Environmental and Ecological Statistics **1**, 303–314 (1994)

Statistical analysis of bioassays, based on hazard modelling

J.J.M. BEDAUX and S.A.L.M. KOOIJMAN

Department of Theoretical Biology, Vrije Universiteit Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

Received January 1993. Revised October 1994.

A stochastic model is proposed to describe time–dependent lethal effects of toxic compounds. It is based on simple mechanistic assumptions and provides a measure for the toxicity of a chemical compound, the so-called killing rate. The killing rate seems a promising alternative for the LC50. The model also provides the no-effect level and the LC50, both as a function of exposure time. The model is applied to real data and to simulated data.

Keywords: killing rate, LC50, maximum likelihood, no-effect level, one-compartment model, quantal assay data, time dependent toxicity

1. Introduction

The analysis of survival data is important in toxicological studies. In many laboratories, bioassays are done routinely to investigate toxicological properties of new chemical compounds. Determination of LC50 values and no-effect concentrations (NEC) is the main objective. The LC50 value of a compound is the concentration expected to cause death of 50% of the population within a fixed time. The no-effect concentration is the maximal concentration having no lethal effect within the duration of the experiment. Both LC50 and NEC depend on the species chosen, the exposure time, the temperature at which the experiment is performed, the age of the experimental animals, etc. In routine experiments, animals are exposed to a compound in a range of concentrations. After a fixed time chosen on the basis of experience and intuition, and depending on the species used, the numbers of survivors are counted for every concentration. The resulting LC50 and NEC estimates only tell something about exposure during that fixed time and as such they have a limited meaning. In more elaborate experiments survivorship is measured after several exposure times. This enables the study of the time-dependence of LC50 and NEC.

For simple experiments, a wide variety of statistical methods is used to estimate LC50 and NEC. Nonparametric methods, such as moving average or (trimmed) Spearman–Kärber, as well as parametric methods, such as probit or logit analysis, are used to estimate LC50; for a review see Hoekstra (1991). Morgan (1988) reviews several extensions of the classical logit and probit models. Estimation methods of NECs can be found in Cox (1987). Kooijman (1981) proposed models to estimate NEC and (time-dependent) LC50 in various experimental designs.

In most parametric procedures, distribution functions are chosen *ad hoc* to describe the stochastic behaviour of the data. Biological knowledge is rarely incorporated in the stochastic model. In this paper we develop a stochastic model based on simple assumptions that are still realistic from a biological point of view. The key assumption is that the hazard rate is proportional to the

1352-8505 © 1994 Chapman & Hall

concentration of the compound in the animal. The idea of relating the hazard rate to the dose is not new. Puri and Senturia (1972), Laurence and Morgan (1989), Morgan (1992, Chapter 5) proposed a stochastic model in which the hazard rate is a function of the concentration of the compound.

A second assumption concerns the kinetics of the compound. We assume a simple linear one-compartment model. The incorporation of the kinetics of the compound in the dose–response model is also not new. Puri and Senturia (1972) constructed a stochastic process underlying the concentration in the animal. Van Ryzin and Rai (1987) used Michaelis–Menten nonlinear kinetics to describe the internal concentration as a function of the external concentration. They only considered steady-state conditions, so their approach cannot be used in the study of time-dependent toxicity.

The usual experimental data sets consist of counts of surviving organisms that have been exposed to a chemical compound at a range of concentrations during a fixed time of exposure. Most statistical methods only provide estimates in simple designs involving a single exposure time. In order to analyse experiments involving several exposure times, the time-dependence of the parameters must be modelled. Kooijman (1981) proposed an extension of the log-logistic tolerance distribution. His model will be compared with the present model. Carter and Hubert (1984) proposed a growth-curve model approach.

Maximum likelihood methods are used to estimate the parameters of the model. To study the statistical properties of the estimators, we applied the method to data obtained by Monte Carlo simulation. In addition, we applied the method to experimental data.

2. Modelling survivorship

The key assumption in this paper is that the hazard rate is proportional to the concentration of the chemical compound in the animal, as far as it exceeds a so-called no-effect level. To be more precise, we assume the hazard $h(t)$ to be proportional to the (positive) difference between the concentration $C(t)$ and the no-effect level C_0. Generally, we do not know the actual concentrations in the animal. We only know the concentration in the environment $c(t)$. (Note that throughout this paper capital C is used to denote the concentration in the animal and lower-case c for the concentration in the environment. Symbols used are shown in Table 1.) We therefore have to make assumptions about the uptake dynamics of the compound in the animal. A simple model, which is still realistic from a biological point of view, is the so-called one-compartment model (Jacquez, 1985).
That is,

$$\frac{\mathrm{d}C(t)}{\mathrm{d}t} = k_\mathrm{u}c(t) - k_\mathrm{e}C(t) \tag{1}$$

where k_e is the elimination rate and k_u the uptake rate. The solution of (1) is easily found to be

$$C(t) = C(0)\exp(-k_\mathrm{e}t) + k_\mathrm{u}\int_0^t \exp(-k_\mathrm{e}(t-\tau))c(\tau)\,\mathrm{d}\tau. \tag{2}$$

Here we consider experimental situations in which $c(t)$ is constant, say $c(t) = c$. We also assume that $C(0)$ is negligibly small. Then (2) reduces to

$$C(t) = \frac{k_\mathrm{u}}{k_\mathrm{e}}c(1 - \exp(-k_\mathrm{e}t)). \tag{3}$$

The ultimate concentration is given by $\lim_{t\to\infty} C(t) = ck_\mathrm{u}/k_\mathrm{e}$, the ratio $k_\mathrm{u}/k_\mathrm{e}$ being known as the bioconcentration factor. If this value is smaller than the no-effect level C_0, i.e. if $c < C_0k_\mathrm{e}/k_\mathrm{u}$, there

Table 1. List of symbols. T, M and L denote the dimensions time, mass and length.

Symbol	*Dimension*	*Interpretation*
h	T^{-1}	hazard rate
t	T	time
C	ML^{-3}	Concentration of the compound in the animal
C_0	ML^{-3}	no-effect level in the animal
c	ML^{-3}	Concentration of the compound in the environment
k_u	T^{-1}	uptake rate
k_e	T^{-1}	elimination rate
c_0	ML^{-3}	no-effect level in the environment
t_c	T	time at which C exceeds the no-effect level
$k_\dagger$	$L^3M^{-1}T^{-1}$	killing rate
λ	T^{-1}	control mortality rate
$\kappa_\dagger$	$L^3M^{-1}T^{-2}$	killing acceleration
c_{L50}	ML^{-3}	ultimate LC50 value
β		slope parameter of the logistic distribution
x_{ij}	–	number of surviving animals at time t_i and concentration c_j
p_{ij}	–	probability of an animal to die between t_{i-1} and t_i, at c_j
n_{ij}	–	number of animals died between t_{i-1} and t_i, at c_j

will be no effect at all, even after long exposure times. This defines the environmental (ultimate) no-effect level $c_0 = C_0 k_e / k_u$.

If $c > c_0$ there is a point in time t_c at which $C(t)$ exceeds C_0 (see Fig. 1). From (3) t_c can be calculated to be

$$t_c = -\frac{1}{k_e} \ln\left(1 - \frac{c_0}{c}\right). \tag{4}$$

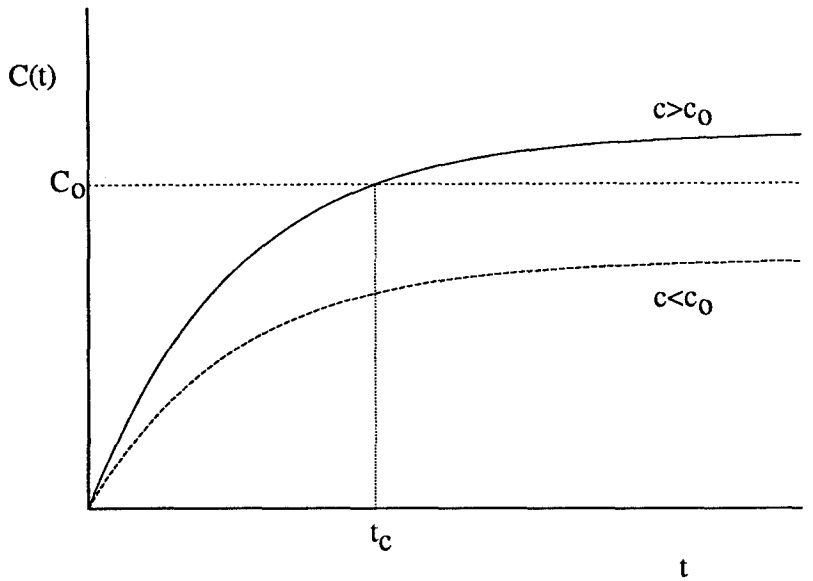

Figure 1. Accumulation curves for two values of the environmental concentration c, one below and one above c_0.

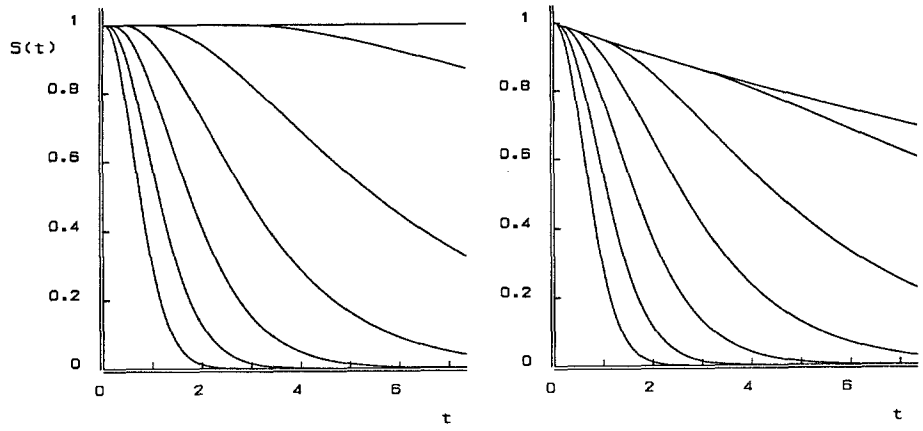

Figure 2. Survivor functions for various values of c (from top downwards 1, 2, 4, 8, 16, 32 and 64). Parameter values are $c_0 = 1.5$, $k_\dagger = 0.1$ and $k_e = 0.5$. Left $\lambda = 0$, right $\lambda = 0.05$.

The hazard rate is now given by

$$h(t;c) \propto C(t) - C_0 = \frac{k_u}{k_e}(c(1 - \exp(-k_e t)) - c_0)_+$$

or

$$h(t;c) \propto (c(1 - \exp(-k_e t)) - c_0)_+, \tag{5}$$

where $(x)_+$ means the maximum of x and 0. This notation will be used frequently in the following. The proportionality constant in (5), written as $k_\dagger$, will be called the killing rate, as proposed in Kooijman (1993, p. 277). It has dimension (concentration $\times$ time)$^{-1}$ and can be viewed as a measure for the toxicity of the compound with respect to survival. The hazard rate can now be written as

$$h(t;c) = k_\dagger(c(1 - \exp(-k_e t)) - c_0)_+.$$

The survivor function $S(t;c)$ of the time of dying caused by the chemical compound at concentration c, is then given by

$$S(t;c) = \begin{cases} \exp\left(\frac{k_\dagger}{k_e} c(\exp(-k_e t_c) - \exp(-k_e t)) - k_\dagger(c - c_0)(t - t_c)\right) & \text{if } c > c_0 \text{ and } t > t_c \\ 1 & \text{otherwise.} \end{cases} \tag{6}$$

In this model formulation, control mortality is readily included. Assuming independence of death caused by the chemical compound and death caused by natural circumstances, we can simply add the corresponding hazard rates. The hazard rate due to control mortality, λ, will be taken constant. This is reasonable because the duration of the experiments is mostly short compared with the mean lifetime of the organisms used. The resulting survivor function is the one obtained in (6) multiplied by $\exp(-\lambda t)$. In Fig. 2 some survival curves are plotted with and without control mortality, for one choice of parameter values.

An interesting special case concerns extremely small elimination rates, so $k_e \to 0$. This occurs for instance with cadmium in some soil arthropods (Janssen *et al.*, 1991). The accumulation process reduces to $(\mathrm{d}/\mathrm{d}t)C = k_u c$, so that $C(t) = k_u ct$ if again the initial concentration in the tissue is negligibly small. The no-effect level (in the environment) now equals 0, because a very small

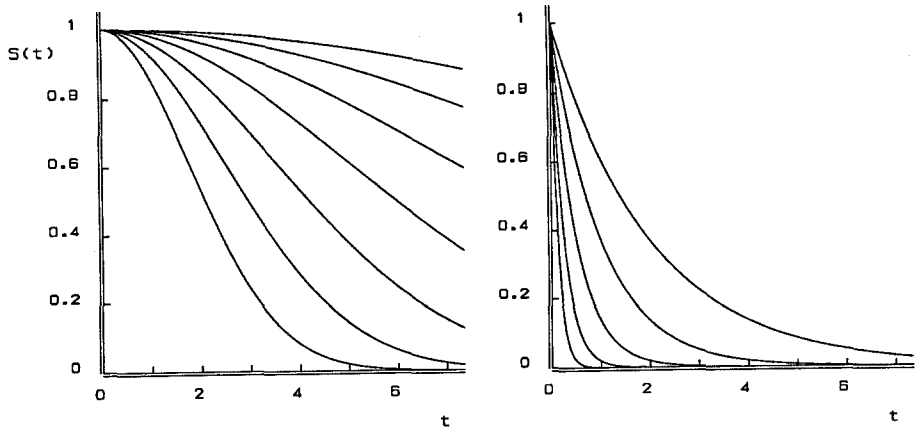

Figure 3. Possible shapes of survivor curves with changing elimination rate. Parameter values ($c_0 = 0$, $k_\dagger = 0.5$, $\lambda = 0$, $k_e = 0.01$) (left) or 100 (right).

concentration in the environment will result ultimately in a very high concentration in the tissue. A no-effect level in the tissue, i.e. the upper boundary of the tolerance range, still exists, of course, and is exceeded at $t_c = C_0(k_u c)^{-1}$. The hazard rate amounts to $h_c = \kappa_\dagger c(t - t_c)_+$. The relationship between the killing acceleration $\kappa_\dagger$ and the killing rate $k_\dagger$ is $\kappa_\dagger = \lim_{k_e \to 0} k_\dagger k_e$. The survival probability is

$$S(t; c) = \exp\{-\tfrac{1}{2}\kappa_\dagger c((t - t_c)_+)^2\}. \tag{7}$$

This represents a Weibull distribution with shape parameter 2.

For very large values of k_e, on the other hand, the survivor function becomes an exponential function

$$S(t; c) = \exp(-k_\dagger t(c - c_0)_+),$$

which can also be seen as a Weibull function, with shape parameter 1. In Fig. 3 some possible shapes of $S(t; c)$ for varying elimination rates are shown.

As an alternative we will discuss the model proposed by Kooijman (1981), which is an extension of the standard log-logistic model. The extension involves a no-effect level and a LC50–time relation consistent with the first-order kinetics (1). It can be summarized as follows. The probability to survive an exposure time t at concentration c is given by

$$S(t; c) = \left(1 + \left(\frac{(c(1 - \exp(-k_e t)) - c_0)_+}{c_{L50} - c_0}\right)^\beta\right)^{-1}, \tag{8}$$

where c_{L50} is the ultimate LC50, i.e. the LC50-value after a very long exposure time. Kooijman (1993, p. 279) compares models (6) and (8). An important difference is the behaviour after long exposure times, if $c > c_0$: in model (6) we have $\lim_{t \to \infty} S(t; c) = 0$ and in (8) $\lim_{t \to \infty} S(t; c) > 0$.

3. Estimation of parameters

In experiments which are set up to evaluate the lethal effect of toxic compounds, the resulting data sets consist of counts x_{ij} of surviving organisms on fixed times t_i, $i = 0, \ldots, r$, exposed to a chemical compound at concentration c_j, $j = 1, \ldots, k$. An example of such a data set is given in Table 2. To fit

Table 2. Number of surviving guppies *Poecilia reticulata* in natural sea water after exposure to the pesticide dieldrin. Data kindly provided by Ms T. Adema (IMW-TNO Laboratories, Delft).

	Concentration of dieldrin ($\mu g l^{-1}$)							
Time (d)	0	3.2	5.6	10	18	32	56	100
0	20	20	20	20	20	20	20	20
1	20	20	20	20	18	18	17	5
2	20	20	19	17	15	9	6	0
3	20	20	19	15	9	2	1	0
4	20	20	19	14	4	1	0	0
5	20	20	18	12	4	0	0	0
6	20	19	18	9	3	0	0	0
7	20	18	18	8	2	0	0	0

the model (6) to such a data set we use the maximum likelihood (ML) method. For this experimental set-up the likelihood function is not the product of the density function of $\underline{t}_\dagger$ in the data points (t_i, c_j), since we do not know exact times of death of the organisms. We only know the time intervals in which death has occurred. Before deriving the likelihood function we have to introduce some new symbols.

The probability p_{ij} of an organism, exposed to concentration c_j, to die between t_{i-1} and t_i is given by $p_{ij} = S(t_{i-1}, c_j) - S(t_i, c_j)$. The number of organisms n_{ij} which died in that period is given by $n_{ij} = x_{i-1,j} - x_{ij}$. The number of organisms surviving at t_r will be denoted by $n_{r+1,j}$. The probability of surviving at t_r is denoted by $p_{r+1,j}$ and equals $S(t_r, c_j)$.

The probability of obtaining the counts x_{ij} can now be written as a product of multinomial probabilities

$$\Pr(\underline{x}_{ij} = x_{ij}) = \Pr(\underline{n}_{ij} = n_{ij}) = \prod_{j=1}^{k} x_{0j}! \prod_{i=1}^{r+1} \frac{p_{ij}^{n_{ij}}}{n_{ij}!}. \tag{9}$$

The log-likelihood function is then given by

$$\ell(\theta; (x_{ij})) = \sum_{i=1}^{r+1} \sum_{j=1}^{k} n_{ij} \ln(p_{ij}) \qquad \text{with } \theta = (c_0, k_\dagger, k_e, \lambda)', \tag{10}$$

where the constant term has been ignored. Maximum likelihood estimates can be found by solving the vector equations

$$G(\theta) = \frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{r+1} \sum_{j=1}^{k} \frac{n_{ij}}{p_{ij}} \frac{\partial p_{ij}}{\partial \theta} = 0. \tag{11}$$

The information matrix $I(\theta)$, defined as minus the expectation of the matrix of second derivatives, can be shown to be

$$I(\theta) = -E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right) = \sum_{j=1}^{k} x_{0j} \sum_{i=1}^{r+1} \frac{1}{p_{ij}} \left(\frac{\partial p_{ij}}{\partial \theta}\right) \left(\frac{\partial p_{ij}}{\partial \theta}\right)'. \tag{12}$$

This matrix can be used to estimate the asymptotic variance–covariance matrix. It can also be used in the so-called method of scoring, an iteration scheme to find the ML estimates:

$$\theta_{i+1} = \theta_i + I^{-1}(\theta_i) G(\theta_i). \tag{13}$$

Table 3. Results of ML estimation on simulated data. In the first horizontal block theoretical values are shown: asymptotic expectations, standard deviations according to (12), multiplied by $\sqrt{x_{0j}}$, and correlation coefficients. In the following blocks simulation results are shown: means, standard deviations ($\times\sqrt{x_{0j}}$) and correlation coefficients of parameter estimates of 1000 simulated data sets.

		Mean	*SD*	$SD\sqrt{x_{0j}}$	*Correlation coefficients*		
Theoretical values	c_0	2	–	1.425	1		
	$k_\dagger$	0.1	–	0.0875	−0.393	1	
	k_e	0.5	–	0.5827	0.755	−0.790	1
$x_{0j} = 5$	c_0	2.095	0.6442	1.441	1		
	$k_\dagger$	0.1453	0.1052	0.2353	−0.377	1	
	k_e	0.5262	0.3721	0.8321	0.431	−0.454	1
$x_{0j} = 10$	c_0	2.033	0.4013	1.269	1		
	$k_\dagger$	0.1179	0.0431	0.1364	−0.335	1	
	k_e	0.5053	0.1916	0.6057	0.635	−0.703	1
$x_{0j} = 20$	c_0	2.018	0.2970	1.328	1		
	$k_\dagger$	0.1100	0.0251	0.1121	−0.308	1	
	k_e	0.4919	0.1286	0.5749	0.655	−0.748	1
$x_{0j} = 50$	c_0	1.995	0.1863	1.318	1		
	$k_\dagger$	0.1042	0.0137	0.0968	−0.354	1	
	k_e	0.4927	0.0815	0.5761	0.695	−0.799	1

As a rough measure of goodness-of-fit we use the deviance of the model (McCullagh and Nelder, 1989). The deviance is defined as twice the difference between the maximum achievable log likelihood and that attained under the fitted model. The maximum achievable log likelihood $\ell_{\sup}((x_{ij}))$ is obtained by estimating each p_{ij} without any constraint, i.e. $\hat{p}_{ij} = n_{ij}/x_{0j}$. Substituting this in (10) we get

$$\ell_{\sup}((x_{ij})) = \sum_{i,j} n_{ij} \ln\left(\frac{n_{ij}}{x_{0j}}\right),$$

where the summand should read 0 if $n_{ij} = 0$. The difference in deviances between nested models can be used as a test statistic. It is equivalent to the usual likelihood-ratio test. The approximate distribution of the test statistic is $\chi^2_{[d]}$, where d is the difference in the number of parameters between the two nested models. The deviance should not be used to test the absolute goodness-of-fit. Usually the asymptotic theory does not apply, because some of the expected numbers $\boldsymbol{E}(\underline{n}_{ij})$ are too small.

Equations (11), (12) and (13) were implemented in a computer program written in APL. The numerical procedure appeared to be sensitive for starting values of the parameters, which cannot be easily found. If $\lambda = 0$, relation (4) can be used to guess starting values for c_0 and k_e.

4. Monte Carlo Simulation

We studied the performance of ML estimation by Monte Carlo simulation. Data sets resembling the data of Table 2 were generated: 8 concentrations (an exponential series, 1, 1.8, 3.2, 5.8, 10, 18, 32, 58), 8 time points $(0, 1, \ldots, 7)$ and fixed parameters ($c_0 = 2$, $k_\dagger = 0.1$ and $k_e = 0.5$) were chosen. Four different values of x_{0j} (5, 10, 20 and 50) were chosen to study the influence of sample size. For every value of x_{0j}, 1000 data sets were generated, by simulating multinomial distributions according to (9) and (6). In every data set we estimated the parameters by solving (11) numerically, where the

Table 4. Results of ML estimation on empirical data given in Table 2. For each parameter, point estimates, standard deviations, and correlation coefficients are given.

Model	*Parameter*	*Units*	*Estimate*	*SD*	*Correlation coefficients*			
(6)	λ	d^{-1}	0	–	–			
	c_0	$\mu g l^{-1}$	2.77	0.303	–	1		
	$k_\dagger$	$l\,\mu g^{-1} d^{-1}$	0.0309	0.00549	–	−0.233	1	
	k_e	d^{-1}	0.727	0.201	–	0.511	−0.790	1
(6)	λ	d^{-1}	0.00835	0.00490	1			
	c_0	$\mu g l^{-1}$	5.20	0.465	0.309	1		
	$k_\dagger$	$l\,\mu g^{-1} d^{-1}$	0.0376	0.00777	0.046	−0.024	1	
	k_e	d^{-1}	0.791	0.281	−0.049	0.281	−0.811	1
(14)	λ	d^{-1}	0	–	–			
	c_0	$\mu g l^{-1}$	0	–	–	–		
	A	$\mu g l^{-1} d$	62.8	4.05	–	–	1	
	β		2.72	0.263	–	–	−0.052	1

'true' parameters were chosen as starting values. The resulting 1000 vectors of parameter estimates $(\hat{c}_0, \hat{k}_\dagger, \hat{k}_e)'$ were analysed by calculating means, standard deviations and correlation matrices. The latter were compared with the theoretical asymptotic values.

Results are shown in Table 3. The parameter estimates behave quite differently with respect to sample size. Estimation of c_0 is accurate, even for 5 animals per concentration, while reliable estimation of $k_\dagger$ apparently needs large sample sizes. The estimates of the standard deviations, which can be compared with the theoretical value after multiplication with $\sqrt{x_{0j}}$, and the estimates of the correlation coefficients indicate that asymptotic theory should only be applied at large values of x_{0j}, say $x_{0j} \geq 20$.

5. Application to experimental data

We have fitted model (6) – with and without control mortality – to the data set of Table 2. Solutions were checked by Monte Carlo searches. We also fitted model (8) to the same data. This led to some numerical problems. All parameters but β grew very small and c_0 even became zero. In the limit situation for small k_e we can then rewrite model (8) to

$$S(t;c) = \left(1 + \left(\frac{ct}{A}\right)^{1/\beta}\right)^{-1} \qquad \text{where } A = \lim_{k_e \to 0} \frac{c_{L50}}{k_e}. \tag{14}$$

Results of parameter estimation are given in Table 4. Inclusion of control mortality in model (14) did not noticeably affect the results.

Data and estimated survivor curves are plotted in Fig. 4. The improvement in fit from inclusion of control mortality in (6) is apparent for $c = 5.6$ and $10\,\mu g l^{-1}$. For $c = 0$ the fit is worse. The deviances are 40.21 and 36.43, respectively. The difference is 3.78, which means that inclusion of control mortality does not lead to a significant improvement ($\alpha = 0.05$). The deviance of model (14) equals 43.57, which is more than the previous values. However, because the models (6) and (14) are not nested, we cannot draw strong conclusions from the deviances.

The second data set that we analysed is given in Table 5. As mortality occurs at $c = 0$ we are forced to include control mortality in the model.

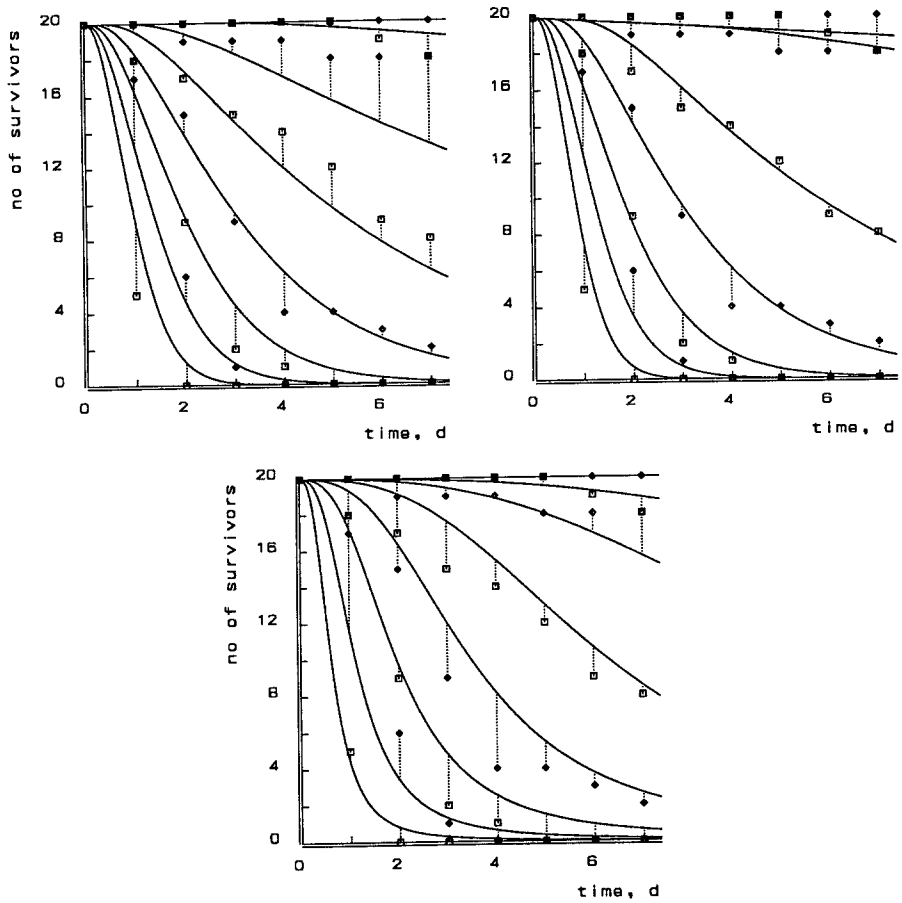

Figure 4. Empirical data of Table 2 with estimated survivor curves. Upper left: model (6) without control mortality. Upper right: model (6) including control mortality. Middle: model (14).

Table 5. Number of surviving daphnids *Daphnia magna* in potassium dichromate. Data were kindly provided by Ms T. Adema (IMW-TNO Laboratories, Delft).

	Concentration $K_2Cr_2O_7$ ($mg\,l^{-1}$)					
Time (d)	0	0.1	0.18	0.32	0.56	1
0	50	50	50	50	50	50
2	50	50	50	50	50	48
5	50	50	50	50	48	36
7	50	50	50	50	48	35
9	49	50	50	50	48	31
12	49	50	50	50	40	15
14	49	50	50	48	32	9
16	49	50	50	47	30	3
19	49	50	50	47	23	0
21	49	50	50	45	16	0

Table 6. Results of ML estimation on empirical data are given in Table 5. For every parameter point estimates, standard deviations and correlation coefficients are given.

Model	*Parameter*	*Units*	*Estimate*	*SD*	*Correlation coefficients*			
(6)	λ	d^{-1}	3.08×10^{-4}	2.89×10^{-4}	1			
	c_0	$mg\,l^{-1}$	0.272	0.0179	0.058	1		
	$k_\dagger$	$l\,mg^{-1}\,d^{-1}$	0.278	0.0414	0.024	0.016	1	
	k_e	d^{-1}	0.214	0.0397	0.029	0.657	−0.501	1
(8)	λ	d^{-1}	3.94×10^{-4}	3.97×10^{-4}	1			
	c_{L50}	$mg\,l^{-1}$	0.161	0.106	0.136	1		
$c_0 = 0$	k_e	d^{-1}	0.0186	0.0143	0.130	0.998	1	
	β		3.848	0.412	0.251	0.526	0.539	1

Fitting model (6) we found several (local) maxima of the likelihood. The deviance at the global maximum equals 35.55. The estimation procedure for model (8) ran into the same numerical problems as encountered with the first data set. Again c_0 became zero. The resulting deviance is 38.38. Results of parameter estimation are given in Table 6 and Fig. 5. The estimates of control mortality are more or less the same in both models, with relatively large standard deviations. The other parameters in model (8) have small standard deviations. In model (8) only β has a small standard deviation. The extreme high correlation between c_{L50} and k_e and the small value of the latter indicate that the resulting model approximates model (14).

6. Discussion

In the present paper we introduce a new model for the analysis of survival data. There are several advantages of our approach. First, our approach provides an alternative measure for the toxicity of a compound with respect to survival, the killing rate $k_\dagger$. The killing rate can be interpreted as the probability of dying, per unit of time and per unit of (environmental) concentration exceeding the no-effect concentration. It does not depend on exposure time, as does LC50. In addition, the model

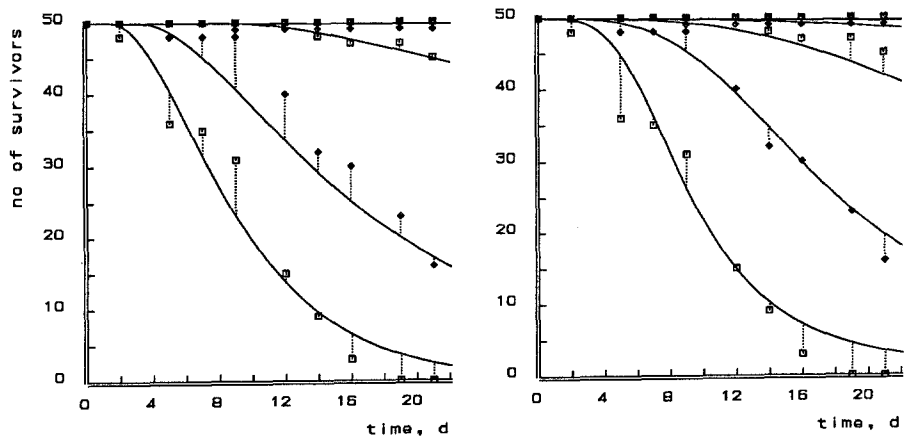

Figure 5. Empirical data of Table 5 with estimated survivor curves. Left: model (6) including control mortality. Right: model (8) including control mortality.

also provides NEC as well as LC50 as functions of time, in a single estimation procedure. The NEC-time relation can be found by calculating c from $C(t) > C_0$ in (3) at a fixed time t, leading to $\text{NEC}(t) = c_0(1 - \exp(-k_e t))^{-1}$. The LC50–time relation can be found by (numerically) solving c from $S(t; c) = 0.5$. Second, we model a survival function based on simple mechanistic assumptions which may, at least in theory, be tested in independent experiments. Most other parametric approaches assume some distribution function without any biological or mechanistic justification.

A key assumption in the model is that the hazard rate is proportional to the concentration of the toxicant. This implies that the lethal effects of a toxicant should disappear as soon as the concentration decreases below the no-effect level, that is, the animals should completely recover instantaneously. A different approach could be to assume that the toxicant causes irreparable damage to the animal, again proportional with the concentration. Diggle and Gratton (1984), Morgan (1992) suggested a comparable idea in their extension to the model of Puri and Senturia (1972). If, in our model, the hazard is taken to be proportional to the total damage, this results in a hazard proportional to the accumulated concentration. In an analogous way Kooijman models aging processes (Kooijman, 1993, pp. 105–12). For small values of k_e, the survivor function approximates a Weibull function with shape parameter 3 instead of 2 as in (7).

The model can, of course, be extended by changing the assumptions. For instance, the assumption about the accumulation process in the animal can be changed in a two-compartment model if a one-compartment model does not make sense. However, this will probably cause estimation problems unless the data set is very detailed. Another extension of our model could be the introduction of stochasticity between animals. As a matter of fact animals are supposed to be identical in the present model: they all have the same kinetic parameters. The only stochastic component is in the process of dying. In the log–logistic model (8) stochasticity is located entirely between animals. There the process of dying is completely deterministic. Therefore both approaches lack reality. To meet this objection we might consider the elimination parameter k_e as a random variable. Stochastic parameters in accumulation models are discussed in Bedaux and Kooijman (1994).

Acknowledgements

This work was supported by Grant PAD 90-18 from the Alternative to Animal Experiments Platform. We thank our colleagues C. Zonneveld and H.A. van den Berg, and three anonymous referees for their comments on the manuscript, and Miss S.L. Yeung for her assistance in the computations.

References

Bedaux, J.J.M. and Kooijman, S.A.L.M. (1994) Stochasticity in deterministic models. In Handbook of Statistics, volume 12: *Environmental Statistics*, G.P. Patil and C.R. Rao, eds., North Holland, pp. 561–81.

Carter, E.M. and Hubert, J.J. (1984) A growth-curve model approach to multivariate quantal bioassay. *Biometrics* **40**, 699–706.

Cox, C. (1987) Threshold dose-response models in toxicology. *Biometrics* **43**, 511–23.

Diggle, P.J. and Gratton, R.J. (1984) Monte Carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society* B **46**, 193–227.

Hoekstra J.A. (1991). Estimation of the LC50, A Review. *Environmetrics* **2**, 139–52.

Jacquez, J.A. (1985). *Compartmental Analysis in Biology and Medicine*. Elsevier Amsterdam.

Janssen, M.P.M., Bruins, A., De Vries, T.H. and Van Straalen, N.M. (1991) Comparison of cadmium kinetics in four soil arthropods species. *Archives of Environmental Contamination and Toxicology* **20**, 305–12.

Kooijman, S.A.L.M. (1981) Parametric analyses of mortality rates in bioassays. *Water Research* **15**, 107–19.
Kooijman, S.A.L.M. (1993) *Dynamic Energy Budgets in Biological Systems. Theory and Applications in Ecotoxicology*. Cambridge University Press.
McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*. Chapman and Hall, London.
Laurence, A.F. and Morgan, B.J.T. (1989) Observations on a stochastic model for quantal assay data. *Biometrics* **45**, 733–44.
Morgan, B.J.T. (1988) Extended models for quantal response data. *Statistica Neerlandica* **42**, 253–72.
Morgan, B.J.T. (1992). *Analysis of Quantal Response Data*. Chapman and Hall, London.
Puri, P.S. and Senturia, J. (1972) On a mathematical theory of quantal response data. *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics*, **4**, 231–47.
Van Ryzin, J. and Rai, K. (1987) A dose–response model incorporating nonlinear kinetics. *Biometrics* **43**, 95–105.

Biographies

Jacques Bedaux is Assistant Professor in Biostatistics at the Department of Theoretical Biology of the Vrije Universiteit Amsterdam. He is involved in several ecotoxicological projects and in the School of Environmental Sciences. He has a long experience as statistical consultant, where most problems relate to environmental research.

Bas Kooijman is Full Professor and head of the Department of Theoretical Biology of the Vrije Universiteit Amsterdam and consultant of the TNO Institute for Environmental Sciences. His scientific contribution to the subject of statistics and modelling in relation with environmental problems include a large number of papers on such topics as the statistical analysis of toxicity and mutagenicity tests, of air pollution and biomonitor data, toxicokinetics and toxic effects, the extrapolation from laboratory to field situations, environmental safety factors. His department collaborates intensively with the National Institute of Public Health and Environmental Protection on risk assessment.