Applied Occupational and Environmental Hygiene

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uaoh20

Lognormal Model for Determining Dose-Response Curves from Epidemiological Data and for Health Risk Assessment

Bernard E. Saltzman

Available online: 30 Nov 2010

To cite this article: Bernard E. Saltzman (2001): Lognormal Model for Determining Dose-Response Curves from Epidemiological Data and for Health Risk Assessment, Applied Occupational and Environmental Hygiene, 16:7, 745-754

To link to this article: http://dx.doi.org/10.1080/10473220121485

Applied Occupational and Environmental Hygiene
Volume 16(7): 745–754, 2001

1047-322X/01 $12.00 + .00

Lognormal Model for Determining Dose-Response Curves from Epidemiological Data and for Health Risk Assessment

Bernard E. Saltzman
Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio

A practical method is proposed for determining human dose-response curves based on reasonable assumptions and simplifications. The epidemiological data needed are the fractions of the population suffering an adverse response from exposures to two or more patterns of fluctuating concentrations of a pollutant, and the statistical parameters of each pattern. The method calculates the two parameters of the threshold type dose-response curve of the pollutant, represented by a cumulative lognormal distribution. This distribution was derived from a reasonable statistical model. The calculation does not require any arbitrary safety factors and yields central values. The dose-response parameters then may be used to calculate the health risk rate of exposure to any other fluctuating concentration pattern. Another method is proposed to select appropriate threshold limit values (TLVs®) using calculations involving these parameters.

Examples are given to illustrate the calculations. Results with hypothetical data gave apparently reasonable results. They showed the importance not only of the geometric mean concentration but also of the geometric standard deviations of both the concentrations and of the dose-response curve, that greatly influence the results. It is believed that results of useful accuracy should be obtained.

Health risk rates are readily understood. They are useful for cost-benefit calculations. Relative rates can be used to compare the hazards of different operations and different plants. The method may make possible the development of standards specifying maximum allowable risk rates.

Keywords Risk Assessment, Risk Analysis, Dose-Response Curve, Threshold Limit Values (TLVs)

Threshold limits imply that the human dose-response curves for many contaminants are of the non-linear threshold type observed in many animal studies. Human dose-response curves obviously would be very useful for establishing such limits, but unfortunately they are scarce. Laboratory clinical studies using toxic exposures are limited by ethical considerations, by small numbers of human subjects, and by their high costs in time and money, and it seems unlikely that the situation will improve soon. Field epidemiological studies can be conducted with substantial numbers of subjects for substantial time periods, but the exposures are usually to fluctuating concentrations of pollutants. Averages of these concentrations have been applied to assumed linear dose-response relationships to determine adverse responses. But they would be inaccurate with non-linear threshold type curves because doses below the threshold would produce very little effect, and those above the threshold would produce disproportionate effects on the curve. No simple way is apparent for converting such data to dose response parameters.

Human dose-response curves also would be very useful for determining health risks. They could eliminate questions about the validity of the common use for this purpose of animal data and the need for the large safety factors that are applied to allow for various uncertainties. Assessment of health risks has become of great current interest. Control policies must be formulated with consideration of costs, health risk rates, and benefits from avoiding these risks. Industrial hygienists must satisfy their managements that the benefits justify the costs. New regulations have been challenged in the courts[1] and in Congress[2] on these issues and on their scientific bases. Problems of assessing environmental health risks have been reviewed by the National Research Council,[3] by the Office of Technology Assessment of the U.S. Congress,[4] in a collection of case studies,[5] and in a comprehensive textbook.[6] Health risk values are readily understood and accurate relative values could be useful for comparing relative hazards of different operations or in different plants.

The practical methods of handling important aspects of these problems that are proposed here are based on experimental and theoretical studies.[7–13] The basic mathematical equations for risk assessment were derived and presented elsewhere.[11] They required reasonable assumptions and simplifications. A lognormal equation with two parameters represented the concentration pattern. A cumulative lognormal equation represented the

dose-response curve of the pollutant. Two generalized parameters were simply derived from these four parameters. The risk calculations were made using the two generalized parameters applied to a graph or a table. The earlier report[11] presented a difficult and complex method for calculating the essential dose-response curves from epidemiological data, that limited its usefulness. A much superior method has been developed here. It and new useful applications are illustrated and explained in Examples 1–4 below. Additional new material is presented in the following section.

SELECTION OF THE PROPER AVERAGING-TIME FOR CONCENTRATION DATA

Fluctuating concentrations are commonly represented as a series of time-averaged numbers. If a concentration peak period (or peak width) is much less than the averaging-time, then the peak height will be greatly attenuated by the averaging process. Those peaks having biologically significant effects must be preserved, whereas those too brief to be biologically significant should be attenuated. This can be accomplished by making a proper choice of an averaging-time that is not too long or too short.

The body also time-averages external concentrations. During a very brief external peak very little pollutant diffuses into the blood, where it is diluted to a relatively large volume to produce a relatively small internal concentration peak; external peaks with the same maximum concentration but with longer periods produce relatively higher concentrations in the blood because more pollutant enters during the longer time. This relationship was studied experimentally and theoretically[12,13] by exposing rabbits to continuously monitored fluctuating concentrations of carbon monoxide and determining the carboxyhemoglobin in numerous small samples of the rabbits' blood taken consecutively in time.

Carbon monoxide diffuses into the blood in the lungs and combines reversibly with hemoglobin with 210–250 times the affinity of oxygen, thus interfering with oxygen transport. Each external carbon monoxide peak produced a delayed and attenuated peak in the internal carboxyhemoglobin concentration.

Simplified pharmacokinetic equations were developed for these processes based on the assumption of rates of diffusion, metabolism, and excretion being proportional to the concentration of the compound (first order relationship). A theoretical S-shaped curve showed that when the ratio of the carbon monoxide peak period to the biological half-life of carboxyhemoglobin was one or less the transmittance of the peak heights by the "biological window" was almost zero; when the ratio was more than 100 the carboxyhemoglobin peak heights were almost 100 percent of the equilibrium value for the peak concentration.

A similar S-shaped curve was developed for the transmittance of the "time-averaged sampling window" versus the ratio of the carbon monoxide peak period to the averaging time. Study of the two curves showed that when the averaging-time of the carbon monoxide fluctuating concentration was no greater than one fourth of the biological half-life of the carboxyhemoglobin, no significant attenuation would occur of peaks that would be transmitted by the biological window. The experimental proof of this was that when the resulting concentration numbers were used in conjunction with the simplified pharmacokinetic equations the calculated concentration pattern of carboxyhemoglobin very closely fitted the experimental values. This demonstrated that no information on biologically significant carbon monoxide peaks was lost, and also that the equations were valid.

For application to health risk calculations it is recommended that the averaging-time be equal to one-half of the biological half-life of the pollutant. This will cause some attenuation of peak heights that will approximate the attenuation produced in the body. Some published data[13] giving approximate values of biological half-lives of pollutants in humans is given in Table I.

DERIVATIONS OF EQUATIONS REPRESENTING THE DOSE-RESPONSE CURVE AND THE POLLUTANT CONCENTRATION DISTRIBUTION

The basic equations used in the proposed method are presented in the next two sections. They were previously presented in full detail elsewhere[11] and are explained and summarized here to provide the material needed to understand and use them.

The model used to represent the dose-response curve may be derived rationally from some basic assumptions. In this derivation "response" means the fraction of a large exposed population showing the adverse effect. Its value ranges from zero to one. It does not mean the magnitude of the effect on one individual. As the levels of the toxic doses are raised, at first a few very sensitive individuals will succumb, next the bulk of them will succumb, and finally the last few highly resistant individuals will succumb. The dose causing each individual to succumb may be equated to the median level, L_{50}, multiplied by a series of N independent random factors, Mk_i, each relative to its median value:

$$L_i = L_{50} M1_i \; M2_i \; M3_i \ldots MN_i \qquad [1]$$

- $M1_i$ may be a function of the individual's weight divided by the same function of the median weight of the individuals in the exposed population
- $M2_i$ a relative function of his or her age
- $M3_i$ of his or her state of health
- $M4_i \ldots MN_i$, of his or her nutrition, activity level, occupational history, and genetic factors; of measurement errors, etc.

Each of these many independent relative factors may differ in a random way for each observed individual.

Taking the logarithms of both sides yields:

$$\log(L_i) = \log(L_{50}) + \log(M1_i) + \log(M2_i) + \log(M3_i) \ldots + \log(MN_i) \qquad [2]$$

TABLE I
Biological half-lives in humans for some common contaminants[13]

Substance	Half-life	Substance	Half-life
Acetone	3 hr	Iron oxide fume (Fe_2O_3)	12 hr
Ammonia	<20 min	Lead[A]	25–40 days
Aniline	2.9 hr	Mercury	5 wks
Benzene	3–5 hr	Methanol	7.0 hr
Benzidine	5.3 hr	Methylene chloride	2.4 hr
Carbon disulfide	0.9 hr	Mineral dust	>6 mo
Carbon monoxide	1–4 hr	Nitrobenzene	86 hr
Carbon tetrachloride	3 hr	Nitrogen dioxide	1 hr
Chlorine	<20 min	Phenol	3.4 hr
Chloroform	15–30 min	p-Nitrophenol	1.0 hr
DDT	1–3 yr	Styrene	0.5–8 hr
Dichlorodifluoromethane	9.4 min	Sulfur dioxide	<20 min
Dimethyl formamide	3.0 hr	Tetrachloroethylene	24–70 hr
Ethyl acetate	2.0 hr	Toluene	12 hr
Ethyl alcohol	1.5–10 hr	1,1,1-Trichloroethane	8.7 hr
Ethyl benzene	5.0 hr	Trichloroethylene	24 hr
Fluorides as F	8 hr	Trichlorofluoroethane	16 min
Hexane isomers	3.0 hr	Vinyl chloride	3 hr
Hydrogen sulfide	<20 min	Xylene	3.8 hr

[A]In soft tissues; 20 yr in bone.

This equates the logarithm of the dose causing each individual to succumb to a term (the logarithm of the median dose) plus a large number of *additive* independent random variables (each being the logarithm of one of the above multiplicative factors). Some of these random terms will be negative if the multiplicative factor is less than one.

This latter equation has the same form as that of a model that can represent the values of an observed variable, on which the usual statistical calculations are made. According to the central limit theorem, the observed variable will have a normal distribution. Therefore the logarithms of the doses will have a normal distribution, and by definition the doses will have a lognormal distribution. Since the dose-response curve represents the sum of individuals succumbing to the stated dose and all lower doses, its distribution must be a cumulative lognormal distribution.

In general, whenever observations are affected by a large number of additive independent random variables a large set will follow a normal distribution. However, when they are affected by a large number of multiplicative independent random factors, the set will follow a lognormal distribution. The lognormal distribution has no negative values. This is a more natural pattern than that of the normal distribution, which theoretically does have negative values. Similar equations can be written for many other types of observations, and lognormal distributions are commonly observed for loudness of sounds, tolerance to drugs, particle sizes, incomes, and so on.[14,15] Observed concentrations of pollutants have been long known to follow lognormal distributions. The multiplicative independent random relative factors in this case could be the number of contaminant sources, the air velocity, the dimensions of the room, the number of open windows, the temperature, the location, the time, and other factors. It appears likely that we live in a lognormal world.

METHOD FOR CALCULATING HEALTH RISK RATE FOR A TIME PERIOD

If the averaging-time is kept constant for the measurements of concentrations, each concentration value can also represent a dose, with the understanding that the exposure time is equal to the averaging-time. Applying each of a series of concentration values to the dose-response curve and summing the fractions of the population responding could give the risk rate of an adverse effect for that period of time under certain common conditions. First, we must assume that the risk rate is low, as it would be for a tolerable or controlled environment. Thus, the total unaffected exposed population remains substantially constant, and only a few isolated high values cause most of the effect.

If the high values are close enough to each other to exert a substantial synergistic effect, then a different method[13] is recommended. It is also assumed that the pool of unaffected very sensitive individuals is not substantially depleted, or that the affected individuals are replaced by others who become ill or injured from other causes so that this pool remains substantially constant and the dose-response curve for the population is not modified during the time period.

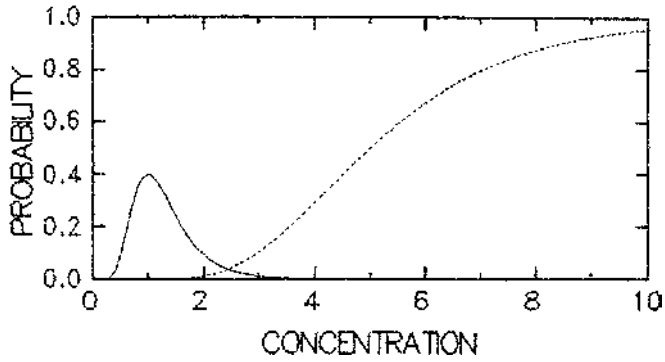

FIGURE 1

Solid curve: Illustrative probability density (lognormal distribution, $\mu_g = 1, \sigma_g = 1.5$) of time-averaged concentrations that may be experienced by an individual. Dashed curve: Illustration of a dose-response curve (cumulative lognormal distribution, $C_{50} = 5, \sigma_R = 1.5$) showing the fraction of population adversely affected. The overlapping area is the region producing health effects.[11]

Instead of the laborious calculations for each of many individual concentrations, a simpler calculation can be made using the distribution curve representing the concentrations, as illustrated by Figure 1. The solid curve on the left represents the probability of each concentration value occurring during one time-averaging period. The dashed curve on the right represents the fraction of the exposed population showing an adverse response to each concentration. The area where the two curves overlap produces the adverse effects. The products of the paired ordinates of the two curves can be integrated to produce the risk, P_{CR}, of an adverse effect over one averaging-time interval. For longer time periods the risk is increased proportionally. If the desired time for expressing the risk rate is n time periods, the rate is:

$$P_n = n P_{CR} \quad [3]$$

where:

n = number of averaging-time intervals
P_{CR} = probability of an adverse effect for 1 interval
P_n = probability of an adverse effect for n intervals

This equation is accurate for values of P_n up to 0.05. It yields a safe upper limit of P_n if the dose-response curve actually is changing during successive intervals because of depletion of sensitive individuals.

The integral of data such as those plotted in Figure 1 is a function of the four parameters of the two curves. It was calculated using the Mathcad computer program in terms of two generalized parameters, V and E, as reported in detail previously.[11] V and E are defined from the four parameters as shown in Eqs. (4) and (5). Eq. (5) also was rearranged for other uses as shown in Eqs. (5a) and (5b):

$$V = \log(\sigma_g)/\log(\sigma_R) \quad [4]$$

$$E = \frac{\log(C_{50}) - \log(\mu_g)}{\log(\sigma_R)} \quad [5]$$

or alternatively: $$\mu_g = C_{50}/(\sigma_R)^E \quad [5a]$$

or alternatively $$\log(C_{50}) = E\log(\sigma_R) + \log(\mu_g) \quad [5b]$$

where:

V = variability ratio
E = exponent in the control factor
C_{50} = concentration producing a response in 50% of the exposed population, a measure of the location of the dose-response curve
σ_R = geometric standard deviation of the dose-response curve, an inverse measure of its slope
$\log(\mu_g)$ = the mean of the logarithms of the exposure concentrations. The antilog, μ_g, is the geometric mean.
$\log(\sigma_g)$ = the standard deviation of the logarithms of the exposure concentrations. The antilog, σ_g, is the geometric standard deviation.

The final results for P_{CR} are given in Table II and are plotted in Figure 2. The variability ratio may range from zero to very rarely as high as 5. Referring to Eq. (5a), the denominator of the right side is the control factor, defined as the ratio of C_{50} to the allowable μ_g for a given allowable probability of an adverse effect. E is the exponent in the control factor. The antilog of a number n is simply 10^n.

The uses of Figure 2 and Table II will be illustrated below in Examples 1–4. Concentrations also represent doses, because of the understanding that their averaging times are also the exposure times of the corresponding doses. Also, the geometric standard deviation parameters are functions of the averaging times. Therefore, the averaging times must be kept constant and must always be specified together with the concentrations.

CALCULATION OF THE PARAMETERS OF THE DOSE-RESPONSE CURVE FROM EPIDEMIOLOGICAL DATA

The two parameters of a dose-response curve for a pollutant can be calculated from epidemiological data of adverse effects from exposures to two or more different fluctuating concentration patterns. The basic calculation method can be understood using Figure 2. The P_{CR} value is calculated from the observed adversely affected fraction, P_n, of the population exposed to the first concentration pattern, using Eq. (3). A horizontal line in Figure 2 at this value will intersect all possible pairs of values

TABLE II
Values of E (Exponent in Control Factor)[11]

Variability ratio	Log(P_{CR})												
	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1.301	−1
0	7.035	6.707	6.362	5.999	5.613	5.201	4.755	4.265	3.719	3.090	2.326	1.645	1.282
0.5	7.863	7.493	7.114	6.707	6.276	5.814	5.315	4.768	4.158	3.455	2.601	1.839	1.433
0.75	8.790	8.381	7.951	7.498	7.017	6.502	5.942	5.331	4.649	3.863	2.908	2.056	1.602
1	9.948	9.480	8.990	8.476	7.934	7.355	6.722	6.031	5.259	4.370	3.290	2.326	1.813
1.5	12.676	12.092	11.482	10.823	10.103	9.345	8.569	7.689	6.705	5.571	4.194	2.965	2.310
2	15.743	14.966	14.204	13.441	12.593	11.590	10.629	9.537	8.316	6.910	5.202	3.678	2.866
2.5	19.021	18.054	17.048	16.131	15.199	14.036	12.799	11.484	10.014	8.321	6.254	4.429	3.451
3	22.374	21.276	20.007	18.866	17.835	16.580	15.032	13.487	11.761	9.772	7.357	5.202	4.053
3.5	25.748	24.579	23.029	21.654	20.477	19.173	17.302	15.524	13.537	11.249	8.468	5.987	4.665
4	29.127	27.929	26.064	24.512	23.118	21.788	19.597	17.585	15.334	12.741	9.592	6.782	5.284
4.5	32.505	31.297	29.097	27.442	25.755	24.411	21.909	19.660	17.144	14.245	10.724	7.582	5.908
5	35.882	34.671	32.119	30.419	28.388	27.035	24.238	21.747	18.963	15.757	11.862	8.387	6.535

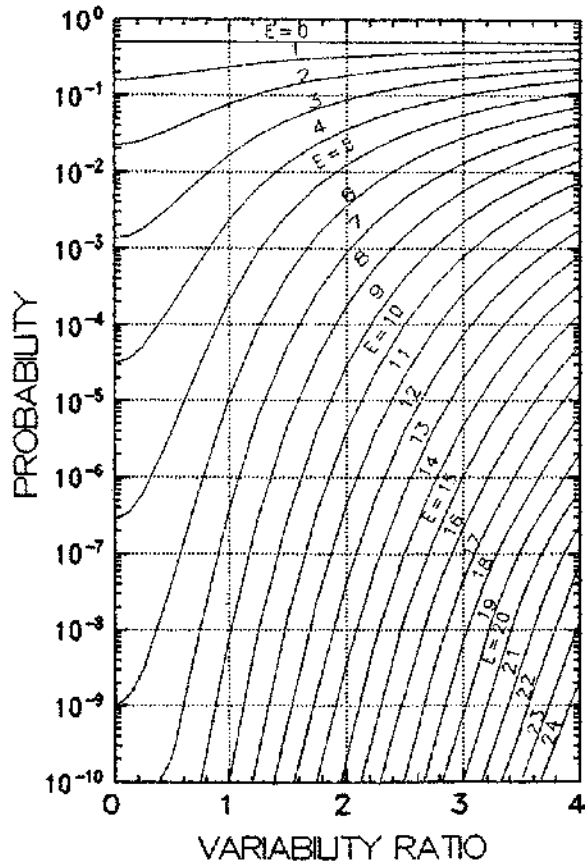

FIGURE 2
Relationships between the probability, P_{CR}, of an effect for a single concentration-averaging time period, the variability ratio, V [defined in Eq. (4)], and the control factor exponent, E [defined in Eq. (5)]. The probability of an effect for n concentration-averaging time periods is nP_{CR}, as shown in Eq. (3).[11]

of V and E that will produce this probability. A series of V values are selected and the corresponding E values are determined along the horizontal line.

The first selected value of V is entered into Eq. (4) together with the observed parameter, σ_g, of the concentration pattern, and the corresponding value of $\log(\sigma_R)$ is calculated. The related value of E is entered into Eq. (5b), together with the above value of $\log(\sigma_R)$ and the observed parameter μ_g, of the concentration pattern. The corresponding parameter, $\log(C_{50})$, is then calculated from this equation. The procedure is repeated for each of the remaining pairs of V and E values. This produces a series of pairs of values of $\log(\sigma_R)$ and $\log(C_{50})$ that can be plotted as shown in Figure 3. This is the locus of all possible pairs of parameters for the dose-response curve that will produce the calculated value of P_{CR} with the observed parameters of this concentration pattern.

The same procedure is followed by similar calculations for a second epidemiological study. The intersection of the two plotted curves gives the values of $\log(C_{50})$ and $\log(\sigma_R)$ that satisfy both studies. If the two curves do not intersect then the two populations must have different dose-response curves. Having greater differences in the values of σ_g will increase the angle between the two curves and improve the precision of the intersection. If more than two studies are available, and their plots do not intersect at the same point, a central point may be selected that minimizes the sum of the squares of the lengths of the perpendiculars from the point to each curve.

The actual calculation procedure requires using interpolated values from Table II for greater accuracy. This method is illustrated in Example 1.

Example 1

In Study A of a pollutant, workers were exposed to a fluctuating concentration having a geometric mean, (μ_g), of 4.3 ppm

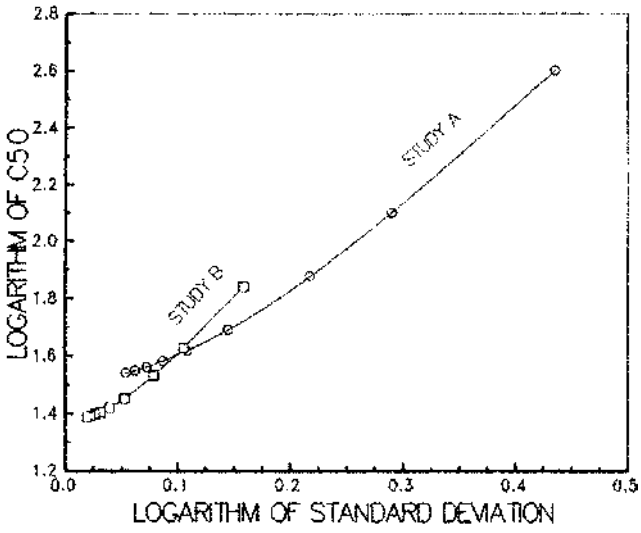

FIGURE 3

Locus of all possible pairs of the two parameters of the dose-response curve that fit the data of Study A, and locus for the data of Study B. The coordinates of the intersection of the two curves give the values of the two parameters fitting both studies listed in Example 1.

(hourly averaged) and a geometric standard deviation, (σ_g), of 1.65. At the end of a year (2000 hr) the total incidence of adverse effects was 5 percent. In Study B, the geometric mean concentration, (μ_g), was 10.0 ppm (hourly averaged), the geometric standard deviation of the concentration pattern, (σ_g), was 1.20, and the yearly total incidence of adverse effects was 0.2 percent. Calculated the parameters of the dose-response curve.

Answer

For Study A the needed initial parameters are: $\log(\mu_g) = 0.6335$, $\log(\sigma_g) = 0.2175$, and $P_{CR} = 0.05/2000 = 2.5 \times 10^{-5}$. The selected values of V are listed in the first column of Table III. The second column lists the paired values of E interpolated from Table II for $P_{CR} = 2.5 \times 10^{-5}$. The interpolation for the first E value corresponding to $V = 0.5$ was made as follows:

$$\log(P_{CR}) = 0.3979 - 5$$

$$E = 4.768 - 0.3979(4.768 - 4.158) = 4.525$$

The third column lists values of $\log(\sigma_R)$ calculated from Eq. (4) by dividing $\log(\sigma_g)$, which is 0.2175, by each selected value of V. Thus, the first value is $0.2175/0.5 = 0.4350$. The fourth column lists the values of $\log(C_{50})$ calculated from Eq. (5b) by multiplying the values of column 2 and 3 and adding $\log(\mu_g)$, which is 0.6335. For example, the first

TABLE III

Calculation of the Parameters of a Dose-Response Curve[A]

	Study A			Study B		
V	E	$\log(\sigma_R)$	$\log(C_{50})$	E	$\log(\sigma_R)$	$\log(C_{50})$
0.5	4.525	0.4350	2.602	5.315	0.1584	1.842
0.75	5.060	0.2900	2.101	5.942	0.1056	1.627
1.0	5.724	0.2178	1.878	6.722	0.07928	1.532
1.5	7.297	0.1450	1.692	8.569	0.05279	1.452
2.0	9.051	0.1087	1.618	10.629	0.03959	1.421
2.5	10.899	0.08699	1.582	12.799	0.03167	1.405
3.0	12.800	0.07250	1.561	15.032	0.02639	1.397
3.5	14.733	0.06214	1.549	17.302	0.02262	1.391
4.0	16.689	0.05437	1.541	19.597	0.01979	1.388

[A]For Example 1.

value is:

$$\log(C_{50}) = 4.525 \times 0.4350 + 0.6335 = 2.602$$

The values of column (4) are then plotted as ordinates against those of column 3 as abscissas, giving the longer curve in Figure 3.

The data for Study B are treated in the same manner, using the same values of V and tabulating the corresponding results in columns 5, 6, and 7. Plotting the values in column 7 against those in column (6) gives the shorter curve plotted in Figure 3. The intersection of the two curves occurs at $\log(C_{50}) = 1.602$, and $\log(\sigma_R) = 0.0986$. The corresponding values are $C_{50} = 40.0$ ppm (hourly averaged) and $(\sigma_R) = 1.255$. These values will satisfy the relationships in both studies.

Note that despite the higher geometric mean concentration in Study B, 10.0 ppm compared with 4.3 in Study A the incidence of adverse effects after a year was only 0.2 percent, compared with 5 percent for Study A. This shows the powerful effect of the lower geometric standard deviation of 1.20 for Study B as compared to 1.65 for Study A. Most of the adverse effects are produced by the uncommon high values of concentration, which occurred much more often in Study A.

CALCULATIONS OF THRESHOLD LIMIT VALUES, HEALTH RISK RATES, AND ALLOWABLE GEOMETRIC MEAN CONCENTRATIONS

A new procedure is proposed here to select threshold limit values (TLVs®) using calculations involving the parameters of a human dose-response curve. The procedure is illustrated in the following example.

Example 2

For the pollutant and the population group in Example 1, calculate the TLV that will guarantee that the probability that

allowable peak concentrations will produce an adverse effect will not exceed: a) 10^{-4}, b) 10^{-5}, and c) 10^{-6}.

Answer

The normal statistical variable Z corresponding to the probabilities of 10^{-4}, 10^{-5}, and 10^{-6} are -3.72, -4.27, and -4.77, respectively. The equation defining the lognormal dose-response curve is:

$$Z = \frac{\log(C) - \log(C_{50})}{\log(\sigma_R)} \qquad [6]$$

Solving for C yields:

$$C = C_{50}(\sigma_R)^Z \qquad [6a]$$

Entering the values for part a: $C = 40.0\ (1.255)^{-3.72}$
$= 17.2$ ppm (hourly averaged)
for part b: $C = 40.0\ (1.255)^{-4.27}$
$= 15.2$ ppm (hourly averaged)
for part c: $C = 40.0\ (1.255)^{-4.77}$
$= 13.5$ ppm (hourly averaged)

TLVs are convenient to use, especially when extensive monitoring has not been conducted. However, they do not consider the concentration distribution pattern. It can be shown theoretically for any lognormal (or normal) distribution that no matter how high the TLV is, it will be exceeded at some times. This could cause substantial adverse effects with some concentration patterns, as will be shown in Example 3.

If sufficient monitoring data are available to determine the parameters of the concentration distribution, the health risk rate can be calculated from the human dose-response parameters. This rate describes the hazard more precisely, enables comparisons of relative hazards, and enables calculation of benefits from reductions in risk.

Example 3

A different exposure to the same pollutant by the same population group as in Example 1, had the following parameters for the concentration distribution pattern: geometric mean concentration, $(\mu_g) = 6.0$ ppm (hourly averaged), geometric standard deviation, $(\sigma_g) = 1.5$.

a) What is the health risk rate per year (2000 hr)?
b) What fraction of the time will a TLV of 13.5 ppm (hourly averaged) be exceeded?

Answer, part a)

First we calculate the general parameters V and E: Eq. (4) yields: $V = \log(1.5)/\log(1.255) = 0.1761/0.0986 = 1.786$
Eq. (5) yields: $E = [\log(40.0) - \log(6.0)]/\log(1.255) = 8.356$

Using this pair of V and E values, Figure 2 is interpolated to give $P_{CR} = 2.7 \times 10^{-5}$. Applying Equation 3, $P_n = 2000 \times 2.7 \times 10^{-5} = 0.054$.

A more precise value may be obtained by interpolating for P_{CR} in Table II in two steps, first for the V value and then for the E value. The V value and the estimated P_{CR} value, above, locate in the table the four surrounding E values that are involved. The E value lies between the values 9.537 and 8.316 listed for $P_{CR} = 10^{-5}$ and $P_{CR} = 10^{-4}$, respectively, and $V = 2$; and the corresponding values 7.689 and 6.705 listed for $V = 1.5$. The interpolated E values for $V = 1.786$ are:

$$\text{for } P_{CR} = 10^{-5}: \quad E = 7.689 + (9.537 - 7.689) \times 0.286/0.5 = 8.746$$
$$\text{for } P_{CR} = 10^{-4}: \quad E = 6.705 + (8.316 - 6.705) \times 0.286/0.5 = 7.626$$

These two values of E are interpolated to yield:

$$\log(P_{CR}) = [(8.746 - 8.356)/(8.746 - 7.626)] - 5 = 0.3482 - 5.$$

Taking the antilog, $P_{CR} = 2.229 \times 10^{-5}$. From Equation 3, $P_n = 2000 \times 2.229 \times 10^{-5} = 0.045$.

Answer, part b)

Applying Eq. (6) to the parameters of the lognormal concentration distribution yields:

$$Z = \frac{\log(C) - \log(\mu_g)}{\log(\sigma_g)} = \frac{\log(13.5) - \log(6.0)}{\log(1.5)} = 2.01$$

$$P_Z = 0.9772, \quad \text{and} \quad 1 - P_Z = 0.0228$$

Thus, in a year (2000 hourly averaged concentrations) 13.5 ppm will be exceeded $2000 \times 0.0228 = 46$ times. Even though this TLV may not be exceeded during one week of monitoring, and the factor of 10^{-6} for it (see Example 2) appears to be adequately safe, as many as 5 percent of the exposed population (see part a) may suffer an adverse effect within one year. If the concentrations were all at the TLV for the entire year, then P_n would equal $2000 \times 10^{-6} = 0.002$, which is 4 percent of 0.045. Thus the 46 uncommon higher concentrations produced more than 96 percent of the adverse effects in this case. The geometric standard deviation of the concentration distribution is an important variable that should not be ignored.

Another type of calculation may be used to determine the maximum allowable geometric mean concentration of a pollutant to ensure compliance with a specified maximum allowable

risk rate, in the case where the geometric standard deviation is known. This is illustrated in Example 4.

Example 4

Assume that the maximum allowable risk rate is 0.1 percent of the exposed population suffering an adverse effect per year (2000 hr), and that the population group, pollutant, and geometric standard deviation of its concentrations are the same as in Example 3. What is the maximum allowable geometric mean concentration for this pollutant?

Answer

From Equation 3: $P_{CR} = 10^{-3}/2000 = 5 \times 10^{-7}$.

V remains unchanged at 1.786. Interpolating Figure 2 using these two values gives $E = 10.0$. From Eq. (5a): $\mu_g = 40.0/(1.255)^{10.0} = 4.13$ ppm (hourly averaged).

A more precise value may be obtained by interpolating E values in Table II between the values of $V = 1.5$ and $V = 2.0$ as follows: $\log(5) = 0.6990$

For $V = 1.5$ $\quad E = 9.345 - 0.6990\,(9.345 - 8.569)$
$= 8.803$

For $V = 2.0$ $\quad E = 11.590 - 0.6990\,(11.590 - 10.629)$
$= 10.918$

For $V = 1.786$ $\quad E = 8.803 + (10.918 - 8.803) \times 0.286/0.5$
$= 10.013$

$\mu_g = 40.0/(1.255)^{10.013} = 4.11$ ppm (hourly averaged)

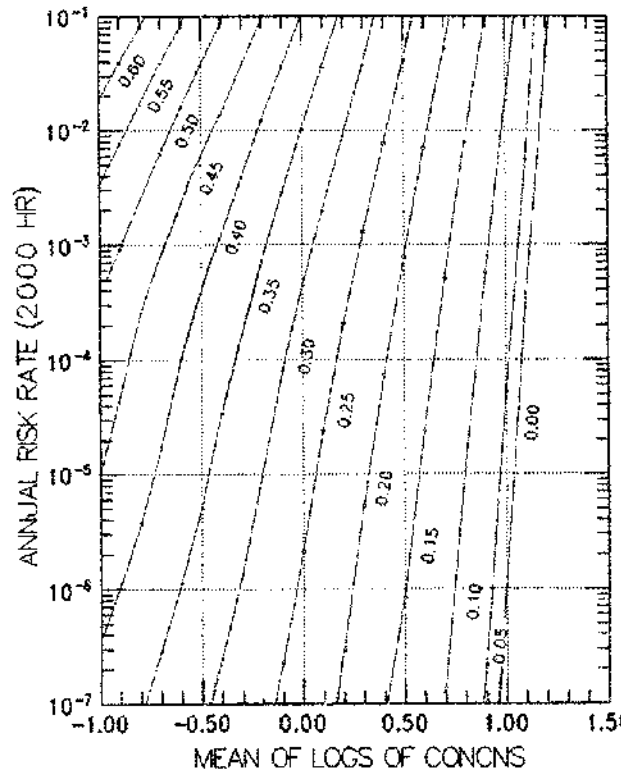

FIGURE 4

Risk assessment chart specifically designed for the dose-response parameters determined in Example 1. The numbers on the curves are values of $\log(\sigma_g)$, which is the standard deviation of the logarithms of the exposure concentrations. The annual risk rate can be determined for any pair of parameters of the measured exposure concentrations, for comparison with a risk standard for this pollutant.

RISK ASSESSMENT CHARTS FOR SPECIFIC POLLUTANTS

Examples 3 and 4 have shown that once the dose-response parameters are known the probability of an adverse response can be calculated for any lognormal concentration pattern. Such relationships can be plotted as shown in Figure 4, which is specifically computed for the pollutant discussed in Example 1, having the dose-response parameters of $\log(C_{50}) = 1.602$, and $\log(\sigma_R) = 0.0986$.

The ordinate is the probability of an adverse effect after exposure for 2000 hr, (P_{2000}). The abcissa is $\log(\mu_g)$, which is the mean of the logarithms of the exposure concentrations. The antilog, μ_g, is the geometric mean exposure concentration. The numbers on the curves are values of $\log(\sigma_g)$, which is the standard deviation of the logarithms of the exposure concentrations. The antilog, σ_g, is the geometric standard deviation of the exposure concentrations. These calculations were carried out using the Mathcad computer program. Such specific plots can be convenient for comparing risks of a measured concentration exposure to a risk standard for the corresponding pollutant.

DISCUSSION

Determination of very low risk rates requires extrapolation of models to far below any possibility of experimental or observational validation. The fit of assumed normal or lognormal distributions to the data depends upon having a very large number of independent random variables. Referring to Figure 1, the accuracy of the extreme right end of the solid curve representing the concentration distribution, and of the extreme left end of the dashed curve representing the dose-response curve is critically important.

In a study[16] that sampled trace metals in the air at 59 sites for a full year each, histograms of the concentrations frequently showed a small second peak at high concentrations. Aerosols are known to exhibit bimodal size distributions. The small second population may originate from different sources or processes. If a histogram of the concentration data shows this effect, a procedure could be developed to subtract out the main peak and determine the parameters of the small high concentration peak. A separate risk calculation could be carried out for the two populations and the two results combined after being weighted

in proportion to the peak areas. A similar problem may arise if the exposed population includes a small group of especially vulnerable people, with few individuals exhibiting intermediate vulnerabilities. This could distort the low concentration end of the dose-response curve.

To be accurate, the concentrations of pollutant inside the body rather than those outside should be used for calculating health risks. If the averaging-time is kept at one-half the biological half-life of the pollutant, the peaks of various periods will be attenuated to approximately the same extent as by the body. Greater accuracy can be achieved by using shorter averaging-times, not exceeding one-fourth of the biological half-life, and applying simple arithmetic step equations[13] to the averaged concentrations. This will convert them to "biologically effective concentrations," values that are at in vivo equilibrium with the internal concentrations. Also, an alternative method of evaluation was proposed, a calculated "biological damage parameter" that takes into account the kinetic process of biological repair.

The extreme effect on the calculated health risk of the variability of susceptibility of the exposed population suggests that toxicological studies with purebred animals, although they may give more reproducible results, may not be properly representative.

The proposed new method for determining human dose-response parameters appears to be based on sound mathematical and scientific principles. But there may be a problem with getting accurate epidemiological data. People in the studied group may have different histories of exposures to many different pollutants and environments as they move around. Also accurate measurements of effects may be difficult. When human dose-effects relationships become available they should be very useful.

The proposed method described in Example 2 for selection of TLVs is based on reasonable mathematical principles. But because TLVs neglect an important variable, the geometric standard deviation of the concentrations, they may not be adequately protective for very variable patterns. Example 3 shows that the TLV selected for a probability of 10^{-6} of adverse effects from allowable peak exposures was inadequate, even though it might not be exceeded during one week of monitoring.

Presently existing occupational health risk rates should be considered if maximum allowable occupational risk rates are to be specified. A companion article[17] presents a perspective on the importance of various occupational hazards, and lists the surprisingly high annual risks of occupational death and nonfatal injury and illness presently being tolerated in some industries. The annual occupational fatality rate for all industries was 4.8 per 100,000 workers, and the 10 worst industries rates ranged from 23.0 to 178.4 per 100,000. The general public tolerates an annual fatality rate from motor vehicle accidents of 15.9 per 100,000. The annual rates for occupational nonfatal injuries and illness were much higher. For the entire private sector this annual rate was 7.4 per *hundred* workers, and the highest 10 industries ranged from 21.0 to 30.3 per *hundred* workers! Thus, the maximum allowable annual risk rate of 0.1 percent cited in Example 4 might be a reasonable beginning rate to specify for occupational nonfatal injuries and illness.

CONCLUSIONS

A practical new method has been presented to determine human dose-effects relationships. Only limited epidemiological data is required for the calculations. The method is based upon reasonable assumptions and approximations. A lognormal distribution is used to fit the fluctuating pollutant concentrations, and a cumulative lognormal distribution to fit the dose-response curve. The equations also may be applied for risk assessment and for selection of TLVs.

The method clarifies several important aspects of the risk assessment. Even though the dose-response curve may appear to have a threshold, the health risk rate can never be reduced to zero, although it can be made very low. This may be emotionally disturbing to some, but absolute safety is not realistic. The calculation takes into account the commonly neglected geometric standard deviations of the fluctuating pollutant concentrations and of its dose-response curve. A high variability of either or both is shown to cause an extreme increase in the risk rate. Some TLVs may be inadequately protective for such patterns.

The required calculations are not difficult or laborious. They do not utilize any arbitrary safety factors and yield a central estimate of the health risk rate. Ranges can be obtained by performing the calculations on appropriate ranges of the four parameters inputted. The hypothetical illustrative calculations appeared to give reasonable answers.

The proposed methods are based on a challenging new and different theoretical framework, worthy of testing to see if it provides results of useful accuracy. They may make possible the development of control standards specifying allowable health risk rates. Such standards should be readily understandable by the general public and would clarify the benefits of compliance.

ACKNOWLEDGMENTS

Partial support was received from NIOSH Education and Research Center Grant No. T42 CCT 510420.

REFERENCES

1. Paustenbach, D.J.: Updating OSHA's Permissible Exposure Limits: Putting Politics Aside. Am Indus Hyg Assoc J 58:845–849 (1997).
2. Kenkerimath, N.: Restoring Reason to Regulation: Applying Scientific Assessment and Cost-Benefit Analysis to Environmental Law. Am Indus Hyg Assoc J 57:791–793 (1996).
3. Committee on Risk Assessment Methodology, National Research Council: Issues in Risk Assessment. National Academy Press, Washington, D.C. (1993).

4. Office of Technology Assessment, U.S. Congress: Researching Health Risks, OTA-BBS-570, U.S. Government Printing Office, Washington, D.C. (1993).
5. Paustenbach, D.J., ed.: The Risk Assessment and Human Health Hazards: A Textbook of Case Studies. John Wiley & Sons, New York (1989).
6. Nelson, D.I.: Risk Assessment in the Workplace. In: The Occupational Environment—Its Evaluation and Control, S.R. DiNardi, ed., Chapter 17. American Industrial Hygiene Association, Fairfax, VA (1997).
7. Saltzman, B.E.: Significance of Sampling Time in Air Monitoring. J Air Poll Cont Assoc 20:660–665 (1970).
8. Saltzman, B.E.: Fourier Analysis of Air Monitoring Data. In Proceedings of the Symposium on Statistical Aspects of Air Quality Data, 1974, pp. 11–13. Publication EPA-654/4-74-038, U.S. Environmental Protection Agency, Research Triangle Park, NC (1974).
9. Saltzman, B.E.: Lognormal Model for Health Risk Assessment of Fluctuating Concentrations. Am Ind Hyg Assoc J 48:140–149 (1987).
10. Saltzman, B.E.: Linear Pharmacokinetic Models for Evaluating Unusual Work Schedules, Exposure Limits, and Body Burdens of Pollutants. Am Indus Hyg Assoc J 49:213–225 (1988).
11. Saltzman, B.E.: Health Risk Assessment of Fluctuating Concentrations Using Lognormal Models. J Air Waste Man Assoc 47:1152–1160 (1997).
12. Saltzman, B.E.; Fox, S.H.: Biological Significance of Fluctuating Concentrations of Carbon Monoxide. Environ Sci Technol 20:916–923 (1986).
13. Saltzman, B.E.: Assessment of Health Effects of Fluctuating Concentrations Using Simplified Pharmacokinetic Algorithms. J Air Waste Man Assoc 46:1022–1034 (1996).
14. Aitcheson, J.; Brown, J.A.C.: The Lognormal Distribution. University Press, Cambridge, UK (1957).
15. Crow, E.L.; Shimizu, K. eds: Lognormal Distributions. Theory and Applications. Marcel Dekker, New York (1988).
16. Saltzman, B.E.; Cholak, J.; Schafer, L.J.; et al.: Concentrations of Six Metals in the Air of Eight Cities. Environ Sci Technol 19:328–333 (1985).
17. Saltzman, B.E.: Recent Risk Rates of Occupational Fatalities, Injuries, and Illness in U.S. Industries and Their Use in Planning Environmental Controls. Appl Occup Environ Hyg 16:723–725 (2001).