Statistica Sinica **20** (2010), 281-301

PESSIMISTIC MODELING IN THE BENCHMARK DOSE METHOD WITH APPLICATION TO BROMATE IN DRINKING WATER

Mendel Fygenson

University of Southern California

Abstract: The problem of low-dose extrapolation is common in risk evaluation of carcinogens. To set safety standards, regulatory agencies often employ the benchmark dose (BMD) method with a default model of low-dose-linearity. They claim this approach is inherently *conservative*, leading to dose levels that are considered protective of the publics health. These dose levels have been historically referred to as Virtually Safe Doses (VSD) and they correspond to doses for which the *upper* bound on the projected lifetime incremental risk is, for example, 1 in 1,000,000. However, for carcinogens that are directly or indirectly beneficial, these VSD may be unpractical and/or excessively protective of the public's health.

This paper extends the framework in Fygenson (2008) to address the question of just how conservative is the current BMD method and provides, for the first time, a *lower* bound on the projected lifetime incremental risk from the so called VSD. The proposed lower bound complements the upper bound derived by the BMD method and can lead to more productive risk/benefit analyses.

Key words and phrases: Benchmark dose, low-dose extrapolation, model uncertainties, pessimistic distributions.

1. Introduction

Establishing safe standards for carcinogenic substances that have direct or indirect benefits is a major challenge for regulatory agencies. It requires coping with scientific uncertainties and with the concerns of parties with conflicting interests. Bromate is one example of a substance that illustrates the associated difficulties. Bromate is a byproduct formed during the water disinfection process of ozonization. It can also be found in some groundwater as a result of industrial or agricultural runoff. The practical and public health benefits of ozonization are confounded by the observation that high doses of bromate cause cancer tumors in laboratory animals. In 2001, the United States Environmental Protection Agency (EPA) compiled a toxicological review of bromate. The review acknowledged limitations due to "knowledge gaps [in the biological mechanisms], uncertainties, quality of data [lack of human data from experimental or observational studies], and scientific controversies" (U. S. EPA, 2001).

In general, the lack of human data on the effects of long-term exposure to very low doses of carcinogen forces regulatory agencies to base their risk evaluations on bioassay data. In bioassays, animals are exposed to doses much higher than those humans are likely to be exposed to, and for much shorter time intervals. Thus, reliance on bioassay data poses two fundamental problems. One is the problem of species conversion: results from test animals need to be converted into implications for people. Another is the problem of low-dose extrapolation (within a species): effects of very low doses must be extrapolated from the much higher dose levels used in the bioassay. This paper is concerned with this second problem, which requires statistical modeling of the dose-response in the bioassay data to anchor the low-dose extrapolations.

1.1. The low-dose extrapolation problem

Throughout the paper we consider the analysis of *quantal* bioassay data where animal groups share the *same* exposure time to the substance under evaluation. Typically in these experiments a total of n animals are randomly assigned to one of $k+1$ dose levels d_i, where $0 = d_0 < d_1 < \ldots < d_k$ and d_0 denotes the control level. At the end of an experiment, the proportion, Y_i/N_i, of animals with the adverse response at every dose level d_i is recorded. Since each animal in the experiment represents an independent Bernoulli trial, the dose-response relationship is captured by the probability for an adverse outcome conditioned on dose d (i.e., $P(d) = P(Y = 1|d)$).

For regulatory purposes, the interest is in the extrapolated probabilities of an adverse outcome from exposure to very low-doses (i.e., $P(d^*), 0 < d^* << d_1$). Specifically, the aim is to identify the dose(s) responsible for an acceptably small increase in the likely number of additional cases in the population. For most carcinogens, the incremental risk acceptable to regulatory agencies ranges from 10^{-4} to 10^{-6}. They therefore define a "virtually" safe dose (VSD) as the dose for which the *upper* bound on the projected lifetime incremental risk is between 1 in 10,000 to 1 in 1,000,000.

For some carcinogens, the difficulty of striking a balance between risks and benefits, combined with scientific uncertainties, elevates the importance of the method used in the analysis. For the last 20 years, the benchmark dose (BMD) method (Crump (1984)) has been the method of choice (e.g., U. S. EPA, 2005).

1.2. The BMD method

Several years ago the EPA greatly increased the accessibility of the BMD method by making available a user-friendly software package, called BMDS. (The software can be downloaded from `http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm`. For an evaluation of the software, see Filipsson and Victorin (2003) or Filipsson, Sand, Nilsson and Victorin (2003).) To outline the

Table 1. Parametric Dose Response Models.

Model	Formulation	Parameters
Quantal-linear	$P(d) = \gamma + (1-\gamma)(1-e^{-\beta d})$	$0 \le \gamma < 1$, $\beta \ge 0$
Logistic	$P(d) = 1 \Big/ (1+e^{-(\alpha+\beta d)})$	$\beta \ge 0$
Log-logistic	$P(d) = \begin{cases} \gamma + \frac{(1-\gamma)}{1+e^{-(\alpha+\beta \ln d)}} & d > 0 \\ \gamma & d = 0 \end{cases}$	$0 \le \gamma < 1$, $\beta \ge 0$
Probit	$P(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha+\beta d} e^{-x^2/2} dx$	$\beta \ge 0$
Log-probit	$P(d) = \begin{cases} \gamma + \frac{(1-\gamma)}{\sqrt{2\pi}} \int_{-\infty}^{\alpha+\beta \ln d} e^{-x^2/2} dx & d > 0 \\ \gamma & d = 0 \end{cases}$	$0 \le \gamma < 1$, $\beta \ge 0$
Weibull	$P(d) = \gamma + (1-\gamma)(1-e^{\beta d^{\delta}})$	$0 \le \gamma < 1$, $\beta \ge 0$, $\delta > 0$
Gamma	$P(d) = \gamma + (1-\gamma)\frac{1}{\Gamma(\delta)} \int_0^{\beta d} x^{\delta-1} e^{-x} dx$	$0 \le \gamma < 1$, $\beta \ge 0$
Linearized MultiStage	$P(d) = \gamma + (1-\gamma)(1-e^{-\sum_{j=1}^{n} \beta_j d^j})$	$0 \le \gamma < 1$, $\beta \ge 0$, n is an integer

BMD method, we retrace a practitioner's three likely steps when using the BMDS program.

The first step is to fit a parametric model to the bioassay data via the maximum likelihood method. The parametric dichotomous dose-response models available on BMDS are listed in Table 1. With the exception of two (logistic and normal), all of these models are based on non-negative distributions through the general formulation

$$P(d) = \gamma + (1-\gamma)F_1(d), \quad d \ge 0, \tag{1.1}$$

where γ is the background risk in the control group, and F_1 is the cdf of tolerances in the animal population.

The second step is to use the fitted curve to get a point estimate of the benchmark dose (BMD), defined as the dose that causes a pre-specified extra risk (referred to as the benchmark risk, BMR) above the risk faced by the control group. The BMR corresponding to a BMD is uniquely defined for a dichotomous response as

$$\text{BMR} = \frac{[P(\text{BMD}) - P(0)]}{(1 - P(0))}. \tag{1.2}$$

Note that for models in (1.1), BMR= F_1(BMD), and BMD is just the BMR-percentile of F_1. To complete the second step, a lower confidence bound for the BMD (referred to as the BMDL) is derived using the asymptotic distribution of the likelihood ratio.

Often, the final step (common in carcinogen evaluations) is to draw a straight line between the BMDL and the origin (zero dose, zero extra-risk). This linear default model is used as an upper bound on the risk from very low-doses that can be deemed as safe (e.g., upper bound for $P(\text{VSD})$).

When using BMDS, risk-assessors must make at least two important input decisions. They must select (*i*) a model from Table 1, and (*ii*) one or more BMR value(s) for which to estimate the BMD and the BMDL.

To aid the user in selecting a parametric model, the BMDS provides a variety of goodness-of-fit tools that include likelihood ratio tests, the Pearson chi-square test for dichotomous data, and the Akaike information criterion (AIC). As a guide in selecting the BMR, the research literature recommends values from 1 to 10% for *quantal* data, with the most common being 1, 5 and 10% (e.g., Gaylor (1989) and Faustman (1996)). According to the EPA, "an excess risk of 10% has generally been the default BMR for quantal data," but "if a study has greater than usual sensitivity, then a lower BMR can be used" (U. S. EPA draft technical guidance document (2000). A method for deriving simultaneous statistical inferences for a number of different BMR is given by Piegorsch, West, Pan and Kodell (2005).

1.3. Statistical and practical challenges

Although the statistical tools of the BMDS are easy to use, and a risk-assessor can justify choosing any of the recommended BMR values, in the big picture of statistical methodology these actions raise some challenging issues common to risk analyses in economics and statistics (e.g., Ryan (2003)). Two fundamental statistical challenges arise from methodological issues: how to account for model uncertainty and how to model extrapolated probabilities in worst-case analyses. The major practical challenge is to objectively assess whether resulting policies are insufficiently or excessively protective of the public's health.

Statistical Challenges. The issue of model uncertainty is especially pronounced and challenging in low-dose extrapolation problems because often different models (F_1) fit the data equally well, yet yield significantly different predictions outside the data range, especially for BMR below 5% (e.g., Sand, Filipsson and Victorin (2002) and Morales, Ibrahim, Chen and Ryan (2006)). Model uncertainty in F_1 can be dealt with, for example, by the Bayesian model averaging approach (e.g., Morales et al. (2006)). or its frequentist version (e.g., Buckland, Burnham and Augustin (1997)). Another possibility is to use the parametric model (from Table 1) that provides the lowest BMDL (i.e., the most conservative model).

However, in low-dose extrapolation problems, model uncertainty in F_1 is not the whole story. In particular, it is not correct to presume that the same model applies both inside and outside the range of positive doses in the bioassay.

Accordingly, it was proposed in Fygenson (2008, p.11-12) that it would be more appropriate to replace (1.1) with a more general model. The appropriate general model for the current setting becomes

$$P(d) = \gamma + (1-\gamma)F(d)I(0 < d < d_1) + (1-\gamma)F_1(d)I(d_1 \le d \le d_k). \qquad (1.3)$$

Here two very different modeling problems are made explicit, highlighting two different types of model uncertainty: uncertainty with respect to F_1 and uncertainty with respect to F – an unknown *non-negative* distribution. For the latter, modeling should not be data driven and, therefore, a different approach is required. This paper extends the approach in Fygenson (2008) to cover *non-negative* distributions and *left-tail* extrapolations. The extended framework provides risk-assessors the opportunity to incorporate their outlooks (i.e., pessimism or optimism) via constrained non-parametric models for F.

Practical Challenges. Most practitioners agree that the BMD method leads to standards that are protective, but at the cost of making "$\cdots$ a number of assumptions, the key one being linearity of the dose response to the origin. In addition, most authorities take a *precautionary* approach in using the upper bound [BMDL] value on the dose for a specified risk." (Fawell and Walker (2006), emphasis added). Consequently, the method is rarely challenged when the dose levels to which humans are presently exposed fall below the VSD. However, when the VSD requires changing the status quo, the question of whether the method leads to overly protective standards is often echoed.

To addresses this question, it would be useful to have an informative lower bound for P(VSD). By estimating the minimum number of extra cases in the population arising from exposure to the VSD, a lower bound would be an important complement to the upper bound on the projected lifetime incremental risk.

The idea for deriving such a lower bound appeared first in Cornfield (1977) but to the best of my knowledge, no such bound has been proposed in the literature. This paper contributes a lower bound on the P(VSD) based on constrained non-parametric models introduced in response to the above statistical and practical challenges.

1.4. Example: incidence of renal cancer upon chronic exposure to bromate

In its evaluation of the carcinogenicity of chronic exposure to bromate, the EPA cites the rodent bioassay studies of Kurokawa, Aoki, Matsushima, Takamura, Imazawa and Hayashi (1986) and DeAngelo, George, Kilburn, Moore and Wolf (1998). Both studies were conducted using an appropriate route of exposure and adequate numbers of animals. However, they report different rates of tumor

incidence and, as a result, the nature of the dose-response is not well defined. The EPA ultimately characterized the cancer risk from bromate using data from DeAngelo et al. (1998) because it used more animals per group than Kurokawa et al. (1986). The EPA noted that "the hazard characterization of bromate suggests that the dose-response assessment should apply a linear extrapolation from data in the observable range to the low-dose region because of the lack of understanding of bromates mode of action and the positive mutagenicity data" (U. S. EPA, 2001). The data from both studies is analyzed via the proposed methodology in Section 3.

1.5. Structure of the paper

The main results of this paper are presented in Section 2 and applied to the analysis of the two rodent bioassay studies in Section 3. Section 4 contains closing remarks and highlights some of the issues needing further attention. The proofs are relegated to an Appendix.

2. Pessimistic Models in Bmds and Associated Confidence Bounds

The EPA recommends the BMD method (often with the linear default model) as "inherently *conservative* of public health, without addition of another factor for human variability" (U. S. EPA 1996, emphasis added). Given the existence of scientific uncertainties and the high value we place on human life, the regulatory agencies are justified in being extra protective of the public's health. However, the word conservative can mean different things and its clarification may have important policy implications.

The use of the BMD method to set standards is part of a decision process. It is therefore natural to view this method within the general framework of modern economic decision models. Within this framework there are only two routes for incorporating cautiousness (i.e., conservativeness) in a decision process: one is through the use of a risk-averse utility function (of the possible outcomes) and the other is through the use of a pessimistic probability model for the possible outcomes (see the discussion in Fygenson (2008)). From the formulation of the low-dose extrapolation problem (i.e., Models (1.1) or (1.3)), it is clear that only the second route (whereby pessimistic models are used outside the data range) applies.

2.1. Two pessimistic (optimistic) families of distributions

By construction (namely, the use of cdf in Models (1.1) or (1.3)), the risk of an adverse outcome is non-decreasing with increasing exposure levels. Therefore constraining the risk to be monotone cannot, in itself, be considered pessimistic. A distribution can be classified as pessimistic, neutral or optimistic based on

monotone patterns in its *extra-risk* mechanism (Fygenson (2008)). For the current discussion it suffices to consider two classifications that are based on the commonly used measures of association known as the *Attributable-Risk* (AR) and the *Odds-Ratio* (OR). These measures describe extra-risk mechanisms and can be defined as functionals of a cdf (F):

$$\begin{aligned} &\text{(i)} \quad AR_F(d^*, d) = F(d^*) - F(d), \\ &\text{(ii)} \quad OR_F(d^*, d) = \frac{[F(d^*)/(1 - F(d^*))]}{[F(d)/(1 - F(d))]}, \end{aligned} \tag{2.1}$$

where d^* and d are any two fixed values such that $d < d^*$, $F(d) > 0$ and $F(d^*) < 1$.

The functionals in (2.1) are particularly appropriate for characterizing a pessimistic outlook in low-dose extrapolation problems because they capture both the *strength* and the *direction* of the relationship between the response (Y) and the exposure levels (D), without requiring explicit knowledge of the formula of F. To apply them, however, one must choose the scale for comparing d^* and d. Fygenson (2008) demonstrates that the *additive* scale (i.e., $d^* = d + \Delta, \Delta > 0$) is most productive in characterizing outlooks of continuous distributions with support on $(-\infty, \infty)$. To extend these results to the class of non-negative continuous distributions (e.g., Table 1), I demonstrate next the importance of using the multiplicative scale (i.e., $d^* = \alpha d, \alpha > 1$). The definition below defines possible outlooks of a distribution with respect to an interval J. For the low-dose extrapolation problem, the interval of interest is the left tail starting at or below the 10^{th} percentile (i.e., the BMD).

Definition 2.1. A distribution function F is inherently *AR*- or *OR*-pessimistic (optimistic) on an interval J if $AR_F(d^*, d)$ or $OR_F(d^*, d)$ is non-decreasing (non-increasing) in d for all $d^* \in J$, respectively. F is said to be inherently AR- or OR-neutral on J if $AR_F(d^*, d)$ or $OR_F(d^*, d)$ are constant in d, for all $d^* \in J$, respectively.

For intuition to the definitions above the reader should consult Fygenson (2008). The main new idea here is the use of the multiplicative scale to achieve important results for the class of non-negative distributions. In what follows, AR-pessimism will be used to provide new insights to the BMD method and OR-pessimism will be used to derive *lower* bounds on the risk associated with the VSD.

When the different parametric models in Table 1 fit the bioassay data equally well, one can be conservative by choosing a pessimistic model. The following theorem provides the necessary classifications.

Theorem 2.1. *Among the non-negative distributions that are available in the BMDS (i.e., in Table 1):*

(i) *Log-normal is the only OR-optimistic distribution in the interval* [0, *Mode*);
(ii) *Log-logistic is the only OR-neutral distribution throughout its support;*
(iii) *Quantile-linear, Weibull, and MLS are OR-pessimistic throughout their support;*
(iv) *Gamma with shape parameter* $\lambda \leq 1$ *is OR-pessimistic throughout its support, while Gamma with* $\lambda > 1$ *is neither OR-pessimistic nor OR-optimistic.*

In summary, while goodness-of-fit tools often fail to discriminate among the distributions in Table 1, choosing one or another to estimate the BMD or the BMDL will produce (for the same data) a hierarchy of values according to the different outlooks inherent in these distributions. Theorem 2.1 provides a (theoretical) explanation for the often-mentioned empirical observation that the log-normals estimate of the VSD is the highest, usually followed by estimates from the log-logistic, gamma, Weibull, and LMS models (e.g., Krewski and Van Ryzin (1981) and Yanagimoto and Hoel (1990)).

2.2. Upper and lower bounds on the risk from low-doses

The notions of AR- and OR-pessimism (as applied here) are part of a nonparametric framework that can provide upper as well as lower bounds on F in (1.3). These bounds, given in the following theorem, are essential for determining the practical consequences of using the BMD method.

Theorem 2.2. *If* F *is AR-pessimistic on the interval* $J = [0, d_p]$*, then*

$$F(d) \leq F_1(d_p)\frac{d}{d_p}, \quad 0 < d \leq d_p; \tag{2.2}$$

If F *is OR-pessimistic on the interval* $J = [0, d_p]$*, then*

$$F(d) \geq (1 + e^{-A}d^{-B})^{-1}, \quad 0 < d \leq d_p < d_q, \tag{2.3}$$

where $d_p < d_q$ *denote a p-percentile and a q-percentile of* F_1*, respectively, and*

$$B = \log \frac{\frac{F_1(d_q)/(1-F_1(d_q))}{F_1(d_p)/(1-F_1(d_p))}}{\log d_q/d_p} \quad \text{and} \quad A = \log \frac{F_1(d_p)}{1 - F_1(d_p)} - B \log d_p. \tag{2.4}$$

Note that I purposely did not substitute p for $F_1(d_p)$ or q for $F_1(d_q)$ in the above theorem. This provides the flexibility of either choosing a value of p (e.g., 10%) and then estimating the corresponding percentile (the EPAs procedure), or choosing a value in the data range and then estimating its corresponding cumulative probability.

Table 2. Upper bounds on percentiles (ξ) with $p = 0.05$ and $q = 0.10$.

	Parameters	$\mathrm{UB}\xi_{10^{-4}}$ ($\xi_{10^{-4}}$)	$\mathrm{UB}\xi_{10^{-5}}$ ($\xi_{10^{-5}}$)	$\mathrm{UB}\xi_{10^{-6}}$ ($\xi_{10^{-6}}$)
Weibull	$\beta = 0.56$	$1.39 \cdot 10^{-3}$	$2.48 \cdot 10^{-4}$	$4.40 \cdot 10^{-5}$
	$\alpha = 1.28$	$(1.19 \cdot 10^{-3})$	$(1.98 \cdot 10^{-4})$	$(3.29 \cdot 10^{-5})$
LMS	$\alpha_1 = 0.380$			
	$\alpha_2 = 0.195$	$4.85 \cdot 10^{-4}$	$6.27 \cdot 10^{-5}$	$8.12 \cdot 10^{-6}$
	$\alpha_3 = 0.002$	$(2.63 \cdot 10^{-4})$	$(2.67 \cdot 10^{-5})$	$(2.63 \cdot 10^{-6})$
LMS_1	$\alpha_1 = 0.390$			
	$\alpha_2 = 0.195$	$4.56 \cdot 10^{-4}$	$5.81 \cdot 10^{-5}$	$7.40 \cdot 10^{-6}$
	$\alpha_3 = 0.000$	$(2.56 \cdot 10^{-4})$	$(2.36 \cdot 10^{-5})$	$(2.56 \cdot 10^{-6})$
LMS_2	$\alpha_1 = 0.460$			
	$\alpha_2 = 0.000$	$2.86 \cdot 10^{-4}$	$3.19 \cdot 10^{-5}$	$3.55 \cdot 10^{-6}$
	$\alpha_3 = 0.086$	$(2.20 \cdot 10^{-4})$	$(2.20 \cdot 10^{-5})$	$(2.20 \cdot 10^{-6})$

Note: distributions parameterized so that the log transformation of the corresponding random variables have a mean zero and standard deviation one.

The importance of imposing OR-pessimism on F when using the BMD method can be judged by the proximity of the lower bound to the true probabilities. Table 2 presents probabilistic lower bounds on the 10^{-4}, 10^{-5} and 10^{-6} percentiles of the standardized Weibull and a variety of standardized linearized multistage (LMS) models. The Weibull and LMS models are the two most pessimistic models among the distributions in Table 1 and are often used in analysis of bioassay data. Note that since the early 80's, and for the following 15 years, regulatory agencies often used the LMS as the *default* model in their risk assessments. For genotoxic substances, the World Health Organization (WHO, 2004) still uses the LMS as the default model with reference risk of 10^{-5} to establish safe doses (i.e., VSD).

Applying the inequalities in Theorem 2.2 to bioassay data requires accounting for the variability in the data. The common approach is to derive a confidence interval for the parameters of interest. The problem of constructing approximate confidence bounds for percentiles in a binary regression setup has been considered by many authors, usually under the standard assumption of a logit or a probit model for F_1 in (1.1). With the exception of one case (where F_1 is a logistic and F is OR-pessimistic), the derivation of the approximate confidence bounds is non-standard and requires a constrained maximum likelihood approach (see Section 3). While other asymptotic approaches exist (e.g., the delta method), the likelihood ratio method has been found to have good theoretical and practical properties. In particular, it is invariant under parameter transformations (Cox and Hinkley (1974)) and, with standard models, it yields coverage probabilities close to their nominal values (e.g., Alho and Valtonen (1995) and Huang (2001)).

Following the BMDS program, we use the likelihood ratio method to derive the approximate $100(1-\alpha)\%$ upper and lower confidence bounds for $F(d)$, $d \in J$.

Proposition 2.1. *If F is AR-pessimistic on the interval $J = [0, d_p]$, then the approximate $100(1-\alpha)\%$ upper confidence bound (UCB) for $F(d)$, $d \in J$ is*

$$U_d = \mathrm{Sup}_{(\boldsymbol{\theta})}\Big\{H(\boldsymbol{\theta}; d, p) : 2(LL(\hat{\boldsymbol{\theta}}) - LL(\boldsymbol{\theta})) \leq \chi^2_{1,1-2\alpha}\Big\}, \tag{2.5}$$

and the approximate $100(1-\alpha)\%$ lower confidence bound (LCB) for any percentile D_γ, $\gamma < p$ is

$$L_\gamma = \mathrm{Inf}_{(\theta)}\Big\{D_\gamma = M(\boldsymbol{\theta}; \gamma, d_p) : 2(LL(\hat{\boldsymbol{\theta}}) - LL(\boldsymbol{\theta})) \leq \chi^2_{1,1-2\alpha}\Big\}. \tag{2.6}$$

If F is OR-pessimistic on the interval $J = [0, d_q]$, then the approximate $100(1-\alpha)\%$ lower confidence bound (LCB) for $d \in J$ is

$$L_d = \mathrm{Inf}_{(\theta)}\Big\{G(\boldsymbol{\theta}; d, p, q) : 2(LL(\hat{\boldsymbol{\theta}}) - LL(\boldsymbol{\theta})) \leq \chi^2_{1,1-2\alpha}\Big\}. \tag{2.7}$$

In the proposition, $d_p = F_1^{-1}(p; \boldsymbol{\theta}) < d_q = F_1^{-1}(q; \boldsymbol{\theta})$ denote a p-percentile and a q-percentile of F_1, respectively. $LL(\boldsymbol{\theta})$ denotes the log likelihood function (of the parameters in F_1), $LL(\hat{\boldsymbol{\theta}})$ denotes its maximum under the parameterization of $(\boldsymbol{\theta})$. The functions in (2.5) through (2.7) are $H(\boldsymbol{\theta}; d, p) = (pd)/[F_1^{-1}(p; \boldsymbol{\theta})]$, $M(\boldsymbol{\theta}; \gamma, d_p) = (\gamma d_p)/[F_1(d_p; \boldsymbol{\theta})]$ and $G(\boldsymbol{\theta}; d, p, q) = 1/(1 + e^{-A(\boldsymbol{\theta})} d^{-B(\boldsymbol{\theta})})$, with $B(\boldsymbol{\theta}) = [\log(q/(1-q)) - \log(p/(1-p))]/[\log(F_1^{-1}(q; \boldsymbol{\theta})) - \log(F_1^{-1}(p; \boldsymbol{\theta}))]$ and $A(\boldsymbol{\theta}) = \log(p/(1-p)) - B(\boldsymbol{\theta}) \log(F_1^{-1}(p; \boldsymbol{\theta}))$.

In closing, note that that seeking the supremum to derive U_d in (2.5) has nothing to do with being pessimistic or conservative. It is merely the proper statistical procedure given that a confidence interval is required to estimate the *upper* bound in (2.2). Also, as with all asymptotic based methodologies, one cannot grantee the small sample performance. Moreover, because in most applications we do not know the true distribution, it is difficult to quantify the "true" relative accuracy of the above confidence bounds. It bears noting however that under (1.3) the magnitude of the confidence bounds depends, to a large extent, on the quality of the (probabilistic) bounds from Theorem 2.1- see Table 2. The extra width of the confidence bound (beyond the corresponding probabilistic bounds) is a function of the number of observations as well as how well they fit the assumed F_1 in (1.3), but does not depend on F in (1.3). In our framework the extrapolation is taking place in F, which is modeled non-parametrically.

2.3. The pessimistic side of the BMD method

The EPA and other regulatory agencies claim that the use of the linear default model with the BMD method is protective of the public's health. This claim is based on the assertion that the dose-response curves are likely to be linear for values in the neighborhood of zero and that $F(d)$ is likely to be sub-linear (i.e., convex) in $d \in (0, \mathrm{BMD}]$. Therefore, in all likelihood, the linear default model provides an upper bound on $F(d)$ for $d \in (0, \mathrm{BMD})$ or, equivalently, a lower bound for the safe low-doses of interest (e.g., VSD).

Based on our framework (see the Appendix), it can be shown that imposing an AR-pessimistic distribution on $F(d)$ for all $d \in (0, \mathrm{BMD}]$ is *equivalent* to assuming that $F(d)$ is sub-linear for all $d \in (0, \mathrm{BMD}]$. As a result, the default linear model recommended by the EPA and other agencies is *equal* to the upper bound on the risk given in (2.2). Thus, the BMD method accounts for model uncertainty in F by imposing the non-parametric constraint of AR-pessimism. How pessimistic is the EPA's approach? It turns out that all continuous unimodal distributions are, by definition, AR-pessimistic on $J \subset (-\infty, \mathrm{Mode}]$.

2.4. A lower bound on the excess-risk of the VSD

The BMD upper bound reflects the doses that are likely to cause maximum additional lifetime risk due to the exposure. Importantly, this does *not* exclude the possibility that the associated VSD can cause no additional risk, in which case it could be construed as overly protective (e.g., Cornfield (1977)). A reasonable response to such possibility is to provide a lower bound on the risk by imposing, once again, a pessimistic constraint on F.

Given that the low dose-response curve is to be modeled by a non-negative distribution, considering a model where F in (1.3) is OR-pessimistic is a reasonable approach since the use of OR-pessimism complements the AR-pessimism currently imposed by the BMD method. The latter constrains the *absolute* excess-risk mechanism to be non-decreasing on an interval, and the former imposes the same constraint but on the *relative* excess-risk mechanism.

Employing both the absolute and the relative constraints has the advantage of covering each mechanism's limitations. For example, an absolute difference in risk of 0.004 may be considered trivial if the risk to the control group is 0.4 (i.e., $P(0) = 0.4$). But, the same difference is substantial for a population with $P(0) = 0.00049$, as $P(d)$ is then more than 9 times larger than $P(0)$. Finally, since OR-pessimism neither implies nor is implied by AR-pessimism, the two provide different kinds of pessimism, but not a hierarchy of degrees of pessimism.

3. Application to Bromate Assay Data

In its evaluation of the carcinogenicity of chronic exposure to bromate, the regulatory agencies cite the rodent bioassay studies of Kurokawa et al. (1986) and

Table 3. Renal Tumor Incidences in Male Rats.

Kurokawa et al. (1986)							
Dose (mg BrO_3^-/kg · day	0	0.7	1.3	2.5	5.6	12.3	33
Tumor Incidence	0/19	0/19	0/20	1/24	5/24	5/20	9/20
DeAngelo et al. (1998)							
Dose (mg BrO_3^-/kg · day	0	1.1			6.1	12.9	28.7
Tumor Incidence	1/45	1/43			6/47	3/39	12/32

DeAngelo et al. (1998). The two studies report different rates of tumor incidence and leave the nature of the dose-response ill-defined. The U. S. EPA chose to characterize the cancer risk of bromate using data from DeAngelo et al. (1998) because that study used more animals per group than Kurokawa et al. (1986). In applying the BMD method, the EPA noted that "the hazard characterization of bromate suggests that the dose-response assessment should apply a linear extrapolation from data in the observable range to the low-dose region because of the lack of understanding of bromate's mode of action and the positive mutagenicity data" (U. S. EPA, 2001).

As an application of the methodology presented in Section 2, in what follows we analyze the renal tumor incidence data from both studies. Our objective is to estimate the VSD for reference risks of 10^{-4}, 10^{-5} and 10^{-6} and to provide statistical lower bounds on the risks from these VSD (i.e., $P(\text{VSD})$). To facilitate comparison between the two studies, we focus on data recorded after rodents were exposed for 100 weeks in the DeAngelo et al. (1998) study, and for 104 weeks in the Kurokawa et al. (1986) study (summarized in Table 3). The studies differ in two major respects important to the key issue of model uncertainty. The DeAngelo et al. (1998) study has more animals per dose level, while the Kurokawa et al. (1986) study extends to lower dose levels. Thus, *a priori*, the latter data set should result in less model uncertainty in the extrapolation range (i.e., with respect to F) and the former should provide a better account for model uncertainty in the data range (i.e., with respect to F_1).

In analyzing the data in Table 3, we follow the steps of a practitioner using the EPA's BMDS program (see Section 1.2). Table 4 provides goodness-of-fit measures for the three models from Table 1 that best fit the data from each study, along with corresponding BMD for 2.5%, 5% and 10% BMR. Notably, all models provide similar quality fits to their respective data sets and all BMD fall well within the data ranges. However, only for the Kurokawa et al. (1986) data are the BMD very similar for all three models. For the DeAngelo et al. (1998) data, BMD from the Quantal-linear model are much smaller than corresponding BMD from the two other models and therefore yield lower values for the VSD.

For ease of comparison, we adopt the Quantal-linear model for both data sets. This model is OR-pessimistic (see Theorem 2.1) and should yield the lowest

Table 4. Results for the three BMDS models that best fit bromate bioassay data.

Model	p-value		AIC	BMD		
	LR test	Pearson χ^2 test		10%	5%	2.5%
	Kuhn and Tucker (1951)					
Quant-lin	0.6996	0.7635	88.7265	4.8785	2.3750	1.1723
Log-probit	0.7805	0.8343	89.3698	4.4136	2.4718	1.4950
Log-logistic	0.6560	0.7464	90.1822	4.5227	2.2885	1.1881
	DeAngelo et al. (1998)					
Quant-lin	0.2610	0.3047	126.486	8.3579	4.0690	2.0084
Log-probit	0.2864	0.2458	126.261	13.4984	8.3580	4.8778
Log-logistic	0.2774	0.2322	126.339	14.5216	9.2437	5.4898

Table 5. VSD and Lower Bounds (LB) on P(VSD) using BMR = 2.5% and 5%.

P(VSD)		Kurokawa, et al. (1986)	DeAngelo, et al. (1998)
$\leq 10^{-4}$	VSD	0.0033016	0.0054606
	LB P(VSD)	$4.3296 \cdot 10^{-5}$	$4.4186 \cdot 10^{-5}$
$\leq 10^{-5}$	VSD	0.0003301	0.0005460
	LB P(VSD)	$4.1491 \cdot 10^{-6}$	$4.2344 \cdot 10^{-6}$
$\leq 10^{-6}$	VSD	0.0000330	0.0000546
	LB P(VSD)	$3.9760 \cdot 10^{-7}$	$4.0577 \cdot 10^{-7}$

estimates of the VSD. To minimize model uncertainty of F in (1.3), we use the 2.5% and the 5% BMD in (2.5) and (2.6) to derive the lower and upper confidence bounds on the low-dose risks, respectively. Estimated VSD corresponding to upper bounds of 10^{-4}, 10^{-5} and 10^{-6} on the extra risk in (1.3) and lower bounds on P(VSD) are presented in Table 5.

The BMD and VSD figures in Tables 4 and 5 were obtained using the EPA's free, user-friendly software package, BMDS. For a given BMR, the software provides the BMD and the BMDL directly. The VSD are easily computed from $10^{-k} \times$ BMDL/BMR for $k = 4, 5$, or 6.

To obtain the lower bound on the P(VSD) (i.e., LBP(VSD)) in Table 5, I used Mathematica. For the MLEs and their log-likelihood values, I used the `FindMinimum` function and for the constrained minimization, I used the `FindRoot` function. For the latter, one must appeal to the Kuhn and Tucker (1951) condition (see Fygenson (2008, p. 20)). For illustration, here are the specific steps I took in the analysis of the DeAngelo et al. data.

Finding the MLEs:

```
F1[b_,d_] := (1-Exp[-b*d])
```

```
F1inv[b_,p_] := Log[1-p]/b

Prob[a_,b_,d_] := a+(1-a)*F1[b,d]

llc[a_,b_,d_,y_,n_] := y*Log[prog[a,b,d] + (n-y)*Log[1-prob[a,b,d]]

ll[a_,b_] := llc[a,b,0,1,45] + llc[a,b,1.1,1,43] +...+ llc[a,b,28.7,12,32]

FindMinimum[-llc[a,b], {a, 0.0000001}, {b, 0.0000001}]
```

$\{61.2432, a- > 0.0187411, b- > 0.0126062\}$ (*output from Mathematica*).

Finding the LBP(VSD).

```
B[b_,p_,q_] := (Log[q/(1-q)]-Log[p/(1-p)])/(Log[F1inv[b,q]]-Log[F1inv[b,p]])

A[b_,p_,q_] := Log[p/(1-p)] - B[b,p,q]*Log[F1inv[b,p]]

G[b_,d_,p_,q_] := 1/(Exp[-A[b,p,q]]*d^(-B[b,p,q]))

FindRoot[{ D[G[b,d,p,q],b]]==0, -ll[a,b]-62.5962==0 }, {b=0.0000001}]
```

$\{b- > 0.00896949\}$ (*output from Mathematica*)

```
G[0.00896949,VSD,p,q]
```

(*gets the* LBP(VSD)).

In summary, the VSD from the DeAngelo et al. (1998) data are about 1.64 times larger than the corresponding VSD from the Kurokawa et al. (1986) data (Table 5). However, the lower bounds on P(VSD) are similar for both data sets and halfway between the upper bounds on P(VSD) and the upper bounds for the preceding reference risks. Thus, the Kurokawa et al. (1986) study makes up for the use of less rodents per dose group by having a lower (positive) dose level and because the Quantile-linear model provides a better fit to its data. Very important to risk/benefit analysis, both studies indicate a similar minimum number of extra cases due to exposure to bromate. Finally, it is important to note that the results in Tables 4 and 5 are based in part on the asymptotic distribution of the likelihood ratio test, and thus the true nominal value in applications with small sample sizes is unknown. For the datasets used in this paper, the sample sizes are relatively large (146 and 206, respectively) and one can safely expect the approximation to be good.

4. Discussion and Conclusions

This paper advocates a new approach to dealing with the well-researched problem of low-dose extrapolation. Ideally one would use biologically based dose-response models for which parameters are calculated independently of curve fitting to the data. Examples include the two-stage models developed by Moolgavkar, Vernon and Knudson (MVK), (see Moolgavkar and Luebeck (1990)).

However, extensive data is required both for building such models and for estimating their parameters. As a result, Crump (1996) writes, "This approach has the same limitation as approaches based on non-biological statistical models: Different assumptions about the dose response from the MVK parameters will fit the tumor bioassay data equally well, but produce enormously different estimates of low-dose risk." (See also National Research Council (1993).)

Since in most cases biologically based models are not practicable, regulatory agencies, like the EPA, recommend the use of the BMD method, combined with conservative inputs, for estimating the VSD. This paper provides insights and tools for arriving at conservative standards when applying the BMD method. In particular, it is argued that the proper way to achieve this goal is to use *pessimistic* models in the extrapolation region. Two broad families of pessimistic distributions – AR-pessimistic and OR-pessimistic – are used. These two forms of pessimism complement each other in the sense that AR-pessimism constrains the *absolute* excess-risk mechanism to be non-decreasing, and OR-pessimism imposes the same constraint on the *relative* excess-risk mechanism.

One way of dealing with model uncertainty in the data range is to use the most pessimistic model among those that fit the data equally well. In Section 2.1, it was established that, among the parametric models commonly used in fitting bioassay data, a few are not at all pessimistic. Characterization of these models with respect to their outlooks thus provides a qualitative explanation of the empirical observation that some models consistently lead to lower VSD.

AR-pessimism provides new insight into the current BMD method. Specifically, it was established that the current BMD method is AR-pessimistic because it involves extrapolating to the VSD by imposing the non-parametric constraint of increasing *absolute* excess-risk on the very low dose-response curve. Furthermore, the use of the BMDL as a point of departure for the extrapolation model, construed by some as enhancing the conservative nature of the BMD method, is actually the only proper statistical procedure, and therefore is neither especially pessimistic nor conservative.

OR-pessimism is instrumental in deriving *lower* bounds on the incremental risk from exposure to the VSD. This lower bound is an important complement to the current BMD method. While having an estimate of the minimum number of extra cases from exposure to levels of carcinogens recognized as safe (i.e., VSD) is of general importance, it is especially so for carcinogens that directly or indirectly benefit the public. In such cases it can help resolve challenges from parties concerned that a particular regulatory standard is excessively protective and therefore unfair and/or unpractical.

The lower bound on $P(\text{VSD})$ was evaluated, as a probability bound in Table 2 and as a statistical (confidence) bound in Table 5, for the data from Section 3. Both were found to be informative. Interestingly, the lower confidence bounds

on $P(\text{VSD})$ in the two bioassay studies in Section 3 were similar even though the reported incidence rate in the two studies were quite different. In general, the results from Tables 2 and 5 are encouraging and strongly suggest that the proposed lower bound should be used with other data and implemented in the EPA's software.

While the software is user friendly and offers advanced features, it remains a work in progress. One issue that requires further attention arises in the estimation of the control groups risk (i.e., $P(0)$), that plays a major role via equation (1.3). In particular, when fitting one of the two unrestricted models (i.e., probit or logit), the BMDS uses model (1.1). Its estimate of $P(0)$ is therefore based on parameters that were estimated from the data (very far away). It would be more appropriate to consider model (1.3) and estimate $P(0)$ non-parametrically. This requires future consideration, especially about the best way of accounting for the variability in the data.

One of the more difficult issues that remain is that of the appropriate point of departure used to anchor the extrapolation of the VSD. Even if the anchor point of choice falls in the data range, extrapolation from the 10^{th} percentile, for example, to the targeted percentile of 10^{-6} is a stretch, and may well be inappropriate. In picking the point of departure there is always a tradeoff between model uncertainty of F_1 (in the data range) and model uncertainty of F (in the extrapolation range). More consideration must be given to this tradeoff. We have seen another aspect of it in the two bioassays of Section 3. The DeAngelo et al. (1998) study used more animals per dose level, while the Kurokawa et al. (1986) study extended to lower (positive) dose levels. *A priori*, the latter data set should result in less model uncertainty in the extrapolation range and the former should provide a better account for model uncertainty in the data range. To find the "right" balance is therefore a design issue as well as a methodological one.

Acknowledgements

I am grateful to an associate editor and two anonymous referees for valuable suggestions and careful reading of the original manuscript.

Appendix

Throughout the Appendix, $g'(x)$ is the derivative of g and $O_F(x) = F(x)/(1 - F(x))$ denotes the odds of a distribution function F.

To prove Theorem 2.1, we require the following result.

Lemma 1. *A distribution function F is inherently OR-pessimistic (optimistic) on the multiplicative scale in interval J if, and only if, $O_F(x)$ is log-convex (log-*

concave) *in* $\log(x)$ *for all* $x \in J$. *Equivalently, if, and only if,* $x \log' O_F(x)$ *is increasing* (*decreasing*) *in* x.

Proof of Lemma 1. $OR_F(\alpha x, x)$ is increasing (decreasing) in $x \Leftrightarrow \log(OR_F(\alpha x, x))$ is increasing (decreasing) in $x \Leftrightarrow d\{\log OR_F(\alpha x, x)\}/dx = \log'(O_F(\alpha x, x)) - \log'(O_F(x)) \geq (\leq) 0 \Leftrightarrow x \log'(O_F(x))$ is increasing (decreasing) in $x \Leftrightarrow \log(O_F(x))$ is convex (concave) in $\log(x)$.

Proof of Theorem 2.1. Parts (i) and (ii) follow from Lemma 1 and Theorem A.3 in Fygenson (2008). Part (iii) can be proven using Lemma 1. Here we illustrate the proof for the LMS, which is the most difficult of the three. (The proofs for the Quantal-linear and the Weibull require following the first four equations in the proof below.)

By Lemma 1, we need to show that, for the LMS model, $x \log' O(x)$ is increasing in x. For the LMS

$$x \log' O(x) = \frac{\sum \beta_i x^i i}{1 - e^{-\sum \beta_i x^i}},$$

and

$$\Big[x \log' O(x)\Big]' = \frac{\sum \beta_i x^{i-1} i^2 - e^{-\sum \beta_i x^i}\Big[\sum \beta_i x^{i-1} i^2 + (\sum \beta_i x^i i)(\sum \beta_i x^{i-1} i)\Big]}{(1 - e^{-\sum \beta_i x^i})^2}. \tag{A.1}$$

The numerator of (A.1) is non-negative if

$$\log\Big(\sum \beta_i x^{i-1} i^2\Big) \geq -\sum \beta_i x^i + \log\left[\sum \beta_i x^{i-1} i^2 + \Big(\sum \beta_i x^i i\Big)\Big(\sum \beta_i x^{i-1} i\Big)\right], \tag{A.2}$$

or if

$$\log\left[1 + \frac{(\sum \beta_i x^i i)(\sum \beta_i x^{i-1} i)}{(\sum \beta_i x^{i-1} i^2)}\right] \leq \sum \beta_i x^i. \tag{A.3}$$

The left hand side of (A.3) is no greater than $(\sum \beta_i x^i i)(\sum \beta_i x^{i-1} i)/(\sum \beta_i x^{i-1} i^2)$. Thus, the inequality in (A.3) holds if

$$\Big(\sum \beta_i x^i i\Big)\Big(\sum \beta_i x^{i-1} i\Big) \leq \Big(\sum \beta_i x^i\Big)\Big(\sum \beta_i x^{i-1} i^2\Big) \quad \text{for all} \quad \beta_i \geq 0 \quad \text{and} \quad x \geq 0. \tag{A.4}$$

Multiplying both sides of (A.4) by x gives

$$\Big(\sum \beta_i x^i i\Big)\Big(\sum \beta_i x^i i\Big) \leq \Big(\sum \beta_i x^i\Big)\Big(\sum \beta_i x^i i^2\Big). \tag{A.5}$$

The left hand side of (A.5) is

$$\Big(\sum \beta_i^2 x^{2i} i^2\Big) + 2\Big(\sum\sum \beta_i \beta_j x^i x^j i j\Big), \tag{A.6}$$

and the right hand side of (A.5) is

$$\Big(\sum \beta_i^2 x^{2i} i^2\Big) + \Big(\sum\sum \beta_i \beta_j x^i x^j (i^2 + j^2)\Big). \tag{A.7}$$

Since $(i^2 + j^2) \geq 2(ij)$, this completes the proof.

To prove part (iv), we need to show (by Lemma 1) that for the Gamma distribution with shape parameter $0 < \alpha < 1$, $x \log' O(x)$ is increasing in $x > 0$ or, equivalently, that

$$\frac{F(x)(1 - F(x))}{xf(x)} \tag{A.8}$$

is decreasing in $x > 0$. Note that (A.8) for the Gamma distribution is

$$\frac{(1/\Gamma(\alpha)) \int_0^x \lambda^\alpha t^{\alpha-1} e^{-\lambda t} dt}{x} \int_x^\infty \Big(\frac{y}{x}\Big)^{\alpha-1} e^{-\lambda(y-x)} dy.$$

Now, with the change of the variables $z = t/x$ and $u = y - x$, the above becomes

$$\begin{aligned}
&\Big(\frac{1}{\Gamma(\alpha)} \int_0^1 \lambda^\alpha (zx)^{\alpha-1} (e^{-\lambda z})^x dz\Big)\Big(\int_0^\infty \Big(\frac{u+x}{x}\Big)^{\alpha-1} e^{-\lambda u} du\Big) \\
&= \frac{1}{\Gamma(\alpha)} \int_0^1 \lambda^\alpha \int_0^\infty (zx)^{\alpha-1} (e^{-\lambda z})^x \Big(\frac{u+x}{x}\Big)^{\alpha-1} e^{-\lambda u} du dz \\
&= \frac{1}{\Gamma(\alpha)} \int_0^1 \lambda^\alpha \int_0^\infty g(u, z; x, \alpha, \lambda) e^{-\lambda u} du dz,
\end{aligned}$$

where $g(u, z; x, \alpha, \lambda) = (zx)^{\alpha-1}(e^{-\lambda z})^x [(u + x)/x]^{\alpha-1}$. Note that $\log' g(u, z; x, \alpha, \lambda) = [(\alpha - 1)/(x + u)] - \lambda z \leq 0$, for all $\alpha \leq 1$. This completes the proof since the above shows that the function $g(u, z; x, \alpha, \lambda)$ is decreasing in $x > 0$ when $0 < \alpha < 1$.

Proof of Theorem 2.2. Part a) follows from Theorem A.1 of Fygenson (2008), which establishes that F is AR-pessimistic (optimistic) on J if, and only if, $F(x)$ is convex (concave) in $x \in J$. Therefore there exists a line βx, say, such that $F(x) \leq \beta x$ and $F(x_p) = \beta x_p$. Setting $F(x_p) = F_1(x_p)$ and solving for β, we get (2.1). To show part b), let $Y = \log(X)$, $G_Y(y) = P(log(X) \leq y)$, and (y_p, y_q) be the p and q percentiles of G_Y, respectively, with $p < q$. From Lemma 1, we know that $\log(O_G(y))$ is convex in $y \in (-\infty, y_q]$. This implies that there exists a line $A + By$, say, such that

$$\log(O_G(y)) \geq A + By, \quad -\infty < y \leq y_p, \tag{A.9}$$

where A and B are the solutions of

$$\log(O_G(y_p)) = A + By_p \ \text{ and } \ \log(O_G(y_q)) = A + By_q.$$

Solving these equations gives

$$B = \frac{(O_G(y_q)) - \log(O_G(y_p))}{y_q - y_p} \quad \text{and} \quad A = \log(O_G(y_p)) - By_p.$$

From (A.9), and since $O_F(x) = O_G(\log(x))$ and $y_i = \log(x_i)$, $i = p, q$, we get

$$O_F(x) = \frac{F(x)}{(1 - F(x))} \geq e^A x^B, \quad 0 < x \leq x_p.$$

To get $F(x) \geq (1 + e^{-A}x^{-B})^{-1}$, simply isolate $F(x)$ on one side of the above equation.

Proof of Proposition 2.1. Under the conditions of the proposition it follows from Theorem 2.1 that $F(x)$ is bounded from above and below by the functions H and G, respectively.

Let $P_i = P(Y_i = 1 \mid X = x_i; \boldsymbol{\theta})$, with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_t)^T \in \Omega$, where the parameter space Ω is an open subset of t-dimensional Euclidean space. Consider the common regularity conditions.

C1. $\lim_{n\to\infty}(n_i/n) = c_i$, $(0 < c_i < 1)$ for all $i = 1, \ldots, K$, where K denotes the number of different X values in the data.

C2. K is at least as large as the number of parameters in the model.

C3. The information matrix $\sum^{-1} = ((\sigma^{\lambda\sigma}))$ with

$$\sigma^{\lambda\sigma} = \sum_{i=1}^{K} c_i \frac{(\partial P_i/\partial\theta_\lambda)(\partial P_i/\partial\theta_\sigma)}{(P_i(1 - P_i))}, \quad (\lambda, \sigma = 1, \ldots, t)$$

is positive definite.

It is not hard to show (Cox and Hinkley (1974)) that, under C1 – C3, the likelihood ratio statistic is asymptotically a χ^2_t, where t is equal to the number of parameters in the model. Using standard arguments on a constrained parameter space (Rao (1973, p.419)) the approximate (1-α)% upper confidence bound for $F(d)$, $d \in J$ is

$$U_d = \mathrm{Sup}_{(\boldsymbol{\theta})}\Big\{H(\boldsymbol{\theta}; d, p) : 2(LL(\hat{\boldsymbol{\theta}}) - LL(\boldsymbol{\theta})) \leq \chi^2_{1,1-2\rho}\Big\}.$$

Using the same arguments we get the approximate (1-α)% lower confidence bound for $F(d)$, $d \in J$, given in (2.5). Note that $\chi^2_{1,1-2\rho}$ is used instead of $\chi^2_{1,1-\rho}$ because a one-sided rather than a two-sided confidence limit is required.

References

Alho, J. M. and Valtonen, E. (1995). Interval estimation of inverse dose-response. *Biometrics* **51**, 491-501.

Buckland, S. T., Burnham, K. P. and Augustin, N. H. (1997). Model selection: a integral part of inference. *Biometrics* **53**, 603-618.

Cornfield, J. (1977). Carcinogenic risk assessment. *Science* **198**, 693-699.

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*. Chapman and Hall, London.

Crump, K. S. (1984). A new method for determining allowable daily intakes. *Fundamental and Applied Toxicology* **4**, 854-871.

Crump, K. S. (1996). The linearized multistage model and the future of quantitative risk assessment. *Human and Experimental Toxicology* **15**, 787-798.

DeAngelo, A. B., George, M. H., Kilburn, S. R., Moore, T. M. and Wolf, D. C. (1998). Carcinogenicity of potassium bromate administered in the drinking water to male B6C3F1 mice and F344/N rats. *Toxicologic Pathology* **26**, 587-594.

Faustman, E. M. (1996). *Review of Noncancer Risk Assessment: Application of the Benchmark Dose Methods* . Prepared for the Commission on Risk Assessment & Risk Management. `http://www.riskworld.com/Nreports/1996/risk_rpt/pdf/faustman.pdf`

Fawell, J. and Walker, M. (2006). Approaches to determining regulatory values for carcinogens with particular reference to bromate. *Toxicology* **221**, 149-153.

Filipsson, A. F. and Victorin, K. (2003). Comparison of available benchmark dose softwares and models using trichloroethylene as a model substance. *Regulatory Toxicology and Pharmacology* **37**, 343-355.

Filipsson, A. F., Sand, S., Nilsson, J. and Victorin, K. (2003). The benchmark dose method - review of available models and recommendations for applications in health risk assessment. *Critical Reviews in Toxicology* **33**, 505-542.

Fygenson, M. (2008). Modeling and predicting probabilities with outlooks (with discussion). *Statist. Sinica* **18**, 9-90.

Gaylor, D. (1989). Quantitative risk analysis for quantal reproductive and developmental effects. *Environmental Health Perspectives* **79**, 243-246.

Huang, Y. (2001). Various methods for interval estimation of the median effective dose. *Commun. Statist.- Simula.* **30**, 99-112.

Krewski, D. and Van Ryzin, J. (1981). Dose-response models for quantal response toxicity data. In *Statistics and Related Topics* (Edited by M. Csorgo, D. A. Dawson, J. N. K. Rao and A. K. Md. E. Saleh), 201-230. North Holland, New York.

Kuhn, H. W. and Tucker, A. W. (1951). Nonlinear programing. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Edited by J. Neyman). University of California Press, Berkeley.

Kurokawa, Y., Aoki, S., Matsushima, Y., Takamura, N., Imazawa, T. and Hayashi, Y. (1986). Dose-response studies on the carcinogenicity of potassium bromate in F344 rats after long-term oral administration. *J. Nat'l. Cancer Inst.* **77**, 977-982.

Moolgavkar, S. H. and Luebeck, G. (1990). Two-event model for carcinogenesis: Biological, mathematical, and statistical considurations. *Risk Analysis* **10**, 323-341.

Morales, K. H., Ibrahim, J. G., Chen, C. and Ryan, L. M. (2006). Bayesian model averaging with applications to benchmark dose estimation for arsenic in drinking water. *J. Amer. Statist. Assoc.: Applications and Case Studies* **101**, 9-17.

National Research Council (1993). Issues in risk assessment. *Committee on Risk Assessment Methodology, Commission on Life Sciences, NRC.* National Academy Press, Washington, D. C.

Piegorsch, W. W., West, R. W., Pan, W. and Kodell, R. L. (2005). Low dose risk estimation via simultaneous statistical inferences. *J. Roy. Statist. Soc. Ser. C* **54**, 245-258.

Rao, C. R. (1973). *Linear Statistical Inferences and its Applications*. Wiley, New York.

Ryan, L. M. (2003). Epidemiologically based environmental risk assessment. *Statist. Sci.* **18**, 466-480.

Sand, S., Filipsson, A. F. and Victorin, K. (2002). Evaluation of the benchmark dose method for dichotomous data: model dependence and model selection. *Regulatory Toxicology and Pharmacology* **36**, 184-197.

U. S. Environmental Protection Agency (1996). Proposed Guidelines for Carcinogen Risk Assessment Office of Research and Development. EPA/600/P-92003C, Washington, D.C..

U. S. Environmental Protection Agency (2000). Draft Technical Guidance Document `http://www.epa.gov/EPA-GENERAL/2000/November/Day-20/g29648.htm`

U. S. Environmental Protection Agency (2001). Integrated Risk Information System `http://www.epa.gov/iris/subst/1002.htm`

U. S. Environmental Protection Agency (2005). Guidelines for Carcinogen Risk Assessment Risk Assessment Forum, Washington, D.C.
`http://www.epa.gov/iriswebp/iris/cancer032505.pdf`

World Health Organization (2004). *Guidelines for Drinking-Water Quality. Recommendations*, Vol.1, 3rd edition. WHO, Geneva.

Yanagimoto, T. and Hoel, D. G. (1990). Comparisons of models for estimation of safe doses using measures of the heaviness of the tail of a distribution. *Ann. Inst. Statist. Math.* **32**, 465-480.

Marshall School of Business, University of Southern California, Los Angeles, CA 90089, U.S.A.

E-mail: mfygenson@marshall.usc.edu

(Received April 2008; accepted November 2008)