

**September 29, 2010**

**REVISED DRAFT**

**MULTISTAGE WEIBULL TIME-TO-TUMOR MODEL  
IN EPA'S BENCHMARK DOSE SOFTWARE (BMDS)**

**METHODOLOGY DESCRIPTION**

**BATTELLE  
505 King Avenue  
Columbus, OH 43201-2693**

**EPA Contract Number EP-C-05-030  
Work Assignment 3-17**

**Prepared for**

**John Fox, Work Assignment Manager  
National Center for Environmental Assessment  
Office of Research and Development  
U.S. ENVIRONMENTAL PROTECTION AGENCY  
Washington, DC 20460**

**Marla Smith, Project Officer  
Engineering and Analysis Division  
Office of Science and Technology  
Office of Water  
U.S. ENVIRONMENTAL PROTECTION AGENCY  
Washington, DC 20460**

### **EPA Disclaimer**

The material in this document has not been subject to Agency technical and policy review. Views expressed by the authors are their own and do not necessarily reflect those of the U.S. Environmental Protection Agency. Mention of trade names, products, or services does not convey, and should not be interpreted as conveying, official EPA approval, endorsement, or recommendation. Do not quote or cite this document.

### **Battelle Disclaimer**

This report is an interim work prepared for the United States Government by Battelle. In no event shall either the United States Government or Battelle have any responsibility or liability for any consequences of any use, misuse, inability to use, or reliance upon the information contained herein, nor does either warrant or otherwise represent in any way the accuracy, adequacy, efficacy, or applicability of the contents hereof.

## Table of Contents

	<u>Page</u>
1.0 INTRODUCTION .....	1
2.0 MODEL DEFINITION.....	2
2.1. Tumor Contexts .....	3
2.2. Log-Likelihood Function .....	4
2.3. Reparameterization of the Multistage Weibull Model.....	5
3.0 ALGORITHM FOR MAXIMUM LIKELIHOOD ESTIMATION .....	7
3.1. Starting Values.....	7
3.2. Estimation of Multistage GEV Model Parameters .....	7
3.3. Estimating the Multistage Weibull Parameter Correlation Matrix and Standard Errors.....	8
4.0 BENCHMARK DOSE (BMD) ESTIMATION .....	8
5.0 PROFILE LIKELIHOOD CONFIDENCE INTERVAL.....	9
5.1. Algorithm for Estimating Profile Likelihood Confidence Intervals .....	10
6.0 OPTIMIZATION FEATURES.....	11
6.1. Analytic Gradients .....	12
6.2. Optimization Parameter Settings .....	12
6.3. Internal Scaling .....	12
7.0 MODELING ISSUES .....	14
7.1. Censoring and Estimability .....	14
7.2. Asymptotic Normality .....	15
7.3. Profile Likelihood Confidence Bounds .....	16
7.4. Interpretation and Estimation of $t_0$ .....	16
8.0 REFERENCES .....	17
Appendix A Algorithm for Multistage GEV Parameter Maximum Likelihood Estimation .....	19
Appendix B Algorithm for Estimation of Profile Likelihood Confidence Interval.....	22
Appendix C Derivatives for Log-Likelihood Functions.....	25
Appendix D Concavity Theorems .....	27
Appendix E Algorithm for Calculating Predicted Counts .....	29

# MULTISTAGE WEIBULL TIME-TO-TUMOR MODEL IN EPA'S BENCHMARK DOSE SOFTWARE (BMDS)

## METHODOLOGY DESCRIPTION

### 1.0 INTRODUCTION

In cancer risk assessments, the multistage Weibull model [Krewski et. al., 1983] is often used to model the probability of a specified tumor response in a test subject exposed to a toxin when some time-related measurement is available on the subject (e.g., time since toxin exposure; subject's age) along with the reported dosage rate of toxin. Here, the term "tumor" refers to a specific type of tumor occurring in a particular tissue or organ of a select sex (and strain of animal, for bioassay data). The multistage Weibull model is a natural extension of the multistage model. It is derived from a model of carcinogenesis where tumor onset results from some fixed number of sequential genetic mutations, with the assumption that the hazard rate for each mutation (i.e., "stage") is homogeneous over time but is dose dependent. The model is generally applicable for non- to moderately fatal cancer types under conditions of low dose exposure, and has been proposed for use with data from both controlled animal experiments [e.g., OSHA 1992] and observational studies of humans [e.g., NRC, 1999].

The U.S. Environmental Protection Agency (EPA) has utilized the commercially-available software TOXRISK<sup>1</sup> to fit the multistage Weibull model to data generated from (animal) time-to-tumor experiments [e.g., USEPA, 2001, 2002]. However, that software is no longer commercially sold or supported. Also, EPA would like to have model parameter estimates reported with standard errors or confidence intervals, and would like to estimate benchmark doses for a wide range of user-specified and additional risks. It is also important to EPA to have a fully documented and tested software utility that calculates confidence intervals using the profile likelihood method (without resorting to linear approximation) and that reports the likelihood profile. Therefore, EPA has pursued the development of a multistage Weibull time-to-tumor modeling capability within its Benchmark Dose Software (BMDS), a tool for risk assessors to estimate a benchmark dose associated with a pre-determined benchmark response by fitting one of several statistical models to dose-response data. The development of a module for incorporating the multistage Weibull model within BMDS also provided EPA with an opportunity to examine and resolve statistical and computational issues associated with the application of this model. This report documents the methodology which has been utilized in implementing a multistage Weibull time-to-tumor model within BMDS.

---

<sup>1</sup> Toxicological Risk Assessment Program (1995). Developed by K. Crump, R. Howe, C. Van Landingham and W. Fuller, Clement International Corporation, Ruston, LA under contract to Electric Power Research Institute, Palo Alto, CA

## 2.0 MODEL DEFINITION

A *time-to-tumor model* describes the probability of a test subject exhibiting a specific tumor-related response by time  $t$  when the subject is exposed to a toxin at dosage rate  $d$ . The tumor-related response takes one of two forms:

- Death of the subject, with death resulting from a tumor (“death from tumor”)
- Onset of a tumor (i.e., the subject develops a tumor of a pathologically detectable size), generally observable only upon examination following death.

The ***k-stage time-to-tumor Weibull model for fatal tumors*** (for some positive integer  $k$ ) characterizes the probability of death from tumor. Under this model, the probability of death from tumor occurring prior to some specified observation time  $t$  upon exposure to a toxin at dose level  $d$  is given by the function

$$F(t, d) = F(t, d, t_0, c, \beta_0, \beta_1, \dots, \beta_k) = \begin{cases} 1 - \exp\left\{-(t - t_0)^c \sum_{i=0}^k \beta_i d^i\right\} & \text{if } t > t_0 \\ 0 & \text{otherwise} \end{cases}$$

where model parameters satisfy the restrictions  $c \geq 1$ ,  $t_0 > 0$ , and  $\beta_i \geq 0$  ( $i = 0, 1, \dots, k$ ). The *shape parameter*  $c$  determines how the risk of death from tumor increases over time. The *location (or induction time) parameter*  $t_0$  is interpreted as the elapsed time that occurs between the onset of a fatal tumor and death from tumor. The *polynomial coefficient scale parameters*  $\beta_0, \beta_1, \dots, \beta_k$ , determine the curvature of the dose-response curve. Assumptions that are implicit to this model are:

- Death from tumor cannot occur prior to time  $t = t_0$ .
- The elapsed time from tumor onset to death from tumor ( $t_0$ ) is constant across subjects.
- Test subjects are tumor free at time  $t = 0$ , and exposure to toxin does not occur prior to  $t = 0$ .

In its current implementation within BMDS, up to a six-stage model (i.e.,  $k = 6$ ) can be fitted. Furthermore, the model can be used to estimate statistics that are related to the probability of achieving either death from tumor (e.g., benchmark dose for fatal risk) or tumor onset (e.g., benchmark dose for incidental risk) over time. Section 4.0 provides more information on estimating benchmark dose.

The ***k-stage time-to-tumor Weibull model for non-fatal tumors*** characterizes the probability of tumor onset. Under this model, the probability of observing tumor onset in a test subject prior to some specified observation time  $t$  upon exposure to a toxin at dose level  $d$  is given by the function

$$G(t, d) = G(t, d, c, \beta_0, \beta_1, \dots, \beta_k) = 1 - \exp\left\{-t^c \sum_{i=0}^k \beta_i d^i\right\}$$

where model parameters  $c$  and  $\beta_i$  have the same constraints as above. Note that this function is equivalent to the time-to-death distribution function for fatal tumors with the location parameter

omitted. The shape parameter  $c$  determines how the risk of tumor onset increases over time, and scale parameters  $\beta_0, \beta_1, \dots, \beta_k$ , determine the curvature of the dose-response curve. Assumptions that are implicit to this model are:

- Any occurrence of death among test subjects is not attributable to the tumor type that is being modeled.
- Test subjects are tumor free at time  $t = 0$ , and exposure to toxin does not occur prior to  $t = 0$ .

## 2.1. Tumor Contexts

In survival/sacrifice studies collecting time-to-tumor data, the test subject is classified within one of the following four outcome categories or *tumor contexts* [Peto et al., 1980]:

- Censored response ( $C$ ). The subject is removed from the study at time  $t$  because it has died or was scheduled to be sacrificed, and upon examination, no tumors of the type being modeled are detected.
- Death from fatal tumor ( $F$ ). The subject dies at time  $t$ , a tumor is detected when the subject is examined, and death is attributed to this tumor. This context is obviously not applicable for non-fatal tumors.
- Incidental tumor ( $I$ ). The subject is removed from the study at time  $t$  because it has died or was scheduled to be sacrificed, and upon examination, a tumor is detected. Death is not attributed to this tumor; otherwise, the subject would be classified as tumor context  $F$ .
- Unknown response observed ( $U$ ). The subject is removed from the study at time  $t$  because it has died (from a cause assumed to be unrelated to the tumor type being modeled) or was scheduled to be sacrificed, but during examination, the presence or absence of the tumor type being modeled cannot be determined (e.g., due to decomposition, or inconclusive necropsy).

Within each of these tumor contexts, the likelihood associated with the  $k$ -stage Weibull is as follows:

$C$ : For fatal tumors, the subject is alive up to time  $t$ , but no tumor is detected up to this time. Therefore, death from tumor can not occur before time  $t + t_0$ :

$$1 - F(t + t_0, d) = \exp \left\{ -t^c \sum_{i=0}^k \beta_i d^i \right\}$$

For non-fatal tumors, the subject is alive up to time  $t$ , but tumor onset has not occurred up to this time:

$$1 - G(t, d) = \exp \left\{ -t^c \sum_{i=0}^k \beta_i d^i \right\}$$

*F*: For fatal tumors, death from tumor occurs at observation time  $t$ :

$$\frac{\partial F(t, d)}{\partial t} = \begin{cases} c(t - t_0)^{c-1} \sum_{i=0}^k \beta_i d^i \exp\left\{-(t - t_0)^c \sum_{i=0}^k \beta_i d^i\right\} & \text{if } t > t_0 \\ 0 & \text{otherwise} \end{cases}$$

For non-fatal tumors, the likelihood is not defined because death from tumor cannot occur.

*I*: For fatal tumors, the subject is alive up to time  $t$ , but a tumor is detected at this time. Therefore, death from tumor is predicted to occur between time  $t$  and  $t + t_0$ :

$$F(t + t_0, d) - F(t, d) = \begin{cases} \exp\left\{-(t - t_0)^c \sum_{i=0}^k \beta_i d^i\right\} - \exp\left\{-t^c \sum_{i=0}^k \beta_i d^i\right\} & \text{if } t > t_0 \\ 1 - \exp\left\{-t^c \sum_{i=0}^k \beta_i d^i\right\} & \text{otherwise} \end{cases}$$

For non-fatal tumors, the subject is alive up to time  $t$ , but a tumor is detected at this time (i.e., tumor onset occurred before time  $t$ ):

$$G(t, d) = 1 - \exp\left\{-t^c \sum_{i=0}^k \beta_i d^i\right\}$$

*U*: For fatal tumors, even though it cannot be determined when or if the animal died due to tumor, it is assumed that any such death would not have occurred up to time  $t$ :

$$1 - F(t, d) = \begin{cases} \exp\left\{-(t - t_0)^c \sum_{i=0}^k \beta_i d^i\right\} & \text{if } t > t_0 \\ 1 & \text{otherwise} \end{cases}$$

For non-fatal tumors, the likelihood is not defined because the time of tumor onset cannot be determined. Thus, test subjects having data with context  $U$  are not considered when modeling time-to-tumor data associated with non-fatal tumors.

## 2.2. Log-Likelihood Function

Assume that test subjects are placed within one of  $D > k$  dosage groups, and let  $d_j$  denote the dosage rate assigned to the  $j^{\text{th}}$  dosage group ( $j = 1, \dots, D$ ). Furthermore, within each dosage group, assume that the subjects are grouped further according to observation time and tumor context. For test subject group  $s$  within dosage group  $j$ , let  $t_{js}$  denote the observation time,  $r_{js}$  denote the tumor context, and  $n_{js}$  denote the number of subjects in the group. Then the log-likelihood for fatal tumors is written as follows:

$$l(t_0, c, \beta_0, \beta_1, \dots, \beta_k) = \ln \left[ \prod_{j=1}^D \left( \prod_{r_{js}=C} [1 - F(t_{js} + t_0, d_j)]^{n_{js}} \times \prod_{r_{js}=F} \left[ \frac{\partial F(t, d)}{\partial t} \right]^{n_{js}} \right. \right. \\ \left. \left. \times \prod_{r_{js}=I} [F(t_{js} + t_0, d_j) - F(t_{js}, d_j)]^{n_{js}} \times \prod_{r_{js}=U} [1 - F(t_{js}, d_j)]^{n_{js}} \right) \right]$$

Because  $C$  and  $I$  are the only tumor contexts associated with non-fatal tumors, the log-likelihood for non-fatal tumors corresponds to the following:

$$l(c, \beta_0, \beta_1, \dots, \beta_k) = \ln \left[ \prod_{j=1}^D \left( \prod_{r_{js}=C} [1 - G(t_{js}, d_j)]^{n_{js}} \times \prod_{r_{js}=I} [G(t_{js}, d_j)]^{n_{js}} \right) \right]$$

Values of  $t_0, c, \beta_0, \beta_1, \dots, \beta_k$  that maximize the log-likelihood function are the *maximum likelihood estimates (MLEs)* of the model parameters.

### 2.3. Reparameterization of the Multistage Weibull Model

Within BMDS, the maximum likelihood (ML) estimation algorithm used for the  $k$ -stage Weibull model is a modification of the ML estimation algorithm for the 3-parameter Weibull distribution [Hirose, 1996]. Both algorithms use a reparametrization of the Weibull to improve computational stability. In the ML algorithm for the  $k$ -stage Weibull model, the reparametrization is:

$$\begin{aligned} \gamma &= 1 / c \\ \mu &= t_0 + c \\ b_i &= \beta_i \times c^c \quad \text{for } i = 0, 1, \dots, k \end{aligned}$$

These new parameters are substituted in the  $k$ -stage Weibull model for  $c, t_0$ , and  $\beta_i$ , respectively, for both the fatal and non-fatal tumor model forms. For fatal tumors, the probability of death from tumor by time  $t$  upon exposure to the toxin at dosage rate  $d$  is

$$F(t, d) = F(t, d, \mu, \gamma, b_0, b_1, \dots, b_k) = \begin{cases} 1 - \exp \left\{ -[1 + \gamma(t - \mu)]^{1/\gamma} \sum_{i=0}^k b_i d^i \right\} & \text{if } t > \mu - \gamma^{-1} \\ 0 & \text{otherwise} \end{cases}$$

For non-fatal tumors, the probability of tumor onset by time  $t$  upon exposure to the toxin at dosage rate  $d$  is

$$G(t, d) = G(t, d, \gamma, b_0, b_1, \dots, b_k) = 1 - \exp \left\{ -[\gamma t]^{1/\gamma} \sum_{i=0}^k b_i d^i \right\}$$

The new parameters satisfy the following constraints:  $0 < \gamma \leq 1$ ,  $\gamma^{-1} < \mu$ , and  $b_i \geq 0$  ( $i = 0, 1, \dots, k$ ). The reparametrized model takes the form of the Generalized Extreme Value (GEV)



distribution [von Mises, 1936] and is therefore called the *k-stage GEV model*. Although the results were limited to ML estimation for uncensored data, a similar GEV reparametrization of the 3-parameter Weibull model has been shown to improve numerical stability for ML estimation, especially when the shape parameter  $c$  is large (i.e., when the data are negatively skewed) [Hirose, 1996].

Note that the reparametrization from  $k$ -stage Weibull to  $k$ -stage GEV parameters allows Weibull parameter MLEs to be expressed as simple analytic functions of GEV parameter MLEs. However, impact of the reparameterization on the estimation procedure is non-trivial, because the conversion from  $k$ -stage Weibull parameters to  $k$ -stage GEV parameters is not one-to-one (i.e., each of the  $k$ -stage GEV parameters is not a function of individual  $k$ -stage Weibull parameters). In particular, while BMDS allows the user to fix the values of any of the model parameters, fixing parameters in the  $k$ -stage Weibull model creates complex constraints on the  $k$ -stage GEV model parameters. Details on the effects of fixing  $k$ -stage Weibull parameters on the  $k$ -stage GEV parameter constraints are provided in Appendix A.

For fatal tumors, the log-likelihood associated with the  $k$ -stage GEV model (“ $k$ -stage GEV log-likelihood”) is expressed as follows:

$$\begin{aligned}
l(\mu, \gamma, b_0, b_1, \dots, b_k) = & \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} [\gamma t_{js}]^{1/\gamma} \sum_{i=0}^k b_i d_j^i \right. \\
& - \sum_{r_{js}=F} n_{js} \left( [1 + \gamma(t_{js} - \mu)]_+^{1/\gamma} \sum_{i=0}^k b_i d_j^i - (\gamma^{-1} - 1) \ln[1 + \gamma(t_{js} - \mu)]_+ - \ln \sum_{i=0}^k b_i d_j^i \right) \\
& - \sum_{r_{js}=I} n_{js} \left( [\gamma t_{js}]_+^{1/\gamma} \sum_{i=0}^k b_i d_j^i - \ln \left[ \exp \left\{ ([\gamma t_{js}]^{1/\gamma} - [1 + \gamma(t_{js} - \mu)]_+^{1/\gamma}) \sum_{i=0}^k b_i d_j^i \right\} - 1 \right] \right) \\
& \left. - \sum_{r_{js}=U} n_{js} [1 + \gamma(t_{js} - \mu)]_+^{1/\gamma} \sum_{i=0}^k b_i d_j^i \right]
\end{aligned}$$

where  $[x]_+$  denotes the maximum of  $x$  and zero (i.e.,  $[x]_+$  equals  $x$  when  $x > 0$  and equals zero otherwise). For non-fatal tumors, the log-likelihood is determined by modifying the above equation as follows:

1. Set the parameter constraint  $\mu = \gamma^{-1}$  (equivalent to setting  $t_0 = 0$  in the  $k$ -stage Weibull model),
2. Remove the second and last terms corresponding to tumor contexts  $F$  (which can occur only for fatal tumors) and  $U$  (which represent missing values for non-fatal tumors), and
3. Replace the third term (corresponding to tumor context  $I$ ) by

$$+ \sum_{r_{js}=I} n_{js} \ln \left[ 1 - \exp \left\{ - [\gamma t_{js}]^{1/\gamma} \sum_{i=0}^k b_i d_j^i \right\} \right]$$

### 3.0 ALGORITHM FOR MAXIMUM LIKELIHOOD ESTIMATION

This section describes the algorithm for MLE in terms of the log-likelihood for the fatal tumor model. For the non-fatal tumor model, a similar algorithm is used which features the same log-likelihood with minor modifications, as described at the end of Section 2.3. Specific details of the algorithm for the fatal tumor model are provided in Appendix A. A general description of the algorithm is provided below.

#### 3.1. Starting Values

Unlike other modules in BMDS, the multistage Weibull module only allows for either all starting values to be specified by the user, or all values to be specified automatically (i.e., no mixing of user-specified and automatic starting values is permitted). User-specified starting values are converted from multistage Weibull to multistage GEV parameter space. Otherwise, the module generates starting values automatically by specifying a search grid over the bounded free multistage GEV parameters ( $\gamma$  and/or  $\mu$ ), and optimizing the objective function, i.e., the log-likelihood, over the remaining free parameters.

The search grid method for finding automatic starting values takes advantage of the multistage Weibull to multistage GEV reparameterization. In addition to gaining numerical stability, the reparameterization bounds the support of the likelihood for an additional model parameter. This allows a finite search grid to be defined over two parameters, thereby simplifying the search problem. In the  $k$ -stage Weibull model, the support of the likelihood for location parameter  $t_0$  is bounded between 0 and  $t_{min} = \min(t_{js} \mid r_{js} = F)$ , or the earliest observation time for which the study coordinator observes a death from tumor. However, the support for the shape parameter  $c$  is not bounded from above. In the  $k$ -stage GEV model, the support of the likelihood for both parameters  $\gamma$  and  $\mu$  are bounded from above and below.

The algorithm specifies a search grid over both  $\gamma$  and  $\mu$ , with the grid centered at the midpoints of the search space. The  $k$ -stage GEV log-likelihood function is maximized with values of  $\gamma$  and  $\mu$  fixed at each grid point. The concavity properties of the log-likelihood function over  $b_i$  ( $i = 0, 1, \dots, k$ ), proven in Appendix D, ensure the convergence of the optimizations to a global maximum under the fixed constraints. Finally, the grid point(s) associated with the largest maximum log-likelihood value(s) under the fixed constraints is (are) selected to provide the starting values for maximizing the unconstrained  $k$ -stage GEV log-likelihood function.

#### 3.2. Estimation of Multistage GEV Model Parameters

After the starting values are determined, the multistage GEV log-likelihood function is maximized by twice executing the *donlp2* optimization (see Section 6.0). This two-step process allows the internal scaling in the optimization to be updated between the two optimization runs. In addition, any parameters that are very close to a parameter boundary after the first optimization run can be fixed at the boundary value for the second run. In cases where multiple

sets of automatic starting values are provided to the estimation procedure, only the set with the largest log-likelihood after the first optimization run proceeds to the second run. Thereby, the second optimization produces refined values of the parameter MLEs efficiently.

### 3.3. Estimating the Multistage Weibull Parameter Correlation Matrix and Standard Errors

BMDS modules use a standard asymptotic method, the inverse of  $(-2 \times)$  log-likelihood Hessian matrix at the model parameter MLEs, to estimate the correlation matrix and standard errors of the parameter MLEs. For the multistage Weibull module, the Hessian of the multistage Weibull log-likelihood has been coded analytically to minimize numerical errors. (In other BMDS modules, the Hessian matrix is approximated using a finite difference approximation.) Subsequently, as with other BMDS modules, the matrix inversion function in the BMDS ‘Assist’ library is used to invert the matrix. As with other BMDS modules, the multistage Weibull module excludes fixed parameters or parameter estimates with values on the parameter boundary from the correlation and standard error calculations. The analytic second order derivatives of the multistage Weibull log-likelihood are shown in Appendix C.

## 4.0 BENCHMARK DOSE (BMD) ESTIMATION

The *benchmark dose (BMD)* is expressed as a function of the model parameters and the *benchmark response (BMR)*, where the functional form is determined based upon whether the *BMR* is interpreted as a measure of additional risk or extra risk.

Added (i.e., Additional) Risk:

For fatal tumors, the *benchmark fatal added risk* is defined as

$$\begin{aligned} BMR &= F(t, BMD) - F(t, 0) \\ &= \exp\left\{-[1 + \gamma(t - \mu)]_+^{1/\gamma} b_0\right\} - \exp\left\{-[1 + \gamma(t - \mu)]_+^{1/\gamma} \sum_{i=1}^k b_i (BMD)^i\right\} \\ &= \exp\left\{-[1 + \gamma(t - \mu)]_+^{1/\gamma} b_0\right\} \left[1 - \exp\left\{-[1 + \gamma(t - \mu)]_+^{1/\gamma} \sum_{i=1}^k b_i (BMD)^i\right\}\right] \end{aligned}$$

The *benchmark dose for fatal added risk* is therefore the (positive) root *BMD* of the following polynomial equation:

$$\sum_{i=1}^k b_i (BMD)^i + [1 + \gamma(t - \mu)]_+^{1/\gamma} \ln\left[1 - BMR \times \exp\left\{[1 + \gamma(t - \mu)]_+^{1/\gamma} b_0\right\}\right]$$

(Note that *BMD* cannot be estimated if  $BMR \geq [1 - F(t, 0)]$ .)

For both fatal and non-fatal tumors, the equivalent equations for incidental risk are determined by substituting  $\mu = \gamma^{-1}$  (which is equivalent to setting  $t_0 = 0$  in the Weibull model). Thus, the *benchmark dose for incidental added risk* is the (positive) root  $BMD$  of the following polynomial equation:

$$\sum_{i=1}^k b_i (BMD)^i + (\gamma t)^{1/\gamma} \ln[1 - BMR \times \exp\{\gamma t\}^{1/\gamma} b_0]$$

#### Extra (i.e., Relative) Risk:

For fatal tumors, the *benchmark fatal extra risk* is defined as

$$\begin{aligned} BMR &= [F(t, BMD) - F(t, 0)] / [1 - F(t, 0)] \\ &= 1 - \exp\left\{-[1 + \gamma(t - \mu)]_+^{1/\gamma} \sum_{i=1}^k b_i (BMD)^i\right\} \end{aligned}$$

The *benchmark dose for fatal extra risk* is therefore the (positive) root  $BMD$  of the following polynomial equation:

$$\sum_{i=1}^k b_i (BMD)^i + [1 + \gamma(t - \mu)]_+^{1/\gamma} \ln[1 - BMR]$$

As with added risk, the equivalent equations for incidental risk, relevant for both fatal and non-fatal tumors, are obtained by substituting  $\mu = \gamma^{-1}$  (which is equivalent to setting  $t_0 = 0$  in the Weibull model). Thus, the *benchmark dose for incidental extra risk* is the (positive) root  $BMD$  of the following polynomial equation:

$$\sum_{i=1}^k b_i (BMD)^i + (\gamma t)^{1/\gamma} \ln[1 - BMR]$$

In each case, for a given value of  $BMR$  and  $t$ , the maximum likelihood estimate of  $BMD$  is calculated by first substituting the model parameters in the polynomial equation by their maximum likelihood estimates. Starting at 0, the polynomial is evaluated by incrementally increasing the value of  $BMD$  by a step-size of the maximum dose ( $d_D$ ) until the polynomial changes sign from negative to positive. The interval where the sign change occurs contains the root. Subsequently, the numerical root of the polynomial is calculated using the algorithm in the BMDS library that searches for the root of a function inside an interval.

## 5.0 PROFILE LIKELIHOOD CONFIDENCE INTERVAL

The profile log-likelihood function  $l_p(\theta)$  for a parameter  $\Theta$  (e.g., extra risk  $BMD$ ) is defined as the maximum value of the log-likelihood when  $\Theta$  is constrained to a fixed value  $\theta$ , i.e.,

$$l_p(\theta) = \max \{l \mid \Theta = \theta\}$$

where  $l$  is the log-likelihood, and the maximization is taken over the free model parameters within the specified constraints. Note that if  $\hat{\theta}$  is the MLE of  $\Theta$ , then  $l_p(\hat{\theta})$  is the maximum of the log-likelihood with no fixed constraints, i.e.,  $l_p(\hat{\theta}) = l(\hat{\theta})$ .

Strictly speaking, the profile likelihood method calculates more general confidence sets, rather than confidence intervals, because the confidence region is defined as those values of the parameter where the profile log-likelihood function is “close” to the maximum. More precisely, for  $0 < \alpha < 1$ , the  $100(1 - \alpha)\%$  profile likelihood confidence region is defined as

$$\{\theta \in \Theta \mid l_p(\theta) \geq l(\hat{\theta}) - \chi_{1,1-\alpha}^2 / 2\}$$

where  $\hat{\theta}$  is the MLE of  $\Theta$ , and  $\chi_{v,p}^2$  is the  $p^{\text{th}}$  percentile of the  $\chi^2$  distribution with  $v$  degrees of freedom (i.e., if  $X$  is a random variable distributed as a  $\chi^2$  with  $v$  degrees of freedom, then  $P(X < \chi_{v,p}^2) = p$ ).

Unless the profile log-likelihood function satisfies some regularity condition, such as concavity, the confidence region is generally not a single interval. For a single parameter (i.e.,  $\Theta \subset \mathbb{R}^1$ ), a two-sided interval that contains the profile likelihood confidence set is

$$\left[ \inf \{ \theta \in \Theta \mid l_p(\theta) \geq l(\hat{\theta}) - \chi_{1,1-\alpha}^2 / 2 \}, \sup \{ \theta \in \Theta \mid l_p(\theta) \geq l(\hat{\theta}) - \chi_{1,1-\alpha}^2 / 2 \} \right]$$

This interval is defined as the *profile likelihood 100(1 -  $\alpha$ )% confidence interval for  $\Theta$* . The interval is likely to be overly conservative, because it can include regions outside the confidence set.

### 5.1. Algorithm for Estimating Profile Likelihood Confidence Intervals

Both *BMD* and slope cannot easily be expressed as one of the parameters in the multistage GEV model. Therefore, the profile log-likelihood for the parameter is also difficult to calculate. For other models with this problem, the BMDS calculates the *BMD* profile log-likelihood by using the SQP method to maximize the log-likelihood with the defining polynomial equation for the parameter as a fixed constraint. Therefore, the same approach is used to evaluate the profile log-likelihood for the  $k$ -stage GEV model. For fatal risk, the constraints in the profile log-likelihood for the  $k$ -stage GEV model are

*BMD for Added (i.e., Additional) Risk:*

$$d(\theta, \gamma, \mu, b_0, b_1, \dots, b_k) = \sum_{i=1}^k b_i \theta^i + [1 + \gamma(t - \mu)]_+^{1/\gamma} \ln \left[ 1 - BMR \times \exp \left\{ [1 + \gamma(t - \mu)]_+^{1/\gamma} \right\} b_0 \right] = 0$$

BMD for Extra (i.e., Relative) Risk:

$$d(\theta, \gamma, \mu, b_0, b_1, \dots, b_k) = \sum_{i=1}^k b_i \theta^i + [1 + \gamma(t - \mu)]_+^{1/\gamma} \ln[1 - BMR] = 0$$

where  $\theta \in \Theta$  is the constrained parameter. (For incidental risk, set  $\mu = \gamma^{-1}$  in the above constraint equations.) The profile likelihood function is alternatively expressed as

$$l_p(\theta) = \max_{d(\theta, \gamma, \mu, b_0, b_1, \dots, b_k) = 0} l(\theta)$$

The profile likelihood confidence interval is estimated by an adaptive search grid algorithm and a binary root search over the constrained parameter  $\theta \in \Theta$ . Basically, the algorithm finds locations at which the profile log-likelihood function  $l_p(\theta)$  crosses the cut-off value of

$$l(\hat{\theta}) - \chi_{1,1-\alpha}^2 / 2$$

where  $\hat{\theta}$  is the MLE of  $\Theta$ , and  $\chi_{v,p}^2$  is the  $p$ th percentile of the  $\chi^2$  distribution with  $v$  degrees of freedom.

To find the lower bound of the profile likelihood confidence interval, the adaptive search grid takes decreasing steps in  $\Theta$ , starting at the MLE  $\hat{\theta}$ . An attempt is made to calculate the profile log-likelihood function at each step, using the values of the optimized free parameters in the previous step as starting values. The step size is reduced if the constrained optimization fails. At any step, if the profile log-likelihood function dips from above to below the cut-off, the results of the current and previous steps (including the parameter value and optimized free parameters) are stored. The adaptive search grid terminates when the constrained parameter reaches the lower bound of the parameter space (the step size reaches its allowable minimum), or when the value of the profile log-likelihood function dips below the cut-off by more than a multiple (currently set at 4) of  $\chi_{1,1-\alpha}^2 / 2$ . Subsequently, a root search algorithm takes the last stored pair of grid points and searches the grid interval for the location where the profile-log likelihood function crosses the cut-off. The result of the root search is the lower bound of the confidence interval.

An equivalent methodology is used to find the upper bound of the confidence interval. Specific details of the algorithm are provided in Appendix B.

## 6.0 OPTIMIZATION FEATURES

Like some other modules in BMDS, the multistage Weibull time-to-tumor module carries out optimizations with donlp2, written by Peter Spellucci.<sup>2</sup> Donlp2 implements a sequential quadratic programming method to optimize twice differentiable functions over domains with non-linear constraints. However, the multistage Weibull time-to-tumor module uses the most recent ANSI C port of the routine,<sup>3</sup> for consistency in the programming language throughout the

<sup>2</sup> Copyright by P. Spellucci (note that donlp2 is restricted to non-commercial use by the copyright)

<sup>3</sup> Version 28/11/2001, obtained from <http://plato.la.asu.edu/donlp2.html>. Donlp2 is no longer available at this web site. Donlp2 must be obtained by requesting it from P. Spellucci at [spellucci@mathematik.tu-darmstadt.de](mailto:spellucci@mathematik.tu-darmstadt.de)

module. The port also has the advantage of being implemented with dynamic memory, leading to potential improvements in computational efficiency. Some minor modifications were carried out within the *donlp2* C code in order to resolve any naming conflicts with the BMDS ‘Assist’ library functions, and to simplify the return of optimization results.

### 6.1. Analytic Gradients

The SQP method implemented in *donlp2* requires the gradient values of the objective function and all non-linear constraint functions. Note that *donlp2* itself does not require the analytic gradient functions to be coded, because the routine includes a user option that automatically calculates and implements various finite difference approximations. For optimization in the BMDS multistage Weibull module, however, all required gradient functions were coded analytically in order to improve computational efficiency and accuracy. The first order derivatives of the GEV log-likelihood are shown in Appendix C. The analytic gradients for the non-linear constraints are simple to calculate and are not included in this document.

### 6.2. Optimization Parameter Settings

The default parameter settings for the *donlp2* optimization are used in the BMDS multistage Weibull module. The convergence criteria specified in the batch file format have not been implemented.

### 6.3. Internal Scaling

The free parameters in the multistage GEV model are scaled internally in order to improve the stability of the *donlp2* optimization. All scaling factors include a common factor, currently set at  $2^{-4}$ .

If the parameter  $\gamma$  is free (i.e., if the parameter  $c$  in the multistage Weibull model is free), then the scaling factor for  $\gamma$  also includes the inverse  $\gamma$  gradient of the multistage GEV log-likelihood at the starting values. Adjustments are made to ensure that the scaling factor is small enough for the first few steps in the optimization to stay within parameter bounds and, if automatic starting values are used, within search grid interval bounds.

If the parameter  $\mu$  is free, then the scaling factor for  $\mu$  depends on the multistage Weibull model parameter  $t_0$ . The scaling factor for  $\mu$  includes the inverse  $\mu$  gradient of the multistage GEV log-likelihood at the starting values, with adjustments for parameter bounds and, if applicable, search grid interval bounds.

The scaling factors for free  $b_i$  parameters are the scaling factor for  $\gamma$  divided by the  $b_i$  gradient of the multistage GEV log-likelihood at the starting values. Adjustments are made to ensure that the scaling factor is small enough (i.e., less than  $2^{-20}$  multiplied by a dose normalization factor

$(d_D)^{-i}$ , where  $d_D$  is the largest dose level in the dataset), due to the highly sensitive nature of the  $b_i$  parameters, especially for large  $i$ .

When calculating the profile log-likelihood function during the adaptive grid search for confidence intervals, examples with test data have shown that an additional scaling factor is required for the defining function in the constrained optimization. The profile log-likelihood function  $l_p(\theta)$  can be specified alternatively with a scaling factor  $S (> 0)$  as

$$l_p(\theta) = \max_{S \times d(\theta, \gamma, \mu, b_0, b_1, \dots, b_n) = 0} l(\theta)$$

Currently, the scaling factor is set using an adaptive algorithm, which restricts  $S$  to stay between a minimum value of  $2^{-8}$  and a maximum value of  $2^{48}$ . At each search grid location, the adaptive scaling factor algorithm attempts to evaluate the profile log-likelihood function by repeatedly halving the scale factor until the constrained optimization converges successfully, or the scaling factor reaches the minimum value of  $2^{-8}$ . If the optimization fails to converge before the scaling factor reaches the minimum, then the evaluation of the profile log-likelihood at the particular grid location is considered to have failed. (Subsequently, the adaptive grid search algorithm either reduces the step size or terminates the search.)

The adaptive scaling factor algorithm is implemented inside the adaptive search grid algorithm as follows:

1. At the first grid point, the initial value of the scaling factor is set at the maximum (i.e.,  $2^{48}$ ) and immediately implements the adaptive scaling factor algorithm.
2. Otherwise, if the profile log-likelihood was successfully evaluated at the previous grid point, then the adaptive search algorithm first tries using the scaling factor value which converged successfully at the previous grid point. If that fails, then the initial value of the scaling factor is set to be the smaller of
  - a. (The previously successful scaling factor value)  $\times 2^4$ , or
  - b.  $2^{48}$  (i.e., the largest allowable scaling factor value),
 and the adaptive scaling factor algorithm is implemented.
3. Otherwise, if the profile log-likelihood was not successfully evaluated at the previous grid point, then the initial value of the scaling factor is set to the initial value at the previous grid point, and the adaptive scaling factor algorithm is implemented.



## 7.0 MODELING ISSUES

### 7.1. Censoring and Estimability

Other than for the fatal tumor context ( $F$ ), the removal of a subject from the study for examination due to death or scheduled sacrifice causes the data to be censored at the time of removal. For the type of data available from the usual cancer bioassays, all time-of-onset observations are censored, because onset can only be determined post-mortem by necropsy. In some cases, subjects are sacrificed only at the end of the study, unless they succumb to interim deaths (e.g., due to other cancers, toxicity, ‘natural’ mortality). Other studies may feature serial sacrifices.

The likelihood specification for each tumor context in Section 2.1 implicitly assumes that the censoring mechanism is conditionally independent from the tumor response (i.e., death for fatal tumors, or onset for non-fatal tumors) given the dose. The sacrifice schedule is controllable and can therefore be designed without violating the conditional independence assumption. Occurrence of interim death, however, cannot be controlled, and the conditional independence assumption could be violated if there is a significant factor (other than dose) that affects both tumor response and interim death. Unfortunately, the conditional independence assumption cannot be assessed effectively with the data or supplemental information that are usually available from time-to-tumor studies. Thus, the appropriateness of these assumptions should be examined carefully on biological grounds.

The level of censoring in the data affects whether some, or even any, of the parameters in the  $k$ -stage Weibull model can be estimated. One particular case where model parameters cannot be estimated for either the fatal or non-fatal tumor model is when **neither fatal ( $F$ ) nor incidental ( $I$ ) tumor contexts are observed** (i.e., no tumors are observed in any of the test subjects). Intuitively, the data in this case do not provide sufficient information to assess features of the model, because all the observations are right-censored (i.e., have not yet occurred). Mathematically in this situation, the MLEs of  $\beta_0, \beta_1, \dots, \beta_k$  all equal zero, leading to a degenerate model.

In the case where **no fatal ( $F$ ) but some incidental ( $I$ ) tumor contexts are observed for a tumor treated as if it would eventually lead to fatality**, the MLE for  $t_0$  in the fatal tumor model is not uniquely defined. Since no fatal ( $F$ ) tumor contexts are observed, the MLE of  $t_0$  is any value greater than  $\max(t_{js} \mid r_{js} \in \{I, U\})$ , or the largest observation time for which the study coordinator observes incidental or unobserved tumor contexts. It is therefore questionable whether the outcome of fitting the model for fatal tumors is useful or meaningful in the absence of any fatal ( $F$ ) observations. The non-fatal tumor model is, however, still viable in this case.

Other cases with estimability problems may also exist, such as when tumor contexts are highly confounded with dosage ( $d$ ) or time ( $t$ ). Model diagnostics may help in identifying potential problems with estimability. **Users are strongly advised to exercise caution when interpreting the results from those cases.**

## 7.2. Asymptotic Normality

Deriving useful statistical properties of the multistage Weibull model is complicated by the dependence of traditional theorems on the asymptotic normality of parameter MLEs. (The censoring stemming from the tumor contexts also adds to the theoretical complexity.) Unfortunately, the multistage Weibull model fails to satisfy standard regularity conditions for asymptotic normality of MLEs [LeCam, 1970] in various realistic situations. Those situations include cases where:

- a. The location (or induction time) parameter  $t_0$  is estimated from the data in the fatal tumor model. The standard regularity condition for asymptotic normality of the MLEs is not satisfied because the support of the likelihood over  $t_0$  depends on the observed data.
- b. In addition to estimating the parameter  $t_0$  as in the previous case, the shape parameter  $c$  lies in the interval  $1 \leq c \leq 2$ , and the data contain observations with a fatal ( $F$ ) tumor context. This condition is considered “irregular” [Hirose, 1996] because the Fisher information matrix for the  $F$  tumor context is infinite [Smith, 1985].
- c. Any of the model parameters fall on the boundary of the parameter space in both the fatal and non-fatal model. Self and Liang (1987) provide general theoretical proof of non-normality in the asymptotic distribution of MLEs when a parameter is on the boundary. Based on this work, and using simulation studies, EPA is currently undertaking a research effort to investigate the impact of parameters on the boundary for many of the models implemented in BMDS [e.g., Sinha et. al., 2007].

The following calculations, in particular, are affected by the cases where asymptotic normality of the MLEs is not proven:

- Correlation matrix and standard error estimates of the parameters. With lack of asymptotic normality, using the inverse Hessian of  $(-2 \times)$  log-likelihood to estimate an asymptotic correlation matrix and standard errors of the model parameters may produce unreliable results. In particular, Wald confidence intervals for the parameters may be inaccurate.
- Profile likelihood confidence intervals of  $BMD$ . The asymptotic  $\chi^2$  property of likelihood ratio tests depends on the asymptotic normality of the MLEs. Therefore, the profile likelihood confidence intervals that are calculated using quantiles of the  $\chi^2$  distribution may be inaccurate.

### 7.3. Profile Likelihood Confidence Bounds

In addition to the questions of asymptotic normality described in Section 7.2, the calculation of profile likelihood confidence intervals has some numerical issues. In particular, the following two features of the profile log-likelihood function present numerical challenges:

- a. The lack of an analytically defined domain for the profile log-likelihood function. The profile likelihood function is defined as the constrained optimum of the log-likelihood, where the constraint is on the parameter of interest. Therefore, the boundary values of the domain (assuming it is an interval) cannot always be calculated analytically.
- b. The lack of unimodality in the profile log-likelihood function. The profile log-likelihood function may not always be unimodal. In that case, the boundaries for the confidence interval may not be the nearest location above and below the parameter MLE where the profile log-likelihood crosses the threshold.

The lack of unimodality in the log-likelihood forces the requirement that the search algorithm search the entire domain in order to find the farthest values above and below the parameter MLE where the profile log-likelihood function crosses the threshold. Currently, the search region for finding a profile likelihood confidence interval attempts to cover the domain, at least for the domain boundaries known to be finite. However, this introduces instability in the code, because the program may fail to execute successfully if the search algorithm tries to evaluate the log-likelihood function outside its numerical domain.

### 7.4. Interpretation and Estimation of $t_0$

The multistage Weibull model for fatal tumors includes a parameter  $t_0$  that represents the time between tumor onset and death from tumor. The assumption that the value of this parameter is fixed across all test subjects yields substantial simplification. However, in reality, the behavior and time course for tumor development is generally more complex than assumed. In particular, the model assumes the following:

- The time between tumor onset and death from tumor is the same across all subjects. As currently specified, the multistage Weibull model does not allow for subjects to exhibit deaths from tumor at varying time intervals after tumor onset.
- The onset of a fatal tumor inevitably leads to death from tumor. The model is implicitly assuming that tumor onset is absolutely fatal.

Estimating  $t_0$  may also present a problem because most time-to-tumor experiments are survival/sacrifice, where surviving test subjects are terminated according to some pre-determined sacrifice schedule. Termination can occur prematurely before death from tumor would occur, leading to censored results over time. Unless a sufficient number of responses categorized as

fatal ( $F$ ) are observed, the data may not contain enough information to obtain a reasonably accurate estimate of  $t_0$ .

A possible solution to this problem may be to fix  $t_0$  at some biologically justifiable value between 0 and  $t_{min}$  (the minimum observation time for  $F$  tumor contexts). Note, however, that the fatal tumor model can become problematic when  $t_0$  is fixed at 0. (In fact, the multistage Weibull model software within BMDS assumes that when the user fixes  $t_0 = 0$ , the non-fatal tumor is being requested.) The problem occurs when the data contain any observations categorized as incidental ( $I$ ) tumor context, because its likelihood component  $F(t + t_0, d) - F(t, d)$  would always be 0. In terms of biological interpretation, setting  $t_0 = 0$  would imply that death from tumor occurs instantaneously at time of tumor onset. Even if such an unlikely form of cancer was possible, the concept of an incidental ( $I$ ) tumor would become meaningless. In any subject where tumor onset is detected, death from tumor should have occurred immediately, which would result in a fatal ( $F$ ) tumor context. Therefore, the parameter  $t_0$  should not be fixed at 0 for the fatal tumor model, especially in the presence of incidental ( $I$ ) tumor contexts.

## 8.0 REFERENCES

- Hirose, H. (1996) Maximum likelihood estimation in the 3-parameter Weibull distribution - a look through the Generalized Extreme-Value Distribution. *IEEE Transactions on Dielectrics and Electrical Insulation*. 3(1):43-55.
- Krewski D., Crump KS., Farmer J., Gaylor DW., Howe R., Portier C., Salsburg D., Sielken RL., and Van Ryzin J. (1983) A Comparison of Statistical Methods for Low Dose Extrapolation Utilizing Time-To-Tumor Data. *Toxicological Sciences* 3:140-160.
- LeCam, L. (1970) On the Assumptions Used to Prove Asymptotic Normality of Maximum Likelihood Estimates. *The Annals of Mathematical Statistics* 41:3:802-828
- NRC (1999) *Arsenic in Drinking Water*. National Research Council, The National Academies Press, Washington, DC.
- OSHA (1992) *Occupational Exposure to Cadmium*. 29 CFR Parts 1910, 1915, 1926 and 1928, Occupational Safety and Health Administration, US Department of Labor, Washington, DC.
- Peto, R., Pike M., Day N., Gray R., Lee P., Parish S., Peto J., Richards S., and Wahrendorf J. (1980) Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. pp. 311-426 In: *Long-Term and Short-Term Screening Assays for Carcinogens: a Critical Appraisal*, IARC Monographs, Annex to Supplement 2, International Agency for Research on Cancer, Lyon, France.

- Self, SG., and Liang, K-Y. (1987) Asymptotic Properties Of Maximum-Likelihood Estimators And Likelihood Ratio Tests Under Nonstandard Conditions. *Journal of the American Statistical Association*. 82: 605-610
- Sinha, B., Kopylev, L., and Fox, J. (2007) Some new aspects of dose-response multistage models with applications. *Technical report at*  
[http://www.math.umbc.edu/~kogan/technical\\_papers/2007/Sinha\\_Kopylev\\_Fox.pdf](http://www.math.umbc.edu/~kogan/technical_papers/2007/Sinha_Kopylev_Fox.pdf)
- Smith, R. L. (1985) Maximum likelihood estimation in a class of non-regular cases. *Biometrika*. 72:67-90.
- USEPA (2001) *Toxicological Review of Quinoline*. CAS No. 91-22-5, in support of summary information on the Integrated Risk Information System (IRIS).
- USEPA (2002) *Health Assessment of 1,3-Butadiene*. Publ. No. EPA/600/P-98/001F. October, 2002. US Environmental Protection Agency, Office of Research and Development.
- von Mises, R. (1936) La distribution del la plus grande de  $n$  valeurs. *Rev. Math. Union Interbalcanique*. 1:141-160.

## Appendix A

### Algorithm for Multistage GEV Parameter Maximum Likelihood Estimation

Initially, restrictions placed on multistage GEV parameters are determined from restrictions on the multistage Weibull parameters. Freeing or fixing the value of Weibull shape parameter  $c$  has the biggest impact on the restrictions on the GEV parameters, because all GEV parameters are functions of  $c$ . The effect that fixing multistage Weibull parameters has on the multistage GEV parameter constraints is as follows:

Multistage Weibull parameter constraints		Multistage GEV parameter constraints	
		Single parameter	Multi-parameter, non-linear
$c$ free		$0 < \gamma \leq 1$	
	$t_0$ free	$1 < \mu$	$0 < \mu - \gamma^{-1}$
	$t_0$ fixed <sup>4</sup>	$1 + t_0 < \mu$	$\mu - \gamma^{-1} = t_0$
	$\beta_i$ free	$0 \leq b_i$	
	$\beta_i$ fixed <sup>5</sup>	$0 \leq b_i$	$\gamma^{1/\gamma} b_i = \beta_i$
$c$ fixed		$\gamma = 1 / c$	
	$t_0$ free	$c < \mu$	
	$t_0$ fixed	$\mu = t_0 + c$	
	$\beta_i$ free	$0 \leq b_i$	
	$\beta_i$ fixed	$b_i = c^c \beta_i$	

Unless specified otherwise, all optimizations for the multistage GEV parameter MLEs will be carried out using the above GEV parameter constraints.

For fixed  $k$ , the algorithm for numerical maximization of the  $k$ -stage GEV log-likelihood involves multiple steps. The following describes the steps when all parameters in the  $k$ -stage GEV model are free:

1. To create a search grid over the open unit interval  $(0, 1)$ , specify a sequence of  $\gamma$  parameter values

<sup>4</sup> These inequality constraints are not necessary in the implementation because parameter  $\mu$  drops out of the GEV likelihood when  $t_0$  is fixed.

<sup>5</sup> For the multistage Weibull constraint  $\beta_i = 0$ , only the single multistage GEV parameter constraint  $b_i = 0$  is required.

$$\gamma[n] = \frac{1 + n/(N+1)}{2}$$

where  $n \in \{-N, -N+1, \dots, N-1, N\}$ . For each  $n$ , also specify a sequence of  $\mu$  parameter values

$$\mu[m, n] = \gamma^{-1}[n] + t_{\min} \frac{1 + m/(M+1)}{2}$$

where  $m \in \{-M, -M+1, \dots, M-1, M\}$ , in order to create a search grid over the open interval  $(\gamma^{-1}[n], t_{\min} + \gamma^{-1}[n])$ , where

$$t_{\min} = \min(t_{js} \mid r_{js} = F)$$

Use  $[m, n]$  as labels for the  $M \times N$  grid points.

2. Maximize the  $k$ -stage GEV log-likelihood with fixed constraints  $\gamma = \gamma[0]$  and  $\mu = \mu[0, 0]$ . The  $k^{\text{th}}$  order polynomial scale parameter (Pk) GEV log-likelihood is the  $k$ -stage GEV log-likelihood with parameters  $\gamma$  and  $\mu$  fixed. The Pk GEV log-likelihood is (non-strictly) concave, according to Theorem 1 in Appendix D, so the optimization will theoretically converge to a global maximum under the fixed constraints. The starting values for this optimization are determined by the following procedure:

- a. Maximize the GEV log-likelihood for each dosage group  $j = 1, 2, \dots, D$ . That is, separately maximize each term inside the summation over  $j$  in the log-likelihood, with fixed constraints  $\gamma = \gamma[0]$ ,  $\mu = \mu[0, 0]$ , and  $b_i = 0$ ,  $i = 1, 2, \dots, k$ , and  $b_0$  free. The scale parameter (P0) GEV log-likelihood is the  $k$ -stage GEV log-likelihood with parameters  $\gamma$  and  $\mu$  fixed, and  $b_i = 0$  for  $i = 1, 2, \dots, k$ .

For dosage groups  $j$  where no incidental (I) or fatal (F) tumor contexts are observed, set  $y_j = 0$ . Otherwise, set

$$y_j = \hat{b}_0[j]$$

where  $\hat{b}_0[j]$  is the global maximum of the P0 GEV log-likelihood for dosage group  $j$ . Note that according to Theorem 2 in Appendix D, the P0 GEV log-likelihood is a strictly concave function for  $b_0 > 0$ ; hence, an incrementally small (positive) starting value for  $b_0$  should be sufficient to ensure convergence of the optimization.

- b. Regress the P0 GEV parameter estimates onto a  $k^{\text{th}}$  degree polynomial of the dosage rates using least-squares. The “data” for this regression consists of pairs  $\{(y_j, d_j); j = 1, 2, \dots, D\}$  and the regression model is of the form

$$y = \sum_{i=0}^k b_i d^i$$

Defining the response vector  $\underline{y} = (y_1, y_2, \dots, y_D)^T$  and the design matrix  $\{X[j, i] = d_j^i; j = 1, 2, \dots, D, \text{ and } i = 0, 1, \dots, k\}$ , the starting values for optimization of the  $Pk$  GEV log-likelihood at  $[0, 0]$  is therefore

$$\underline{\beta}^0 = (X^T X)^{-1} X^T y$$

3. Maximize the  $Pk$  GEV log-likelihood for  $m = 0$ , with fixed constraints  $\gamma = \gamma[n]$  and  $\mu = \mu[0, n]$ . Carry out the optimizations in sequence order  $n = 1, 2, \dots, N$ , followed by  $n = -1, -2, \dots, -N$ . If the  $Pk$  GEV log-likelihood attains a value of  $\hat{l}[0, n]$  at maximum  $\hat{b}_i[0, n]$ ,  $i = 0, 1, \dots, k$ , then the starting values for the optimization at grid points  $[0, n]$  are

$$\begin{aligned} (\gamma[n])^{1/\gamma[n]} \times (\gamma[n-1])^{1/\gamma[n-1]} \times \hat{b}_i[0, n-1], i = 0, 1, \dots, k & \quad \text{for } n > 0, \text{ and} \\ (\gamma[n])^{1/\gamma[n]} \times (\gamma[n+1])^{1/\gamma[n+1]} \times \hat{b}_i[0, n+1], i = 0, 1, \dots, k & \quad \text{for } n < 0 \end{aligned}$$

4. Maximize the  $Pk$  GEV log-likelihood for  $m \neq 0$ , with fixed constraints  $\gamma = \gamma[n]$  and  $\mu = \mu[m, n]$ . For each  $n$ , carry out the optimizations in sequence order  $m = 1, 2, \dots, M$ , followed by  $m = -1, -2, \dots, -M$ . If the  $Pk$  GEV log-likelihood attains a value of  $\hat{l}[m, n]$  at maximum  $\hat{b}_i[m, n]$ ,  $i = 0, 1, \dots, k$ , then the starting values for the optimization at grid points  $[m, n]$  are

$$\begin{aligned} \hat{b}_i[m-1, n], i = 0, 1, \dots, k & \quad \text{for } m > 0, \text{ and} \\ \hat{b}_i[m+1, n], i = 0, 1, \dots, k & \quad \text{for } m < 0. \end{aligned}$$

5. Select grid point(s)  $[m, n]$  that attain(s) the largest value of the  $Pk$  GEV log-likelihood. Maximize the  $k$ -stage GEV log likelihood, with the selected gridpoint values as starting values in the first run of the optimization. Carry out a second run of the optimization to refine the maximization. (If multiple sets of starting values are used, only carry out a second run with the set that attains the largest value of the  $k$ -stage GEV log-likelihood after the first run.)

Only minor modifications to the above steps are required to account for fixed model parameters (and lower stage models).



## Appendix B

### **Algorithm for Estimation of Profile Likelihood Confidence Interval**

To evaluate the profile log-likelihood function  $l_p(\theta)$  for a parameter  $\theta \in \Theta$ , the SQP method is used to maximize the  $k$ -stage GEV log-likelihood, with the defining polynomial equation for  $\theta$  as a constraint in the optimization:

#### BMD for Fatal Added Risk

$$\sum_{i=1}^k b_i \theta^i + [1 + \gamma(t - \mu)]^{1/\gamma} \ln[1 - BMR \times \exp\{[1 + \gamma(t - \mu)]^{1/\gamma}\} b_0] = 0$$

#### BMD for Fatal Extra Risk

$$\sum_{i=1}^k b_i \theta^i + [1 + \gamma(t - \mu)]^{1/\gamma} \ln[1 - BMR] = 0$$

(For incidental added and extra risk, set  $\mu = \gamma^{-1}$  in the above equations.)

Before describing the algorithm, the following definitions and labels are required:

- Label the parameter search interval as  $(\theta_{\inf}, \theta_{\sup})$ , where the algorithm will search values between  $\theta_{\inf}$  and  $\theta_{\sup}$  for the profile likelihood confidence interval<sup>6</sup>:

#### BMD Added Risk

$$(\theta_{\inf}, \theta_{\sup}) = (0, 1).$$

#### BMD Extra Risk

$$(\theta_{\inf}, \theta_{\sup}) = (0, m\hat{\theta}), \text{ where integer } m > 2 \text{ is selected by the user.}$$

- Label the free parameters in the multistage GEV model as  $\underline{\omega} \in \Omega$ , so that the elements of  $\underline{\omega}$  represents a subset of the parameters  $\{\gamma, \mu, b_0, b_1, \dots, b_k\}$ .
- For any integer value  $n$  and positive integer value  $N$ , define the following sequence:

$$\theta[n, N] = \hat{\theta} \times (1 + n2^{-N})$$

Note that  $\theta[0, N] = \hat{\theta}$ , the MLE of  $\theta$ . Label the minimum and maximum allowable values of  $N$ , which determines the maximum and minimum allowable stepsizes, as  $N_{\min}$  and  $N_{\max}$ , respectively. In the current setting,  $N_{\min} = 2$  and  $N_{\max} = 22$ .

---

<sup>6</sup> The domains of profile log-likelihood functions are not necessarily well defined. The intervals given here are selected to ensure a fairly complete search of the entire domain for the bounds.

- Define  $\hat{\omega}[n, N]$  as the optimized values of the free parameters when the profile log-likelihood is evaluated at  $\theta[n, N]$ . Note that  $\hat{\omega}[0, N]$  are the MLEs of the free parameters.
- The *cut-off value*  $\Psi$  is defined as

$$\Psi = \hat{l}(\hat{\theta}) - \chi_{1,1-\alpha}^2 / 2$$

where  $\chi_{v,p}^2$  is the  $p$ th percentile of the  $\chi^2$  distribution with  $v$  degrees of freedom.

Note that the  $100(1 - \alpha)\%$  profile likelihood confidence region is defined as

$$\{\theta \in \Theta \mid \hat{l}(\theta) > \Psi\}$$

Then the profile likelihood  $100(1 - \alpha)\%$  confidence interval bounds for the  $k$ -stage Weibull time-to-tumor model are calculated using the following algorithm:

1. To determine an appropriate initial step size, start with  $N = N_{\min}$ , attempt to calculate the profile log-likelihood at  $\theta[-1, N]$ , using the MLEs  $\hat{\omega}[0, N]$  as starting values. Increment  $N$  by 1 up to  $N_{\max}$ , until the constrained optimization for  $l_p(\theta[-1, N])$  converges. Continue incrementing  $N$  by 1 up to  $N_{\max}$ , until the constrained optimization for calculating  $l_p(\theta[1, N])$ , using  $\hat{\omega}[0, N]$  as starting values, also converges. Label the final value of  $N$  as  $N_{\text{init}}$ . If convergence of the constrained optimization cannot be achieved at both  $\theta[-1, N]$  and  $\theta[1, N]$  before reaching the minimum stepsize (i.e.,  $N_{\text{init}} = N_{\max}$ ), then abort the algorithm and return an error message.
2. To calculate the lower profile likelihood confidence bound:
  - a. Start by initializing  $n = 2$ , and  $N = N_{\text{init}}$ .
  - b. Attempt to calculate the profile log-likelihood at  $\theta[-n, N]$ , using  $\hat{\omega}[-n+1, N]$  as starting values. Halve the stepsize (i.e., increase  $N$  by 1 up to  $N_{\max}$ ), and let  $n = 2n - 1$ , until the constrained optimization for calculating  $l_p(\theta[-n, N])$  converges.

If the constrained optimization for  $\hat{l}(\theta[-n, N])$  converges, then

- A. If  $\hat{l}(\theta[-n, N]) < \Psi$ , and  $\hat{l}(\theta[-(n-1), N]) \geq \Psi$  (i.e., the profile log-likelihood crosses the cut-off value), then set

$$\{\theta_{\min}, \hat{\omega}_{\min}\} = \{\theta[-n, N], \hat{\omega}[-n, N]\}$$

and

$$\{\theta_{\max}, \hat{\omega}_{\max}\} = \{\theta[-n+1, N], \hat{\omega}[-(n-1), N]\}.$$

- B. If  $N \leq N_{\max}$ , and  $\theta[-n, N] > \theta_{\inf}$ , then increment  $n$  by 1, and redo step b. Otherwise, proceed to step c.

Otherwise proceed to step c.

- c. If the last calculated value of the profile log-likelihood (i.e., constrained optimization converged is less than the cut-off value  $\Psi$ ), then carry out a root search of the function  $\hat{l}(\theta) - \Psi$  in the interval  $(\theta_{\min}, \theta_{\max})$ . The root is the lower confidence bound.

(Note: Because BMDS uses a binary search algorithm, the profile log-likelihood is calculated at the midpoints of an interval by using the mean of the optimized free parameters at the endpoints as starting values in the constrained optimization. For example, to calculate  $\hat{l}([\theta_{\min} + \theta_{\max}]/2)$  in the first step of the binary search, use  $(\underline{\omega}_{\min} + \underline{\omega}_{\max})/2$  as the starting values in the constrained optimization.)

Otherwise, the lower confidence bound is less than the final value of  $\theta[-n, N]$  at which the profile log-likelihood was calculated (i.e., the constrained optimization converged).

3. To calculate the upper profile likelihood confidence bound, carry out steps similar to the lower profile likelihood confidence bound in step 2, except that the root search is carried out for values of  $\theta[n, N] < \theta_{\sup}$ , rather than  $\theta[-n, N] > \theta_{\inf}$ .

## Appendix C

### Derivatives for Log-Likelihood Functions

#### C.1 First-Order Derivatives of the Multistage GEV Log-likelihood

$$\begin{aligned}
 \frac{\partial l}{\partial \gamma} = & \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} \frac{(\gamma t_{js})^{1/\gamma} (1 - \ln[\gamma t_{js}])}{\gamma^2} \sum_{i=0}^k b_i d_j^i \right. \\
 & - \sum_{r_{js}=F} n_{js} \left( \frac{g_{js}^{1/\gamma} (\gamma[t_{js} - \mu]/g_{js} - \ln g_{js})}{\gamma^2} \sum_{i=0}^k b_i d_j^i + \frac{\ln g_{js}}{\gamma^2} - \frac{(\gamma^{-1} - 1)(t_{js} - \mu)}{g_{js}} \right) \\
 & - \sum_{r_{js}=I} n_{js} \left( - \frac{G_{js}^{-1} (\gamma t_{js})^{1/\gamma} (1 - \ln[\gamma t_{js}])}{\gamma^2} + \frac{(1 + G_{js}^{-1}) g_{js}^{1/\gamma} (\gamma[t_{js} - \mu]/g_{js} - \ln g_{js})}{\gamma^2} \right) \sum_{i=0}^k b_i d_j^i \\
 & \left. - \sum_{r_{js}=U} n_{js} \frac{g_{js}^{1/\gamma} (\gamma[t_{js} - \mu]/g_{js} - \ln g_{js})}{\gamma^2} \sum_{i=0}^k b_i d_j^i \right] \\
 \\
 \frac{\partial l}{\partial \mu} = & \sum_{j=1}^D \left[ \sum_{r_{js}=F} n_{js} \left( g_{js}^{1/\gamma-1} \sum_{i=0}^k b_i d_j^i + \frac{\gamma-1}{g_{js}} \right) + \sum_{r_{js}=I} (1 + G_{js}^{-1}) n_{js} g_{js}^{1/\gamma-1} \sum_{i=0}^k b_i d_j^i + \sum_{r_{js}=U} n_{js} g_{js}^{1/\gamma-1} \sum_{i=0}^k b_i d_j^i \right] \\
 \\
 \frac{\partial l}{\partial b_r} = & \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} (\gamma t_{js})^{1/\gamma-1} d_j^r - \sum_{r_{js}=F} n_{js} \left( g_{js}^{1/\gamma} - \left[ \sum_{i=0}^k b_i d_j^i \right]^{-1} \right) d_j^r \right. \\
 & \left. - \sum_{r_{js}=I} n_{js} \left( - G_{js}^{-1} (\gamma t_{js})^{1/\gamma-1} + (1 + G_{js}^{-1}) g_{js}^{1/\gamma} \right) d_j^r - \sum_{r_{js}=U} n_{js} g_{js}^{1/\gamma} d_j^r \right] \quad \text{for all } r = 0 \text{ to } k
 \end{aligned}$$

where  $g_{js} = 1 + \gamma(t_{js} - \mu)$ , and  $G_{js} = \exp\{[(\gamma t_{js})^{1/\gamma} - g_{js}^{1/\gamma}] \sum_i b_i d_j^i\} - 1$ .

#### C.2 Second-Order Derivatives of the Multistage Weibull Log-likelihood

$$\frac{\partial^2 l}{\partial c^2} = \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} t_{js}^c (\ln t_{js})^2 \sum_{i=0}^k \beta_i d_j^i - \sum_{r_{js}=F} n_{js} \left( \tau_{js}^c (\ln \tau_{js})^2 \sum_{i=0}^k \beta_i d_j^i + c^{-2} \right) \right]$$

$$\begin{aligned}
& - \sum_{r_{js}=I} n_{js} \left( \tau_{js}^c (\ln \tau_{js})^2 \sum_{i=0}^k \beta_i d_j^i - \frac{t_{js}^c (\ln t_{js})^2 - \tau_{js}^c (\ln \tau_{js})^2}{G_{js}} \sum_{i=0}^k \beta_i d_j^i + (G_{js} + 1) \left[ \frac{t_{js}^c (\ln t_{js})^2 - \tau_{js}^c (\ln \tau_{js})^2}{G_{jk}} \sum_{i=0}^k \beta_i d_j^i \right]^2 \right) \\
& - \sum_{rs=U} n_{js} \tau_{js}^c (\ln \tau_{js})^2 \sum_{i=0}^k \beta_i d_j^i \Big]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial c \partial t_0} &= \sum_{j=1}^D \left[ \sum_{r_{js}=F} n_{js} \left( \tau_{js}^{c-1} (1 + c \ln \tau_{js}) \sum_{i=0}^k \beta_i d_j^i - \tau_{js}^{-1} \right) \right. \\
& + \sum_{r_{js}=I} n_{js} \left( \tau_{js}^{c-1} (1 + c \ln \tau_{js}) (1 + G_{js}^{-1}) \sum_{i=0}^k \beta_i d_j^i - c \tau_{js}^{c-1} (t_{js}^c \ln t_{js} - \tau_{js}^c \ln \tau_{js}) (G_{js} + 1) \left[ \frac{1}{G_{js}} \sum_{i=0}^k \beta_i d_j^i \right]^2 \right) \\
& \left. + \sum_{r_{js}=U} n_{js} \tau_{js}^{c-1} (1 + c \ln \tau_{js}) \sum_{i=0}^k \beta_i d_j^i \right]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial c \partial \beta_r} &= \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} t_{js}^c (\ln t_{js}) d_j^r - \sum_{r_{js}=F} n_{js} \tau_{js}^c (\ln \tau_{js}) d_j^r \right. \\
& - \sum_{r_{js}=I} n_{js} \left( \tau_{js}^c \ln \tau_{js} - \frac{t_{js}^c \ln t_{js} - \tau_{js}^c \ln \tau_{js}}{G_{js}} + (G_{js} + 1) \frac{(t_{js}^c - \tau_{js}^c)(t_{js}^c \ln t_{js} - \tau_{js}^c \ln \tau_{js})}{G_{js}^2} \sum_{i=0}^k \beta_i d_j^i \right) d_j^r \\
& \left. - \sum_{r_{js}=U} n_{js} \tau_{js}^c (\ln \tau_{js}) d_j^r \right] \quad \text{for all } r = 0 \text{ to } k
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial t_o^2} &= \sum_{j=1}^D \left[ - \sum_{r_{js}=F} n_{js} \left( c(c-1) \tau_{js}^{c-2} \sum_{i=0}^k \beta_i d_j^i + (c-1) \tau_{js}^{-2} \right) - \sum_{r_{js}=I} n_{js} \left( c(c-1) \tau_{js}^{c-1} (1 + G_{js}^{-1}) \sum_{i=0}^k \beta_i d_j^i + (G_{js} + 1) \left[ \frac{c \tau_{js}^{c-1}}{G_{js}} \sum_{i=0}^k \beta_i d_j^i \right]^2 \right) \right. \\
& \left. - \sum_{r_{js}=U} n_{js} c(c-1) \tau_{js}^{c-2} \sum_{i=0}^k \beta_i d_j^i \right]
\end{aligned}$$

$$\frac{\partial^2 l}{\partial t_0 \partial \beta_r} = \sum_{j=1}^D \left[ \sum_{r_{js}=F} n_{js} c \tau_{js}^{c-1} d_j^r - \sum_{r_{js}=I} n_{js} \left( c \tau_{js}^{c-1} (1 + G_{js}^{-1}) + (G_{js} + 1) \frac{c \tau_{js}^{c-1}}{G_{js}} (t_{js}^c - \tau_{js}^c) \sum_{i=0}^k \beta_i d_j^i \right) d_j^r - \sum_{r_{js}=U} n_{js} c \tau_{js}^{c-1} d_j^r \right] \quad \text{for all } r = 0 \text{ to } k$$

$$\frac{\partial^2 l}{\partial \beta_r \partial \beta_u} = \sum_{j=1}^D \left[ - \sum_{r_{js}=F} n_{js} \left( \sum_{i=0}^k \beta_i d_j^i \right)^{-2} d_j^r d_j^u - \sum_{r_{js}=I} n_{js} (G_{js} + 1) \left( \frac{t_{js}^c - \tau_{js}^c}{G_{js}} \right)^2 d_j^r d_j^u \right] \quad \text{for all } r = 0 \text{ to } k, u = 0 \text{ to } k$$

where  $\tau_{js} = t_{js} - t_0$ , and  $G_{js} = \exp\left\{\left[t_{js}^c - \tau_{js}^c\right] \sum_i \beta_i d_j^i\right\} - 1$

## Appendix D

### Concavity Theorems

**Theorem 1.** The GEV log-likelihood function with  $\mu$  and  $\gamma$  fixed in the domain is a (non-strictly) concave function of  $b_0, b_1, \dots, b_k$ .

Proof: The proof uses the following well-known (directional-derivative) theorem for convexity (concavity):

Let  $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$  be a twice-differentiable function over the convex domain  $D(f) \subset \mathfrak{R}^n$ . Then  $f$  is convex (concave) iff

$$\frac{d^2 f(\underline{x} + a\underline{x}')}{da^2} \geq (\leq) 0$$

for all  $\underline{x} \in \mathfrak{R}^n, \underline{x}' \in \mathfrak{R}^n / \{0\}, a \in \mathfrak{R}$ , such that  $\underline{x}, \underline{x} + a\underline{x}' \in D(f)$ .

Now, let  $\underline{b} = (b_0, b_1, \dots, b_k)$  and  $\underline{b}' = (b'_0, b'_1, \dots, b'_k)$  such that  $b_i, b'_i \geq 0$  for  $i = 0, 1, \dots, k$ . Then

$$\begin{aligned} \frac{dl(\underline{b} + a\underline{b}')}{da} &= \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} (\gamma t_{js})^{1/\gamma} \sum_{i=0}^k b'_i d_j^i \right. \\ &\quad - \sum_{r_{js}=F} n_{js} \left( [1 + \gamma(t_{js} - \mu)]^{1/\gamma} - \left[ \sum_{i=0}^k (b_i + ab'_i) d_j^i \right]^{-1} \right) \sum_{i=0}^k b'_i d_j^i \\ &\quad - \sum_{r_{js}=I} n_{js} \left( [1 - \gamma(t_{js} - \mu)]^{1/\gamma} - h_{js} \left[ \exp \left\{ h_{js} \sum_{i=0}^k (b_i + ab'_i) d_j^i \right\} - 1 \right]^{-1} \right) \sum_{i=0}^k b'_i d_j^i \\ &\quad \left. - \sum_{r_{js}=U} n_{js} [1 + \gamma(t_{js} - \mu)]^{1/\gamma} \sum_{i=0}^k b'_i d_j^i \right] \\ \Rightarrow \frac{d^2 l(\underline{b} + a\underline{b}')}{da^2} &= \sum_{j=1}^D \left[ - \sum_{r_{js}=F} n_{js} \left( \sum_{i=0}^k b'_i d_j^i / \sum_{i=0}^k (b_i + ab'_i) d_j^i \right)^2 \right. \\ &\quad \left. - \sum_{r_{js}=I} n_{js} \left( h_{js} \sum_{i=0}^k b'_i d_j^i / \left[ \exp \left\{ h_{js} \sum_{i=0}^k (b_i + ab'_i) d_j^i \right\} - 1 \right] \right)^2 \exp \left\{ h_{js} \sum_{i=0}^k (b_i + ab'_i) d_j^i \right\} \right] \\ &\leq 0 \quad \text{if } b_i + ab'_i \geq 0 \text{ for } i = 0, 1, \dots, k \\ \text{where } h_{js} &= (\gamma t_{js})^{1/\gamma} - (1 + \gamma[t_{js} - \mu])^{1/\gamma}. \blacksquare \end{aligned}$$

**Theorem 2.** The GEV log-likelihood with  $\mu$  and  $\gamma$  fixed (with  $\gamma^{-1} < \mu$ ) and  $b_i = 0$  for  $i = 1, 2, \dots, k$  is a strictly concave function for  $b_0 > 0$  if fatal ( $F$ ) or incidental ( $I$ ) tumor contexts are observed.

Proof:

$$\begin{aligned}
\frac{dl}{db_0} &= \sum_{j=1}^D \left[ - \sum_{r_{js}=C} n_{js} (\gamma t_{js})^{1/\gamma} - \sum_{r_{js}=F} n_{js} ([1 + \gamma(t_{js} - \mu)]^{1/\gamma} - b_0^{-1}) \right. \\
&\quad \left. - \sum_{r_{js}=I} n_{js} [1 + \gamma(t_{js} - \mu)]^{1/\gamma} - \frac{h_{js}}{\exp\{h_{js}b_0\} - 1} - \sum_{r_{js}=U} n_{js} [1 + \gamma(t_{js} - \mu)]^{1/\gamma} \right] \\
\Rightarrow \frac{d^2l}{db_0^2} &= \sum_{j=1}^D \left[ - \sum_{r_{js}=F} n_{js} b_0^{-2} - \sum_{r_{js}=I} n_{js} \left( \frac{h_{js}}{\exp\{h_{js}b_0\} - 1} \right)^2 \exp\{h_{js}b_0\} \right] \\
&< 0 \\
\text{for } b_0 > 0, \text{ where } h_{js} &= (\gamma t_{js})^{1/\gamma} - (1 + \gamma[t_{js} - \mu])^{1/\gamma} > 0, \text{ for } \gamma^{-1} < \mu. \blacksquare
\end{aligned}$$

## **Appendix E**

### **Algorithm for Calculating Predicted Response Counts**

The probability of response under the Fatal tumor model is given by

$$P_F(t, d) = P(t, d, t_0, c, \beta_0, \beta_1, \dots, \beta_k) = \begin{cases} 1 - \exp\left\{-(t - t_0)^c \sum_{i=0}^k \beta_i d^i\right\} & \text{if } t > t_0 \\ 0 & \text{otherwise} \end{cases}$$

and the probability of response under the Incidental tumor model is given by

$$P_I(t, d) = P(t, d, c, \beta_0, \beta_1, \dots, \beta_k) = 1 - \exp\left\{-t^c \sum_{i=0}^k \beta_i d^i\right\}$$

Predicted response counts at any given dose for Fatal/Incidental tumor models, are calculated by the following equation:

$$\text{Pred}(d) = \sum_{i=1}^n N(d, t_i) \times P(t_i, d, c, \beta_0, \beta_1, \dots, \beta_k)$$

In this equation,  $N(d, t_i)$  specifies the total number of observations at a given dose  $d$  and time  $t_i$ , and is given by

$$N(d, t_i) = N_F(d, t_i) + N_I(d, t_i) + N_C(d, t_i)$$

where,  $N_F(d, t_i)$ ,  $N_I(d, t_i)$ , and  $N_C(d, t_i)$  are the number of Fatal, Incidental, and Censored response tumor contexts at dose  $d$ , and time  $t_i$  respectively. Animals classified as having an unknown outcome (context “U”) are not counted.

The probability of tumor response,  $P(t_i, d, c, \beta_0, \dots, \beta_k)$ , is given by either  $P_F(t, d)$  or  $P_I(t, d)$ , depending on tumor model specified by the user (Fatal/Incidental) within the input (\*.d) file.

Note that  $P_I(t, d)$  is the cumulative probability of tumor “onset” up to time  $t$  and therefore includes both animals for which the tumor caused death (classified as “F”) and animals with tumors observed incidental to another cause of death (classified as “I”), i.e., all animals having the tumor.



### Example calculation

Input data for Fatal tumor model

DOSE CLASS TIMEN

6	C	23	1
6	C	66	4
6	I	66	6
6	F	74	1

Number of tumor contexts at dose = 6 is calculated as follows,

$$N(d = 6, t = 23) = 1$$

$$N(d = 6, t = 66) = 10 (6+4)$$

$$N(d = 6, t = 74) = 1$$

Predicted Response at dose = 6 is give by

$$\begin{aligned} \Pr(d = 6) = & N(d = 6, t = 23) * P(\text{Fatal} \mid d = 6, t = 23) + \\ & N(d = 6, t = 66) * P(\text{Fatal} \mid d = 6, t = 66) + \\ & N(d = 6, t = 74) * P(\text{Fatal} \mid d = 6, t = 74) \end{aligned}$$