle|\hat{\theta}_k, M_k\right)g(\hat{\theta}_k|M_k), \quad (10)$$

where Σ_k is the negative inverse Hessian matrix, $\hat{\theta}_k$ is the MAP estimate, $\ell\left(D\middle|\hat{\theta}_k, M_k\right)$ is the likelihood of the model given the data D, and $g(\hat{\theta}_k|M_k)$ is the prior density for θ_k.

For each model M_k, one calculates the MAP and calculates I_k using equation (2). The posterior probability of the model is

$$\pi_k(M_k|D) = \frac{g(M_k)I_k}{\sum_{i=1}^{9} g(M_k)I_k},$$

where $g(M_k)$ is the prior probability of model M_k (e.g., 1/9 if each of 9 models is treated as equally plausible *a priori*).

Computation of the BMDL

The posterior density of the BMD is then:

$$g_{ma}(BMD) = \sum_{k=1}^{9} \pi_k(M_k|D)g_k(BMD|M_k, D), \quad (11)$$

From this quantity, 100(1-α)% confidence limits on the BMD are estimated by integrating (11), that is one finds the value BMD_α such that:

$$\alpha = \int_{-\infty}^{BMD_\alpha} \Pr(BMD \mid D)dBMD, \quad (12)$$

where the integral is approximated using the method of profile likelihood discussed in Fletcher and Turek [14]. Given that the profile likelihood is used, the method is fast taking less than 1 second to complete the estimation. Additionally, one can solve (12) for (1-α) to compute the upper bound of the benchmark dose or the BMDU.

Testing of Methods

To test the performance of the proposed method, Monte Carlo simulations were run using thirty-four different dose response curves (i.e., simulation templates) assuming an experimental condition designed to mimic chronic bioassays. In this setting, simulated datasets consisted of

four dose groups with 50 observations per group with geometric spacing between doses (0, 0.25, 0.5, and 1.0); 2000 simulated datasets were analyzed, investigating coverage, bias (% of true BMD), and BMD/BMDL ratio. As comparisons to the proposed methods, three other approaches were also applied to the simulated datasets:

1. the recommended US EPA approach defined in the Agency's BMD Technical Guidance [11];
2. an competing Bayesian model averaging method from Shao and Shapiro [17] (this methodology fit models (1)-(4) and (6)-(8) using the same priors as defined above and a model averaging approach as defined in that manuscript); and
3. the non-parametric method described in Guha et al. [18].

Materials TO BE Provided for Review

The following materials will be provided for distribution to reviewers.

- Report describing development and testing of method, which will consist of a draft manuscript (28 pages) and supplemental material (9 pages)
- Model source code file (createSim3.r; the software is written in R)
- An Excel based program that will allow for running the model averaging software with user selected priors
- Charge to reviewers (EPA will provide specific questions for the reviewers to address)

III. Scope of Work

Task 1. Prepare the Proposal

The contractor shall submit a fixed price cost proposal.

Task 2. Conference Calls with EPA Technical Project Officer (TPO)

Within three working days after proposal approval and TO award, the contractor shall convene a conference call with the TPO and appropriate contractor staff to clarify outstanding questions and confirm the schedule. The contractor shall provide weekly reports on progress to the TPO by telephone, and shall initiate additional communication with the TPO should developments arise that will affect the conduct or schedule of this peer review.

Task 3. Conflict of Interest (COI) Analysis and Certification

a. Prior to selecting expert reviewers, the contractor shall perform an evaluation to determine the existence of an actual or potential conflict of interest (COI) for each proposed reviewer. The contractor shall incorporate the Attachments to this PWS, listing yes/no questions and requests for supporting information, into its established process to evaluate and determine the presence of an actual or potential COI.

b. The contractor shall resolve issues of actual or potential conflicts of interest and panel composition before selecting the reviewers. As each situation must be evaluated on a case-by-case basis after consideration of specific circumstances, the contractor may consult with the TPO in carrying out these responsibilities. Consultation between the contractor and the TPO must be documented and provided to the TPO to assure transparency in the process and full disclosure, if questions arise concerning COI.

c. The contractor shall provide a written basis (see Section V, *Conflict of Interest*, and Attachment 2, *Conflict of Interest Analysis and Certification – Questions and Supporting Information*) and signed certification (see Attachment 1) that might be made public for concluding that there are no unresolved actual or potential conflicts of interest issues among the reviewers, and to ensure that the set of reviewers selected is suitably balanced with respect to any actual or apparent bias.

d. The contractor shall require each reviewer to provide a signed declaration (Attachment 1) that the reviewer is not arranging any new professional relationship with, or obtaining new financial holdings in, an entity which is not yet reported to the contractor or could be viewed as related to the topic under discussion and its associated stakeholders. Note: *signed* certifications and declarations are required by EPA only for those reviewers selected; there is no need to have candidate reviewers sign the declaration.

e. The contractor shall provide the TPO with resumes for all reviewers. These resumes might later be made public.

Task 4. Identify and Screen Proposed Reviewers

Within 7 working days after proposal approval, the contractor shall supply the TPO with a draft list of at least five (5) proposed reviewers, with a summary of their experience and qualifications and potential conflicts of interest and a CV.

The contractor may inform candidate reviewers and potential candidates about the nature of the task at hand, including the charge to reviewers, and background information in this statement of work, but should not send printed material or copies of review materials to prospective candidates.

EPA and the contractor will confer by telephone regarding the proposed reviewers on a day mutually convenient but expected to be 2-3 work days after EPA receives the lists of proposed reviewers. The TPO may provide comments to the contractor regarding any concerns about a known or potential conflict of interest, or concerns about a proposed reviewer's expertise, and may ask for a response from the contractor detailing how the concern will be resolved. When the TPO has agreed with the contractor on a list of proposed reviewers, he will notify the contractor to proceed with Task 5.

<u>Experience and Qualifications</u>

Reviewers should be experienced in the use of dose-response models for chemical risk assessment, and familiar with EPA benchmark dose (BMD) methods. In particular, reviewers should have experience or expert knowledge of model averaging methods and Bayesian statistics as they apply to dose-response analysis. Reviewers should understand maximum *a posteriori* estimation and "profile likelihood" methods. Ideally, at least one reviewer will have experience in developing or modifying such models, developing computer programs, and reviewing and testing such programs for quality assurance purposes (e.g., to validate accuracy of results and the correct implementation of algorithms used for statistical estimation and optimization).

The reviewers shall have scientific credentials equivalent to a Ph.D. or M.D. and shall be regarded as experts in their field. Experience and accomplishments shall be demonstrated by authorship of peer-reviewed publications and by professional activities (including, but not limited to, serving on governmental and non-governmental advisory panels or committees, organizing or chairing symposia, developing publicly-available software, and recognition by

professional societies). The contractor shall make a diligent effort to obtain reviewers who are nationally-recognized and respected experts.

EPA shall have an opportunity to provide comments to the contractor regarding known or potential conflicts of interest, or concerns about their expertise, before final selection occurs. If EPA determines that the reviewers are scientifically unqualified for the specific review, the contractor shall have three days to submit substitutes for review and selection by the TPO. The contractor shall respond in writing to any comments from EPA regarding potential COI or reviewer qualifications, describing any further inquiries or research conducted by the contractor regarding COI or qualifications and any actions taken as a result (e.g., dropping a proposed reviewer, proposing one or more additional reviewers).

Task 5. Select External Reviewers

When the TPO has agreed with the contractor on the list of proposed reviewers, he will notify the contractor to proceed with this task. Then, within 3 working days, the contractor shall select and secure arrangements with three (3) experts from among those on the list of proposed reviewers. After a list of proposed reviewers is agreed upon, selection of three reviewers from the list will be made by the contractor, without direction or influence from EPA.

The contractor shall ascertain whether, or not, reviewers have a conflict of interest (as defined by FAR subpart 9.5). Attachment 2 identifies the information that must be collected from each reviewer. The contractor shall forward a copy of the conflict of interest certifications (see Section V and Attachment 1of this PWS), to the EPA Project Officer and TPO, within 3 work days after selecting and engaging the services of the reviewers.

Task 6. Begin the Peer Review

Within 5 working days of receiving the TPO's agreement to the list of proposed reviewers, the contractor will engage the services of 3 reviewers and send them the charge and the review materials. The reviewers' written comments will be due no later than 4 weeks after the contractor dispatches review materials to the reviewers. However, this 4-week time frame may be extended as needed by EPA in written technical direction if a change is necessary to allow time for obtaining a sufficient number of qualified reviewers.

Each peer reviewer will be required to provide responses to all the questions in the charge. These responses must be more than perfunctory, providing details sufficient to indicate the factual basis for the reviewer's responses.

The following review materials will be provided by EPA to the contractor:

- Report describing development and testing of method, which will consist of a draft manuscript (28 pages) and supplemental material (9 pages)
- Model source code file (createSim3.r; the software is written in R)
- An Excel based program that will allow for running the model averaging software with user selected priors
- Charge to reviewers (EPA will provide specific questions for the reviewers to address)

The focus of the review should be on the model report and reviewing the statistical methods used to calculate the model averaged BMDs and BMDLs. The review charges do not require a reviewer to run the programs or evaluate the software source code. However, because some

reviewers may want to run the programs and evaluate the source code, EPA will provide executable files and source code files along with the review materials.

Note: The contractor and the reviewers must treat these materials as confidential. Reviewers must not retain copies of the materials or share them with others. Reviewers must return the materials promptly to the contractor after their work is completed, or affirm that the materials have been destroyed. The contractor shall take reasonable measures (including securing advance agreement by reviewers) to insure that these materials are not shared, released, made accessible to persons (including Contractor staff) not directly responsible for conducting the review, or made publicly accessible.

Task 7. Conduct Telephone Conference

The contractor will arrange for a telephone conference of the reviewers, selected EPA staff, and staff from the EPA contractor which developed the software. The conference will occur approximately 1-2 weeks after reviewers receive their review materials. Three hours maximum (with scheduled breaks) should be allowed for the conference, though it may well end in 1-2 hours. Reviewers are expected to have become familiar with the materials by the time of the teleconference. The conference will provide an opportunity for reviewers to ask questions and get clarifications from EPA and EPA's contractor about the materials and the models. It will also provide an opportunity for reviewers to exchange ideas and to benefit from one another's insights.

The contractor will take notes of the discussions and provide a written summary to EPA within 5 working days after the conference.

Task 8. Deliver the Letter Reviews to EPA

The contractor shall transmit electronic copies of the letter reviews to the TPO by email within 3 working days of receipt by the contractor. The contractor shall deliver the individual letter reviews in legible and printable electronic format(s). Acceptable formats include PDF files (preferred), MS/Word documents, and Excel spreadsheets (because software supporting these formats is currently provided to EPA staff and is thus available to the TPO). Other document formats should be converted, or else printed legibly and then scanned into a PDF file. Original copy submitted by reviewers (electronic or paper copy) will be conveyed to the EPA TPO and will become the permanent property of EPA. The contractor shall insure that reviewers' electronic documents have been purged of hidden data that reviewers did not intend to submit.

IV. Schedule of Deliverables

Copies of all deliverables shall be sent to the PO and the TPO in an electronic format, preferably PDF (PDF, MSWord, and Excel formats are acceptable).

Schedule of Deliverables Elapsed times are in normally scheduled working days, unless otherwise noted	
Task 1. Conference Calls with EPA TPO	
Submit cost proposal	Within 2 weeks after receipt of PWS
Task 2. Conference Calls with EPA TPO	
Conference call with the TPO, others as appropriate Weekly updates on progress for TPO	Within 1 week of TO award weekly, TBA
Task 3. COI Analysis and Certification	concurrent with Task 4
Task 4. Identify and Screen Proposed Reviewers	
provide EPA with list of at least 5 proposed reviewers & their qualifications	Within 2 weeks of TO award
EPA-Contractor teleconference to discuss proposed reviewers	Within 1 week of sending list of proposed reviewers
Task 5. Select External Reviewers	
Select and secure reviewers	1 week after agreement on proposed reviewers
Send EPA certification *per* Section V and Attachment 1 of this PWS	1 week after agreement on proposed reviewers
Task 6. Begin Peer Review	
Engage reviewers and send charge and review materials	Within 2 weeks of agreement on proposed reviewers
Task 7. Conduct Telephone Conference	
Arrange and conduct telephone conference; provide summary notes to EPA	To be determined, but not more than 3 weeks after sending materials for review
Task 8. Deliver the Letter Reviews to EPA	
Reviewers' written comments will be due no later than 5 weeks* after the contractor sends review materials to the reviewers. Transmit letter reviews to EPA within 3 days of receipt. (*subject to change by written technical direction).	To be determined, but not more than 5 weeks after sending materials for review

V. Conflict 0f Interest

The contractor shall warrant that, to the best of the contractor's knowledge and belief, that there are no relevant facts or circumstances which could give rise to a conflict of interest, as defined in FAR subpart 9.5, or that the contractor has disclosed all such relevant information.

The contractor agrees to notify the EPA Project Officer immediately, that to the best of its knowledge and belief, no actual or potential conflict of interest exists or to identify to the EPA Project Officer any actual or potential conflict of interest the contractor may have.

The contractor agrees that if an actual or potential conflict of interest is identified during the performance, the contractor will immediately make a full disclosure in writing to the Project Officer. This disclosure shall include a description of actions, which the contractor has taken or proposes to take, after consulting with the EPA Project Officer, to avoid, mitigate, or neutralize

the actual or potential conflict of interest. The contractor shall continue performance until notified by the EPA Project Officer of any contrary action to be taken.

1. Certifying and describing analysis and conclusion

The contractor shall provide the EPA PO and TPO written certification, within the time specified in the Task #5 of the PWS after an award, that:

a. The contractor has resolved all conflict of interest issues, either by eliminating a particular reviewer from the panel or by determining that the interest will not impair the individual's objectivity nor create an unfair competitive advantage for any person or organization.
b. The contractor recognizes its continuing obligation to identify and report any conflicts of interest arising during performance of the peer review.
c. All personnel who perform work under this task order or relating to this project have been informed of their obligation to report any conflict of interest to the contractor who shall, in turn, report to the EPA PO.

2. Ongoing Compliance Review during contract performance

a. The contractor shall require advanced notification from panelist concerning changes to information disclosed under Task #2.
b. The contractor shall inform the PO and TPO of any change in financial or professional relationships that may create either an actual or potential conflict of interest or bias during the period of performance.
c. The contractor shall consult the CO and PO concerning available options in cases where actual or potential conflict of interest is determined.

3. Disclosure of Information Used in Conflict of Interest Evaluation

The financial and professional information obtained by the contractor as part of the evaluation to determine existence of actual or potential conflict of interest is considered private and non-disclosable to EPA or outside entities, except as required by law or requested as part of a formal investigation by the EPA Office of Inspector General, General Accountability Office, or Congressional Committee.

VI. Special Conditions

Travel

No travel is anticipated under this task order.

Management Controls

- The Contractor shall certify there is no conflict of interest. The contractor shall provide the following conflict of interest certification in the proposal:

> I certify that, to the best of my knowledge and belief, no actual, apparent, or potential organizational or individual conflicts of interest related to this work assignment exist. Personnel who perform work under this work assignment, or relating to the work assignment, have been informed of their obligation to report personal and organizational interests. All actual, apparent or potential organizational or individual conflicts of interest related to this work assignment have been reported to the contracting officer or are attached, if applicable.

- The contractor shall be responsible for obtaining a conflict of interest certification for subcontractors performing peer review services.

Notice Regarding Guidance Provided Under this Task Order:

Guidance is strictly limited to technical and analytical support. The contractor shall not engage in activities of an inherent governmental nature such as the following:

(1) Formulation of Agency Policy

(2) Selection of Agency priorities

(3) Development of Agency regulations

Should the contractor receive any instruction from an EPA staff person that the contractor ascertains to fall into any of these categories or goes beyond the scope of the contract of work assignment, the contractor shall immediately contact the PO or TPO.

The contractor shall also ensure that work under this task order does not contain any apparent or real personal or organizational conflict of interest. The contractor shall certify that none exist at the time the proposal is submitted to the EPA.

VIII. EPA CONTACT INFORMATION

Copies of all correspondence pertaining to the performance of this Task Order shall be sent to the PO and TOPO.

Task Order Project Officer (TOPO):

Jeffrey S. Gift, Ph.D.
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment
Telephone: 919-541-4828
Fax: 919-541-0245
e-mail: gift.jeff@epa.gov

Mailing Address:
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment
(MC B-243-01)
Research Triangle Park, NC 27711

Alternate TOPO:

Allen Davis
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment
Office Phone: 513-569-7024
AWL Phone: 615-730-6693
Cell: 205-422-0655
e-mail: davis.allen@epa.gov

Mailing Address:
U.S. EPA, Office of Research and Development,
National Center for Environmental Assessment
26 West Martin Luther King Drive
Cincinnati, OH 45268, MC A110

Charge to External Peer Reviewers

Subject: External Review (Letter) of EPA/NCEA Dichotomous Dose-Response Model Averaging using Bayesian Maximum a-posteriori Estimation

The primary purpose of this review is to obtain external expert opinions on EPA's proposed model averaging approach for the derivation of BMD and BMDL values from dichotomous dose-response datasets. In addition to responding to the specific charge questions listed below, the contractor shall encourage reviewers to offer suggested improvements or enhancements to the software in response to the final charge question, "Additional Recommendations." However, the contractor shall convey to the reviewers that additional recommendations are not a requirement of this review and will not necessarily be addressed in the initial version of the software.

1. Are the documentation materials describing the proposed method, including the description of the principles and advantages of BMA in general, accurate and clear?
2. Are the methods described adequate for the derivation of BMDLs that are reasonable for use as points of departure for use in EPA risk assessments? In particular, with respect to:
    (a) Use of approximations such as profile posterior density (PPD, a Bayesian analogue of profile likelihood) for model-specific posterior BMD distributions, and Laplace approximation for integrated likelihood (marginal density of the data).
    (b) The possibility of having more parameters than dose groups in a given application, for a single model. The EPA states in the report provided that incorporation of prior information for model parameters allows application to data with fewer dose groups than parameters.[1] In current dose-response modeling practice, EPA does not use a model when the number of parameters exceeds the number of dose groups.
        i. Do the reviewers agree with EPA that the proposed BMA methodology is reasonable to use, when the number of parameters exceeds the number of dose groups for individual models?
        ii. Is additional research suggested, e.g., for some cases that may be problematic?
        iii. Related to (i), the total number of parameters combined for all models is expected to frequently exceed the number of dose groups (typically 3-5). The EPA has concluded that if the approach proposed for individual models is reasonable, so are the BMA results. Will the large number of fitted parameters result in "overfitting"?2
    (c) The approach for the derivation of BMD point estimators.3 The currently-proposed estimator is the weighted average of the maximum a-posteriori point estimates from individual models, weighting by model posterior weights. Do the reviewers suggest one or more alternative or additional BMD point estimators for the model averaging context?

---

[1] The EPA is aware of general Bayesian literature which suggests that informative priors can address identifiability issues.

[2] This term is sometimes used to indicate that a model is very flexible resulting in a relatively complicated fit with features that may not be repeatable.

[3] Possible uses of BMD point estimates include comparative and meta-analyses, common use of the ratio BMD [point estimate]/ BMDL ratio as in indication of the quality of the model results.

(d) The proposed default model parameter priors defined in the draft manuscript (described as "Prior 1" in Appendix 3 simulation results). The "Diffuse Hill" condition described in Appendix 3 of the draft manuscript support material uses the same prior as the proposed approach, except the hill model's prior is more diffuse. Can you comment on if this prior is preferable to the proposed "Prior 1" set of priors?

(e) The use of equal model weights (described as "Even" model weighting in Appendix 3 simulation results). As discussed in Section 4 of the draft manuscript, in an effort to account for problematic conditions in the literature, we increased the quantal linear default weighting ("MAQ approach"; described as "QL = 0.5" model weighting in Appendix 3 simulation results), and this resulted in better results with little evidence of deleterious performance for other models. Can you comment on if this prior weighting should be used in place of equal weights?

3. Do you agree with the particular models selected for BMA or do you recommend a different set of models?

4. Was adequate testing of the methods performed? In particular,

   (a) What additional steps, if any, are recommended to build confidence in the profile posterior density and Laplace approximations? Are any special situations evident, where these approximations may work particularly poorly?

   (b) Is the Monte Carlo testing approach taken for the sensitivity analyses an appropriate tool for evaluating the method?

   (c) Are the "templates" adequate or is additional testing recommended in order to evaluate other aspects of study design such as numbers per dose group, or dose spacing.

5. What output other than the BMD, BMDL, posterior weights and plots, would be necessary to provide enough information to users for the purposes of quantitative risk assessment? What fit statistics would be necessary to assess model/method performance/fit (e.g., global goodness-of-fit p-values, scaled residuals, posterior p-value)?

6. Is the USEPA proposal to implement this methodology with default priors reasonable in practice, given the likely user of BMDS who have limited familiarity with Bayesian methods? If yes,

   (a) How does this methodology compare with current methods, with regard to likelihood that non-statisticians will use it appropriately and accurately?

   (b) What situations may be envisioned where default priors would be over-ridden, and what measures would help to make sure this is done appropriately (if it is needed)?

7. Additional Recommendations: Are there any additional aspects of software development and testing, or model documentation, or reporting of model results that give you special cause for concern? If so, please describe your concerns and recommendations.

ATTACHMENT 1

Reviewer Certification

Please sign below to certify (1) that you have fully and to the best of your ability completed this disclosure form, (2) that you will update your disclosure form promptly by contacting the IRIS peer review manager if relevant circumstances change, (3) that you are not currently arranging new professional relationships with, or obtaining new financial holdings in, an entity (related to the subject of this review) which is not yet reported, and (4) that the certification below, based on information you have provided, and your CV may be made public for review and comment (see note * below).

Signature _____ Date _____

(Print name)_____

(***The financial and professional information obtained by the contractor as part of the evaluation to determine existence of actual or potential conflict of interest is considered private and non-disclosable to EPA or outside entities, except as required by law or requested as part of a formal investigation by the EPA Office of Inspector General, General Accountability Office, or Congressional Committee.**)

Contractor Certification

The contractor has reviewed the information provided on the Conflict of Interest Disclosure form for by _____ and certifies that to the best of the contractor's knowledge and belief, there are no relevant facts or circumstances which could give rise to a conflict of interest, as defined in FAR subpart 9.5, or that the contractor has disclosed all such relevant information.

Disclosure (if applicable):

_____

_____

Signature _____

Date _____

Contractor's Senior Peer Review Manager

ATTACHMENT 2

Conflict of Interest Analysis and Certification – Questions and Supporting Information

a. To the best of your knowledge and belief, is there any connection between the matter under review and any of your and/or your spouse's compensated or uncompensated employment, including government service, during the past 24 months? Yes _____ No _____
b. To the best of your knowledge and belief, is there any connection between the matter under review and any of your and/or your spouse's research support and project funding, including from any government, during the past 24 months? Yes _____ No _____
c. To the best of your knowledge and belief, is there any connection between the matter under review and any consulting by you and/or your spouse, during the past 24 months? Yes _____ No _____
d. To the best of your knowledge and belief, is there any connection between the matter under review and any expert witness activity by you and/or your spouse, during the past 24 months? Yes _____ No _____
e. To the best of your knowledge and belief, have you, your spouse, or dependent child, held in the past 24 months, any financial holdings (excluding well-diversified mutual funds and holdings, with a value less than $15,000) with any connection to the matter under review? Yes _____ No _____
f. Have you made any public statements or taken positions on or closely related to the matter under review under review? Yes _____ No _____
g. Have you had previous involvement with the development of the review materials you have been asked to review? Yes _____ No _____
h. To the best of your knowledge and belief, is there any other information that might reasonably raise a question about an actual or potential personal conflict of interest or bias? Yes _____ No _____
i. To the best of your knowledge and belief, is there any financial benefit that might be gained by you or your spouse as a result of the outcome of this review? Yes _____ No _____
j. Compensated <u>and non-compensated employment</u> (for panel member and spouse): list sources of compensated and uncompensated employment, including government service, for the preceding two years, including a brief description of work.
k. <u>Research Funding</u> (for panel member): list sources of research support and project funding, including from any government, for the preceding two years for which the panel member served as the Principal Investigator, Significant Collaborator, Project Manager or Director. For panel member's spouse, provide a general description of research and project activities in the preceding two years.
l. Consulting (for panel member): compensated consulting activities during the preceding two years, including names of clients if compensation provided 15% or more of annual compensation. For panel member's spouse, provide a general description of consulting activities for the preceding two years.
m. <u>Expert witness activities</u> (for panel member): list sources of compensated expert witness activities and a brief description of each issue and testimony. For panel member's spouse, provide a general description of expert testimony provided in the preceding 2 years.
n. <u>Assets: Stocks, Bonds, Real Estate, Business, Patents, Trademarks, and Royalties</u> (for panel member, spouse and dependent children): specific financial holdings that collectively had a fair market value greater than $15,000 at any time during the preceding 24-month period

(excluding well-diversified mutual funds, money market funds, treasury bonds and personal residence).

o. Liabilities (for panel member, spouse and dependent children): liabilities over $10,000 owed at any time in the preceding twelve months (excluding a mortgage on personal residence, home equity loans, automobile and consumer loans).
p. Public Statements: A brief description of public statement and/or positions on or closely related to the matter under review by the panel member.
q. Involvement with document under review: A brief description of any previous involvement of the panel member in the development of the document (or review materials) the individual has been asked to review.
r. Other potentially relevant information: A brief description of any other information that might reasonably raise a question about actual or potential personal conflict of interest or bias.

References

1. Raftery, A.E., Madigan, D. and Hoeting, J.A., 1997. Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*(437), pp.179-191.
2. Buckland, S.T., Burnham, K.P. and Augustin, N.H., 1997. Model selection: an integral part of inference. *Biometrics*, pp.603-618.
3. Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T., 1999. Bayesian model averaging: a tutorial. *Statistical science*, pp.382-401.
4. Claeskens, G. and Hjort, N.L., 2008. *Model selection and model averaging* (Vol. 330). Cambridge: Cambridge University Press.
5. Bailer, A.J., Noble, R.B. and Wheeler, M.W., 2005. Model uncertainty and risk estimation for experimental studies of quantal responses. *Risk Analysis*, *25*(2), pp.291-299.
6. Wheeler, M.W. and Bailer, A.J., 2007. Properties of model-averaged BMDLs: a study of model averaging in dichotomous response risk estimation. *Risk Analysis*, *27*(3), pp.659-670.
7. Wheeler, M.W. and Bailer, A.J., 2009. Comparing model averaging with other model selection strategies for benchmark dose estimation. *Environmental and Ecological Statistics*, *16*(1), pp.37-51.
8. Shao, K. and Gift, J.S., 2014. Model uncertainty and Bayesian model averaged benchmark dose estimation for continuous data. *Risk Analysis*, *34*(1), pp.101-120.
9. Simmons, S.J., Chen, C., Li, X., Wang, Y., Piegorsch, W.W., Fang, Q., Hu, B. and Dunn, G.E., 2015. Bayesian model averaging for benchmark dose estimation. *Environmental and Ecological Statistics*, *22*(1), pp.5-16.
10. Wheeler, M.W. and Bailer, A.J., 2013. An empirical comparison of low-dose extrapolation from points of departure (PoD) compared to extrapolations based upon methods that account for model uncertainty. *Regulatory Toxicology and Pharmacology*, *67*(1), pp.75-82.
11. US EPA, 2012, Benchmark Dose Technical Guidance Document, US EPA Research Triangle Park NC.
12. Faes, C., Aerts, M., Geys, H. and Molenberghs, G., 2007. Model averaging using fractional polynomials to estimate a safe level of exposure. *Risk Analysis*, *27*(1), pp.111-123.
13. Wheeler, M. and Bailer, A.J., 2012. Monotonic Bayesian semiparametric benchmark dose analysis. *Risk Analysis*, *32*(7), pp.1207-1218.
14. Greig, D.M., Porteous, B.T. and Seheult, A.H., 1989. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp.271-279.
15. Fletcher, D. and Turek, D., 2012. Model-averaged profile likelihood intervals. *Journal of agricultural, biological, and environmental statistics*, *17*(1), pp.38-51.
16. Efron, B. and Tibshirani, R.J., 1994. *An introduction to the bootstrap*. CRC press.
17. Kan's paper
18. Guha N., Roy, A., Kopylev, L., Fox J., Spassova, M. and White, P., 2013. Nonparametric Bayesian methods for benchmark dose estimation. *Risk Analysis*, *33*(9), pp. 1608-1619.