Published in final edited form as:
Environ Ecol Stat. 2009 March 1; 16(1): 63–73. doi:10.1007/s10651-007-0073-5.

Bootstrap methods for simultaneous benchmark analysis with quantal response data

R. Webster West,
Department of Statistics, Texas A&M University, College Station, TX 77843, USA

Daniela K. Nitcheva, and
Division of Biostatistics, South Carolina Department of Health and Environmental Control, Columbia, SC 29201, USA

Walter W. Piegorsch
Department of Mathematics and BIO5 Institute, University of Arizona, Tucson, AZ 85721, USA

Abstract

A primary objective in quantitative risk assessment is the characterization of risk which is defined to be the likelihood of an adverse effect caused by an environmental toxin or chemcial agent. In modern risk-benchmark analysis, attention centers on the "benchmark dose" at which a fixed benchmark level of risk is achieved, with a lower confidence limits on this dose being of primary interest. In practice, a range of benchmark risks may be under study, so that the individual lower confidence limits on benchmark dose must be corrected for simultaneity in order to maintain a specified overall level of confidence. For the case of quantal data, simultaneous methods have been constructed that appeal to the large sample normality of parameter estimates. The suitability of these methods for use with small sample sizes will be considered. A new bootstrap technique is proposed as an alternative to the large sample methodology. This technique is evaluated via a simulation study and examples from environmental toxicology.

Keywords

Benchmark dose; Bootstrap; Multistage model; Quantal data; Quantitative risk assessment; Simultaneous inferences

1 Introduction

Dose-response studies are conducted in the quantitative risk assessment of environmental toxins or other chemical agents in an attempt to specify the probability of an adverse response as a function of the amount of exposure to the agent. The risk, $R(d)$, is defined as the probability that a subject exposed to a dose, d, of a hazardous agent will develop a particular adverse outcome. Potential models for $R(d)$ depend on the nature of the outcomes measured in the study. For quantitative outcomes such as weight loss, standard regression models coupled with a definition of an adverse outcome lead to straightforward models for risk (Kodell and West 1993). For quantal outcomes, such as whether or not a subject develops a tumor, the choice of risk model is somewhat more subjective. A very common model is $R(d) = 1 - \exp\{-\theta(d)\}$ which is a simplified version of the well-known Armitage-Doll multistage model (Armitage and Doll

e-mail: west@stat.tamu.edu.

1954). In this formulation, $\theta(d)$ is typically chosen as a low-order polynomial in d such as the quadratic function, $\theta(d) = \beta_0 + \beta_1 d + \beta_2 d^2$ (Guess and Crump 1976; Krewski and van Ryzin 1981; Bailer and Smith 1994). Since it can correspond to a simple two-stage progression to the observed adverse effect, this particular model is often times called the two-stage model. The parameters in the model β_j, $j = 0, 1, 2$, are typically required to be non-negative so that risk is a monotone increasing function of dose when $d \geq 0$.

Background risk is defined as the risk for the control population, $R(0)$, which has no exposure to the agent. The *extra risk function* is defined as the risk above this background level corrected for non-response in the unexposed population: $R_E(d) = \{R(d) - R(0)\}/\{1 - R(0)\}$. Under the two-stage model, $R_E(d) = 1 - \exp\{-\beta_1 d - \beta_2 d^2\}$. Statistical methods for obtaining upper confidence limits on extra risk at a given dose are of interest. Also of interest are methods for obtaining lower confidence limits on the dose level at which a certain excess risk, called a benchmark risk (BMR), is achieved. This dose level is known as a *Benchmark Dose*, or BMD (Crump 1984), while a lower confidence limit on the benchmark dose is called the benchmark dose lower limit, or BMDL (Crump 1995). Identifying BMDLs is of specific interest for regulatory purposes, as these values are sometimes used for setting occupational or environmental exposure criteria. Often a risk regulator has a range of BMRs in mind for a given data set; e.g., one may desire BMDLs at a variety of BMR values for excess risk between 0.01 and 0.10 (Faustman and Bartell 1997; Schlosser et al. 2003). The corresponding sequence of BMDLs at each of these values of extra risk will have a joint confidence level which may be far below the nominal level for each individual BMDL. Using the sequence of BMDLs without adjustment is an improper inferential technique and leads to a suspect regulatory decision.

For the two-stage model, Al-Saidy et al. (2003) proposed a method to obtain simultaneous upper confidence bands for $R_E(d)$ over an interval $0 < d < B$ where the value of $B > 0$ is specified a priori. The method employs a modification of Scheffé's S-method for building confidence bands (Pan et al. 2003), and hence is referred to as the S-method. Nitcheva et al. (2005) performed a formal comparison of the S-method with various other techniques for constructing multiplicity-adjusted inferences on sets of one to five BMR values. The S-method should in this case be conservative since it is designed to provide coverage over an entire interval of dose values. In their study, the methods were compared in terms of coverage probability and the median absolute difference between the BMDLs and the true benchmark doses. The S-method and a Bonferroni-adjusted likelihood ratio technique exhibited superior performance in Nitcheva et al.'s comparisons, in that both methods tended to provide a coverage probability above the nominal level and relatively tight BMDLs. While the likelihood ratio technique often had slightly tighter BMDLs, the S-method possessed the distinct advantage that its BMR values could be chosen post hoc. Both methods, however, tended to provide very conservative coverage probabilities.

Guess and Crump (1976) argue that the maximum likelihood estimators (MLEs) for the two stage model parameters have an asymptotic normal distribution when $\beta_j > 0$, for all j. The S-method appeals to the asymptotic normality of the parameter estimates in that estimates of the asymptotic variances and covariance are required for the construction of the confidence band. The conservatism of the S-method described above appears in large part due to an inadequate asymptotic approximation when applied to small samples. Herein, new methods to address this small-sample conservatism will be proposed. In Sect. 2, a new bootstrap technique is developed in an attempt to better incorporate small sample variation into the confidence band construction. The properties of this method are evaluated in Sect. 3 via a simulation study, and the method is then applied to several real examples in Sect. 4. A short discussion is provided in Sect. 5.

2 A bootstrap approach

At the *i*th dose d_i ($i = 1, \ldots, n$) the number of subjects exhibiting an adverse effect, Y_i, is recorded. Assume that the Y_i s are independent, binomial variates with parameters N_i and $R(d_i)$, where N_i is the number of subjects tested at dose d_i and $R(d_i)$ models the unknown probability that a subject will respond adversely. Under the two-stage model the MLEs, $\mathbf{b} = [b_0 b_1 b_2]^T$, of the unknown parameters, $\beta = [\beta_0 \beta_1 \beta_2]^T$, are constrained optimization of the corresponding binomial likelihood function. The MLEs of risk and extra risk are then simply $\widehat{R}(d) = 1 - \exp\left\{-b_0 - b_1 d - b_2 d^2\right\}$ and $\widehat{R}_E(d) = 1 - \exp\left\{-b_1 d - b_2 d^2\right\}$, respectively.

With a benchmark analysis, a 100(1 - α)% upper confidence band on $R_E(d)$ is desired over a range of dose values. As in Al-Saidy et al. (2003), it will be assumed that the dose range of interest is $0 < d < B$. Note that the extra risk is a function of the linear predictor $\eta(d) = \beta_1 + \beta_2 d$; i.e., $R_E(d) = 1 - \exp\{-d\eta(d)\}$. Hence, an upper confidence band on $R_E(d)$ can be easily constructed from an upper confidence band on $\eta(d)$. In other words, if a function $\eta_U(d)$ can be defined such that

$$\Pr\left(\eta(d) \le \eta_U(d); 0<d<B\right) = 1 - \alpha, \tag{1}$$

then

$$\Pr\left(R_E(d) \le 1 - \exp\left\{-d\eta_U(d)\right\}; 0<d<B\right) = 1 - \alpha.$$

Therefore, $1 - \exp\{-d\eta_U(d)\}$ serves as a 100(1 - α)% upper confidence band on $R_E(d)$ for $0 < d <$ B. This upper confidence band-can be inverted to obtain a lower confidence band on benchmark dose over a range of BMR values. At any specific BMR value, the corresponding 100(1 - α)% BMDL is defined as the smallest positive solution of

$$d\eta_U(d) + \log(1 - \mathrm{BMR}) = 0.$$

Focusing on obtaining an upper confidence band on $\eta(d)$, define $\eta(d,\alpha)$ as any pointwise upper confidence limit on $\eta(d)$ that satisfies $\Pr(\eta(d) \le \eta(d,\alpha)) = 1 - \alpha$. For example, $\eta(0, \alpha_1)$ represents a 100(1 - α_1)% upper confidence limit on $\eta(0)$. Similarly, $\eta(B, \alpha_2)$ represents a 100(1 - α_2)% upper confidence limit on $\eta(B)$. The line connecting these two upper confidence limits is given by

$$\eta_U(d) = \eta(0,\alpha_1) + \left(\eta(B,\alpha_2) - \eta(0,\alpha_1)\right) d/B. \tag{2}$$

Since $\eta_U(d)$ is linear, $\eta(d) \le \eta_U(d)$ over $0 < d < B$ holds if and only if $\eta(0) \le \eta_U(0)$ and $\eta(B) \le \eta_U(B)$. If α_1 and α_2 lie between 0 and 1 such that $\alpha_1 + \alpha_2 = \alpha$, we have that

$$\Pr\left(\eta(d) \le \eta_U(d); 0<d<B\right) = \Pr\left(\beta_1 \le \eta(0,\alpha_1), \beta_1+\beta_2 \le \eta(B,\alpha_2)\right) \ge 1 - (\alpha_1+\alpha_2) = 1 - \alpha$$

by Bonferroni's inequality. Thus, this definition of $\eta_U(d)$ leads to an upper confidence band that achieves at least a minimal 1 - α confidence level.

Note that the above approach is quite similar to that of Bowden and Graybill (1966), where proportional width confidence bands for linear regression models were first developed. In their pioneering work, Bowden and Graybill made use of normality assumptions to derive exact 100

(1 - α)% confidence bands for a simple linear regression model. With the two-stage model, however, the normality of maximum likelihood estimates is not guaranteed for smaller samples. Indeed, for smaller samples, the MLEs of β_1 and β_2 often appear very non-normal, due to the parameter constraints which lead to sampling distributions with point masses at zero.

Since the exact distributions of the parameter estimates are not known in this case, consider a bootstrap approach for defining the upper confidence limits on $\eta(0)$ and $\eta(B)$. Let $\left\{Y_{1j}^*, \ldots, Y_{nj}^*\right\}$ for $j = 1, \ldots, K$ denote a sequence of K independent boot-strap samples taken with replacement from the observed data. In other words, each Y_{ij}^* is a sample taken with replacement from the Y_i adverse responses and N_i Y_i nonadverse responses observed at the dose d_i. Thus, Y_{ij}^* is a pseudo-binomial random variable with parameters N_i and Y_i/N_i, where Y_i/N_i is the observed proportion of subjects at dose d_i that respond adversely. For each bootstrap sample, one can compute the maximum likelihood estimates of β_1 and β_2 along with the corresponding estimate of the predictor $\widehat{\eta}_j^*(d)$. Thus, the output of the bootstrap sampling process consists of K bootstrap estimates of the linear predictor, $\left\{\widehat{\eta}_1^*(d), \ldots, \widehat{\eta}_K^*(d)\right\}$. The goal is to define a curve which lies above 100(1 - α)% of these bootstrapped predictor estimates over the range from 0 to B.

Using a bootstrap percentile approach, an approximate 100(1 - α)% upper confidence limit on $\eta(d)$ at a specific dose d, $\eta^*(d, \alpha)$, is the 100(1 -α)th percentile of $\left\{\widehat{\eta}_1^*(d), \ldots, \widehat{\eta}_K^*(d)\right\}$. The approximate 100(1 - α)% bootstrap upper confidence band on $\eta(d)$ over $0 < d < B$ is then given by

$$\eta_U(d) = \eta^*(0, \alpha_1) + (\eta^*(B, \alpha_2) - \eta^*(0, \alpha_1))\, d/B. \tag{3}$$

The bootstrap approach outlined above is nonparametric in that model information is not taken into account when generating the bootstrap samples. This approach can be problematic in practice if none or all of the responses at a particular dose level are adverse, since it leads to observed proportions of zero or one, respectively. In either case, there will be no variability in the bootstrap samples at that dose value. To incorporate more of the underlying variability, a parametric bootstrap where the sample proportion of adverse responses is replaced with the estimated risk from the fitted model can be used in these special cases. This modified, semi-parametric approach does not guarantee a nondegenerate bootstrap sample, but it does make it less likely to occur. While further study on it is warranted, the approach may be more robust than other strategies which, e.g., replace sample proportions of zero and one with values that are arbitrarily close to these endpoints.

3 Simulation results

The simultaneous upper limits on $R_E(d)$ from the bootstrap approach are expected to contain the true $R_E(d)$ for all $d \in (0, B)$ at least 100(1 -α)% of the time. In small samples, however, the coverage characteristics of the method are less certain. To evaluate the bootstrap approach, a Monte Carlo simulation study was conducted where the small-sample empirical coverage was estimated for a variety of two-stage quantal response models. The suite of eight models considered was the same as those used by Al-Saidy et al. (2003) to validate the S-method. The background risk, $R(0)$, in these models ranged between 1% and 30%, and the corresponding risk at the highest dose took values between 5% and 90%. The specific parameter values for the models are given in the first three columns of Table 1. Four dose levels, $d = 0, 0.25, 0.5, 1$, with equal numbers of subjects, $N_i = N$, per dose-group were used in the simulations, corresponding to a common design in cancer risk experimentation (Portier 1994). Values of

N ranged between 25 and 500. The value of *B* required to implement the bootstrap method was set to the largest dose, $d = 1$. For each model configuration, 2,000 pseudo-binomial data sets were simulated, and 2,000 bootstrap samples were generated for each data set.

Table 1 displays the simulation study results. The empirical coverage rates displayed under each sample size in Table 1 were computed by determining the number of times out of the 2,000 simulation runs that the bootstrap upper confidence band on extra risk based on (3) was above the true extra risk function over the entire interval from 0 to *B*. The standard value of $\alpha = 0.05$ was used in the construction of the bands, so that the nominal confidence level associated with the bootstrap confidence bands should be at least 0.95. Notice then that the approximate standard error of the estimated coverage is $\sqrt{(0.05)(0.95)/2000} \approx 0.005$, and it never exceeds $\sqrt{(0.5)(0.5)/2000} = 0.011$. In constructing the upper confidence limits, $\eta^*(0, \alpha_1)$ and $\eta^*(B, \alpha_2)$, the values of $\alpha_1 = \alpha_2 = 0.025$ were used.

Overall, Table 1 shows that the coverage probabilities are very close to the nominal 0.95 level. The most prominent exception is the coverage probability for the first model at $N = 25$. Upon closer inspection, this model was found to be problematic in that for small sample sizes, a high proportion of the simulated responses were often zero across all four doses. Clearly, in practice such data sets would not be subjected to a dose-response analysis. Indeed, Al-Saidy et al. (2003) also found similar coverage instabilities with this sample size for this model in their evaluation of the S-method. In the remaining cases, the bootstrap method produced far more stable and often much less conservative results on average than the S-method as studied by Al-Saidy et al. (2003). The fourth model is the most obvious example of the less conservative nature of the bootstrap method in comparison to the S-method. For this model, the coverage probabilities of the S-method reported by Al-Saidy et al. (2003) were never below 0.978 and were above 0.998 for three of the five sample sizes considered. The bootstrap method only has a high coverage probability at the smallest sample size, with the coverage probabilities being much closer to 0.95 for sample sizes of 50 or larger.

Note that since our simultaneous BMDLs are built by direct inversion of the simultaneous confidence bands based on (3), the results reported in Table 1 for simultaneous coverage of extra risk also represent empirical simultaneous coverages for bounding the BMDs. In order to make a full comparison with the S-method results from Al-Saidy et al. (2003), however, the simultaneous coverage probability associated with individual BMDLs at the commonly used benchmark risk levels of 0.01, 0.05 and 0.10 were also evaluated and are shown in Table 2. For the more shallow risk functions displayed in the first three rows of Table 2, we see very conservative simultaneous coverage rates that are all larger than 0.98. For the more steep risk functions displayed in the remainder of Table 2, however, we see much less conservative coverage rates. For each model, there is no discernable trend in coverage rate across sample size.

4 Examples

The simulation study in Sect. 3 illustrates that the simultaneous coverage of the bootstrap confidence bands tends to be less conservative than that of the S-method. This does not imply, however, that the BMDLs obtained from using the bootstrap method are larger than those from the S-method. It could be the case that the bootstrap band is higher than the S-method at low doses but lower at high doses where coverage violations might tend to occur. In fact, since the bootstrap band is based on a linear limit on $\beta_1 + \beta_2 d$ while the S-method band is based on a hyperbolic limit on $\beta_1 + \beta_2 d$, the two bands on R_E will intersect in at most two points past the obvious intersection at $d = 0$. To further evaluate this matter, the 95% bootstrap upper band on extra risk and the 95% S-method band on extra risk were computed for three liver carcinogenesis data sets found in the literature. The value of *B* was taken as the largest dose in

each study. For the bootstrap method, the upper confidence limits at 0 and *B* were constructed using $\alpha_1 = \alpha_2 = 0.025$.

The first data set considered was the oft-cited Aflatoxin B_1 data on liver carcinogenicity of the mycotoxin contaminant reported by Wogan et al. (1974). The data set consists of the number of male rats exhibiting hepatocellular tumors at each of six doses (measured in μg/kg/day). The raw proportions at the dose values of 0, 0.04, 0.2, 0.6, 2 and 4 are 0/18, 2/22, 1/22, 4/21, 20/25 and 28/28, respectively. (Since the observed proportion at dose zero is zero, we will apply a parametric resampling at this dose level in the analysis below.) The second data set, as reported by Crump et al. (1977), contains the proportions of mice exhibiting liver tumors after exposure to the pesticide dieldrin, at each of four dose levels (measured in ppm). The raw proportions at doses of 0, 1.25, 2.5 and 5 are 17/156, 11/60, 25/58 and 44/60, respectively. The third data set, reported by Janardan (1995), contains the proportions of mice exhibiting liver tumors after exposure to the pesticide DDT, at each of five dose levels (measured in ppm). The raw proportions at doses of 0, 2, 10, 50 and 250 are 4/111, 4/105, 11/124, 13/104 and 60/90, respectively.

The upper band on extra risk for each data set was inverted to compute the associated BMDLs at BMRs ranging from 10^{-5} to 10^{-1}. Results from the analysis of each data set, including comparisons with Al-Saidy et al.'s S-method BMDLs, are shown in Table 3. For the Aflatoxin data, the bootstrap BMDLs are less conservative (i.e., higher) than the S-method BMDLs at all BMR values. The same is true for the dieldrin data with the exception of the highest BMR value of 0.01. For the DDT data, the opposite is true in that the S-method BMDLs are higher than those from the bootstrap method. The BMDL values highlight the fact that the upper confidence bands for both methods are approximately linear over the low dose range. For all three data sets, the upper band for the bootstrap method (not shown) is slightly below the S-method at low doses (very low doses in the case of DDT) and above the S-method band at higher doses.

5 Discussion

Bootstrap methodology for use in quantitative risk assessment has a rich history. Crump and Howe (1985) discussed the use of bootstrap methods for calculating pointwise lower confidence limits in low dose extrapolation for risk assessment. Similar or associated recommendations for use of bootstrap-based inferences in quantitative risk analysis were given by Bailer and Smith (1994), Smith and Sielken (1988), Schulz and Griffin (1999), and Brand et al. (2001).

The methodology discussed herein expands the use of bootstrap methods to the case of simultaneous confidence bands. For the general case of a linear model, Hall and Pittelkow (1990) constructed bootstrap confidence bands under the assumption of a symmetric error distribution. They specifically evaluated both S-method bands and fixed-width bands in conjunction with simple linear predictors. By contrast with the approach proposed herein, they used model information to base their bootstraps on resampling from the residuals of the original model fit. The bootstrap procedure developed herein employs information from the binomial/multi-stage model only when the observed sample proportions are either 0 or 1, and in this sense it has more of a semi-parametric character.

As developed in (3), the linear confidence band on the predictor, $\eta(d)$, is a proportional-width form. Its shape will typically be narrower on the lower end of the dose interval and wider on the higher end of the dose interval. (Here 'width' is defined as the distance from the upper band to the predictor.) Since the S-method is designed to have minimal width at the mean dose level, the proportional-width feature will naturally lead the bootstrap band to be somewhat tighter

than the S-method band over the low dose region. This greater precision, combined with the closer-to-nominal simultaneous coverage probabilities observed in our simulations, makes the bootstrap procedure an enticing alternative to the S-method for quantal response benchmark analysis.

The coverage characteristics are intriguing, since the procedure is based on a conservative Bonferroni adjustment. However, upon closer inspection of our simulations, the linear predictor rarely exceeded both the upper limit at 0 and the upper limit at B. Typically, lines that rise above the upper limit at 0 are different from those that are above the upper limit at *B*. Indeed, in the example data from Sect. 4, such simultaneous conflicts never occurred for the lines constructed for the bootstrap samples. These empirical results suggest that the coverage probability for the bootstrap procedure will often be very close to nominal.

While only the case of $\alpha_1 = \alpha_2 = \alpha/2$ has been considered here, the proportional width feature of the bootstrap band can be further manipulated to produce even narrower limits for doses close to 0: simply make α_1 larger than α_2. Consequent tightness in the band close to 0 will be balanced with greater 'proportional' width at higher doses, although this is problematic since increasing α_1 may increase the probability of making erroneous inference about BMD at lower risk values. One could also alter the overall shape of the band by including intermittent dose values between 0 and *B* in the band construction. This results in a piecewise linear function that connects the upper confidence limits at each of the selected doses. To provide a total coverage of α, adjust the confidence levels at each of the intermittent doses in similar fashion to that employed in (2). Note that addition of too many intermittent points is inadvisable, since it likely will result in a band that takes the highest bootstrap predictor value at each dose, and will clearly be very conservative.

Of course, in practice the two-stage risk model may or may not properly characterize the true risk described by the actual data at hand. Indeed, it is an open question as to how well our bootstrap procedure generalizes to other forms of risk. It may be worthwhile to examine its performance under other risk models, particularly those under which extra risk cannot be expressed as a function of a linear predictor. Along these lines, Piegorsch et al. (2006) studied the effects of model misspecification on the performance of S-method and proportional width bands. The proportional width bands were seen to be more resilient to certain forms of model misspecification and we take from this that a similar analysis for the proposed bootstrap procedure could be a fruitful area of future research.

One could also consider using the bootstrap procedure with a band function that is not linear. Indeed, a hyperbolic function similar to that used in the construction of the S-method band could be considered. One would attempt to identify such a function that covered 95% of the bootstrapped linear predictors over 0 to *B*. As is the case for linear band functions, this function need not be unique, so some form of optimality consideration could be used to define the band. Associated with this, an interesting question for future research would be to define the surface that is "closest" to the estimated linear predictor over the entire range from 0 to B that contains 95% of the bootstrap samples. This would be a piecewise linear function constructed from segments of the bootstrapped predictor values over this range. The attractive aspect of this approach is that the data would be used to determine the shape of the band rather than a band function being imposed a priori. Construction of such a surface is a very complex computational problem, but it may lead to very interesting insights.

Acknowledgments

The authors wish to thank the three referees for their helpful comments. This work was initiated while all the authors were with the University of South Carolina. It was funded under grant #R01-CA76031 from the US National Cancer

Institute and grant #RD-83241901 from the US Environmental Protection Agency. Its contents are solely the responsibility of the authors and do not necessarily reflect the official views of these agencies.

Author Biographies

R. Webster West is a professor in the Department of Statistics at Texas A&M University. He received his Ph.D. in Statistics from rice University in 1994. Since that time, he has been actively developing new statistical methods for application to dose response models used in toxicology.

Daniela K. Nitcheva received a Ph.D. degree in Statistics from the University of South Carolina. In the Fall of 2003, she joined the Department of Epidemiology and Biostatistics at the University of South Carolina as a Research Assistant Professor. She is currently in the Division of Biostatistics, SC DHEC. Her primary research interest is quantitative risk estimation under the multistage model.

Walter W. Piegorsch is the Director of the Graduate Interdisciplinary Program in Statistics at the University of Arizona, Tucson, AZ, and a member of the Research Faculty of the University's BIO5 Institute. He studies modeling and analysis for environmental data, with emphasis on environmental hazards and risk assessment. He also has research interests in geo-spatially referenced disaster informatics, simultaneous inferences, generalized linear models, and the historical development of statistical thought as prompted by problems in the biological and environmental sciences. Among other activities, he has served on the Board of Scientific Counselors for the U.S. National Toxicology Program (2000-2004), on the Council of the International Biometric Society (2002-2005), and as Chairman of the American Statistical Association Section on Statistics & the Environment (2004). He earned his Ph.D. in Statistics at the Biometrics Unit, Cornell University, Ithaca, NY in 1984, after which he spent nine years as a practicing statistician with the US National Institute of Environmental Health Sciences in Research Triangle Park, NC, and then 13 years as a faculty member with the Department of Statistics at the University of South Carolina.

References

Al-Saidy OM, Piegorsch WW, West RW, Nitcheva DK. Confidence bands for low-dose risk estimation with quantal response data. Biometrics 2003;59:1056–1062. [PubMed: 14969485]

Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. Brit J Cancer 1954;8:1–12. [PubMed: 13172380]

Bailer AJ, Smith RJ. Estimating upper confidence limits for extra risk in quantal multistage models. Risk Anal 1994;14:1001–1010. [PubMed: 7846307]

Bowden DC, Graybill FA. Confidence bands of uniform and proportional width for linear models. J Am Stat Assoc 1966;61:182–198.

Brand KP, Catalano PJ, Hammitt JK, Rhomberg L, Evans JS. Limitations to empirical extrapolation studies: the case of BMD ratios. Risk Anal 2001;21:625–640. [PubMed: 11726017]

Crump KS. A new method for determining allowable daily intake. Fund Appl Toxicol 1984;4:854–871.

Crump KS. Calculation of benchmark doses from continuous data. Risk Anal 1995;15:79–89.

Crump, KS.; Howe, R. A review of methods for calculating confidence limits in low dose extrapolation. In: Clayson, DB.; Krewski, D.; Munro, I., editors. Toxicological risk assessment, volume I: biological and statistical criteria. CRC Press; Boca Raton, FL: 1985. p. 187-203.

Crump KS, Guess HA, Deal KL. Confidence intervals and tests of hypotheses concerning dose response relations inferred from animal carcinogenicity data. Biometrics 1977;33:437–451. [PubMed: 911968]

Faustman EM, Bartell SM. Review of noncancer risk assessment: applications of benchmark dose methods. Hum Ecol Risk Assess 1997;3:893–920.

Guess HA, Crump KS. Low-dose extrapolation of data from animal carcinogenicity experiments—analysis of a new statistical technique. Math Biosci 1976;32:15–36.

Hall P, Pittelkow YE. Simultaneous bootstrap confidence bands in regression. J Stat Comput Sim 1990;37:99–113.

Janardan KG. Environmental risk assessment model for low dose extrapolation. Commun Stat-Theor M 1995;24:2621–2634.

Kodell RL, West RW. Upper confidence limits on excess risk for quantitative responses. Risk Anal 1993;13:177–182. [PubMed: 8502790]

Krewski, D.; van Ryzin, J. Dose response models for quantal response toxicity data. In: Csörgö, M.; Dawson, DA.; Rao, JNK.; Saleh, AKME., editors. Statistics and related topics. North-Holland; Amsterdam: 1981. p. 201-231.

Nitcheva DK, Piegorsch WW, West RW, Kodell RL. Multiplicity-adjusted inferences in risk assessment: benchmark analysis with quantal response data. Biometrics 2005;61:277–286. [PubMed: 15737104]

Pan W, Piegorsch WW, West RW. Exact one-sided simultaneous confidence bands via Uusipaikka's method. Ann I Stat Math 2003;55:243–250.

Piegorsch WW, Nitcheva DK, West RW. Excess risk estimation under multistage model misspecification. J Stat Comput Sim 2006;76:423–430.

Portier CJ. Biostatistical issues in the design and analysis of animal carcinogenicity experiments. Environ Health Persp 1994;102(Suppl 1):5–8.

Schlosser PM, Lilly PD, Conolly RB, Janszen DB, Kimbell JS. Benchmark dose risk assessment for formaldehyde using airflow modeling and a single-compartment, DNA-protein cross-link dosimetry model to estimate human equivalent doses. Risk Anal 2003;23:473–487. [PubMed: 12836840]

Schulz TW, Griffin S. Estimating risk assessment exposure point concentrations when the data are not normal or lognormal. Risk Anal 1999;19:577–584. [PubMed: 10765423]

Smith LA, Sielken RL. Bootstrap bounds for safe doses in the multistage cancer dose-response model. Commun Stat-Simul C 1988;17:153–175.

Wogan GN, Paglialunga S, Newberne PM. Carcinogenic effects of low dietary levels of Aflatoxin B1 in rats. Food Cosmet Toxicol 1974;12:681–685. [PubMed: 4375655]

Table 1

Empirical simultaneous coverage rates on $0 < d < 1$ for upper limits on extra risk under the multistage model $R(d) = 1 - \exp\{-\beta_0 - \beta_1 d - \beta_2 d^2\}$; rates based on 2,000 simulated data sets, nominal $\alpha = 0.05$

β_0	β_1	β_2	*N*				
			25	**50**	**100**	**300**	**500**
0.0101	0.0183	0.0229	0.9135	0.9840	0.9570	0.9345	0.9405
0.0101	0.0215	0.0738	0.9720	0.9535	0.9640	0.9350	0.9420
0.0101	0.0250	0.1881	0.9540	0.9545	0.9405	0.9415	0.9365
0.0513	0.0153	0.0388	0.9865	0.9520	0.9445	0.9590	0.9545
0.1054	0.0563	0.1260	0.9505	0.9385	0.9425	0.9480	0.9595
0.1054	0.0823	0.1690	0.9605	0.9340	0.9460	0.9530	0.9545
0.1054	0.2200	1.9772	0.9270	0.9475	0.9435	0.9455	0.9450
0.3567	0.4865	0.5431	0.9515	0.9570	0.9440	0.9520	0.9565

Table 2

Empirical simultaneous coverage rates for extrapolated lower limits on benchmark doses corresponding to benchmark risk levels of 0.01, 0.05 and 0.10 under the multistage model $R(d) = 1 - \exp\{-\beta_0 - \beta_1 d - \beta_2 d^2\}$; rates based on 2,000 simulated data sets

β_0	β_1	β_2	*N*				
			25	**50**	**100**	**300**	**500**
0.0101	0.0183	0.0229	1.0000	0.9950	0.9865	0.9860	0.9885
0.0101	0.0215	0.0738	0.9910	0.9820	0.9905	0.9830	0.9850
0.0101	0.0250	0.1881	0.9875	0.9910	0.9910	0.9850	0.9815
0.0513	0.0153	0.0388	0.9920	0.9735	0.9705	0.9775	0.9740
0.1054	0.0563	0.1260	0.9635	0.9625	0.9700	0.9765	0.9780
0.1054	0.0823	0.1690	0.9760	0.9565	0.9745	0.9765	0.9830
0.1054	0.2200	1.9772	0.9590	0.9655	0.9685	0.9735	0.9750
0.3567	0.4865	0.5431	0.9620	0.9755	0.9675	0.9750	0.9785

Table 3

Simultaneous BMDLs for both the bootstrap method and the S-method for three data sets. BMDLs were computed by inverting a 95% upper confidence band in each case. The bootstrap confidence band was based on 2,000 simulated data sets

Substance	BMR	BMD	Bootstrap BMDL	S-method BMDL
Aflatoxin B_1	10^{-5}	0.000301	0.000024	0.000018
	10^{-4}	0.002922	0.000242	0.000183
	10^{-3}	0.023701	0.002414	0.001840
	10^{-2}	0.124259	0.023606	0.018430
	10^{-1}	0.482164	0.202236	0.186968
Dieldrin	10^{-5}	0.000195	0.000055	0.000053
	10^{-4}	0.001944	0.000549	0.000530
	10^{-3}	0.019190	0.005489	0.005297
	10^{-2}	0.172437	0.054648	0.053126
	10^{-1}	1.102334	0.527424	0.547405
DDT	10^{-5}	0.007617	0.002571	0.002858
	10^{-4}	0.076123	0.025714	0.028582
	10^{-3}	0.756938	0.257151	0.285895
	10^{-2}	7.189874	2.572926	2.866312
	10^{-1}	53.99410	25.93560	29.41630