

Changepoint Alternatives to the NOAEL

R. Webster WEST and Ralph L. KODELL

The onset-of-trend changepoint model is applied to the dose–response setting encountered in regulatory toxicology experiments with continuous responses. An investigation of the least squares criterion over each of the dose intervals provides insights into the distribution of the changepoint estimate for small dose group sample sizes and a simple computational procedure for finding this estimate. A new accessible proof of the consistency of this estimate as dose group size increases is developed. For normally distributed data, an asymptotic procedure is developed for constructing a lower confidence limit on the changepoint. The performance of the changepoint estimate and the lower confidence limit are compared directly to the no-observed-adverse-effect-level (NOAEL) via a large simulation study and an example dataset. The lower confidence limit on the changepoint is shown to be a better regulatory quantity than the NOAEL from a statistical perspective.

Key Words: Onset of trend; Profile likelihood.

1. INTRODUCTION

Dose–response toxicity studies involving rodents are used to study substances that are potentially dangerous to humans. After observing the responses of rodents exposed to various doses of the substance over some time frame, researchers must use this information to identify a potentially safe dosage for humans. Frequently, researchers compute the no-observed-adverse-effect-level (NOAEL) which is the *experimental* dose level just below the lowest dose level with responses that are significantly different from control. The goal of the NOAEL approach is to identify the largest no effect level for the toxin. The NOAEL is then divided by safety factors when identifying a safe dose for humans. The use of the NOAEL in regulatory toxicology has come under criticism (Kimmel and Gaylor 1988). Leisenring and Ryan (1992) discussed several undesirable properties of the NOAEL including its inverse relationship to sample size, its being limited to only discrete experimental doses and its failure to use any information from the dose–response relationship.

In this article, a changepoint dose–response model is developed as an alternative to the NOAEL for the case of continuous responses with constant variance. Related work includes

R. Webster West is Associate Professor, Department of Statistics, University of South Carolina, Columbia, SC 29208. Ralph L. Kodell is Director, Division of Biometry and Risk Assessment, National Center for Toxicological Research, Jefferson, AR 72079.

©2005 American Statistical Association and the International Biometric Society
Journal of Agricultural, Biological, and Environmental Statistics, Volume 10, Number 2, Pages 197–211
DOI: 10.1198/108571105X46525

the estimation of thresholds for quantal endpoints in toxicology (Cox 1987; Ulm 1999; Carroll, Roeder, and Wasserman 1999) and the estimation of changepoints for continuous endpoints (Julious 2001; Hirotsu and Marumo 2002; Chen 1998). For the model to be developed, the changepoint is the largest dose level that has the same mean response as control, so the changepoint is exactly what the NOAEL is trying to estimate. A least squares procedure for estimating the changepoint will be developed which uses dose-response information. The asymptotic properties of the least squares procedure are used to show the consistency of the estimate and to develop an asymptotic lower confidence limit on the changepoint for normally distributed data. This new methodology will be compared to the NOAEL for real and simulated data as it is evaluated from a regulatory perspective.

2. THE CHANGEPOINT DOSE-RESPONSE MODEL

Seber and Wild (1989) provided an extensive background in nonlinear regression models, and Feder (1975a, 1975b) provided a general introduction to segmented regression models. The models developed herein are more specific to the toxicological setting. Suppose n subjects are exposed to a toxin at each of g different dose levels. Let X_{ij} be a continuous random variable which denotes the response of the j th subject, $j = 1, \dots, n$, exposed to dose d_i , $i = 1, \dots, g$. It will be assumed that d_1, \dots, d_g represents an ordered sequence of dose values, $d_1 < d_2 < \dots < d_g$. The formulation of the onset of trend model in this dose-response setting is

$$X_{ij} = \begin{cases} \alpha + \epsilon_{ij} & d_i \leq d^* \\ \alpha^* + \beta d_i + \epsilon_{ij} & d_i > d^*, \end{cases} \quad (2.1)$$

where ϵ_i represents a sequence of independent and identically distributed random variables with mean zero and variance σ^2 . The model above allows for a jump in the mean response at the changepoint, d^* . From a biological perspective, this seems unreasonable for most settings encountered in toxicological experiments, so it will be assumed that the mean of the model is continuous at the changepoint. This leads to the relationship, $\alpha^* = \alpha - \beta d^*$, or

$$d^* = \frac{\alpha - \alpha^*}{\beta}. \quad (2.2)$$

The changepoint model then has only three free parameters as formulated above.

2.1 ESTIMATING d^* BY LEAST SQUARES

Least squares estimates of d^* , α , α^* , and β must minimize

$$S(d^*, \alpha, \beta) = \sum_{i: d_i \leq d^*} \sum_{j=1}^n (X_{ij} - \alpha)^2 + \sum_{i: d_i > d^*} \sum_{j=1}^n (X_{ij} - \alpha^* - \beta d_i)^2$$

subj
invo
and
of S

and

so th
poin

1, ...
 $\hat{\alpha}(d_k)$
conv
estim

Ulm 1999; continuous model to be response as least squares response to show the limit on the error to the value.

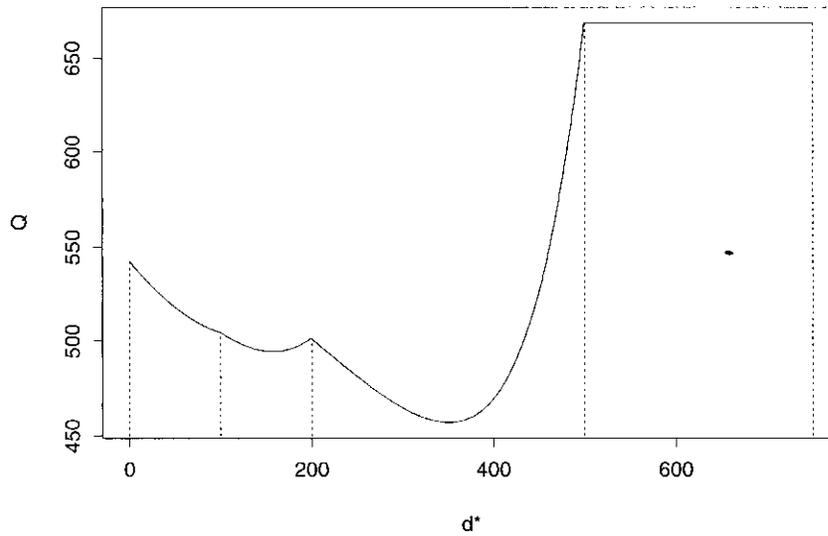


Figure 1. Profile of Q versus d^* for the aconiazide data.

ression model regression fitting. Suppose X_{ij} be a $i = 1, \dots, n$, an ordered end model

subject to (2.2). Clearly, S is not a differentiable function of d^* because this quantity is involved in the indices of the above summations. For fixed d^* , least squares estimates of α and β can be obtained by substituting $\alpha - \beta d^*$ for α^* and then setting the partial derivatives of S with respect to α and β equal to 0. This yields

$$(2.1) \quad \hat{\alpha}(d^*) = \frac{1}{ng} \sum_{i=1}^g \sum_{j=1}^n X_{ij} - \hat{\beta}(d^*) \frac{1}{ng} \sum_{i:d_i > d^*} \sum_{j=1}^n (d_i - d^*),$$

and

$$(2.2) \quad \hat{\beta}(d^*) = \frac{\sum_{i:d_i > d^*} \sum_{j=1}^n (d_i - d^*) \left(X_{ij} - \frac{1}{ng} \sum_{i=1}^g \sum_{j=1}^n X_{ij} \right)}{n \sum_{i:d_i > d^*} (d_i - d^*)^2 - \frac{n}{g} \left(\sum_{i:d_i > d^*} (d_i - d^*) \right)^2}$$

so that the minimum value of the least squares criterion as a function solely of the change-point can be expressed as

$$(2.3) \quad Q(d^*) = S(d^*, \hat{\alpha}(d^*), \hat{\beta}(d^*)).$$

It is clear that $\hat{\alpha}(d^*)$ and $\hat{\beta}(d^*)$ are continuous on each dose interval, $[d_k, d_{k+1})$, $k = 1, \dots, g - 2$. Likewise, Q is clearly continuous on any such dose interval. As $\Delta t \downarrow 0$, $\hat{\alpha}(d_k - \Delta t)$ and $\hat{\beta}(d_k - \Delta t)$ converge to $\hat{\alpha}(d_k)$ and $\hat{\beta}(d_k)$, respectively, and $Q(d_k - \Delta t)$ converges to $Q(d_k)$. Therefore, Q is a continuous function of d^* over the entire dose range.

A profile plot of Q versus d^* provides some interesting insights into the least squares estimation procedure. Such a plot is shown in Figure 1 for the aconiazide data analyzed by

Kodell and West (1993). These data represent the body weight gain (or loss) in grams for Fischer 344 rats treated with various doses of aconiazide over a 14-day period. For these data, there are 5 dose levels, 0, 100, 200, 500, and 750 mg/kg body weight with 10 animals per dose group. The complete dataset is available online at <http://lib.stat.cmu.edu/jasadata/fine-b.dat>. The vertical dashed lines in the plot are drawn from the curve to the horizontal axis at each dose value to emphasize the behavior of the function over each dose interval. A minimum value of the function over the first dose interval, $[0, 100)$, does not exist. A minimum value over the second dose interval, $[100, 200)$, is located at approximately 160 mg/kg body weight. The overall minimum value occurs in the third dose interval, $[200, 500)$, corresponding to a changepoint estimate of approximately 350 mg/kg body weight. Note that a changepoint in the last dose interval is not identifiable.

Although a profile search over the full range of the dose values could be used to estimate d^* , the value of d^* that minimizes Q on each dose interval can be obtained in closed form if a minimum value exists on the interval. To find the minimum value over the k th dose interval, $[d_k, d_{k+1})$, one must first minimize the adjusted least squares criterion,

$$S_k(\alpha, \alpha^*, \beta) = \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \alpha)^2 + \sum_{i=k+1}^g \sum_{j=1}^n (X_{ij} - \alpha^* - \beta d_i)^2,$$

which reflects this assumption about the location of the changepoint.

Let $\hat{\alpha}_k$, $\hat{\alpha}_k^*$, and $\hat{\beta}_k$ denote the parameter values that minimize S_k . Because S_k is clearly a differentiable function of these parameters for any $k = 0, \dots, g-2$, standard techniques can be used to show

$$\hat{\alpha}_k = \frac{1}{nk} \sum_{i=1}^k \sum_{j=1}^n X_{ij}, \quad (2.4)$$

$$\hat{\alpha}_k^* = \bar{X}_{k+1} - \hat{\beta}_k \bar{d}_{k+1}, \quad (2.5)$$

and

$$\hat{\beta}_k = \frac{\sum_{i=k+1}^g \sum_{j=1}^n (d_i - \bar{d}_{k+1}) X_{ij}}{n \sum_{i=k+1}^g (d_i - \bar{d}_{k+1})^2}, \quad (2.6)$$

where

$$\bar{X}_{k+1} = \frac{\sum_{i=k+1}^g \sum_{j=1}^n X_{ij}}{n(g-k)} \quad \text{and} \quad \bar{d}_{k+1} = \frac{\sum_{i=k+1}^g d_i}{g-k}.$$

The least squares estimation procedure in this case uses the responses at doses less than or equal to d_k to estimate the mean level prior to the changepoint, and the responses at doses greater than d_k to estimate the linear trend after the changepoint.

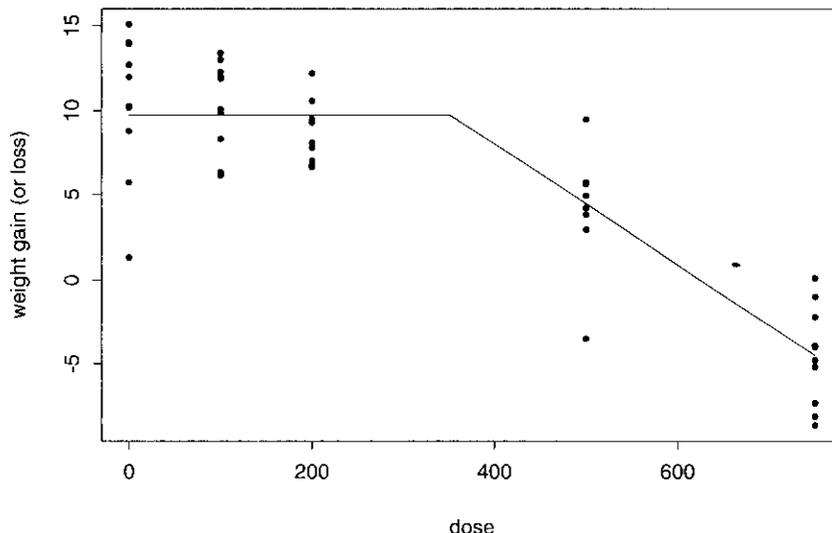


Figure 2. Fitted changepoint model for the aconiazide data.

Clearly, S_k is a convex function with a unique minimum value at the parameter estimates given above. There is no guarantee, however, that the changepoint estimate defined by the constraint in (2.2) will be in the k th dose interval at this minimum value. Let the resulting estimate of d^* be defined by

$$\hat{d}_k^* = \frac{\hat{\alpha}_k - \hat{\alpha}_k^*}{\hat{\beta}_k}. \tag{2.7}$$

If $\hat{d}_k^* < d_k$, then Q is an increasing function of d^* on $[d_k, d_{k+1})$, and the minimum value of Q on this interval must occur at d_k . Likewise, if $\hat{d}_k^* > d_{k+1}$, then Q is a decreasing function on $[d_k, d_{k+1})$, which will not have a minimum value on this half open interval. As discussed earlier in reference to Figure 1, this occurs for the first dose interval for which \hat{d}_1^* was found to be 124.6 mg/kg body weight which is not in $[0, 100)$. Let the minimum value of Q over $[d_k, d_{k+1})$ be defined by Q_k if a minimum exists. The overall minimum value of Q over the dose range, $[d_0, d_{g-1})$, can then be found by simply comparing Q_1, \dots, Q_{g-2} . This greatly reduces the number of calculations required with a profiling technique. The least squares estimate of the changepoint, \hat{d}^* , is the dose value, \hat{d}_k^* , corresponding to this minimum value. For the aconiazide data, \hat{d}^* was found to be $\hat{d}_3^* = 351.437$ mg/kg body weight. A plot of the fitted changepoint model for these data is shown in Figure 2.

The above discussion indicates that the changepoint estimate follows a mixture distribution where $\hat{d}^* = \hat{d}_k^*$ with some probability p_k . The small sample behavior of this distribution will be explored in more detail via a simulation study discussed in Section 4. The next section discusses the asymptotic behavior of the changepoint estimate.

2.2 THE CONSISTENCY OF \hat{d}^*

The consistency of \hat{d}^* is not guaranteed by standard asymptotic theory since the least

squares criterion in this case is not a "smooth" function of d^* . This section provides a unique tractable proof of the consistency of \hat{d}^* that uses the interval-wise aspects of the estimation procedure discussed earlier.

The variances of $\hat{\alpha}_k$, $\hat{\alpha}_k^*$, and $\hat{\beta}_k$ clearly converge to 0 as the dose group size, n , goes to infinity. Each of these random variables then converges in probability to their respective expected values. These expected values, however, depend on the true location of d^* . For the remainder of this discussion, it will be assumed that the true value of d^* is contained in the m th dose interval, $[d_m, d_{m+1})$.

For $k < m$, observations at doses greater than d_m are contaminated with observations which do not follow the appropriate regression model. Therefore, for $k < m$, $\hat{\alpha}_k \xrightarrow{P} \alpha$, but $\hat{\alpha}_k^*$ and $\hat{\beta}_k$, will not converge in probability to α^* and β , respectively. For $k < m$,

$$\hat{\alpha}_k^* \xrightarrow{P} \frac{(m-k)\alpha + (g-m)(\alpha^* + \beta\bar{d}_{m+1})}{g-k} - \beta\bar{d}_{k+1} \sum_{i=m+1}^g a_{ik}(d_i - d^*), \quad (2.8)$$

and

$$\hat{\beta}_k \xrightarrow{P} \beta \sum_{i=m+1}^g a_{ik}(d_i - d^*), \quad (2.9)$$

where

$$a_{ik} = \frac{d_i - \bar{d}_{k+1}}{\sum_{i=k+1}^g (d_i - \bar{d}_{k+1})^2}$$

Using the above results, Appendix A establishes the fact that \hat{d}_k^* converges in probability to a value that is greater than d_{k+1} for $k < m$. Applying the results of Section 2.1, this implies that asymptotically Q will be a decreasing function over $[d_1, d_m)$ since Q_k will not exist for $k < m$.

Likewise, for $k > m$, observations at doses less than d_m are contaminated with observations that do not have mean α . Therefore, for $k > m$, $\hat{\alpha}_k^* \xrightarrow{P} \alpha^*$ and $\hat{\beta}_k \xrightarrow{P} \beta$, but $\hat{\alpha}_k$, will not converge in probability to α . For $k > m$,

$$\hat{\alpha}_k \xrightarrow{P} \frac{m\alpha + (k-m)\alpha^* + \beta \sum_{i=m+1}^k d_i}{k}$$

Appendix B establishes the fact that \hat{d}_k^* converges in probability to a value less than d_k , for $k > m$. Applying the results from Section 2.1 again, this implies that asymptotically Q will be an increasing function over $[d_{m+1}, d_g)$.

Let \hat{m} denote the dose interval containing the least squares estimator of the changepoint so that $\hat{d}^* = \hat{d}_{\hat{m}}^*$. The above arguments combined with the continuity of Q show that $\hat{m} \xrightarrow{P} m$. This implies that $\hat{d}^* \xrightarrow{P} \hat{d}_m^*$. The estimates of $\hat{\alpha}_m$, $\hat{\alpha}_m^*$ and $\hat{\beta}_m$ converge in probability to α , α^* and β , respectively, so $\hat{d}_m^* \xrightarrow{P} d^*$. Thus, it has been shown that $\hat{d}^* \xrightarrow{P} d^*$. In other words, asymptotically the estimate corresponding to the true interval containing the changepoint

is selected consistent

Altho limit is mc confidence dependent of normal

Because th that T_m hc

The varia unbiased e

One can t freedom.

The rc will be the by applyin limit. $\hat{d}_{L,1}^*$

where t_p (freedom. limit on d lower con

is selected with probability one by the least squares criterion, and this estimate is itself a consistent estimator of the true changepoint.

3. LOWER CONFIDENCE LIMITS FOR d^*

Although a point estimate of the changepoint is certainly important, a lower confidence limit is more useful from a regulatory perspective. In this section, an asymptotic $(1-p)100\%$ confidence interval for d^* will be developed. The results from the previous section are not dependent on the data having a particular distribution, but in this section, the assumption of normality will be made. Let

$$T_m = \hat{\alpha}_m - \hat{\alpha}_m^* - \hat{\beta}_m d^*. \tag{3.1}$$

Because the m th dose interval contains the true changepoint, it is straightforward to show that T_m has a Normal distribution with $E(T_m) = 0$ and

$$\text{var}(T_m) = \sigma^2 \left(\frac{g}{nm(g-m)} + \frac{(\bar{d}_{m+1} - d^*)^2}{n \sum_{i=m+1}^g (d_i - \bar{d}_{m+1})^2} \right). \tag{2.9}$$

The variance of T_m can be estimated by replacing σ^2 in the above expression with its unbiased estimator

$$s_m^2 = \frac{\sum_{i=1}^m (X_{ij} - \hat{\alpha}_m)^2 + \sum_{i=m+1}^g (X_{ij} - \hat{\alpha}_m^* - \hat{\beta}_m d_i)^2}{gn - 3}. \tag{3.3}$$

One can then easily formulate a ratio which has a t -distribution with $gn - 3$ degrees of freedom.

The results of the previous section also guarantee that the asymptotic distribution of $T_{\hat{m}}$ will be the same as that for T_m . Asymptotically, one can find a lower confidence limit on d^* by applying an inverse regression procedure using $T_{\hat{m}}$. The $(1-p)100\%$ lower confidence limit, $\hat{d}_{L,1-p}^*$, is defined as the smallest root of

$$T_{\hat{m}}^2 = t_p(gn - 3) s_{\hat{m}}^2 \left(\frac{g}{n\hat{m}(g - \hat{m})} + \frac{(\bar{d}_{\hat{m}+1} - d^*)^2}{n \sum_{i=\hat{m}+1}^g (d_i - \bar{d}_{\hat{m}+1})^2} \right), \tag{3.4}$$

where $t_p(gn - 3)$ represents the p th quantile of a t -distribution with $gn - 3$ degrees of freedom. Using this procedure with the aconiazide data yielded a 95% lower confidence limit on d^* of 242.69 mg/kg body weight. The small sample coverage probability of the lower confidence limit is studied in the simulation study outlined in the next section.

4. SIMULATION STUDY

A simulation study was conducted with three main goals in mind. The first goal was to examine the performance of the changepoint methodology developed herein for small samples. The second goal was to examine the relative performance of the changepoint procedure to the NOAEL approach. Lastly, the study was designed to study the performance of both approaches under different parameterizations of the changepoint model. The key parameters were thought to be the location of the true changepoint and the slope of the regression model after the changepoint. The interval-wise characteristics of the changepoint estimation procedure developed might be affected by a changepoint that falls on the boundary of a dose interval. Also, the variance of the changepoint estimate for the interval containing the true changepoint can be shown to be inversely proportional to β^2 . Thus, the strength of the changepoint model is greatly determined by the rate at which the mean response changes at doses greater than d^* .

The simulated model had the five dose values of 0, 1, 2, 3, and 4. Sample sizes of 10, 20, and 50 per dose group were considered since these sample sizes are typically encountered in toxicological experiments. Two locations were considered for the true changepoint, $d^* = 1$ and $d^* = 1.5$, the first location being on the edge of a dose interval and the second being between experimental doses. Values of -10 , -3 , and $-.1$ were used for β going from a stronger to a weaker changepoint model. Because each value of β is negative, an adverse effect is defined by a mean response lower than the control mean. For each model, α was chosen to be 3, and the value of α^* was adjusted according to (2.2) while the error variance was held constant at 1.

For each parameter configuration, 1,000 realizations of the model defined in (2.1) were generated. For each realization the changepoint estimate, the lower confidence limit on the changepoint and the NOAEL were computed. The first two quantities were computed using the methodology described in Sections 2 and 3. The NOAEL was computed by conducting multiple lower tail two sample t tests of each dose group against control at the 5% significance level. Leisenring and Ryan (1992) used a similar approach to make pairwise comparisons to controls for establishing the NOAEL for quantal data. Statistically more sophisticated strategies can be used such as the procedures of Dunnett (1964) and Williams (1971), as reviewed by Tamhane, Hochberg, and Dunnett (1996). Also, improved statistical procedures have been developed more recently for determining the minimum effective dose or the maximum safe dose (and, consequently, the NOAEL) (Bauer, Rohmel, Maurer, and Hothorn 1998; Anraku 1999; Bretz and Hothorn 2000; Hothorn and Hauschke 2000; and Tamane, Dunnett, Green, and Wetherington 2001). Although these more advanced procedures correct for multiple testing, in practice, the simple procedure used here is most often employed by toxicologists and risk assessors. Correcting for multiple testing leads to a larger NOAEL while the simple procedure is more conservative. This makes the primary comparison of interest between the NOAEL and the lower confidence limit on the changepoint, that is, the proportion of times the true changepoint is not exceeded, somewhat more fair to the NOAEL. For the aconiazide example, the NOAEL was computed to be 200 mg/kg

| d^* | β |
|-------|---------|
| 1 | -10 |

-3

-.1

1.5 -10

-3

-.1

body wei
the previc
to zero in

The
denote th
the table.
asymptoti
Dividing
each quar
NOAEL i
in a regul
relative to

Even
probabilit
show tha
different p
anti-conse
location o
heavily in
at 1, the N
whereas i
at 1.5 wh

Table 1. Results of 1,000 Simulations at Each Combination of d^* , β , and n

| d^* | β | n | $\hat{d}_{L,.95}^* \leq d^*$ | $NOAEL \leq d^*$ | $NOAEL \leq \hat{d}_{L,.95}^*$ | $\hat{d}_{L,.95}^* < NOAEL \leq d^*$ | $d^* < NOAEL$ |
|-------|---------|-----|------------------------------|------------------|--------------------------------|--------------------------------------|---------------|
| 1 | -10 | 10 | 950 | 1000 | 89 | 561 | 350 |
| | | 20 | 963 | 1000 | 71 | 560 | 369 |
| | | 50 | 960 | 1000 | 73 | 573 | 354 |
| | -3 | 10 | 962 | 1000 | 73 | 557 | 370 |
| | | 20 | 943 | 1000 | 108 | 528 | 364 |
| | | 50 | 958 | 1000 | 93 | 553 | 354 |
| | -.1 | 10 | 955 | 13 | 6 | 28 | 966 |
| | | 20 | 946 | 28 | 5 | 40 | 955 |
| | | 50 | 943 | 70 | 17 | 55 | 928 |
| 1.5 | -10 | 10 | 936 | 1000 | 1000 | 0 | 0 |
| | | 20 | 959 | 1000 | 1000 | 0 | 0 |
| | | 50 | 954 | 1000 | 1000 | 0 | 0 |
| | -3 | 10 | 956 | 942 | 803 | 135 | 62 |
| | | 20 | 953 | 998 | 994 | 4 | 2 |
| | | 50 | 951 | 1000 | 1000 | 0 | 0 |
| | -.1 | 10 | 956 | 6 | 4 | 30 | 966 |
| | | 20 | 949 | 17 | 9 | 36 | 955 |
| | | 50 | 952 | 33 | 13 | 60 | 927 |

body weight which is less than the lower 95% confidence limit for the changepoint given in the previous section. Also, for the purposes of this comparison, the NOAEL was set equal to zero in cases where it did not exist.

The simulation results are displayed in Table 1. The first three columns in the table denote the values of d^* , β , and n that were used for the results displayed in each row of the table. The next two columns in the table provide the number of realizations where the asymptotic 95% lower confidence limit (labeled $\hat{d}_{L,.95}^*$) and the NOAEL were below d^* . Dividing these numbers by 1,000 would provide an estimate of “coverage probability” for each quantity where coverage is defined as the quantity being below d^* . Even though the NOAEL is an estimate and not a confidence limit its coverage probability is very important in a regulatory setting. The remaining three columns provide the location of the NOAEL relative to the changepoint estimator and the lower confidence limit.

Even for the relatively small sample sizes used in the simulation study the coverage probability for $\hat{d}_{L,.95}^*$ is right on target at the nominal level of .95. The simulation results show that the coverage probability for the NOAEL, however, varies greatly across the different parameter configurations; from very conservative for steeper slopes to extremely anti-conservative for shallower slopes. When the NOAEL coverage probability is high, the location of the NOAEL relative to the changepoint estimate and the lower confidence limit is heavily influenced by the location of the true changepoint. When the changepoint is located at 1, the NOAEL tends to fall between the two values when the coverage probability is high whereas it tends to fall below the lower confidence limit when the changepoint is located at 1.5 when the coverage probability is high. When the coverage probability is low, the

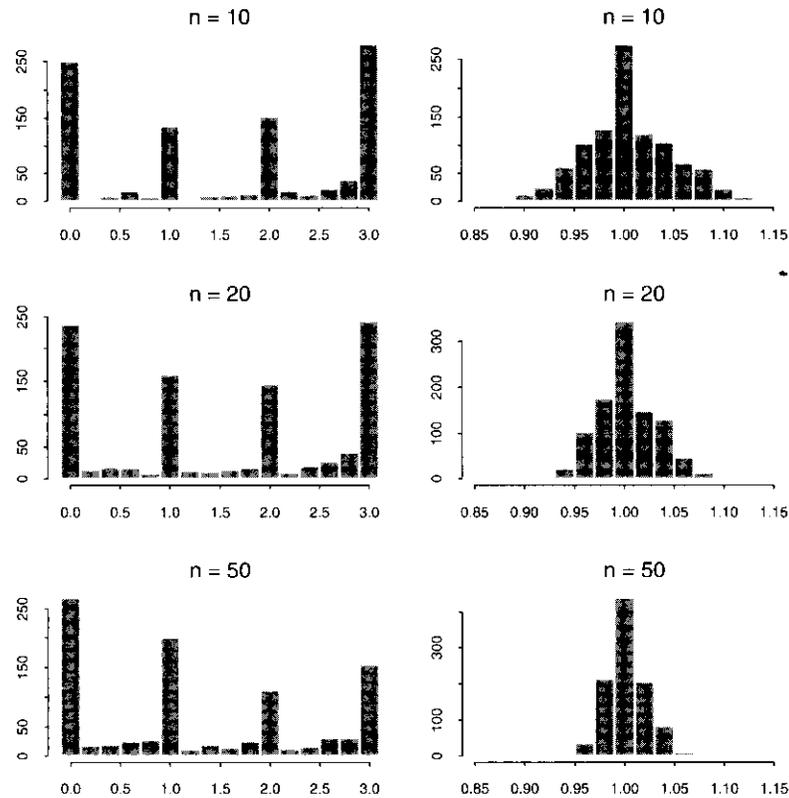


Figure 3. Histogram of changepoint estimates for $d^* = 1$. The left hand column contains estimates for each dose group size with $\beta = -1$. The right hand column contains estimates for each dose group size with $\beta = -10$.

NOAEL is not only larger than the lower confidence limit but in most cases it is also larger than the point estimate.

The discrete nature of the distribution of the changepoint estimate is demonstrated in the left column of Figure 3 for the configuration of $d^* = 1$ and $\beta = -1$. The histogram of the values for the estimate shows clear spikes at dose values of 0, 1, 2, and 3. The large number of zero estimates is indicative of cases where a very shallow slope is combined with a changepoint that is close to zero. Indeed, for models that are truly linear with no changepoint, the changepoint estimate and the lower confidence limit are almost always zero. The distribution of the changepoint estimate was much less discrete for larger (absolute) values of the slope as shown in the right hand column of Figure 3 for the configuration of $d^* = 1$ and $\beta = -10$. For the $\beta = -1$ configuration, the distribution is almost uniformly distributed across the observed doses at a dose group size of 10 with only a slight shift to the lower end of the dose range for a dose group size of 50. For the $\beta = -10$ configuration, the distribution is symmetric about the true changepoint value of 1. In both configurations, as dose group size increases, the variability in the estimate decreases. The variability, however, decreases at a much slower rate for the smaller value of β . The lower confidence limit for the configuration in the left column of Figure 3 was almost always 0 while the NOAEL

was almost a case since the interval contained the conservative

In the simulation applied to a test were also considered different features of aconiazide distribution mean

This model forced to decrease model and the model has a

The quality with a more realistic that one would misspecification values for the estimate and The average average lower size of 50. The lack of fit between linear model remaining due to NOAEL when group size of 115 for a dose

The mean is approximately a dose than the characteristic is being applied the NOAEL. For a group size of 10, the change in the overall distribution directly consequence always the confidence of fit.

Revisiting the presumption

was almost always 4. The location of the lower confidence limit is very interesting in this case since the estimate used in its computation in many cases does not correspond to the interval containing the true changepoint. This indicates that the lower confidence limit may be conservative for small sample sizes.

In the situation discussed above, the changepoint approach and the NOAEL were both applied to a true changepoint model. In a follow up simulation study, the two methodologies were also compared for a model where the true underlying dose response relationship had different features. The model chosen for the study was the best fit quadratic model for the aconiazide data displayed in Figure 2. This model is defined by the following quadratic mean

$$10.24753 + .000534211d - 2.658048 \times 10^{-5}d^2. \quad (4.1)$$

This model provides a nice fit to the aconiazide data, and in this case a researcher would be forced to decide which dose-response model to use. In comparing the models, the quadratic model and the changepoint model have the same number of parameters, but the changepoint model has a smaller sum of square error.

The quadratic model which was simulated features a very flat response at low levels with a more rapid change at higher doses. These features are somewhat similar to the features that one would expect from a changepoint model. In order to study the effects of model misspecification, 1,000 realizations of the quadratic model were simulated using the 5 dose values for the aconiazide data with dose group sizes of 10, 20, and 50. The changepoint estimate and its lower limit were computed for each realization along with the NOAEL. The average changepoint estimate for each dose group size was approximately 316 with the average lower limit increasing from 278 for a dose group size of 10 to 301 for a dose group size of 50. The fitting procedure for the changepoint model is forced to balance between lack of fit before and after the changepoint. The procedure adapts in this case by fitting a linear model to the responses from the last two dose groups and using information from the remaining dose groups to estimate the constant level of the changepoint model. The average NOAEL when it existed decreased from 138 for a dose group size of 10 to 100 for a dose group size of 50. The number of realizations when the NOAEL did not exist increased from 115 for a dose group size of 10 to 277 for a dose group size of 50.

The mean response for the quadratic model dips below the control mean at approximately a dose of 20, so in terms of estimating this value, the NOAEL does a much better job than the changepoint model. This should not be surprising because the changepoint model is being applied inappropriately. The lack of dose response information works in favor of the NOAEL in this case. It should be stated, however, that even for the small dose group size of 10, the changepoint model would have been rejected by simply examining a residual plot in the overwhelming majority of cases. This simulation study does demonstrate the possible dire consequences that can be encountered by blindly fitting the changepoint model. As is always the case in model fitting, diagnostic procedures should be used to assess goodness of fit.

Revisiting the actual aconiazide data (Figure 2), the 95% lower confidence limit on the presumed changepoint was estimated to be 242.69 mg/kg, even though doses as low

as 20 mg/kg could actually have been associated with an adverse effect according to the best fitting quadratic model (i.e., the fitted mean response at 20 mg/kg is below the fitted mean response at 0 mg/kg). However, this does not necessarily imply that the two models would give dramatically different results in practice. Based on the same best fitting quadratic model produced with these data, Kodell and West (1993) calculated 178 mg/kg as a 95% lower confidence limit on a 1% benchmark dose for body weight loss. The 1% benchmark dose, like the changepoint, is another quantity that has been proposed as an alternative to the NOAEL (NRC 2000). Hence for these aconiazide data, the changepoint lower limit (242.69), the 1% benchmark dose lower limit (178) and the NOAEL (200) are in fairly good agreement. This close correspondence of alternative methods is reassuring in the present application, and it addresses to some degree the issue of uncertainty in model selection, which is often ignored when carrying out statistical inferences.

5. CONCLUSIONS

Because of the recognized limitations of the NOAEL, Crump (1984) proposed the use of a benchmark dose as an alternative where the benchmark dose was defined to be the dose corresponding to a low level of excess risk above background in the range of 1 to 10%. Although the benchmark dose approach has gained acceptance over the years as a "point of departure" for setting exposure levels, there is still much support among toxicologists and risk assessors for using the NOAEL, because of its association with a presumed zero or negligible level of risk. The changepoint methodology discussed in this article should also be appealing to the regulatory community for the same reason. In addition, the statistical properties of this methodology (the consistency of estimate and the asymptotic validity of the lower confidence limit procedure) are far superior to the NOAEL when the true underlying model is a classic threshold model. Also, as demonstrated in the simulation study, the lower confidence limit on the true changepoint does not necessarily lead to a more conservative regulatory quantity than the NOAEL which is a point estimate of the true changepoint. Combining the changepoint and benchmark dose methodologies may lead to a more unified approach to toxicological risk assessment. The process would begin with the application of the changepoint method to determine if a nonzero no effect level can be determined in the experimental dose range. If the lower confidence limit on the changepoint is found to be zero, the risk assessor could then apply the benchmark dose method to determine a dose associated with an acceptable amount of risk. For a unified approach to be widely applicable, a larger repertoire of changepoint models must be developed in this setting other than the simple onset-of-trend model discussed herein. Models to be developed should consider other functional forms for the mean response before and after the true changepoint and the effects of nonconstant variance across the dose range. In toxicology experiments, however, the small number of dose values typically used will be problematic in terms of estimating parameters for models which allow for more curvature in the response after the changepoint.

Rec
the proo
Usi

where

Note tha

$\sum_{i=}$

Adding
result yi

where

It will no
of the ab

so $f_{u,v} 1$

f_r

APPENDIX A

Recall the assumption that the true changepoint, d^* , is in the interval, $[d_m, d_{m+1}]$. For the proof given below, it is assumed that $k < m$.

Using (2.10)–(2.9) with Slutsky’s theorem,

$$\hat{d}_k \xrightarrow{P} \bar{d}_{k+1} - \frac{\sum_{i=k+1}^g (d_i - \bar{d}_{k+1})^2}{(g - k)(f_{m,g}(d^*) - \bar{d}_{k+1})}, \quad (A.1)$$

where

$$f_{u,v}(d) = \frac{\sum_{i=u+1}^v d_i(d_i - d)}{\sum_{i=u+1}^v (d_i - d)}, \quad u = 1, \dots, v \quad v = 1, \dots, g.$$

Note that

$$\begin{aligned} \frac{\sum_{i=k+1}^g (d_i - \bar{d}_{k+1})^2}{g - k} &= \frac{(\bar{d}_{k+1} - d_{k+1}) \sum_{i=k+1}^g (d_i - d_{k+1} + d_{k+1} - \bar{d}_{k+1})^2}{\sum_{i=k+1}^g (d_i - d_{k+1})} \\ &= (\bar{d}_{k+1} - d_{k+1}) (f_{k,g}(d_{k+1}) - \bar{d}_{k+1}). \end{aligned}$$

Adding and subtracting d_{k+1} from the right hand side of (A.1) and substituting the above result yields

$$\hat{d}_k \xrightarrow{P} d_{k+1} + c_k,$$

where

$$c_k = \frac{(\bar{d}_{k+1} - d_{k+1}) (f_{m,g}(d^*) - f_{k,g}(d_{k+1}))}{f_{m,g}(d^*) - \bar{d}_{k+1}}.$$

It will now be shown that $c_k > 0$ by carefully considering the numerator and denominator of the above expression. The derivative of $f_{u,v}$ with respect to d ,

$$f'_{u,v}(d) = \frac{(v - u) \sum_{i=u+1}^v (d_i - \bar{d}_{u+1})^2}{\left(\sum_{i=u+1}^v (d_i - d) \right)^2} \geq 0,$$

so $f_{u,v}$ is an increasing function of d . Since $d_{k+1} \leq d^*$ for $k < m$,

$$f_{m,g}(d_{k+1}) = \frac{\sum_{i=m+1}^g d_i(d_i - d_{k+1})}{\sum_{i=m+1}^g (d_i - d_{k+1})} \leq \frac{\sum_{i=m+1}^g d_i(d_i - d^*)}{\sum_{i=m+1}^g (d_i - d^*)} = f_{m,g}(d^*). \quad (A.2)$$

Also, because $d_u < d_i \leq d_v$, for $i = u + 1, \dots, v$,

$$f_{k,m}(d_{k+1}) = \frac{\sum_{i=k+1}^m d_i(d_i - d_{k+1})}{\sum_{i=k+1}^m (d_i - d_{k+1})} < d_{m+1} \leq \frac{\sum_{i=m+1}^g d_i(d_i - d^*)}{\sum_{i=m+1}^g (d_i - d^*)} = f_{m,g}(d^*). \quad (\text{A.3})$$

Adding the cross-products of the inequalities in (A.2) and (A.3) leads to

$$f_{k,g}(d_{k+1}) = \frac{\sum_{i=k+1}^g d_i(d_i - d_{k+1})}{\sum_{i=k+1}^g (d_i - d_{k+1})} < \frac{\sum_{i=m+1}^g d_i(d_i - d^*)}{\sum_{i=m+1}^g (d_i - d^*)} = f_{m,g}(d^*).$$

Also, note that $\bar{d}_{k+1} - d_{k+1} > 0$, so that the numerator of c_k is positive. Using Chebyshev's inequality,

$$f_{m,g}(d^*) = \frac{\sum_{i=m+1}^g d_i(d_i - d^*)}{\sum_{i=m+1}^g (d_i - d^*)} \geq \frac{\left(\sum_{i=m+1}^g d_i\right) \left(\sum_{i=m+1}^g (d_i - d^*)\right)}{(g - m) \sum_{i=m+1}^g (d_i - d^*)} = \bar{d}_{m+1}.$$

Since $\bar{d}_{m+1} > \bar{d}_{k+1}$, $f_{m,g}(d^*) - \bar{d}_{k+1} > 0$, so that the denominator of c_k is also positive.

APPENDIX B

Using (2.10),

$$\hat{d}_k^* \stackrel{P}{\rightarrow} \frac{md^* + \sum_{i=m+1}^k d_i}{k}.$$

Because $d^* < d_k$ for $k > m$ and $d_i < d_k$ for $i = m + 1, \dots, k$,

$$md^* < md_k \quad \text{and} \quad \sum_{i=m+1}^k d_i < (k - m)d_k.$$

Combining the two inequalities above and dividing each side by k ,

$$\frac{md^* + \sum_{i=m+1}^k d_i}{k} < d_k.$$

The estimate \hat{d}_k^* for $k > m$ converges in probability to the expression on the left hand side of the above inequality as n goes to infinity.

[Received April 2003. Revised September 2004.]

- Anraku, K. (1999) 141–152.
- Bauer, P., Rohme Active Con
- Bretz, F., and Hc Dose-Resp 135–154.
- Carroll, R. J., R 44–54.
- Chen, J. (1998), and Metho
- Cox, C. (1987),
- Crump, K. S. (1 Toxicology
- Dunnett, C. W. (
- Feder, P. I. (1975 The Annal
- (1975b),
- Hirotsu, C., and Journal of
- Hothorn, L. A., Journal of
- Julious S. A. (2 Statistical
- Kimmel, C. A., 8, 15–20.
- Kodell, R. L., an Risk Analy
- Leisenring, W., a cology, 15
- NRC (National B ington, DC
- Seber, G. A. F.,
- Tamhane, A. C., 52, 21–37
- Tamhane, A. C Identifying
- Williams, D. A. pared With
- Ulm, K. (1999) Sciences, 8

REFERENCES

- (d^*). (A.3)
- Anraku, K. (1999), "An Information Criterion for Parameters Under a Simple Order Restriction," *Biometrika*, 86, 141–152.
- Bauer, P., Rohmel, J., Maurer, W., and Hothorn, L. (1998), "Testing Strategies in Multi-Dose Experiments Including Active Control," *Statistics in Medicine*, 17, 2133–2146.
- Bretz, F., and Hothorn, L. A. (2000), "A Powerful Alternative to William's Test with Application to Toxicological Dose-Response Relationships for Normally Distributed Data," *Environmental and Ecological Statistics*, 7, 135–154.
- Carroll, R. J., Roeder, K., and Wasserman, L. (1999), "Flexible Measurement Error Models," *Biometrics*, 55, 44–54.
- Chen, J. (1998), "Testing for a Changepoint in Linear Regression Models," *Communications in Statistics—Theory and Methods*, 27, 2481–2493.
- Cox, C. (1987), "Threshold Dose Response Models in Toxicology," *Biometrics*, 43, 511–523.
- Crump, K. S. (1984), "A New Method for Determining Alternative Daily Intakes," *Fundamental and Applied Toxicology*, 4, 854–871.
- Dunnnett, C. W. (1964), "New Tables for Multiple Comparisons With Control," *Biometrics*, 20, 482–491.
- Feder, P. I. (1975a), "On Asymptotic Distribution Theory in Segmented Regression Problems—Identified Case," *The Annals of Statistics*, 3, 49–83.
- (1975b), "The Log-Likelihood Ratio in Segmented Regression," *The Annals of Statistics*, 3, 84–97.
- Hirotsu, C., and Marumo, K. (2002), "Changepoint Analysis as a Method for Isotonic Inference," *Scandinavian Journal of Statistics*, 29, 125–138.
- Hothorn, L. A., and Hauschke, D. (2000), "Identifying the Maximum Safe Dose: A Multiple Testing Approach," *Journal of Biopharmaceutical Statistics*, 10, 15–30.
- Julious S. A. (2001), "Inference and Estimation in a Changepoint Regression Problem," *Journal of the Royal Statistical Society, Ser. D*, 50, 51–61.
- Kimmel, C. A., and Gaylor, D. W. (1988), "Issues in Qualitative and Quantitative Risk Analysis," *Risk Analysis*, 8, 15–20.
- Kodell, R. L., and West, R. W. (1993), "Upper Confidence Limits on Excess Risk for Quantitative Responses," *Risk Analysis*, 13, 177–182.
- Leisenring, W., and Ryan, L. (1992), "Statistical Properties of the NOAEL," *Regulatory Toxicology and Pharmacology*, 15, 161–171.
- NRC (National Research Council) (2000), *Methods for Developing Spacecraft Water Exposure Guidelines*, Washington, DC: National Academy Press.
- Seber, G. A. F., and Wild, C. J. (1989), *Nonlinear Regression*, New York: Wiley.
- Tamhane, A. C., Hochberg, Y., and Dunnnett, C. W. (1996), "Multiple Test Procedures for Dose Finding," *Biometrics*, 52, 21–37.
- Tamhane, A. C., Dunnnett, C. W., Green J. W., and Wetherington, J. D. (2001), "Multiple Test Procedures for Identifying the Maximum Safe Dose," *Journal of the American Statistical Association*, 96, 835–843.
- Williams, D. A. (1971), "A Test for Differences Between Treatment Means When Several Dose Levels are Compared With a Zero Dose Control," *Biometrics*, 27, 103–117.
- Ulm, K. (1999), "Nonparametric Analysis of Dose-Response Relationships," *Annals of New York Academy of Sciences*, 895, 223–231.
- thebyshev's
- ft hand side