BELLE Newsletter Vol. 2, No. 1, August 1993

THE SECOND ANNUAL BELLE CONFERENCE: A REVIEW

Leonard Sagan
Electric Power Research Institute, Palo Alto, CA

Approximately 100 participants attended the second annual BELLE Conference held in Crystal City Virginia on April 26, 27, 1993. Greetings were expressed by both Edward Calabrese of the University of Massachusetts who is also Director of BELLE, and Arthur Wykes, President of the Association of Government Toxicologists. Calabrese delivered a review of the development and goals of BELLE. He emphasized the need to "Let the data lead us" without imposing constraints on interpretation.

There were two opening statements on the general subject of scientific paradigms. Kenneth Schaffner, a physician and philosopher of science at George Washington University, discussed the concept of scientific paradigms, as developed by Thomas Kuhn in his 1970 book, *The Structure of Scientific Revolutions*. Kuhn had suggested that science operates on models ("paradigms") which are revised only when contrary information becomes overwhelming and a "scientific revolution" occurs. Schaffner discussed some developments in Kuhn's concept of paradigms and their relevance to the BELLE meeting.

In his presentation, Leonard Sagan (Electric Power Research Institute) turned attention to the radiation paradigm, its history, and the assumptions inherent in that paradigm. The radiation paradigm, which has subsequently been expanded to include effects of exposures to chemicals at low doses, implies the absence of a threshold and deleterious (but non detectable) effects at very low exposures. Sagan suggested that the development of this model was very much influenced by the environmental ethic of the 1950's and 1960's. He also suggested that there are many sectors of society (e.g., lawyers, regulators, radiation scientists) who now have a stake in maintaining the model, even though there is little scientific validation and growing economic costs and little health benefit.

George Milo (Ohio State University) discussed the inconsistencies/consistencies of expression of biological endpoints with the expression of molecular mutations when subacute non-toxic dosages of activated environmental xenobiotics activate oncogenes and/or other molecular mutations in human cells. He discussed the issue of whether humans fit the present linear animal model of progression used to explain malignant progession. The over interpretation of animal data has not shed a great deal of light on the complexity of the biological carcinogenic endpoint when human cells were exposed to either sub-acute or chronic dosage levels of exposure to activated environmental xenobiotics.

Three speakers examined the statistical basis for the low-dose non-threshold model. Mike Davis (Environmental Protection Agency) reported a systematic review of a sample of 1,800 articles containing dose-response data culled from the toxicology literature. Applying certain criteria, 147 of these articles were selected for more detailed analysis. Of these, 22 (15 percent) showed evidence of a "U-shaped" dose-response relationship. Davis illustrated some of the conceptual and statistical difficulties in ascertaining the incidence of non-monotonic relationships.

David Gaylor (National Center for Toxicological Research) suggested that there may be an optimal exposure level for all environmental agents. Hormetic responses will only appear, he said, if the optimal dose is greater than the naturally occuring environmental level. If the optimal level is less than the environmental level, then there will be no stimulatory or hormetic effect from increasing exposure levels. Gaylor also reported his own review of some 3,000 animal studies carried out for carcinogenesis. In none of them did he find evidence of hormesis, although he admitted that these animal studies, carried out at relatively high doses and with relatively few animals, are less than ideal for the examination of a beneficial effect.

Peter Groer (University of Tennessee) described Bayesian techniques for examining low dose data, searching for "change points" in the data. In both beagle and human data, he demonstrated evidence of such change points. The human data were derived from studies of women watch dial painters with body burdens of radium 226, and with bone cancer as an end point.

Four speakers addressed the question of mechanisms: how could stimulatory effects at low doses be explained? Are observations of phenomena at low doses consistent with the linear paradigm? Harihara Mehendale (Northeast Louisiana State University) sees stimulated tissue repair as central to an understanding of the response to a toxic

agent. If the initial exposure is large, then the repair mechanism is unable to respond, whereas with smaller doses the repair response appears and becomes instrumental in recovery. Exposure to a second inhibitory agent or event can also interfere with repair resulting in much higher toxicity.

Colin Hill (University of Southern California) spoke of a class of phenomena becoming known as an "adaptive response". There are two types of adaptive response that appear to be elicited by radiation. One is a response during a cell's life cycle, usually inducible, that reduces the toxic effects of radiation insult, thus resulting in more cells surviving. The other is a heritable adaptive response that can be passed on from one generation to another. Again, the end result is, at least for a short term, an increase in survival. However, Hill showed that, at least in the later type of adaptive response, the long term effect of an inducible change may be genetically deleterious. Although adaptive responses are now considered real phenomenon, there is still some confusion as to the their exact nature. He concluded that understanding of adaptive responses and their incorporation into risk assessment will only occur when some systematic use of models is adopted.

Joan Smith-Sonnenborn of the University of Wyoming focused on the role of stress proteins as pivotal in the adaptive response (see Vol. 1, No. 3, March 1993 BELLE newsletter). She noted that different patterns of proteins are synthesized in response to a wide variety of stressors, and speculated that these proteins may play a role in a variety of protective mechanisms including DNA repair and increased cell survival.

Angelo Turturro (National Center for Toxicological Research) described the series of experiments done at NCTR to investigate the mechanisms underlying the well known increase in survival and decrease in carcinogenesis of caloric restricted animals. Many physiological parameters, including energy metabolism, circadian rhythms, and DNA repair, are altered in these animals. From these observations, it would appear that dietary factors may play an important role in modifying dose response relationships, perhaps making them U-shaped.

Although unable to attend, James Trosko's (Michigan State University) submitted paper raised some interesting questions about the role of radiation in carcinogenesis. Within the context of the current paradigms of the multi-stage concept of carcinogenesis, stem-cell theory of cancer and the oncogene/tumor suppressor gene theory, the fact that ionizing radiation induces mostly deletion mutations and chromosomal rearrangements calls into question acute, low dose exposures of radiation as a "complete carcinogen" and non-threshold models. He noted that the

initiation/promotion/progression stages of carcinogenesis involve multiple genetic and epigenetic events and are mechanistically different. Different mechanisms can lead to the initiation, promotion, or clonal expansion of a stem cell and its ultimate conversion to a metastatic cell. Therefore, Trosko pointed out that while radiation is associated with the appearance of cancers after exposure, several questions have to be answered, such as: "Does ionizing radiation affect all or just some of these stages?; Does ionizing radiation only contribute to promotion by cytotoxicity-induced hyperplasia and to the initiation/progression stages by epigenetic alteration of gene expression?"

Three epidemiologists presented views on human responses to low doses of chemical agents and radioactivity. Ethel Gilbert (Pacific Northwest Laboratory) reported on her studies of workers exposed to occupational levels of radiation. In her view, epidemiological studies do not have sufficient precision to resolve the issue of low dose effects, whether harmful or protective. The principal use of these studies will be to provide an upper bound on risks.

Robert Miller (National Cancer Institute), from a review of the literature, found that the lowest doses at which effects of radiation have been demonstrated are: 0.04 Gy for thyroid cancer among Israeli children given radiotherapy for ringworm of the scalp; 0.2 Gy for somatic cell mutations detectable by the glycophorin A test; 0.1 - 0.0.2 Gy for small head size after in utero exposure to the atomic bomb at 4 - 17 wk of gestational age; 0.2 Gy for long-lasting chromosomal aberrations; 0.2 - 0.5 Gy for leukemia and breast cancer among Japanese A-bomb survivors, with the highest relative risk among those exposed under 15 years of age; and 0.6 Gy for severe mental retardation, with brain damage found through Medical Res Image after in utero exposure to radiation from the atomic bomb at 8 - 15 wk of gestational age. Diagnostic x-ray exposures later in pregnancy have been related to a 1.5-fold excess of childhood leukemia in Great Britain and New England.

As has been known for some time, cancer has been noted in occupationally exposed human populations working with chemicals. In some cases however, as pointed out by Ralph Cook (Dow Corning), there is evidence of a protective effect. This was illustrated with data from dioxin exposed populations, an observation which is consistent with laboratory animal data.

Several speakers were asked to comment on the proceedings of the meeting. One of these was John Graham (Harvard School of Public Health). Graham noted that during the two day symposium, he heard several examples of agents which appeared to be inconsistent with the non-threshold paradigm. They were: alcohol, diet, dioxin, and

aspirin, all of which produced paradoxical (and beneficial) effects at low doses. He suggested that scientists consider several tactics for achieving a new paradigm. The first was "Don't overreach", by which he meant that scientists should be cautious about generalizing from weak data. Secondly, scientists should be open to exceptions to rules, and should be bold in disseminating those exceptions. Thirdly, we must continue to conduct the mechanistic studies necessary to build a base of knowledge about low dose effects. Finally, he advised that there be a symmetry in our standards of proof, being just as tough on those observations which appear to prove the rule (e.g., the non-threshold model) as on those which seem to be exceptions to the rule (hormesis).

Both Roger McClellan (Chemical Industry Institute of Toxicology) and John Higginson (Georgetown University) expressed concern about the dogmatic application and misapplication of the non-threshold model. Higginson proposes that rigid risk assessment, based on a linear model, be abandoned, and that a more pragmatic common sense judgmental approach be utilized based on a weight of the total evidence analysis, including mechanisms. McClellan closed the meeting, emphasizing the need to conduct studies of carcinogenesis at realistic levels of exposure, rather than at near toxic levels.

BELLE Book and Proceedings

To order your copy of the book "Biological Effects of Low Level Exposures to Chemicals and Radiation" (ISBN 0-87371-665-5) or to order proceedings from the Second Annual BELLE Conference entitled "New Perspectives on Dose-Response Relationships and Low Level Exposures", April 26 and 27, 1993. Please call Lewis Publishers, Chelsea, MI at 1-800-272-7737.

Modeling Hormesis

A recent article was published concerning the modeling on hormetic responses. Since this work has direct relevance to BELLE, the published report was sent to five individuals for independent reviews. Based on the reviews a decision was made to request permission to reprint the paper in the BELLE newsletter. What follows is the original article, followed by the five reviewers' comments and the response of the original authors to the reviewer comments.

ECOTOXICOLOGY AND ENVIRONMENTAL SAFETY 25, 25-32(1993)
Calculation of the EC50 and Its Confidence Interval
When Subtoxic Stimulus Is Present

P.H. Van Ewijk* and J.A. Hoekstra**
**Solvay Duphar BV P.O. 900, 1380 DA Weesp. The Netherlands: and **National Institute of Public Health and Environmental Protection, P.O. Box 1, 3720 BA Bilthoven, The Netherlands*

Received January, 30 1992

In many experiments in toxicology and ecotoxicology with continuous response, like growth, a stimulus is found for low doses, while higher doses are toxic. This subtoxic stimulus is called hormesis. The standard logistic model cannot be used when hormesis occurs. A model is given which can describe this phenomenon. The model has the EC50 as one of its parameters. This has the advantage that the EC50 and its confidence interval can be determined directly by nonlinear regression. Also, the tests for the presence of hormesis and lack of fit are described.

INTRODUCTION

In ecotoxicology and toxicology many experiments are done to establish the dose-response relationship for a certain compound and organism. The results of these experiments are often analyzed by a logistic model and summarized in the form of the EC50, the concentration or dose which gives a 50% effect. The logistic model can be applied to dichotomous data such as survival or death and continuous data, e.g. Weight or biomass. The authors restricted the study to the continuous case.

For continuous responses a stimulus for low doses of otherwise toxic compounds is found in many of these experiments (Stebbing, 1982). This stimulus is called hormesis. If hormesis occurs, the standard logistic model does not fit. Common practice is then to use the model anyway or drop part of the data. A better solution was proposed by Brain and Cousens (1989) who extended the logistic model. This extension naturally implements hormesis in the logistic model.

However, nonlinear regression with their model does not give the EC50 explicitly. Therefore, the researchers propose a reparameterization of this model, which will give the EC50 and its confidence interval directly, and describe how to test for hormesis.

A MODEL FOR HORMESIS

Extension of the Logistic Model

Often dose-response toxicity data follow a sigmoidal curve. This curve can be described by the logistic model. The general form of the logistic model is

$$y = \frac{k}{1 + e^{b(\log x + g)}} \quad \text{or} \quad y = \frac{k}{1 + e^{bg} x^{b}} \quad \text{or} \quad y = \frac{k}{1 + (x/x_0)^{b}},$$

where x_0 = EC50 = exp(-*g*).

This can be written as

$$\text{Log} = \frac{y}{k - y} = -b(\log x + g)$$

In model (1) k stands for the value of y at x=0 and g=-log (EC50). The parameter k is estimated from the complete set of data, not just from the control, as are the other parameters. If b>0, the response is monotonically decreasing and the x-axis an asymptote, as curve 1 in Fig. 1. The parameter -b stands for the slope of the line on logit-log-scale. On the original scale (y vs x), b relates to the slope of the tangential line in the point of inflection, where

$$X = EC50: \ [dy/dx]_{x=EC50} = -b * k/(4x_0$$

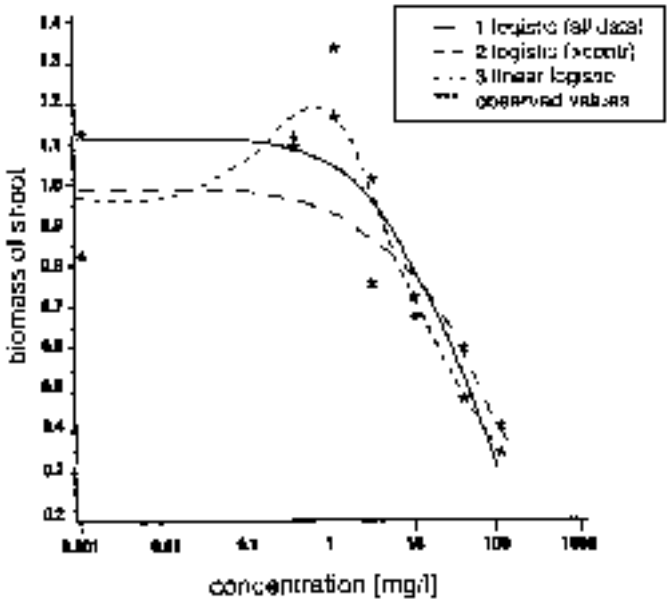

Figure 1. Fitted curves by nonlinear regression for case study isobutylalcohol in nutrient solution. Curve 1, standard logistic model fitted to all available data: curve 2, standard logistic model fitted to data, after exclusion of concentrations with mean response>control response; curve 3, linear logistic model (5) fitted to all available data.

For the data in Table 1 and Fig. 1 the logistic model obviously shows a lack fit. If this model is fitted to the data anyway, the control response will probably be over-estimated and the EC50 estimate can be misleading. To allow for hormesis, Brain and Cousens (1989) generalized model (1) into

$$y = \frac{k(1 + fx)}{e^{bg} x^{b}}$$

Table Results of an Experiment with Isobutylalcohol in Nutrient Solution
1. (Hulzebos et al., in Preparation)

		Shoot weight	
	Replication		
Concentration	1	2	Mean
0	1.126	0.833	0.980
0.32	1.096	1.106	1.101
1	1.163	1.336	1.250
3.2	0.985	0.754	0.870
10	0.716	0.683	0.700
32	0.560	0.488	0.524
100	0.375	0.344	0.360

This model can describe an increase for low levels of x, cf. Fig. 1, curve 3. In the AIDS literature, model (2) is known as the linear-logistic model, cf. Isham (1988, 1989). There, however, generally it is assumed that b<0, which results in a curve with an exponential shape for low values of x and a straight line for high values.

In model (2) the intuitive meaning of the parameters is lost, apart from the interpretation of the parameter k, which still stands for the response at x=0. The EC50 (X0) should be calculated from the equation

$$e^{bg}x_0^b - 2fx_0 - 1 = 0.$$

This equation has no explicit solution for X0. So, the estimation of the EC50 involves two steps:

first, find the parameter estimates for model (2);
second, solve Eq. (3) numerically.

The construction of a confidence interval for the EC50 also needs additional calculations.

The EC50 as a Parameter in the Model
Since it is very cumbersome to find the EC50 and its confidence interval from model (2), the model is rewritten so that the EC50 is included as a parameter. Then, no extra step is needed and an approximate confidence interval can be constructed from the estimated standard error which is in general supplied by nonlinear regression procedures. First Eq. (3) is solved for g, leading to

$$e^{bg} = \frac{2fx_0 + 1}{x^b}$$

Substitute (4) in (2):

$$y = \frac{k(1 + fx)}{1 + (2fx_0 + 1) * (x/x_0)^b}$$

The parameter k stands for the response at x=0; f stands for homesis: if f>0 the curve shows an increase for low doses. The parameter x_0 is the EC50. The parameter b loses its simple interpretation.

Numerically it is better to replace x_0 by (μ)=log(x_0). This means one has to replace x_0 in formula (5) with $e^{(\mu)}$.

The Fit of the Model
Because of the restriction to a continuous response, nonlinear regression procedures will fit model (5) to the data, e.g., PROC NONLIN from SAS (SAS Institute Inc., 1990) and FITNONLINEAR from GENSTAT (Payne et al., 1987). Some extra input is needed, however.

Most packages require initial estimates or a grid over which starting values for the parameters are searched. The following procedure was adopted to find initial estimates. The parameter k can be easily estimated by the mean of the control responses. The parameter b is found by linear regression of log(y/(max(y)‹y) on log(x). Only data with response lower than the mean response (k) for the control are included. The slope of this regression line is an estimate of -b. The ratio -intercept/slope from the same regression can be used as initial estimate for (μ).

A more robust and therefore more reliable estimate for (μ) is

$$\mu = \sum w_i * \log(\text{concentration}i) / \sum w_i$$

$$\text{with } w_i = | y_i - 0.5*k |^{-1} .$$

If the response lies near k/2, than W_i will be great, while concentrations with responses far from k/2 will hardly contribute to the estimation.

While fitting a nonlinear model the derivatives of the function to the parameters have to be calculated. This can be done numerically, but it is better to supply these derivatives in formula. The derivatives can be found in the Appendix in SAS code.

Test for Hormesis and for Lack of Fit
Since hormesis is described by only one parameter, the test for hormesis is formed by the ratio of the extra explanation of the linear logistic model and the residual variation,

$$F_{\text{hormesis}} = (\text{SSE}_{\text{logistic}} - \text{SS}_{\text{linlogistic}}) / \hat{\sigma}^2$$

where

$\text{SSE}_{\text{logistic}}$ = residual sum of squares for logistic model, $\text{SSE}_{\text{linlogistic}}$ = residual sum of squares for logistic model, ^{2} = estimate of residual variation, e.g., SSE linlogistic/ residual degrees of freedom. This statistic follows approximately a F(1, df) distribution, where df stands for the degrees of freedom of ^{2}.

When a design is replicated, i.e., more than one result per dose, it is possible to test for lack of fit for the total model. The procedure is comparable to the test for hormesis.

$$F_{\text{lof}} = \{(\text{SSE} - \text{SS}_{\text{rep}}) / (dfe - df_{\text{rep}})\} / \sigma^2$$

in which
SSE = residual sum of squares for the current model. SS_{rep} = sum of squares betwen replications. dfe = residual degrees of freedom for current model. df_{rep} = degrees of freedom between replications, ^{2} = estimate of residual

variation. e.g. SS_{rep}/df_{rep} . This F1of should be compared with the F(dfe‹df_{rep}, df_{rep}) distribution.

A CASE STUDY: ISOBUTYLALCOHOL IN NUTRIENT SOLUTION

Table 1 lists the results of an experiment in which isobutylalcohol was dissolved in nutrient solution in which lettuce plants were grown. The biomass of the shoot was determined after 21 days. The standard logistic model was fitted to these data (fit 1) and to the data after exclusion of concentrations with mean response greater than the mean control response (fit 2). The third fit included all data and used the linear logistic model (5).

Table 2 gives the parameter estimates of various fits. The parameter for control weight (k) is high for the standard logistic model. This lack of fit is illustrated in Fig, 1. The EC50 estimate from the second fit (logistic model, part of the data excluded) is 43% higher than for the standard logistic model, while the linear logistic model (fit 3) gives an intermediate result, 23% higher.

Table 2. Results for Isobutylalcohol Experiment: Parameter Estimates

Fit	Data	Model	k	b	x0	f
1	All data	Standard	1.10	0.78	29	
		Logistic	(0.93, 1.29)	(0.25, 1.34)	(12, 66)	
2	<Control	Standard	0.99	0.69	41	
		Logistic	(0.82, 1.16)	(0.24, 1.14)	(17, 97)	
3	All data	Linear	0.97	1.28	35	1.7
		Logistic	(0.79, 1.14)	(1.17, 1.40)	(13, 97)	(1.3, 4.6)
Initial Estimate		0.98	0.24	68	0	

					99a	

Note. Model (1) was used for fit 1 and 2. Model (5) for fit 3. For fit 2 concentrations with mean response greater than mean control response excluded. For each parameter the estimate is given and the 95% confidence interval (in parentheses). The parameter k stands for the initial response (X=0), b for the slope, X0 for the EC50, and f for hormesis. [a]Robust estimate $\mu = \ldots w_i * \log(\text{concentration}_i)/\ldots w_i$ with $w_i = |y_i < 0.5*k|^{-1}$.

Table 3 presents the sums of squares and statistics derived from these.

Table 3. Results for Isobutylalcohol Experiment: Sums of Squares

Fit	Data	Model	SSE	dfe	MSE	p value 1of
1	All data	Standard	0.242	11	0.0220	0.035
		Logistic				
2	<Control	Standard	0.076	7	0.0109	0.26
		Logistic				
3	All data	Linear	0.125	10	0.0125	0.18
		Logistic				
Between replicates ($\square^2$rep)						
All data			0.065	7	0.0092	
<Control		0.045	5	0.0089		

Note. Model (1) was used for fit 1 and 2, Model (5) for fit 3. Fot fit 2 concentrations with mean response greater than

mean control response were excluded. SSE stands for the residual sum of squares (error). Dfe for the residual degrees of freedom. MSE, the residual mean square is the ratio of the two. The last column (P value, 1of) gives the result of the lack of fit test.

For all fits a calculation was made for a lack of fit test, based on the variation between replicates, (σ^2rep). For fit 1, the logistic model, it is found that:

$$F_{\text{lof}} = \{(\text{SSE} - \text{SS}_{\text{rep}}) / (dfe - df_{\text{rep}})\} / \sigma^2_{\text{rep}}$$
$$= \{(0.242 - 0.065) / (11 - 7)\} / 0.0092 = 4.81.$$

Since this is greater than the 5% critical value for a F(4, 7) statistic, 4.12, the lack of fit is significant. The P value is 0.035.

For fit 2, the MSE is much lower than that for model (1) and the lack of fit test not significant. This is misleading, however, since the doses with mean response greater than the control response were excluded in the calculation of the MSE and the lack of fit test. For these doses Fig. 1 shows a clear lack of fit.

Fit 3, the linear logistic model (5), gives no significant lack of fit (P=0.18). Hence, the linear logistic model describes the data satisfactorily.

A test can be made for the presence of hormesis:

$$F_{\text{hormesis}} = (\text{SSE}_{\text{logistic}} - \text{SS}_{\text{linlogistic}}) / \text{MSE}_{\text{linlogistic}}$$
$$= (0.242 - 0.125) / 0.0125 = 9.38.$$

Thus, the hormesis is significant (P = 0.012).

DISCUSSION

The linear logistic model can be implemented into a nonlinear regression procedure. The author reparameterization looks a bit more complicated but has the advantage of a straightforward estimation of the EC50.

If there is no hormesis, the corresponding parameter f in model (5) equals 0. As described in the case study it is simple to test for hormesis. When model (5) was fitted to 72 datasets (see Van Ewijk and Hoekstra, in preparation), it was found, that f can be negative, even significantly. Then f describes the asymmetry in the model and indicates the inappropriateness of the standard logistic model.

As usual in nonlinear regression, numerical problems can arise if the data do not contain enough information to estimate all parameters. Of course, the extra parameter f enlarges this risk, especially in those cases where no hormesis exists.

APPENDIX: SAS-CODE FOR THE FIT OF THE LINEAR LOGISTIC MODEL

```
TITLE 'SAS-code for fit of linear logistic model';
INFILE 'final76.dat' FIRSTOBS=2;
INPUT dos resp;
IF dos >0.01 THEN ldos=LOG(dos);
ELSE ldos=LOG(0.01); *the first part of the file finds initial estimates:

*-------------------------------------------------------------

PROC MEANS NWAY DATA=horm; *calculation of mean for control:

VAR resp;
CLASS dos;
OUTPUT OUT=hormmn MEAN=respmn;
DATA contrmn;
```

```
SET hormmn;
IF dos=0;
k=respmn;
*k = mean of control;
DATA hormcomb;
MERGE horm hormn ;
IF -N-=1 THEN SET contrmn;
logitr=LOG(resp/(k-resp));
CALL SYMPUT('inik',k); *initial estimate for k;
PROC REG DATA=hormcomb OUTEST=estlin; *linear regression to find;
MODEL logitr=ldos; *initial estimate for b and xo;
RUN;
DATA -NULL- ;
SET estlin;
CALL SYMPUT('inib',-ldos);
CALL SYMPUT('inix0',EXP(-INTERCEP/ldos));
%LET inif=0; *now the initial estimates are known, nonlinear regression can be performed;

*-----------------------------------------------------------------------

PROC NLIN DATA=horm;
PARMS k=&inik x0=&inix0 b=&inib f=&inif; *initial estimates
IF dos > 0.0001 THEN DO; *specification of derivatives:
xx0b=(dos/x0)**b;
fx0=2*f*x0+1
denom=1+ fx0 * xx0b;
y= k*(1 + f*dos) / denom;
DER.k = y/k;
DER.b = -y / denom * fx0 * LOG(dos/x0) * xx0b;
DER.x0 = y /denom * xx0b * (-2*f + fx0 * b/x0);
```

```
DER.f = k* dos / denom - y/denom * 2 * x0 * xx0b;
END;
ELSE DO; *derivatives for control need;
denom=1; *special care;
y=k;
DER.k= 1;
DER.b= 0;
DER.f= 0;
DER.x0=0;
END;
MODEL resp = y;
RUN; *fit of the model;
```

REFERENCES

Brain P., and Cousens. R. (1989). An equation to describe dose responses where there is stimulation of growth at low doses. Weed Res. 29, 93-96.

Hulzebos, E.M., Adema, D.M.M., Dirven-Van Breeman, E.M., Henzen, L., Baarselman. R., and Van Gestel, C.A.M. Phytotoxicity studies with Lactuca Sativa in soil and nutrient solution. Environ Toxicol. and Chem. In press.

Isham, V. (1988). Mathematical modeling of the transmission dynamics of HIV infection and AIDS: A review. J.R. Stat. Soc. A 151, 5-49.

Isham, V. (1989). Estimation of the incidence of HIV infection. Phil. Trans. R. Soc. Lond. B 325, 113-121

Payne, R.W., et al (1987). GENSTAT 5 Reference Manual. Clarendon Press, Oxford.

SAS INSTITUTE, Inc. (1990). SAS Language: Reference, Version 6, 1st ed. SAS Institute, Inc., Cary, NC.

Stebbing, A.R.D. (1982). Hormesis The situation of growth by low levels of inhibitors. Sci Total Environ. 22, 213-234.

Van Ewijk, P.H., and Hoekstra, J.A. Curvature measures and confidence intervals for the linear logistic model, in preparation.

REVIEW COMMENTS ON THE PAPER A "CALCULATION OF THE EC50 AND ITS CONFIDENCE INTERVAL WHEN SUBTOXIC STIMULUS IS PRESENT" BY VAN EWIJK AND HOEKSTRA

By Rick Cothern
Environmental Protection Agency
Washington, D.C.

The idea of modifying the logistic model to include hormetic effects seems a good one. However this paper appears to be a use of someone else's modification, the prose are difficult to understand and it adds little to current understandings. A more complete application of the ideas to a wider range of data would be useful and should be conducted. In addition there are some errors. For example in Figure 1 it is not possible to plot x=0 on a log scale, the confidence limits for x0 in Table 2 do not look right (likely a typo) and the attached SM coding does not reflect the numerical replacement of x0 by mu = log (x0) as claimed in the middle of page 27. Also Table 1 has no unit.

By Tom Downs
School of Public Health
University of Texas, Houston, TX

The logistic function with dose as argument is often used to model a continuous response. The logistic function per se is symmetric about its midpoint value, EC50. The authors have modified the standard logistic model to allow for

hormetic (non-symmetric) responses by incorporating into it an asymmetry parameter. In the formulation of the logistic model used, the response is enhanced, unchanged, or inhibited at low doses according as this asymmetry parameter is positive, zero, or negative.

The authors also give a large-sample F-statistic for testing the hypothesis that the asymmetry parameter is zero. Calculation of this statistic for a given dose-response data set requires nonlinear fitting of both the original and modified logistic models to the data. An algorithm for using the SAS statistical software program to do this is provided. The methods are illustrated with a numerical example using concentrations of isobutylalcohol, dissolved in nutrient solutions in which lettuce plants were grown, as doses, and the biomass of the resulting shoots as responses.

The development of a variety of models suitable for studying hormesis is needed. Currently used dose-response models are inadequate to study hormesis since virtually all of them mathematically force the response to be a monotonic increasing function of the dose‹thus excluding any consideration of hormesis. The paper by Van Ewijk and Hoekstra is therefore a welcome addition to the scientific study of hormetic phenomena.

By David W. Gaylor
National Center for Toxicological Research
Jefferson, AR

P.H. Van Ewijk and J.A. Hoekstra (EcotoKicology and Environ. Safety 25: 25-32, 1993) present a statistical test for the presence of hormesis in a dose-response model. They correctly use the reduction in the sum of squares of error, due to adding a hormetic term to the dose-response model, to test for hormesis. This is a standard regression model fitting technique that has been in use for a number of years (see eg, O. Kempthorne, The Design and Analysis of Experiments, John Wiley and Sons, Inc., N.Y., 1952). The validity of the test depends upon a constant variance across dose groups. Examination of the differences in results between replicates generally shows less consistency at the low doses. Perhaps, less weight should be given to the low dose data. That is, the statistical significance of the hormetic effect is weaker than the reported value. Nevertheless, the test for hormesis, based upon an improved fit of dose-response data, provides a more convincing demonstration of hormesis than simply comparing the result at a low dose to the controls.

By David J. Svendsgaard
Health Effects Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Introduction

Van Ewijk and Hoekstra make a solid contribution to the area of risk assessment. They have shown how to test for hormesis and how to obtain estimates and confidence limits of the EC50 using a logistic model extended to model hormesis (Weed Res. 29:93-96, 1989).

However, there is a need for more research in the area of low exposure-level risk assessment modeling. Consider two issues needing research:

Models for testing hypotheses need to be developed because even the most sensitive statistical tests require too many animals to detect small effects at low dose. For example, in comparing treatment to control means, Dunnett's test is often used, (Biometrics 20, 482-491, 1964) because it protects against multiple comparison error. This test was used in trying to detect U-shaped dose response curves, (BELLE Newsletter 1:1, 1992), in 58 data sets from NTP (National Toxicology Program) developmental toxicology studies, using average-litter fetal weight as the endpoint. A power analysis indicated that 97 dams per treatment group would be required to detect a 5% change from control when type I and II error are 5% and 20%, respectively. Only rarely are such large sample sizes used. The incorporation of models into hypothesis testing should require fewer animals because the models may be "focused" to detect certain features such as hormesis.

Crump has suggested the benchmark approach to risk assessment (Fund. Appl. Tox. 4:854-871, 1984). Here, one uses a model to estimate the EC10 (effective concentration yielding a 10% change from control), and to obtain the lower 95% confidence limit on this concentration. The Van Ewijk model could be modified to estimate the EC10 when hormesis is present.

First, we will discuss using models to test for hormesis and then using models to estimate effective doses, such as the EC50.

Using the Van Ewijk Model to Test for Hormesis

The Van Ewijk model could have been used to fit the NTP fetal weight data with no modification required. However, other responses may require model modification. For example, endpoints such as enzyme activity may have U-shaped response curves that first decrease and then increase up to some saturation level. Another feature of some types of response data is that toxicity may mask some dose-response curves and cause otherwise monotonically increasing dose-response curves to decrease at high dose. Stead et al developed a method for modifying dose-response models for such masking (Mutation Research, 85:13-27, 1981).

Devidas et al proposed a modification to the logistic model to provide improved fit at low dose, (Statistics in Medicine, 12:881-892, 1993), and it could be applied to the Van Ewijk model:

$$E[y] = \frac{k(1 + fx)}{[1 + (2fx0 + 1)\ (x/x_0)^b]}$$

where [Ey] is the mean value of y, the notation follows the article, and the modification involves a new parameter □ to be estimated. This model fits the Van Ewijk data slightly better (MSE=0.0122 rather than 0.0125), and yields a test for hormesis that remains significant (F1,9 = 4.84). Had the test not been significant, the usefulness of the Van Ewijk model to test for hormesis would have been questionable. The point is that in using models to test the hormesis hypotheses, it is going to be difficult to select the model for the null hypothesis. The best-fitting monotonic dose-response model is often a strong contender, but the class of pasrimonious models will certainly vary among data sets.

Estimation of the Effective Dose

If the purpose of the model is really estimation of the EC50, other models may fit the data better, or be easier to use. An alternative class of models is a low order polynomial in log dose. For example, one could fit:

$$E[y] = B_0 \text{ if } x=0$$

$$E[y] = B_0 k + B_1 \log(x_1/x_0) + B_2 [\log(x_1/x_0)]^2 + B3[\log(x_1/x_0)]^3 \text{ if } x > x^*$$

where is the lowest nonzero dose that was used, so this model makes no inference about the shape of the curve in the interval between the control and the lowest nonzero dose. A quadratic rather than a cubic model should be sufficient when the response is weight or height, but a cubic may be necessary to describe a U-shaped curve at low dose and toxicity masking the response at high dose. The parameter x0 is the dose(s) yielding the mean response equal to k times the control mean.

The quadratic and cubic models were fit to the Van Ewijk data. The quadratic (and even a cubic) didn't fit as well as the Van Ewijk model. However, the 95% confidence intervals on the EC50 were comparable (10 to 81 for the cubic and 18 to 103 for the quadratic). These intervals were obtained by using their suggestion to estimate the log of the EC50 rather than the EC50, itself.

The polynomial models are easy to fit and to modify, but have poor extrapolation charcteristics. For example, below the lowest nonzero dose, the cubic model fit to the Van Ewijk data would indicate that the response approaches negative infinity as the dose approaches zero. So the graph of the fitted response curve of the cubic model should not be extended far from the interval containing the nonzero doses.

A problem occurs in fitting the logistic when the response means are all above 50% of the control mean. Either the EC50 is estimated poorly (with large standard errors), or the assumption that zero is the right-end asymptote is questionable. Usually the data are not adequate to obtain a good estimate of this right-end asymptote when it is a model parameter. Alternatively, one could retain the assumption that the right-end asymptote is zero, and one could consider reparameterizing the model to estimate the EC10 rather than the EC50.

In the risk assessment context, the use of the EC10 involves a value judgement. For example, one of their response means is 20% higher than the control mean, but the EC10 focuses on the dose yielding only a 10% decrease. This is justified if we view increases as healthy and decreases as unhealthy. However, a better approach to risk assessment would be to view the apparent improvement in health as evidence of the test chemical's biological activity. The other endpoints should be more carefully examined in this dose-interval for adverse effects. These effects may have small magnitude but correct direction.

Should the EC50 be based on half the control mean or on half the peak mean? Consider a developmental toxicology study of an essential nutrient such as zinc. The zinc has at least a U-shaped dose-response curve when the response is malformations at birth. Usually, we equate zero exposure to the normal situation. In this case, however, normal is probably the peak of the U-shaped curve. So rather than basing the EC50 on half the control weight, it should be based on half the peak weight. Perhaps, the EC50 needs to be redefined for other essential chemicals having U-shaped dose-response curves, as well.

Whether one uses an EC50 or an EC10, there are going to be data sets where the toxicant is so ineffective that poor estimates of these doses result. So, the toxicity of the chemical is apparently below the detection limit of the experimental procedure. In these cases, it may be best to assume that the true effective dose is censored on the high side. A censored dose indicates that x0 is some unknown dose larger than the maximum dose used, but it may be infinite (there is no dose that will yield the effect). This viewpoint allows us to use some of the many statistical procedures for dealing with censored data in subsequent analyses.

Conclusions

The Van Ewijk model appears to be useful in testing for hormesis. When the data permit the use of other models for this test, they should be considered. One such model is Devidas modification to the logistic model. The Van Ewijk model can also be used in some instances to estimate effective dose (i.e., EC50) when there is hormesis present. However, modifications need to be developed for other cases. In model selection, one needs to consider simplicity, ease of use, and the purpose of the modeling.

Modeling can lead to an indeterminate conclusion about the effective dose as when the standard errors are too large. One needs to be aware of the detection limits of the experimental protocol. This involves the consideration of

error due to the experimental procedure and due to lack of knowledge of the true shape of the model. Finally, the theory of censored data may be helpful in subsequent analyses using the estimated effective doses as a dependent variable.

Disclaimer
This document has been reviewed in accordance with U. S. Environmental Protection Agency policy and approved for publication. The views expressed do not necessarily reflect the policy of the Agency and no official endorsement should be inferred. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

By Jan Zielinski and Dan Krewski
Health and Welfare Canada
Ottawa, Ontario
In bioassays when the response (death, toxic reaction, etc.) may vary with a covariate (dose) probability of the binary event is often modeled using logistic function p=1/(1+exp(-d)). There is no overriding biological rationale for use of the logistic model, and its widespread application to the analysis of dose-response data can only be supported on the ground of mathematical convenience. When such a model is fitted the response variable is assumed to follow binomial distribution, and parameters are estimated using maximum likelihood estimation. The nonlinear regression procedure in SAS can be used to search for maximum likelihood estimates (by implementing an iteratively reweighed least-squares algorithm). This procedure is customarily called logistic regression. The assumption of binomial distribution is critical for justification of the procedure.

The logistic function can be also used to model continuous (non binary) response variables. In order to justify use of non-linear regression for finding parameters of this function, other distributional assumptions have to be made. In particular errors are assumed to be independent and normally distributed.

The paper, in fact is concerned with the latter modeling situation. However this is not clearly stated in the text. This might mislead occasional user of statistical methods, and it is a main flaw of the paper.

In technical terms, the paper describes fitting of linear-logistic function for modeling dose response curve. The model

is used to fit the data on the isobutylalcohol dissolved in nutrient solution and the biomass of the lettuce shoots to which this solution was applied.

The example is informative and well presented. The inclusion of the SAS programs implementing the procedures described in the paper may be very useful to any researcher who wants to model non-monotone dose-response relationships.

The paper is of satisfactory quality and its reprint might prove useful.

Authors Response

P.H. Van Ewijk* and J.A. Hoekstra**
***Solvay Duphar BV P.O. 900, 1380 DA Weesp.**
The Netherlands: and **National Institute of Public
Health and Environmental Protection, P.O. Box 1,
3720 BA Bilthoven, The Netherlands

We thank the reviewers for their thorough reviews and the editor for the opportunity to respond to the comments given.

We agree with Dr. Svensgaard, that the linear logistic model is one of the many possible to describe these kind of data. We choose this model since the commonly used logistic model is one of its submodels. Formula (2) was taken from the Brain & Cousens article, the reparameterization in (3) -(5) is ours (comment Dr. Cothern). Dr. Cothern's comment on the inconsistency between table 2 and the SAS-program is highly appreciated. We apologize for this error. The program in the appendix is ok, but does not estimate μ=LOG(x0) but x0 itself. This has no implications for this dataset, apart from the confidence interval for x0 in table 2. This interval was derived from the confidence interval for μ, by backtransforming the confidence limits. If one wants to follow our suggestion and use LOG(x0) the following modifications in the program are necessary:

- remove the word EXP in "CALL SYMPUT ('inix0;EXP(INTERCEP/ldos));"

- replace in the PARMS-statement x0=&inix0
  with 1x0=&inix0;
- replace the statement
  DER.x0 = y / denom *xx0b* (-2*f+fx0*b/x0)
  with
  DER.lxo = xo*y / denom*xx0b* (-2*f+fx0*b/x0)
- add after "IF dos > 0.001 THEN DO;" the statement
  x0=EXP(1x0);

Both Dr. Gaylor and Dr. Zielinski & Krewski comment on the assumptions for nonlinear regression. We indicated in the paper that this method was meant for the continuous case. We forgot to include the usual assumptions for (non) linear regression: independent, identically, normally distributed errors. We agree with Dr. Gaylor that a tendency for lower spread in the high concentrations is not uncommon and maybe present in our data. Modifications are possible then (cf eg Seber and Wild (1989), section 2.8), but we did not want to make our paper more complicated and used the simple approach, which we think is satisfactory in this case.

The logistic model is commonly used, but does not accommodate non-monotonic data (cf comment Dr. Downs); the linear logistic does. Comparing the fit of these two models seems a logical and consistent procedure. Dr. Svensgaard discusses some alternative models. It is good to have alternatives, but one has to be careful for data-dredging and the implicit optimism (Efron, 1986) when looking for the best fitting one. Therefore, we use the standard logistic model as null-model. The only sound way to confirm hormesis, is replication of the experiment.

The extension of the linear model with a power in the nominator (Devidas modification from Dr. Svensgaard) is attractive. In practice, problems can arise while fitting this 5-parameter model, as can with our 4-parameter model when the data do not contain enough information.

The risk assessment has to be done in a biological context. For xenobiotic substances the 0 concentration seems a logical reference. Whether hormesis is a risk, is a complicated, biological/ecological question. The choice of reference for the EC50 (Dr. Svensgaard) also has to be answered in a biological setting. Again, for xenobiotic

substances the 0 seems logical, but not for essential minerals. The discussion on hazard identification based on statistical tests (Dr. Svensgaard) has been a subject of interest for us as well. We agree with Dr. Svensgaard, that power of these tests is a major issue here. We favor the benchmark approach. Dose response models can be used here or model-free methods (cf Crump, 1986, Hoekstra & Van Ewijk (1993)).

We apologize for the omission of dimensions in table 1. These should be [mg/l] for concentration and [g] for biomass.

Efron, Bradley, "How biased is the apparent error rate of a prediction rule?", Journal of the American Statistical Association, June 1986, Vol. 81, No. 394, pp. 461-470 Hoekstra, J.A. and Ewijk, P.H. Van, "Alternatives for the no observed-effect level", Environmental Toxicology and Chemistry, Vol. 12, 1993, pp. 187-194 Seber, G.A.F. And Wild, C.J. (1989) Nonlinear Regression, New York: John Wiley & Sons.