REGULATORY TOXICOLOGY AND PHARMACOLOGY 25, 121–129 (1997)
ARTICLE NO. RT961079

Categorical Regression of Toxicity Data: A Case Study Using Aldicarb[1]

Michael L. Dourson,* Linda K. Teuschler,†[2] Patrick R. Durkin,‡ and William M. Stiteler§

*Toxicology Excellence for Risk Assessment, 4303 Kirby Avenue, Cincinnati, Ohio 45223; †National Center for Environmental Assessment, U.S. Environmental Protection Agency, Cincinnati, Ohio 45268; ‡Syracuse Environmental Research Associates, Inc., Fayetteville, New York 13066; and §Syracuse Research Corporation, Syracuse, New York 13210

Received September 25, 1995

Categorical regression is a mathematical tool that can be adapted to estimate potential health risk from chemical exposures. By regressing ordered categories of toxic severity or pathological staging on exposure dose, this method can estimate the likelihood of observing any of the categories of severity at any dose level. Depending on the nature of the available data, these estimates can take the form of incidence rates for any of the categories in an exposed population or the probability of a new study conducted at a specified dose level being classified as one of the categories. Categorical regression is illustrated using toxicity data on aldicarb. For aldicarb, the data fall into three different groups: human clinical studies, dietary exposures in experimental animals, and accidental human exposure by contaminated crops. The U.S. EPA has assessed this literature and developed a reference dose (RfD) of 0.001 mg/kg-day. The results of applying categorical regression to data from human clinical studies suggests a maximum likelihood risk estimate of adverse effects of 0.008% at a 10-fold higher dose than the RfD when blood cholinesterase inhibition is not considered as an adverse effect. When blood cholinesterase inhibition of 20% or more is considered as an adverse effect, a maximum likelihood risk estimate of adverse effects is 0.1% at a dose 10-fold higher than the RfD. © 1997 Academic Press

INTRODUCTION

Expert work groups throughout the world support the use of a "safe" dose concept in noncancer risk assessment (Lu and Sielken, 1991). For example, the United States Environmental Protection Agency currently expresses risk assessments for noncancer toxicity as reference doses (RfDs) (Barnes and Dourson, 1988) or reference concentrations (RfCs) (Jarabek *et al.*, 1990). The RfD is an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.

An RfD is estimated by dividing an experimental or epidemiologic no-observed-adverse-effect level (NOAEL) of the critical effect (i.e., generally the first adverse effect or its known precursor as dose or concentration increases) by one or more uncertainty factors. In selecting the critical effect and its NOAEL, the available observations are sometimes grouped into one of four ordered categories of severity: no effects, nonadverse effects, adverse effects, and frank effects (see for example, Dourson *et al.*, 1985). While this selection of critical effect considers all of the available data to estimate an RfD, it does not provide any explicit estimate of error or variability in the RfD. Furthermore, the RfD approach does not provide a method for expressing or measuring the potential consequences of excess exposures.

Hertzberg and Miller (1985) and Hertzberg (1989) have proposed the use of a categorical regression procedure (McCullagh, 1980) to provide some expression of the risk of adverse health effects from chemical exposures in excess of the RfD. In addition to incorporating different end points and levels of severity, the regression accommodates both quantal and continuous data, and data that are reported both quantitatively and qualitatively. It is also possible to incorporate additional explanatory (independent) variables such as exposure duration (see, for example, Guth *et al.*, 1991). Thus, potential exists for developing models to estimate risk under a variety of different exposure scenarios.

The interpretation of the risk estimate is dependent upon the unit of input to the categorical regression model. For many compounds, the available toxicity data provide information at the dose group level and

[1] The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of the U.S. Environmental Protection Agency.

[2] To whom correspondence and reprint requests should be addressed at Office of Research and Development, National Center for Environmental Assessment, U.S. Environmental Protection Agency, Cincinnati, Ohio 45268.

0273-2300/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

not at the individual experimental animal or human level. When the available data only contain dose group information, then the effects for each dose group are classified into one of the four severity categories. The risk estimate, then, may be interpreted as the probability that an experimental group exposed to a new dose will exhibit a certain category of response. When individual animal or human data are available regarding the frequency of a particular response within a dose group, then the risk estimate may be interpreted as the probability that an individual animal or human exposed to a new dose will exhibit a particular category of response.

The purpose of this paper is to illustrate the application of categorical regression to chemicals of concern to the regulatory community (see for example Rao *et al.*, 1993; or Farland and Dourson, 1992, for arsenic). Aldicarb has been selected here because it has been reviewed by the U.S. EPA (1991) and WHO (1991). The RfD for this compound has undergone revision by the U.S. EPA RfD/RfC Work Group (U.S. EPA, 1993), and interest is high in potential risks above the aldicarb RfD based on residues in food crops. This analysis is not intended to serve as a reevaluation of the aldicarb RfD. Instead, taking the toxicologic evaluations and judgements made in the derivation of the RfD, this analysis shows one approach to estimating potential risks at exposure levels above the RfD.

METHODS

Definitions

Definitions used throughout this paper are consistent with the parlance of the U.S. EPA (1996). These definitions are meant for illustration only; other terms are used in different organizations and countries. These definitions include:

Adverse effects—Biochemical changes, functional impairments, or pathological lesions that impair performance and reduce the ability of the organism to respond to additional challenge. An exposure level at which there are statistically or biologically significant increases in frequency or severity of adverse effects between the exposed population and its appropriate control group is often referred to as an adverse effect level (AEL).

Critical effect—The first adverse effect, or its known precursor, that occurs as the dose rate increases.

Frank effects—Unmistakable adverse effects, such as irreversible functional impairment or mortality. An exposure level that produces frank effects at a statistically or biologically significant increase in frequency or severity between an exposed population and its appropriate control group is often referred to as a frank effect level (FEL).

Lowest observed adverse effect level (LOAEL)—The lowest exposure level at which there are statistically or biologically significant increases in frequency or severity of adverse effects between the exposed population and its appropriate control group.

No-observed-effect level (NOEL)—An exposure level at which there are no statistically or biologically significant increases in the frequency or severity of any effect between the exposed population and its appropriate control group.

No-observed-adverse-effect level (NOAEL)—An exposure level at which there are no statistically or biologically significant increases in the frequency or severity of adverse effects between the exposed population and its appropriate control; some effects may be produced at this level, but they are not considered adverse, nor precursors to adverse effects.

Reference Dose (RfD)—An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious effects during a lifetime.

Uncertainty factor (UF)—One of several, generally 10-fold factors used in operationally deriving the reference dose (RfD or RfC) from experimental data. UFs are intended to account for (i) the variation in sensitivity among the members of the human population, (ii) the uncertainty in extrapolating animal data to the case of humans, (iii) the uncertainty in extrapolating from data obtained in a study that is of less-than-lifetime exposure, (iv) the uncertainty in using LOAEL data rather than NOAEL data, and (5) the inability of any single study to adequately address all possible adverse outcomes in man.

The Categorical Regression Model

The ordered severity classification used in this analysis is the four-component system used in many risk assessments: no effects, nonadverse effects, adverse effects, frank effects. This four component system is useful in categorizing a spectrum of toxicities, often continuous measurements, into a framework for assessing risk. However, categorical regression can be conducted using more or fewer categories. For example, the Agency for Toxic Substances and Disease Registry routinely uses five categories of severity in the development of its hazard profiles. For each human dose group exposed to aldicarb, the site and nature of the observed effect(s) were classified as shown in Table 1 using a scheme adapted from Goldman *et al.* (1990a).

In developing the categorical regression model, it is useful to begin with the special case in which the response falls into one of only two categories with or without ordering assumed. For convenience, let S represent the response, with the two possible outcomes of S de-

TABLE 1
Severity Assignments Used in Analysis of Aldicarb Data

Severity grade	Site	Effect
Frank effects	Cholinergic effects: Severe (Level 4)[a]	Abdominal pain, nausea and/or vomiting, diarrhea
	Cholinergic effects: Severe (Level 3)[a]	Seizures, disorientation or confusion, excitation
	Whole body	Mortality
Adverse effects	Brain, whole blood or RBC acetylcholinesterase	Inhibition[b]
	Cholinergic effects: Mild (Level 2)[a]	Muscular weakness or twitching
	Cholinergic effects: Mild (Level 1)[a]	Blurred vision and/or watery eyes, pinpoint pupils, excess salivation, sweating or clamminess
	Nervous system	Hyperactivity or altered patterns of locomotion
Nonadverse effects	Plasma, whole blood or RBC acetylcholinesterase	Inhibition[b]: Whole blood and RBC were classified as NOAEL only for alternative analysis
No effects	All	No effect

[a] Classification of cholinergic effects adapted from Goldman *et al.* (1990a).
[b] Inhibition of 20% or greater.

noted by 1 (no effect) and 2 (effect). Then a model can be formulated to express the probability that $S = 1$, denoted as $P = \Pr[S = 1]$, and that $S = 2$ as $\Pr[S = 2] = 1 - P$.

It is assumed that the response is related to one or more explanatory variables according to some specified functional relationship called a link function. Suppose there are $k \geq 1$ explanatory variables represented by the vector $\boldsymbol{x}' = (x_1, x_2, \cdots, x_k)$. Because P is a probability, a necessary constraint is that as $\boldsymbol{x}$ varies, the values generated for P must remain between 0 and 1. While a number of different link functions satisfying the required constraint could be assumed, one which has been widely applied and is used in this analysis is the logistic function given by Eq. (1).

$$P = \frac{\exp(\alpha + \beta'\boldsymbol{x})}{1 + \exp(\alpha + \beta'\boldsymbol{x})}, \tag{1}$$

where $\alpha + \beta'\boldsymbol{x} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$. One reason for the popularity of the logistic function is that it is easily converted to a linear relationship by the logit transformation:

$$L = \log(P/(1 - P)) = \alpha + \beta'\boldsymbol{x}. \tag{2}$$

The logit in this case is the log of the odds associated with $S = 1$.

When the response variable represents several ordered categories, the above model can be generalized. Let the response S represent the four severity categories, no effects, nonadverse effects, adverse effects, and frank effects with assigned values 1, 2, 3, and 4, respectively. Note that the severity categories are ordered such that the increase in severity is represented as the assigned value increases. Let the probabilities associated with the four severity outcomes be denoted by $P_1(x)$, $P_2(x)$, $P_3(x)$ and $P_4(x)$, respectively. Then the logit transformation described above can be generalized for application to the cumulative probabilities associated with the ordered severity classes. These cumulative logits are defined as

$$\begin{aligned}
L_1 &= \log\left(\frac{p_1(x)}{1 - p_1(x)}\right) = \log\left(\frac{p_1(x)}{p_2(x) + p_3(x) + p_4(x)}\right) \\
L_2 &= \log\left(\frac{p_1(x) + p_2(x)}{1 - (p_1(x) + p_2(x))}\right) = \log\left(\frac{p_1(x) + p_2(x)}{p_3(x) + p_4(x)}\right) \\
L_3 &= \log\left(\frac{p_1(x) + p_2(x) + p_3(x)}{1 - (p_1(x) + p_2(x) + p_3(x))}\right) \\
&= \log\left(\frac{p_1(x) + p_2(x) + p_3(x)}{p_4(x)}\right).
\end{aligned} \tag{3}$$

The ith cumulative logit is the log of the odds associated with $S \leq i$, i.e., associated with the severity being less than or equal to category i.

Following the two-category example, assume a model relating the cumulative logits to the explanatory variable(s) $\boldsymbol{x}$ as

$$L_i(x) = \alpha_i + \beta'\boldsymbol{x}. \tag{4}$$

This model specifies that the relationship between the cumulative logits and the explanatory variable(s) is determined by a series of parallel straight lines (or planes when there is more than one x variable, i.e., when $k > 1$). The intercepts (α_i's) are called cutpoint parameters.

An interesting result of assuming this particular model is that the log odds ratio for two different values, say $\boldsymbol{x}$ and $\boldsymbol{x}^*$, of the explanatory variable(s) is proportional to the difference between $\boldsymbol{x}$ and $\boldsymbol{x}^*$ with the same proportionality constant for every category (severity

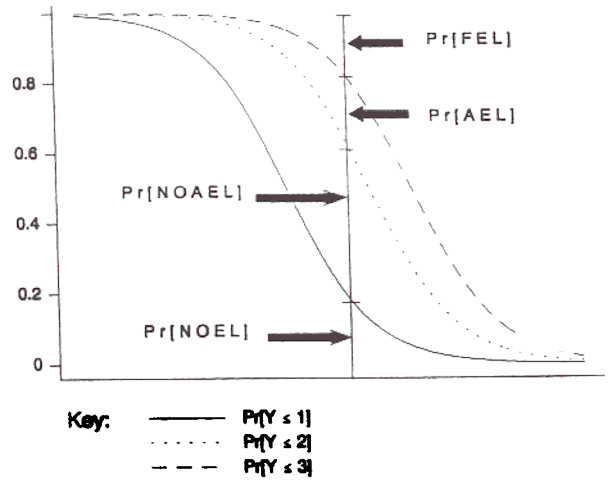

FIG. 1. Relationship of probabilities for various categories in the hypothetical categorical regression model.

class). This proportionality can be tested for and must be present for the modeling results to be valid.

Figure 1 shows a hypothetical proportional odds model relating the severity classes to a single explanatory variable. Each curve represents $P(S \leq i)$, the probability of observing effects of severity class i (for $i = 1$, 2, 3) or less as a function of dose (x). These curves decrease with increases in dose. For example, the $P(S \leq 2)$, or the probability that the observed severity is less than or equal to a nonadverse or minor effect, will decrease as the chemical insult increases. This curve is generated using Equation [4] with $i = 2$. The probability of observing a particular severity class is then calculated by subtraction, e.g., the probability of observing nonadverse effects is the difference between two curves: Pr[nonadverse effects] = Pr[nonadverse effects or less] − Pr[no effects].

Confidence limits are obtained for the probabilities estimated from a categorical regression model by calculating the appropriate confidence limits for the linear predictors $L_i = \alpha_i + \beta' x$ and then back transforming with the logistic function.

The Aldicarb RfD

The current RfD for aldicarb is 0.001 mg/kg/day. This value and supporting text are currently found in the U.S. EPA's Integrated Risk Information System (U.S. EPA, 1996). This RfD is based on a human NOEL for cholinergic effects (e.g., sweating) of 0.01 mg/kg and an uncertainty factor of 10 to account for sensitive humans. As detailed in U.S. EPA (1996) and other reviews (U.S. EPA, 1991; WHO, 1991), aldicarb has been associated with a variety of adverse effects. Cholinergic effects, however, were selected by U.S. EPA as the critical endpoint for derivation of an RfD. The available data fall into three general classes: human experimental studies, reports of human poisonings attributed to the consumption of produce contaminated with aldicarb, and studies of acetylcholinesterase inhibition in experimental animals.

Human Experimental Studies

The two clinical studies available on aldicarb, Haines (1971) and Wyld *et al.* (1992), are summarized in Table 2. These studies are cited as Union Carbide (1971) and Rhone-Poulenc (1992) in U.S. EPA (1996). Both studies had similar experimental designs.

In the study by Haines (1971), analytical grade aldicarb (99.2% pure) was consumed as a bolus in distilled water by groups of four males at doses of 0.025, 0.05, and 0.1 mg/kg between 9:00 AM and 9:15 AM. The individuals engaged in normal business activities during the day except for postexposure hours 1, 2, 3, 4, and 6 when clinical observations and blood and urine collections were made. A substantial decrease in blood AChE activity defined as >20% was noted in all groups and in all individuals for at least one postexposure observation. In the high-dose group, all four individuals evidenced signs of cholinergic effects (see Table 1 for a summary) that included weakness and sweating; two individuals reported nausea. In the mid-dose group one individual had a runny nose. In the low-dose group, one individual reported apprehension. The duration and time to onset of all these symptoms except the runny nose are consistent with aldicarb toxicity and were attributed to aldicarb by the U.S. EPA (Sette, 1992).

Information on the more recent study by Wyld *et al.* (1992) is taken from Sette (1992) with additional statistical analyses by Pettigrew (1992) and Kahn and Jacobs (1992). Aldicarb (99.0% pure) was administered to 36 subjects in 200 ml of orange juice during breakfast. Unlike the Haines study, Wyld used control groups of 16 males and 6 females. Not all exposures were concurrent but rather occurred on 8 different days. Four males served both in control groups on 1 day and in exposed groups on another day. Each subject received orange juice, with or without aldicarb, during breakfast at approximately 8:30 a.m. The subjects were instructed to sip the orange juice throughout the breakfast period. After breakfast, all subjects were kept recumbent for 6 hr. Cholinesterase measurements were taken at 1, 2, 4, 6, 8, and 21 hr postexposure. The key clinical observations and number of persons with plasma and/or red blood cell (RBC) cholinesterase inhibition in excess of 20% of 1-hr pretest values which were attributed to aldicarb exposure are also summarized in Table 2. Note that headaches were reported in the 0.01 mg/kg males but were not attributed to aldicarb exposure because the time of onset was not always consistent with aldicarb administration and the effect

TABLE 2
Frequency of Responders with Clinical Signs or Blood Cholinesterase Inhibition in Humans from Clinical Studies on Alidcarb Ingestion

Study	Dose (mg/kg/day)	Group size	Responders with clinical signs	Responders with blood cholinesterase inhibition[a]
Haines (1971)	0.025	4	1 Apprehension	4 Whole blood
	0.05	4	1 Runny nose[c]	4 Whole blood
	0.10	4	4 Weakness and sweating, nausea in 2 individuals	4 Whole blood
Wyld *et al.* (1992)[b]	0	22	0	0 Plasma and 0 RBC
	0.010	8	2 Headaches[c]	0 Plasma and 0 RBC
	0.025	12	1 Sweating	12 Plasma and 3 RBC
	0.050	12	1 Sweating	12 Plasma and 11 RBC
	0.06	1	1 Sweating	1 Plasma and 1 RBC
	0.075	3	1 Lightheadedness	3 Plasma and 3 RBC

[a] Inhibition at or exceeding 20% at 1 hr postexposure when compared to individual pretest values. Haines (1971) measured inhibition in whole blood and 1-hr pretest values were used for comparison. Values for Wyld *et al.* (1992) are the maximum of plasma and RBC inhibition.
[b] Male and female toxicity data are combined from the Wyld *et al.* (1972) study. Haines (1971) only used males.
[c] In determining the RfD, headaches and runny nose were discounted as being related to aldicarb exposures (U.S. EPA, 1996).

did not occur at higher doses (Sette, 1992). Females did not evidence any of the cholinergic effects seen in the males. All groups of females did not evidence a drop in blood pressure compared to baseline. In addition, the 0.025 mg/kg group of females evidenced a statistically significant reduction in blood pressure compared to the female control group.

Poisoning Incidents

Episodes of poisoning in human populations associated with the consumption of contaminated produce have been reported in the literature (Goldman *et al.*, 1990a; Hirsch *et al.*, 1987). Goldman *et al.* (1990b) have estimated doses associated with reported poisoning episodes based on age, sex, average body weights, self-reported symptoms and consumption, and aldicarb sulfoxide residues from watermelons and cucumbers. As illustrated in Fig. 2, the doses estimated from these poisoning incidents range between 0.002 and 0.1 mg/kg. While recognizing the uncertainties associated with the dose estimation methods and self-reporting, the U.S. EPA (1996) concluded that these dosage estimates are regarded as valid and reasonable estimates of the potency of aldicarb. The evidence presented suggests that clinical manifestations of cholinesterase poisoning occurred at doses between 0.002 and 0.1 mg/kg.

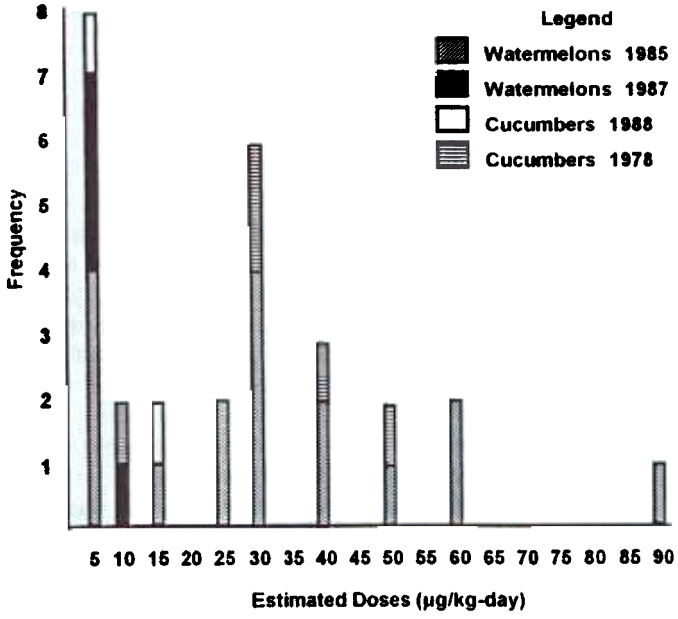

FIG. 2. Number of cases of human poisonings associated with the consumption of contaminated watermelons or cucumbers. Data from Goldman *et al.* (1990b).

In evaluating the significance of poisoning incidents in deriving the RfD, conflicting evaluations of the Goldman *et al.* (1990b) study had to be resolved. Rothman *et al.* (1991) and Peterson *et al.* (1991) both reviewed the Goldman *et al.* (1990b) study. These reviewers concluded that the Goldman study is not appropriate for deriving an RfD because of uncertainties in the exposure assessment and in the ability to clearly demonstrate that each case of poisoning was caused by aldicarb. Rothman *et al.* (1991) took particular note of the "bimodal" distribution of cases (see Fig. 2) and suggested that cases in the low-dose region may have been reported as a consequence of the publicity surrounding the poisoning episodes and the reported effects may not have been caused by aldicarb. Milbey (1992), in reviewing these criticisms, concluded that while the studies are limited, they do provide an indication that adverse effects can occur at doses as low as 0.002 mg/kg/day, a factor of two above the RfD.

In developing the RfD of 0.001 mg/kg/day, the U.S. EPA has considered both positions. On the one hand,

TABLE 3
Frequency of Categories of Effect Associated with Aldicarb Exposure in Humans

Study	Dose (mg/kg/day)	Group size	Frequency of responders within categories[a]			
			No effects	Nonadverse effects	Adverse effects	Frank effects
Wyld	0.0	22	22 (22)	0 (0)	0 (0)	0 (0)
Wyld	0.010	8	8 (8)	0 (0)	0 (0)	0 (0)
Wyld	0.025	12	0 (0)	8 (11)	4 (1)	0 (0)
Haines	0.025	4	0 (0)	0 (3)	4 (1)	0 (0)
Wyld	0.050	12	0 (0)	1 (11)	11 (1)	0 (0)
Haines	0.050	4	0 (0)	0 (4)	4 (0)	0 (0)
Wyld	0.075[b]	4	0 (0)	0 (2)	4 (2)	0 (0)
Haines	0.10	4	0 (0)	0 (0)	2 (2)	2 (2)

[a] Data are summarized from Table 2. Numbers in parentheses are frequencies when whole blood (Haines, 1971) or RBC (Wyld *et al.*, 1992) cholinesterase inhibition of 20% or greater are not considered as adverse effects and are thus categorized as NOAELs. Numbers without parentheses are frequencies when whole blood (Haines, 1971) or RBC (Wyld *et al.*, 1992) cholinesterase inhibition of 20% or greater are considered as adverse effects and are thus categorized as AELs.

[b] The data of the 0.060 mg/kg/day dose group were combined with those of the 0.075 mg/kg/day dose group.

U.S. EPA (1996) found that exposures as low as 0.002 mg/kg/day may be associated with adverse effects, presumably in sensitive individuals. On the other hand, as described above, the RfD of 0.001 mg/kg/day is based on the clinical data. This is not necessarily contradictory; if the dose of 0.002 mg/kg/day is seriously regarded as the LOAEL for sensitive individuals, the RfD could well be its NOAEL. One common interpretation of the RfD is that it represents the NOAEL of the sensitive human population.

Categorical regression can be used to model these data. However, since incidence data are not available, the results are not directly comparable to those of the clinical studies and were, therefore, not included in the regression analysis given in this paper.

Animal Studies

As described in U.S. EPA (1996) a medium confidence data base exists on the toxicity of aldicarb in experimental animals and humans. Chronic toxicity in animals is manifested at doses roughly comparable to those that produce acute toxicity, and these doses are somewhat higher when compared to the information described for humans in this paper. Nearly all hazards from aldicarb exposure are associated with cholinesterase inhibition, even though other effects have been monitored.

For example, an LD_{50} value in rats is given as 0.8 mg/kg/day (Klaassen *et al.*, 1986), whereas chronic studies in rats show NOAELS for cholinesterase inhibition, as well as a host of other effects, of 0.3 mg/kg/day (U.S. EPA, 1996). Reproductive studies over two generations in rats demonstrate only decreased pup body weight at 0.7 mg/kg/day and a NOAEL for this effect at 0.3 mg/kg/day. Parental toxicity was limited to cholinesterase inhibition and decreased body weights at 0.7 mg/kg/day (U.S. EPA, 1996).

Other species or bioassays show effects in a similar range of dose rate. For example, chronic studies in dogs demonstrate LOAELs for plasma cholinesterase inhibition as low as 0.028 mg/kg/day (U.S. EPA, 1996) in the absence of clinical signs; other effects were not seen, however, at doses as high as 0.25 mg/kg/day. Developmental toxicity bioassays in rats show conflicting NOAELs and LOAELs for developmental effects in the range of 0.125 to 1.0 mg/kg/day. Developmental toxicity studies in rabbits illustrate a NOAEL of 0.25 mg/kg/day and a LOAEL of 0.5 mg/kg/day for these endpoints (U.S. EPA, 1996).

RESULTS

Regression Results

Incidence data are available on both blood cholinesterase inhibition and clinical cholinergic effects of aldicarb from two studies (summarized in Table 2). The reported laboratory and clinical signs of cholinergic inhibition are readily categorized using the classification scheme given in Table 1. RBC and whole blood inhibition >20% were categorized as either adverse effects or, for the alternative analysis, as nonadverse effects (see Table 3), resulting in two regressions.

Categorical regression was performed on the combined Haines (1971) and Wyld (1992) data sets for the two different classifications of cholinesterase inhibition. The final model used the log of the dose as the explanatory variable, rather than dose, as the proportional odds assumption was not met under the latter treatment. The control group data was excluded from

TABLE 4
Probability of a Category of Effect Input to Model: Inhibition >20% = AEL

Dose (mg/kg/day)	Probability			
	No effects	No adverse effects	Adverse effects	Frank effects
0.01	0.94	0.05	0.001	0.0
0.02	0.13	0.73	0.14	1E-5
0.04	0.001	0.05	0.95	0.001
0.06	5E-5	0.002	0.97	0.03
0.08	1E-5	4E-4	0.88	0.12
0.10	0.0	9E-5	0.6	0.4

the analysis for this reason, but also because four of the controls in the Wyld (1992) study served as control and treatment subjects on different days of the study. For both regressions, all parameter estimates were significantly different from zero ($P < 0.001$) and the χ^2 tests for the contribution of dose to the model were significant ($P = 0.0001$). At the $P = 0.05$ level of significance, the χ^2 test for testing the proportional odds assumption was acceptable ($P = 0.09$) for the regression where cholinesterase inhibition >20% was classified as an adverse effect, but was not significant ($P = 0.04$) for the other model. Although the goodness-of-fit for this regression is marginal, but not significant, the P values for the two models are comparable. We chose to show the results of both regressions because of existing debate within the scientific community on whether increases in RBC and whole blood cholinesterase inhibition are adverse effects or not.

The results of these two regressions are shown in Tables 4 and 5 and in Fig. 3. "Risk," as it is shown in Table 4 and in Fig. 3, is the probability that a person would exhibit a certain category of effect given exposure to a specific dose of aldicarb. In Table 5, the risk is the combined probability of an adverse or a frank effect, P(AE or FE), for each analysis, along with the 95% upper confidence limit for that risk, given exposure to a specific dose of aldicarb. For both analyses, the separate and combined probabilities associated with adverse or frank effects at the RfD (0.001 mg/kg/day) are estimated to be nil. As shown by a comparison of the curves in Fig. 3, the differences in the results of these two analyses lie in the point of marked ascent and maximum height of the dose–response curves.

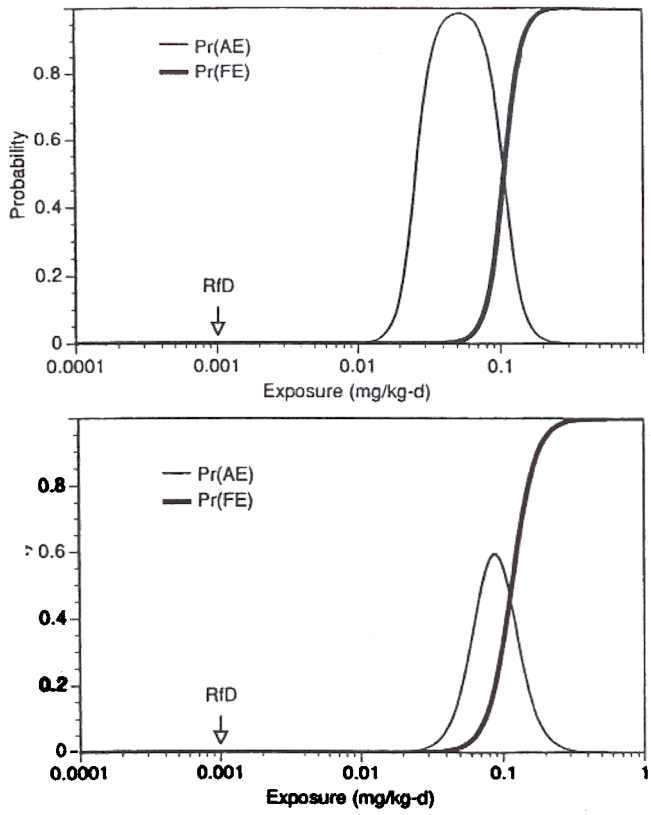

FIG. 3. Probability of either an adverse effect (AEL) or frank effect (FEL) with aldicarb exposure, where whole blood or RBC cholinesterase inhibition of 20% or greater is considered to be an AEL (top) or a NOAEL (bottom). See Table 1 for categories of effect and text for discussion.

TABLE 5
Probability of an Adverse or Frank Effect

Dose (mg/kg)	Inhibition > 20% = nonadverse effect[a]		Inhibition > 20% = adverse effect	
	P(AE or FE)	Upper 95%CL	P(AE or FE)	Upper 95%CL
0.001	—	0.00001	—	0.00001
0.003	—	0.00015	—	0.0007
0.01	0.00008	0.004	0.0014	0.04
0.015	0.0008	0.015	0.03	0.17
0.02	0.002	0.03	0.14	0.36
0.025	0.008	0.06	0.44	0.67
0.03	0.02	0.11	0.79	0.93
0.035	0.04	0.15	0.89	0.97
0.04	0.07	0.21	0.95	0.99
0.10	0.88	0.97	0.99	1.00

[a] Model failed to meet proportional odds assumption. Probabilities are shown only to allow for a comparison of the two approaches to data categorization.

Estimated Risks

The two curves in the top of Fig. 3 illustrate the P(AE) and the P(FE), given dose, when whole blood and RBC cholinesterase inhibitions of 20% or greater are considered as adverse effects. For example, the probability of effect levels at 10-fold of the RfD (0.01 mg/kg/day) are estimated as P(AE) = 0.1% and P(FE)

= 0.0% and at 100-fold of the RfD (0.1 mg/kg/day) are estimated at $P(AE) = 60\%$ and $P(FE) = 40\%$. These probabilities are also given in Table 4, which shows the probability of observing any of the four categories given dose. The combination of small numbers of study subjects and the scatter observed in the dose–response patterns suggest that the error associated with these estimates may be large. Table 5 shows the combined P(AE or FE) and the corresponding 95% upper confidence limits. For the RfD and for 10-fold and 15-fold above the RfD, the P(AE or FE) values are 0.0, 0.1, and 3.0%, respectively. For these same doses, the upper 95% confidence limits on the P(AE or FE) are 0.001, 4, and 17%, respectively.

The bottom of Fig. 3 and Table 4 show the regression results when whole blood and RBC cholinesterase inhibitions are not considered as adverse effects. As expected, the risks are lower than those of the other regression, since one group of adverse effect (i.e., whole blood or RBC inhibition) has been reclassified as nonadverse.

DISCUSSION AND CONCLUSIONS

This paper illustrates how categorical regression, which incorporates toxicological judgment and statistical analyses, can be used to assess risks above the RfD. In so doing, the "quality" of the RfD is, at least indirectly, evaluated. Similar evaluations using categorical regression have been conducted with different chemicals (Guth *et al.*, 1991; Rao *et al.*, 1993). A related method has also been published by Renwick and Walker (1993). A feature related to both of these methods is that the magnitude and severity of the effects used to determine the RfD are related to estimations of risk above it.

As illustrated in this analysis, sufficient data are available in human clinical studies to assess the risk above the RfD on aldicarb. Judgments differ concerning the proper classification of the observed responses in the clinical studies. The analysis presented in this paper suggests that these make a difference in the assessment of risk above the RfD. For example, the combined risks of adverse effects or frank effects at 10-fold of the RfD are either 0.1 or 0.008% when whole blood and RBC cholinesterase inhibition are considered to be adverse effects, or they are not, respectively (see Table 5). This is a 12-fold difference in the risk estimate.

The dose–response relationship for aldicarb is extremely steep, and the logistic model indicates a very low level of risk near the RfD. Based on maximum likelihood estimates, substantial levels of risk (e.g., 0.1%) are not encountered until the RfD is exceeded by a factor of 10 or more (see Fig. 3). Because of the small number of subjects in these experiments, 95% upper confidence limit estimates of risk might be more appropriate. These upper limit estimates of risks of either an AEL or an FEL at three-fold the RfD are 0.07% when whole blood and RBC inhibition are considered as adverse and are 0.01% when these effects are not considered adverse.

A second issue to be considered in interpreting this analysis for aldicarb is the effect of toxicological judgment on the model results. For example, consider the judgment that headaches reported at the 0.010 mg/kg dose group of the Wyld *et al.* (1992) study were not attributed to aldicarb exposure (Sette, 1992). If headaches were judged as related to aldicarb exposure, the risks above the RfD would increase. On the other hand, another possible reclassification would be to judge the "apprehension" reported at the 0.025 mg/kg dose level in the Haines study as not related to aldicarb exposure. This individual became nervous when he overheard reports of adverse effects in the high-dose group. The apprehension reported was transitory and may have been unrelated to aldicarb exposure. This latter classification would serve to decrease the estimation of risk above the RfD in Fig. 3.

The third, and perhaps the largest, issue is the interpretation and use of the human poisoning incidents. The results from the human poisoning incidents are not inconsistent with the RfD and can be used to corroborate the results of the categorical regressions shown in Fig. 3. However, it is difficult to quantitatively compare the results of Figs. 2 and 3, and categorical regression cannot be used directly to resolve this lack of comparability. The type of data represented by the reports of poisoning cannot be directly incorporated into the categorical regression procedure shown here because information on the entire population at risk is not available. That is, data on the exposed humans showing no effects or failing to report effects are missing. Alternatively, not using the incidence data in the clinical studies allows a comparison with the poisoning incidents, but the resulting confidence in the assessment is not as great as with the incidence data of the clinical studies alone.

The experimental animal data support the observations seen in the human studies. Moreover, categorical regression of these animal data was also done (results not shown), and is generally consistent with the results of clinical human data. A more definitive comparison between humans and experimental animals was not attempted because not all animal studies yielded incidence data, and the human clinical studies were considered sufficient to draw some conclusions. Such comparative analysis may be an area for future research as more chemicals are analyzed using a categorical regression approach.

As with any empirical modeling effort, this application of categorical regression is somewhat model dependent and does not fully express all of the uncertainties

or resolve all of the judgmental issues. In addition, the estimations of risk above the RfD are model dependent especially as estimations are made in the lower dose ranges away from the actual data. Because of shortcomings in the reported data, this method cannot be used to directly interpret and use the reports of human poisoning. Nonetheless, the technique is a promising additional tool that appears complementary to the RfD methodology. Although it gives different information than the RfD and does not necessarily address or evaluate uncertainties associated with the RfD method, it offers a way to use more of the available toxicity data when making risk management decisions.

ACKNOWLEDGMENTS

The authors thank Drs. Robert Beliles, Harlal Choudhury, Penny Fenner-Crisp, Thomas Crisp, Reto Engler, Richard Hertzberg, Amal Mahfouz, Pat Murphy, Ed Ohanian, and Roy Sjoblad with the U.S. Environmental Protection Agency for helpful comments and insight into the numerous risk assessment issues on this chemical.

REFERENCES

Anderson, J. A., and Philips, P. R. (1981). Regression discrimination, and measurement models for ordered categorical variables. *Appl. Statistics* **30**, 22–31.

Barnes, D. G., and Dourson, M. L. (1988). Reference Dose (RfD): Description and use in health risk assessment. *Reg. Toxicol. Pharmacol.* **8**, 471–486.

Dourson, M. L., Hertzberg, R., Hartung, R., and Blackburn, K. (1985). Novel methods for the estimation of acceptable daily intake. *Toxicol. Ind. Health.* **1**(4), 23–33.

Farland, W., and Dourson, M. L. (1992). Noncancer health endpoints: Approaches to quantitative risk assessment. In *Comparative Environmental Risk Assessment* (R. Cothern, Ed.) Lewis, Boca Raton, LA.

Goldman, L. R., Smith, D. F., Neutra, R. R., *et al.* (1990a). Pesticide food poisoning from contaminated watermelons in California, 1985. *Arch. Environ. Health* **45**(4), 229–236.

Goldman, L. R., Belelr, M., and Jackson, R. J. (1990b). Aldicarb food poisonings in California, 1985–1988: Toxicity estimates for humans. *Arch. Environ. Health* **45**(3), 141–147.

Guth, D. J., Jarabek, A. M., Wymer, L., and Hertzberg, R. C. (1991). *Evaluation of Risk Assessment Methods for Short-Term Inhalation Exposure.* Presented at the 4th Annual Meeting of the Air & Waste Management Association, Vancouver, British Columbia, June 16–21.

Haines, R. G. (1971). *Ingestion of Aldicarb by Human Volunteers: A Controlled Study of the Effect of Aldicarb on Man.* Unpublished report. In EPA Pesticide Petition No. 1F1008.

Hertzberg, R. C. (1989). Extrapolation and scaling of animal data to humans: Fitting a model to categorical response data with application to species extrapolation of toxicity. *Health Phys.* **57**,(Suppl. 1), 405–409.

Hertzberg, R. C., and Miller, M. (1985). A statistical model for species extrapolating using categorical response data. *Toxicol. Ind. Health* **1**(4), 43–63.

Hirsch, G. H., Mori, B. T., Morgan, G. B., Bennett, P. R., and Williams, B. C. (1987). Report of illnesses caused by aldicarb-contaminated cucumbers. *Food Add. Contam.* **5**(2), 155–160.

Jarabek, A. M., Menache, M. G., Overton, J. H., Jr., Dourson, M. L., and Miller, F. J. (1990). The U.S. Environmental Protection Agency's inhalation RfD methodology: Risk assessment for air toxics. *Toxicol. Ind. Health* **6**(5), 279–301.

Kahn, H. D., and Jacobs, H. L. (1992). *Memorandum: Analysis of Aldicarb Human Study Data.* U.S. Environmental Protection Agency, Washington D.C., Sept. 2, 1992.

Klaassen, C. D., Amdur, M. O., and Doull, J. (1986). Casarett and Doull's Toxicology: The Basic Science of Poisons, 3rd ed., p. 539. Macmillan, New York.

Lu, F. C., and Sielken, R. L. (1991). Assessment of safety/risk of chemicals: Inception and evolution of the ADI and dose-response modeling procedures. *Toxicol. Lett.* **59**, 5–40.

McCullagh, P. (1980). Regression models for ordinal data. *J. R. Stat. Soc. B* **42**, 109–142.

Milbey, T. H. (1992). Review of Goldman et al., 1990 a&b. U.S. Environmental Protection Agency, Washington DC, February 11, 1992.

Peterson, H. D. (1991). *Preliminary Results of Rhone-Poulenc's Review of Goldman et al. 1990 a&b.* U.S. Environmental Protection Agency, Washington DC, Dec. 18, 1991.

Pettigrew, H. M. (1992). *Memorandum: Aldicarb: Statistical Analysis of Data from a Safety and Tolerability Study of Aldicarb at Various Dose Levels in Healthy Male and Female Volunteers.* U.S. Environmental Protection Agency, Washington DC, Sept 2, 1992.

Rao, V. R., Levy, K., and Lustik, M. (1993). Logistic regression of inhalation toxicities of perchloroethylene—Application in noncancer risk assessment. *Reg. Toxicol. Pharmacol.* **18**, 233–247.

Renwick, A. G., and Walker, R. (1993). An analysis of the risk of exceeding the acceptable or tolerable daily intake. *Reg. Toxicol. Pharmacol.* **18**, 463–480.

Rothman, K., Pastides, H., and Cole, P. (1991). *Epidemiological Review of Goldman et al.* Paper. Submitted to Rhone-Poulenc Ag Company. U.S. Environmental Protection Agency, Washington DC, Oct. 30, 1991.

Sette, W. F. (1992). *Memorandum: Joint OPPT/OW/ORD Review of 1992 Aldicarb Human Study.* U.S. Environmental Protection Agency, Washington DC, Sept. 4, 1992.

U.S. EPA. (1991). *Drinking Water Criteria Document for Aldicarb.* Prepared by the Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office, Cincinnati, OH for the Office of Water, Washington, DC.

U.S. EPA. (1996). *Integrated Risk Information System (IRIS).* Online. National Center for Environmental Assessment, Cincinnati, OH.

WHO (World Health Organization). (1991). *Environmental Health Criteria 121: Aldicarb.* Geneva, World Health Organization. 130 p.

Wyld, P. J., Watson, C. E., Nimmo, W. S., and Watson, N. (1992). *A Safety and Tolerability Study of Aldicarb at Various Dose Levels in Healthy Male and Female Volunteers.* Study submitted to EPA by Rhone-Poulenc Company. As summarized by Sette (1992).