P. Meyer.

721560

On the Likelihood of Decompression Sickness

P.K. WEATHERSBY, L.D. HOMER, and E.T. FLYNN
Naval Medical Research Institute
Bethesda, Maryland 20814

REPOSITORY LLNL
COLLECTION PAUL MEYER'S PAPERS
BOX No. HUMAN RESEARCH NAVY/DECOMP. SICKNESS EXPER. STUDY FILES
FOLDER N/A

Running head: Likelihood and Decompression Sickness

Send all correspondence to:

P.K. Weathersby, Ph.D.
Naval Medical Research Institute
Hyperbaric Medicine Program Center
Bethesda, Maryland 20814
(202) 295-5898

J. Appl. Physiol.
In press, 1984

1157454

ABSTRACT

The occurrence of decompression sickness (DCS) in animals and men is characterized by the extreme variability of individual response. Nevertheless, models and analyses of decompression results have generally used a critical value approach to separate safe and unsafe decompression procedures. Application of the principle of maximum likelihood provides a formal and consistent way to quantify decompression risk and to apply models to data on decompression outcome. Using the maximum likelihood principle, a number of models were fit to such data from dose-response and maximum pressure reduction experiments with both rats and men. Several different formulations of two- and three-parameter models described the data well. In addition to summarizing data sets, the analyses provide a way to maximize the value of experimental observations, test theoretical predictions, estimate uncertainty in conclusions, and recommend safe practices.

Key words: decompression sickness, maximum likelihood, risk, rat, confidence regions, safety

1157455

Decompression sickness (DCS or the "bends") is a random event. Of a group of men experiencing the same diving conditions, seldom do all demonstrate signs of DCS. Despite the manifest variability of the effects of decompression among men, human decompression theories and trials are pursued according to deterministic viewpoints that predict the inevitable occurrence of DCS if a controlling parameter exceeds a critical value and the certain absence of DCS if the critical value is not exceeded. This critical parameter may be a bubble formation rate (20), a gas volume (13), a tissue supersaturation (21), a number of bubbles (22), or another possibility. A probabilistic approach that deals explicitly with variable outcome is clearly needed. Mathematical tools exist that deal with variable binary outcomes and allow theories to be fit to binary data. One very successful approach is the principle of maximum likelihood (10). Stated simply, maximum likelihood applies a theory or model to data and adjusts the parameters of the model until the theoretical predictions and the actual experimental data are in the closest possible agreement.

In this paper we will introduce the principle of maximum likelihood and demonstrate its utility with simple examples. Formal models of DCS will be postulated and evaluated using this principle in both animal and human studies.

LIKELIHOOD CRITERION

Assume that we have detailed records of the pressure exposures of a group of divers and their reaction to decompression. We accept that the probability of DCS is a specific function of certain dive characteristics:

$$\underline{p}(\text{DCS}) = \underline{f}(\text{dive}) \qquad \langle 1 \rangle$$

The actual functionality in Eq. 1 can range from a detailed physical model of bubble dynamics to a one-line empirical equation. If we are only interested in the occurrence of a specific response, each individual either demonstrates

1157456

symptoms or he does not. Thus, the sum of the probability of DCS and the probability of no symptoms of DCS in the dive must equal one:

$$\underline{p}(\text{no DCS}) = 1.0 - \underline{p}(\text{DCS}) \qquad \langle 2 \rangle$$

Eq. 2 can also be generalized to consist of multiple responses.

In examining the data, the probability of the actual outcome of a single man-dive can be defined using Eq. 2. Let X_i, the outcome variable, equal one if DCS occurred, and zero if no DCS occurred. Then, using the binary criterion of Eq. 2, the probability of the actual result of dive (i) is:

$$\underline{p}(\text{obs i}) = \underline{p}(\text{DCS,dive i})^{X_i} \quad \underline{p}(\text{no DCS,dive i})^{(1 - X_i)} \qquad \langle 3 \rangle$$

Here, $\underline{p}$(obs i) is the probability of the observed outcome of dive number (i).

The calculation of $\underline{p}$(obs) for one dive can be repeated using calculations for all other divers whose decompression conditions are known. Specifically, if the results of an entire trial of $\underline{n}$ decompression outcomes are known, and if the results of each trial can be considered independent from all the others, the likelihood of the entire trial outcome is defined as the product of the probabilities of each observation:

$$L\ (\text{trial}) = \underline{p}(\text{obs 1})\ \underline{p}(\text{obs 2})\ \underline{p}(\text{obs 3}) \ldots p(\text{obs } \underline{n}) \qquad \langle 4 \rangle$$

L is the "likelihood function" that yields the theoretical prediction of the entire outcome of the trial. The probability of the actual outcome of the trial of the first diver, whether or not he was stricken with DCS, is $\underline{p}$(obs 1). Each successive term is calculated as either $\underline{p}$(DCS) or $\underline{p}$(no DCS) for each subject in the trial according to Eq. 3. There is no requirement that each observation (each man-dive) replicate the others, only that each dive has a calculable outcome, as defined by Eqs. 1 and 2. Because L is the product of many numbers less than one, of which many may be very small, the natural logarithm of L, or LL (log likelihood) is commonly used. Since L or LL is a

1157457

direct measure of how well theory fits the entire trial, data analysis consists of adjusting the unknown constants (parameters) of the theory to maximize LL (16). In other words, constants are chosen to maximize agreement of the theory in question with the outcome of the entire series of man-dives. After adjusting constants, a good theory (Eq. 1) will have a large likelihood; a bad theory will have a small likelihood.

A simple example demonstrates the construction of a likelihood function. Suppose that in a series of man-dives on the same decompression schedule the following results occurred: N, N, Y, N, N, N, N, N, Y, Y (where N is no symptom and Y is a case of DCS). The simplest probability model is that a single constant probability, $\underline{c}$, exists:

$$\underline{p}(\text{bend}) = \underline{c} \qquad \langle 5 \rangle$$

Applying Eqs. 2, 3, 4, and 5 to that model and to the data, we obtain:

$$L = (1-\underline{c})\ (1-\underline{c})\ (\underline{c})\ (1-\underline{c})\ (1-\underline{c})\ (1-\underline{c})\ (1-\underline{c})\ (1-\underline{c})\ (\underline{c})\ (\underline{c})$$

$$L = (1-\underline{c})^7\ (\underline{c})^3$$

If $\underline{c}$ is allowed to vary over its possible range of zero to one, LL varies as shown in Fig. 1. The maximum likelihood occurs at $\underline{c} = 0.3$. This is the value we would choose intuitively for three cases out of 10 trials. A choice of $\underline{c}$ other than 0.3 lowers the value of L or LL. The graph of LL as a function of $\underline{c}$ is extremely flat, however, indicating many other possible values of $\underline{c}$ have nearly as great a likelihood. Thus, the conclusion that $\underline{c} = 0.3$ is very uncertain.

Fig. 1 also depicts a curve of LL versus $\underline{c}$ in the case of the same raw incidence, but in a tenfold larger trial: for example, 30 Ys and 70 Ns. This curve has a sharper maximum, and thus a greater confidence level for the conclusion that the value of $\underline{c}$ is near 0.3. The sharpness in the plot of LL vs. a parameter contains information about the precision of an estimated

1157458

parameter and this curvature is called the "information" in some statistics texts (10).

For simple, one-parameter models and small data sets, the calculation of LL can be performed by hand. For larger problems we use a computer program modified from a Marquardt nonlinear, least-square search routine to find the maximum likelihood (1). The program also estimates asymptotic standard errors of the parameters from calculations of the partial derivative of LL with respect to all of the estimated parameters (16) on the assumption that the log likelihood surface is shaped parabolically near the maximum. These asymptotic standard errors are reported in the tables of results.

DECOMPRESSION MODELS

This section addresses the construction of probabilistic models of DCS, i.e., formulations of f(dive) in Eq. 1. Probabilistic models of quantized responses are encountered frequently in pharmacology when the dose-response behavior of a drug is studied. In our case, the dose is the pressure history of the diver, and the response is DCS or no DCS. Candidate models for f should be well behaved, that is, they should predict a p(DCS) of zero when no decompression occurs, and should rise smoothly toward a near certainty of DCS when the pressure reduction is very large.

First, we will use another specific numerical example. Table 1 shows the hypothetical results of a study similar to an actual published report concerning human decompression sickness (9). Six men were exposed individually to compressed air in a chamber and then quickly decompressed to the surface. A simple model can be proposed that allows for variable depth, D, but ignores time and repeated use of the same subjects. The model does not correspond precisely to the views of the authors of the study (9), but has been chosen mainly for simplicity:

1157459

$$\underline{p}(DCS) = D / (D + D50) \qquad <6>$$

D50, the only parameter in the model, is the depth at which 50% of a diver population would be expected to suffer DCS. Eq. 6 is more complicated than Eq. 5, but still has only one parameter, D50. Note that Eq. 6 predicts only a small chance of DCS for a very shallow dive. At the other extreme, DCS is highly probable for a dive much deeper than D50. The (log) likelihood function for this model:

$$LL = \ln[34/(34+D50)] + \ln\{1 - [34/(34+D50)]\} + \ldots + \ln[38/(38+D50)]$$

is shown graphically in Fig. 2. The maximum occurs at a D50 of 47 fswg, but the value of D50 is uncertain as may be seen from the width of the peak. Thus, the approximate 95% confidence limits are still rather wide; the estimate of D50 lies between 15 and 250 fswg.

We can now postulate a series of more complicated models and examine their applicability to more extensive data sets. We will break the $\underline{f}$(dive) of Eq. 1 into two steps: the definition of decompression stress (or dose), and the linking of this stress to outcome by a dose-response function. The first dose-response function we will use is the general Hill equation, sometimes used in biology to describe, for example, radioimmunoassay standard curves or the oxyhemoglobin binding curve:

$$\underline{p}(DCS) = R^n / (R^n + R50^n) \qquad <7>$$

R is the dose or measure of decompression stress; R50 is the dose corresponding to a 50% incidence of DCS; and the exponent n is the order of the Hill equation that controls the steepness of the central portion of the sigmoid curves produced. Eq. 6, above, is a Hill equation with an exponent of one. The risk function is another model sometimes used in engineering failure or disease survival analysis (15):

$$\underline{p}(DCS) = 1.0 - \exp(-kR^n) \qquad <8>$$

1157460

The two parameters, k and n, are purely empirical numbers to be estimated with data. Eq. 8 has a natural extension to complicated risks by replacing R with a specific, time-varying function. Thus, it could be used directly in model dives with pressure histories that vary with time. Only one-step decompressions from saturation dives were considered in the present work; gas uptake and elimination kinetics and other time effects were ignored.

The choice of R depends on the decompression theory considered. In formulating mechanistic models, R may be the calculated bubble-nucleation rate, the bubble volume, the number of bubbles, etc. In purely empirical or phenomenological models, simpler measures can be used. Thus, we can choose the amount of pressure reduction, ΔP, for the sudden decompression of a tissue having the inert gas dissolved at tension $P1^1$ to a lower ambient pressure P2 (both pressures in absolute units):

$$R = P1 - P2 = \Delta P \qquad <9>$$

We will now examine the applicability of the above model equations to two actual data sets, one from rats and the other from men.

RAT AIR DECOMPRESSION DOSE RESPONSE

To obtain a detailed dose-response curve, 516 male Sprague-Dawley rats were exposed to one of nine pressures of compressed air for a period of 2 hr. This time was chosen as the essential saturation time when additional exposure time appeared to not increase the occcurence of DCS, although shorter times have previously been used for rat studies (3,5). The rats weighed an average of 272 g (range: 209-343 g), but the average rat weight at each exposure depth was not kept uniform. The rats were exposed 10 at a time in a steel pressure chamber, where pressure was increased to the maximum at a rate of 150 fsw/min. At the end of 2 hr, the pressure level was reduced to atmospheric pressure in 6-10 sec, and the rats were observed for symptoms of DCS (11). Only serious

1157461

symptoms (hindlimb paralysis and death) were included in the diagnosis of DCS. Nine different depths were used: 90, 100, 120, 150, 175, 200, 225, 250, and 300 fswg.

The results are summarized in Table 2 and Fig. 3. The fraction of rats with observable serious symptoms is plotted vs. alveolar nitrogen tension in Fig. 3. One animal each in the 175 and 200 fswg groups developed minor symptoms that clouded the binary diagnosis. These two observations were excluded from the calculation of raw incidence, but were treated as outcomes of 0.5 in the likelihood work (Eq. 3). The 95% confidence limits for raw incidence of DCS at each depth were obtained from standard tables for binomial sampling (8), and are plotted and tabulated. As expected, the fraction of symptomatic animals rises smoothly from zero to one over a finite range of decompression stress.

Several combinations of the models discussed earlier were applied to describe the dose-response curve. Differences in log likelihood between different models fit to the same data were used to assess the applicability of the different models. [The absolute value of the likelihood is analogous to the sum-of-squared-errors in analysis of variance (10)]. For a basis of comparison, suppose that instead of a continuous dose-response model, we calculate the average raw incidence for each of the nine depths independent of the data for other depths. We can then calculate a log likelihood according to Eq. 4 for this discontinuous, multistep model -- the perfect fit to these data. This LL, the first entry of Table 3, is -145.52, and no continuous model can have a greater likelihood.

The remaining entries in Table 3 are from continuous models of dose response. We first use the general Hill equation (Eq. 7) for dose response, and a pressure difference (Eq. 9) to define dose or DCS stress. Fitting to the

1157462

data produces estimates of a 50% dose of 134 fsw and a Hill-order parameter, n, of 8.8. (By comparison, using the Hill equation for the O_2-hemoglobin association curve generally gives an n of 2-3.) The next entry in Table 3 shows that forcing the data to a Hill exponent of 1.0 provides a poor fit to the data (see Fig. 3) and has a much lower likelihood. When the exponent is estimated from the data, the sigmoid is steeper in the central region, visibly fits the data better, and significantly increases the likelihood. It does so, however, at the cost of a more complex model and an additional parameter that must be obtained from fitting the data.

Is the two-parameter Hill equation as good as the multistep model? Such questions can be answered formally by using a likelihood ratio test statistic in a manner similar to the F statistic in least-square fitting (10). The likelihood ratio test statistic uses the ratio of the likelihoods of two hypotheses. Because differences of logarithms are equivalent to ratios, the difference in LL can be used. Suppose we have a general form of a model that has been fitted to a data set, and a more restricted hypothesis that can be expressed as a set of fixed parameters in the general model. The likelihood ratio (LR) test statistic is calculated:

$$LR = 2\ [LL(\text{general formulation}) - LL(\text{restricted hypothesis})] \qquad \langle 10 \rangle$$

Because LL is the log of the likelihood for each of the two theoretical possibilities considered, LR is a measure of the improvement achieved in the likelihood arising from the selection of one model over the other. The likelihood ratio is distributed as a chi-square variable whose degrees of freedom (df) correspond to the number of constrained parameters involved in the comparison (10,16).

For comparison of the multistep and two-parameter Hill model, the log likelihood improvement from Table 3 is 145.52 - (-150.30), or 4.78, from which

1157463

LR = 9.56. The multistep model estimates nine parameters. The degrees of freedom for the comparison is 7, and the tabulated value of chi-square for 7 df at the 0.05 confidence level is 14.07. Because the LR is less than the applicable chi-square, there is no significant difference between the two models, and the two-parameter Hill equation accurately summarizes these data.

Some other models also do well. The next two entries in Table 3 summarize the use of the risk function, Eq. 8, for dose-response. These two entries (lines 4 and 5 in Table 3) show that for the same number of estimated parameters, a risk formulation is better than a Hill function for describing the shape of the dose-response curve. The one-parameter model (n fixed at 1.0; compare entries 3 and 5) improves the LL from -286 to -270, while the two-parameter, variable-order fit (n allowed to vary; compare entries 2 and 6) ranges from -150 to -146. Since these are rather small differences, the preference for the risk model is not too strong.

The final entry in Table 3 demonstrates that still another dose-response function can be successfully applied. The pressure difference was used as the random variable in the cumulative, normal (Gaussian) distribution, sometimes called the probit model. The fit is excellent for this two-parameter model (LL = -148.13) with a mean depth of DCS of 170 fswa, comparable to the 50% response of other models. Several other one- and two-parameter models were also evaluated with maximum likelihoods within the range shown in Table 3. In summary, Table 3 shows that two one-parameter models fail to fit the data, but that several two-parameter formulations are excellent descriptions of rat decompression dose response. We do not assert that any of these models is "correct;" rather, our analysis showed that at least three models agree well with the data.

1157464

HUMAN HELIUM DECOMPRESSION

Human decompression experiments of the same scope as rat experiments are not available. It is possible, however, to combine observations from several sources to allow some testing of the hypotheses using the principle of maximum likelihood. Data were accumulated from the United States and the United Kingdom naval laboratories on human pressure-reduction experiments with HeO_2 mixtures. The following criteria were applied in accepting observations:

1. Subjects must have been saturated breathing HeO_2 for at least 24 hr at P1 prior to decompression.

2. Following rapid decompression, subjects must have remained at pressure P2 for at least 4 hr before being declared free of DCS symptoms. For example, if three subjects were therapeutically decompressed to P2 and one subject had symptoms at 1 hr for which all were recompressed, the data was taken as 1 bend + 0 no bends. This procedure probably introduces a slight bias since the subjects who would eventually not demonstrate symptoms are ignored.

3. No subject was accepted if he had DCS or even the marginal or equivocal symptoms known as "niggles" earlier in the dive. Most of the data were obtained during complex exposures lasting many days, and consideration of only the first demonstration of symptoms could also introduce a slight bias.

The accumulated data on 310 man-dives are recorded in Table 4. Sources are the Royal Navy Physiological Laboratory (RNPL), in the United Kingdom, deep-dive experiments from 1971-1974, (personal communication from J. Vorosmarti, partially reported in references 2 and 19) and Experimental Dive Unit (EDU) deep dives in the United States (personal communications from W. Spaur and E. Thalmann, partially reported in references 12 and 17). Since four values of $P_{I_{O_2}}$ were used, the definition of P1 for analysis was a calculated

1157465

alveolar PHe. That is consistent with the common, though unproven, assumption that inspired oxygen does not contribute to the risk of decompression (5,21). The approximate 1 ATA of nitrogen present in the pressure chamber atmosphere during the EDU experiments was treated as helium. Decompression rates were approximately 3 fsw/min in the RNPL data, and 30-60 fsw/min in the EDU data. In addition to the categories of DCS and no-DCS, a marginal category is also included. This includes cases that defied definite diagnosis by the attending medical officers. Depending on the purpose of the analysis, these dives could all be considered safe, or all resulting in DCS, or an intermediate outcome such as using X = 0.5 in Eq. 3.

The data in Table 4 are difficult to portray graphically. All the experiments are plotted in Fig. 4. The coordinates define the conditions used in the experiments of Table 4 as pressure reduction (P1-P2) versus final pressure after decompression (P2). The same symbol is used for each point because actual portrayal of the outcome would require a third dimension to record multiple values of 0 or 1 (or 0.5).

The paucity of detailed dose-response data is illustrated in Fig. 5. A subset of Table 4 was selected that is similar to the rat study design of Table 2: five series of exposures completed at a P2 of 154 fswg. The replicate observations from Table 4 were pooled to estimate raw incidence and 95% binomial confidence levels as in the rat study. The error bars are enormous, and the two model predictions plotted (see below) are well within the limits of experimental error. The wide latitude of reasonable theories is obvious.

Fig. 5 can also be used to illustrate a sometimes disturbing feature of binary data and maximum likelihood. When working with least-squares fitting to continuous data, a graph showing data points, together with the model curve,

1157466

can be used to visually estimate the goodness of fit. On a plot of experimental outcome versus the independent variable such as that in Fig. 5 (exposure pressure in this case), the raw data consist of points with ordinate values of either zero or one. A smooth function through a region bounded by individual data points on parallel lines does not provide much visual information. Statistical comparisons among models provide more useful information on relative goodness of fit.

Table 5 provides the results of analyzing all the data of Table 4 with marginal cases assigned an X of 0.5 in Eq. 3. (The effect of this assignment will be discussed later.) First, instead of the pressure-dependent models presented, suppose that we used the constant probability model, Eq. 5. That entry in Table 5 shows that a maximum LL of -78.11 is obtained. That result can be compared to more appealing models such as used for the rat data. The second entry in Table 5 is the first-order ($n = 1.0$) Hill equation using pressure difference, ΔP, as the dose. Its LL is scarcely different from the constant p result, which combined with the high and imprecise value of R50, implies that this model is not well suited to the data. Using an estimated exponent in the two-parameter Hill equation (not tabulated) does not significantly increase the likelihood, and increases R50 to several thousand feet with an imprecision of several hundred percent.

Numerous reports suggest that the simple difference of P1 and P2 is not a constant measure of the DCS dose, but that a "safe dose" increases with depth (5,13,22). One way to account for such a possibility is to define decompression stress as a ratio of pressures. Another aspect that can be included is the possibility of an absolute threshold of pressure reduction below which no chance of DCS exists. These two possibilities lead to an alternative definition of dose:

1157467

$$R = (P1 - P2 - Pt) / P2 \quad <11>$$

The threshold pressure, Pt, could be interpreted physically as a stable supersaturation, or (with a different definition of P1) as tissue gas undersaturation produced by metabolic functions. (R = 0 whenever (P1-P2) is less than Pt.)

Lines 3 and 4 in Table 5 use the ratio dose as defined by Eq. 11. With the threshold suppressed (Pt = 0.0) and the Hill equation held at n = 1, the one parameter of R50 is a fairly precise 4.2, and the likelihood is a major improvement over that obtained with a simple pressure difference. The likelihood improves significantly if a nonzero value of P_t is allowed (compared to $P_t = 0$, LR = 7.06; $P<0.05$ with 1 df). In fact, the threshold safe decompression of almost 25 ft, and the R50 ratio of 2.1 are in good agreement with values commonly used to specify safe human decompression (13,14,21). In Eq. 11 we used P2 in the denominator. For a data set containing only a single value of P2, such as the rat dose-response study, both difference and ratio definitions of dose necessarily provide identical likelihoods. For data sets where both P1 and P2 vary, the results may be different. We found that for the data in Table 4, however, use of either P1 or P2 in the denominator of Eq. 11 yielded the same general result.

Substitution of Eq. 11 in the risk function, Eq. 8, (line 5 in Table 5) also provided a reasonably good description of the data with $P_t = 0$. Again, a significant improvement in fit was achieved when P_t was allowed to assume a nonzero value (line 6 in Table 5). We do not imply that these results prove the existence of a region of completely safe decompression. Finding that a nonzero threshold parameter improves the description of these data conflicts with numerous anecdotal reports of DCS incidence following mild pressure exposures, and with opinion regarding some of the suggested mechanisms of DCS.

1157468

The 95% confidence limits around the predicted threshold conditions (6 to 44 fsw) make it impossible to deny that if a safe region exists, it may be exceedingly small.

An alternate way to account for the depth effect is to expand the definition of R50 in the Hill equation (Eq. 7). The conditions of 50% incidence of DCS, expressed as a pressure difference, are expanded as a polynomial in the final pressure, P2:

$$R50 = a + (b-1)P2 + c\ P2^2 \qquad \langle 12 \rangle$$

We still use the same pressure-reduction definition of dose:

$$R = P1 - P2 \qquad \langle 9 \rangle$$

Eq. 12 replaces the constant pressure difference in Eq. 8, where a, (b-1), c are the estimated intercept, slope, and curvature, respectively, of all points of 50% incidence of DCS on a plot such as Fig. 4. (On the more frequently used coordinates of P1 vs. P2, the 50% line has a slope of b.) The final two entries in Table 5 are the result of this model. Line 7 has a slope parameter significantly greater than one, and thus supports an effect of depth. No strong evidence of curvature in the 50% incidence line can be seen, however, since the quadratic coefficient, c, is not significantly different from zero and the LR comparison of Line 7 (no curvature) to Line 8 (best curvature) is not statistically significant. Not tabulated in Table 5 are the results of fitting several other two- and three-parameter models whose LL are in the same range as models 3-8.

Two plots show the results of fitting. Fig. 4 shows the location of the data of Table 4 and model-derived lines (entries 6 and 8 from Table 5) of constant 5% incidence of DCS. The lines are obtained from straightforward substitution of estimated model parameters and p (DCS) = 0.05 in the dose-response functions. Lines of 50% incidence are not shown because they

1157469

would be off scale, far above the data points. Such behavior is expected where data are not obtained under conditions where 50% of the human subjects are injured. Despite the different mathematical features of the two models plotted, the isoincidence lines are rather close.

A direct dose-response plot is presented in Fig. 5 for the most highly replicated subset of the Table 4 data. Because of the low number of total observations (31 total compared to 516 used for Fig. 3) the binomial sampling uncertainty is enormous. Clearly, either of the two model curves is consistent with the data.

Several models seem to have equivalent ability to describe the specific dives tabulated, but do they all make similar predictions of the answers to practical questions? For example, what do the models predict for a 5% incidence of DCS upon reduction of pressure to 1 ATA? The nonthreshold models (models 3,5,7,8 in Table 5) predict a PHe before decompression of 44-55 fswa, while the threshold models (lines 4 and 6 in Table 5) give estimates of 60-65 fswa. This correspondence is reassuring because the data actually have a raw incidence of about 7%, and all formulations that describe the data equally well should make equivalent predictions near that incidence range. The difference between the threshold and nonthreshold model predictions is also understandable. The former predicts absolutely no DCS over a certain range of small decompressions, while the latter predicts a small incidence over the same range. Thus, as is expected, the two classes of models have divergent predictions for low incidences (below 5%), but good agreement at incidences corresponding to decompressions reasonably higher than the threshold values. If statistically significant decisions need to be made about specification of an absolute threshold, further experiments must be performed because the available data are insufficient to answer the question. Data from hundreds of

1157470

dives would be required to reliably estimate the frequency of events that only occur about 1% of the time.

A frustrating aspect of human decompression experiments is the substantial fraction of cases in which diagnosis is uncertain. Concern arose that interpretation of marginal cases could significantly affect model performance and the conclusions of the present study. To address this concern two models used in Table 5 were applied to three different data sets. All data sets came from Table 4, but marginal cases were interpreted differently: all safe (X = 0), all DCS (X = 1), or all X = 0.5. The likelihood values shown in Table 6 cannot be compared across the three treatments of marginal cases because they are actually different data, but the estimated values of the parameters vary systematically. When all questionable outcomes are considered symptom free, somewhat higher polynomial coefficients for R50 or lower k are obtained, while considering any questionable outcomes as DCS makes the 50% incidence line include lower values of decompression dose. The predictions of 5% bends are similarly changed. The changes are not too great in magnitude, however, indicating that arbitrary interpretation of these marginal cases does not have a controlling influence on the analysis. Choice of model actually makes a bigger difference.

Two final features of the use of likelihood are illustrated in Figure 6. The coordinates are identical to those in Fig. 4. The solid line is the current U.S. Navy safety limit for human diving under these conditions with assumed inspired oxygen of 0.4 ATA (18). The two shaded bands are the 1 SE confidence regions around predicted 5% and 10% incidence lines, estimated with one of the models studied (Table 5, line 8). The uncertainty in the predictions is rather severe and becomes huge at greater depths. If 95% confidence bands were plotted, the overlap between 5% and 10% predictions would

1157471

be much greater. Thus, specification of any region defined as safe is quite difficult. Although subject to substantial uncertainty, the current Navy allowance appears to be riskier at the shallow end, and safer at the greater depths.

DISCUSSION

Studies of human decompression sickness are frustrated by two aspects of the disease. First, as in many biological processes, outcome is highly variable. Thus, if one of a small group of divers is stricken with DCS, it is impossible to decide whether the decompression schedule is unsafe, or whether the diver has an isolated case of bad luck with a generally safe schedule. The frequent practice of accepting a new procedure as safe if the first 10 subjects are free of symptoms only means [by the binomial distribution at $p<.05$ (8)] that the actual true incidence is unlikely to exceed 31 out of the next 100 man-dives. The second problem with human DCS is that barring misadventure, the actual incidence figures will always be much lower than optimal for estimation of parameters. Human diving has evolved into a relatively safe practice and the safe outcome data, or zeros, greatly outnumber the cases of DCS, or ones. Direct resolution between decompression schedules having 1% and 2% incidence requires approximately 1200 replicate observations. Such voluminous data will never become available. By contrast, resolution of the twofold difference between 30% and 60% incidence might be accomplished with just 20 dives.

The necessity for mathematical models to guide the prediction of safe decompression is well accepted. Many of the currently used decompression tables were obtained by informal use of empirical parameters such as the risk parameters defined in Eqs. 9 and 11. The important aspect of modeling in this study is the prediction of a specific probability of DCS rather than a

1157472

deterministic trigger. Further work in probabilistic models can now be anticipated.

An obvious and important extension of the present work here is the combination of risk event and kinetic terms. Numerous gas-exchange models are available in which the calculation of tissue-inert gas tension may be combined with predicting the risk of excess gas either empirically or mechanistically. Such a combined theory could be applied using maximum likelihood data from "bounce" (unsaturated tissue gas) diving. Formal recognition must be paid to the time-varying nature of such exposures, most simply by replacing the R in Eq. 8 with a specific time function such as a hazard function (15).

Whatever models are chosen, they involve certain unknown parameters. Evaluation and comparison of models requires an estimation procedure for the parameters that can: a) recognize that the occurrence of DCS under conditions possible for human experimentation is very uncertain, b) pool information from experiments with little or no replication and c) estimate uncertainty in the estimated parameters and allow statistical testing of generalized models. Maximum likelihood is a well established theory of estimation that we believe is well suited for these needs. The information produced in human diving trials is obtained under dangerous conditions and this technique can guide experiments so that a minimum number of subjects are used, but reliable answers to specific questions are still obtained.

A major lesson of the present exercise is the appreciation that the tabulated 310 human dives provide very limited information for statistical comparison of models. The human data in Table 4 have been examined repeatedly by authors promoting theories of decompression, but none have included an explicit treatment of its variability (13,20,22). A fair summary of our analyses is that these human saturation diving data justify about two

1157473

adjustable parameters for a theory and that several models can fit the data equally well.

The methodology is suitable for estimating risk in human diving, and for planning operations with an acceptable risk. Uncertainty in choosing models, uncertainty in the estimated parameters, and uncertainty in predictions such as Fig. 6 can be combined systematically to aid in making a judgment of acceptability. Recognition of the uncertainty, no matter how disturbing, is preferable to treatment of limited data sets with narrowly defined critical limits and no estimate of uncertainty. Fortunately, the data in the present study are sufficient to reliably choose safe limits within a few tens of feet at a given incidence level, regardless of the specific model used to analyze the data.

Although Buckles (7), Flynn (11), and Hills (14) recognized variable decompression results by reporting the distribution of outcome, variability was not included seriously in experimental design and interpretation until recently with the work of Berghage (3,4,5,6). The new aspect reported here is specific modeling of the response distribution by use of explicit dose-response models and a systematic procedure to apply the models to the data. Though statistically appropriate to binary outcome experiments, maximum likelihood may be difficult for some to accept because it forces a specific consideration of probabilities. These may be less satisfying intuitively than presuming that decompression sickness can be treated as a deterministic problem. DCS is not manifestly an all-or-nothing response to environmental pressures, however, and use of this method should improve our intuitive reaction to both decompression models and decompression data.

1157474

ACKNOWLEDGEMENTS

We are indebted to E.E.P. Barnard, W.H. Spaur, E.D. Thalmann, and J. Vorosmarti for access to their original observations on human decompression. We also thank S. Survanshi for computer programming assistance and M.A. Darmody for editorial support. This work was supported by the Naval Medical Research and Development Command under Work Units M0099PN.01A.0005 and MF58524.01D.0001. The opinions and assertions contained herein are the private ones of the writers and are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large.

The experiments supported herein were conducted according to the principles set forth in the "Guide for the Care and Use of Laboratory Animals," Institute of Labortary Animals Resources, National Research Council, DHEW, Pub. No. (NIH) 78-23.

1157475

FOOTNOTES

1. Several other papers analyzing these types of data have used various definitions of Pl. We defined Pl as alveolar inert gas by using the alveolar air equation, rather than attempting another normalization for the effect of differences in inspired oxygen. The calculation of alveolar conditions assumes saturated water vapor, pCO_2 of 40 mm Hg and a respiratory quotient of 0.9, but in fact gives numerical values of Pl very close to the partial pressure of inspired inert gas.

1157476

REFERENCES

1. BAILEY, R.C., and L.D. HOMER. An analogy permitting maximum likelihood estimation by a simple modification of general least squares algorithms. Bethesda, Md.: NMRI Technical Report 77-55, 1977. (NTIS No. AD A046957)

2. BARNARD, E.E.P. Fundamental studies in decompression from Steady-State Exposures. In: Underwater Physiology V. Proceedings of the Fifth Symposium on Underwater Physiology, edited by C.J. Lambertsen. Bethesda, Md.: Federation of American Societies for Experimental Biology, 1976, p. 263-271.

3. BERGHAGE, T.E., F.W. ARMSTRONG, and K.J. CONDA. Relationship between saturation exposure pressure and subsequent decompression in mice. Aviat. Space Environ. Med. 46: 244-247, 1975.

4. BERGHAGE, T.E., and T.M. MCCRACKEN. Equivalent air depth: fact or fiction. Undersea Biomed. Res. 6: 379-384, 1979.

5. BERGHAGE, T.E., J.A. GOMEZ, C.E. ROA, and T.R. EVERSON. Pressure reduction limits for rats following steady-state exposures between 6 and 60 ATA. Undersea Biomed. Res. 3: 261-271, 1976.

6. BERGHAGE, T.E., J.M. WOOLEY, and L.J. KEATING. The probabilistic nature of decompression sickness. Undersea Biomed. Res. 1: 189-196. 1974.

7. BUCKLES, R.G., and E. HARDENBERGH. The role of inert gas exchange and population statistics in studies of decompression sickness. In: Chemical Engineering in Medicine. Advances in Chemistry Series 118, edited by D.D. Reneau. Washington, DC: American Chemical Society, 1973, p. 17-34.

8. DIEM, K. (Editor), Documenta Geigy Scientific Tables. (6th edition) Ardsley, NY: Geigy Chem Corp., p. 85-103, 1962.

9. DUFFNER, G.J., and H.H. SNIDER. Effects of exposing men to compressed air and helium-oxygen mixtures for 12 hours at pressures of 2-2.6 atmospheres.

1157477

Washington, DC: Experimental Diving Unit Report 1-59, 1959. (NTIS No. AD 206733)

10. EDWARDS, A.W.F. Likelihood (an account of the statistical concept of likelihood and its application to scientific inference) Cambridge: University Press, 1972.

11. FLYNN, E.T., and C.J. LAMBERTSEN. Calibration of inert gas exchange in the mouse. In: Underwater Physiology. Proceedings of the Fourth Symposium on Underwater Physiology, edited by C.J. Lambertsen. New York: Academic Press, 1971, p. 179-191.

12. FLYNN, E.T., and W.H. SPAUR. Recent U.S. Navy experiments in saturation-excursion diving. UMS Workshop on development of decompression procedures for dives in excess of 400 feet, Bethesda, MD: Undersea Medical Society, Inc., 1975.

13. HENNESSY, T.R. An examination of the critical released gas concept in decompression sickness. Proceedings of the Royal Society. (London) B197: 299-443, 1977.

14. HILLS, B.A. The variation in susceptibility to decompression sickness. Int. J. Biometeor. 12: 343-348, 1968.

15. KALBFLEICH, J.D., and R.L. PRENTICE. The Statistical Analysis of Failure Time Data. New York: Wiley, 1980, p. 1-19 and p. 189-198.

16. KENDALL, M.G., and A. STUART. The Advanced Theory of Statistics. (Volume 2, 4th edition) London: Haffner, 1979, p. 38-81 and p. 240-280.

17. SPAUR, W.H., E.D. THALMANN, E.T. FLYNN, J.W. ZUMRICK, T.W. REEDY, and J.M. RINGELBERG. Development of unlimited duration excursion tables and procedures for helium-oxygen saturation diving. Undersea Biomed. Res. 5: 159-177, 1978.

18. U.S. Navy Diving Manual. Vol. 2, Chap. 14. Washington, D.C.: Navy

1157478

Department. Change 1, 1975. (NAVSEA 0094-LP-001-9010)

19. VOROSMARTI, J., R. deG. HANSON, and E.E.P. BARNARD. Further studies in decompression from steady-state exposure to 250 meters. In: Underwater Physiology VI. Proceedings of the Sixth Symposium on Underwater Physiology, edited by C.J. Lambertsen. Bethesda, MD: Undersea Medical Society, Inc., 1978, p. 435.

20. WEATHERSBY, P.K., L.D.HOMER, and E.T. FLYNN. Homogeneous nucleation of gas bubbles in vivo. J. Appl. Physiol.: Respirat. Environ. Exercise Physiol. 53: 940-946, 1982.

21. WORKMAN, R.D. Calculation of air saturation decompression tables. Washington, D.C.: Experimental Diving Unit Report 11-57, 1957. (NTIS No. AD 627906)

22. YOUNT, D.E. Application of a bubble formation model to decompression sickness in rats and humans. Aviat. Space Environ. Med. 50: 44-50, 1979.

23. YOUNT, D.E., and R.H. STRAUSS. Bubble formation in gelatin: a model for decompression sickness. J. Appl. Physics. 47: 5081-5088, 1976.

115479

TABLE 1.

Hypothetical Decompression After Breathing Compressed Air

	Diver Number					
Depth (fswg)	1	2	3	4	5	6
34	Y*	N†	N	N	N	N
36	–‡	N	Y	Y	N	N
38	–	Y	–	–	Y	Y

* Y = DCS symptoms.

† N = no DCS symptoms.

‡ – = no dive.

1157480

TABLE 2.

Rat Decompressions to 1 ATA From Compressed Air Exposures*

Depth (fswg)	Alveolar pN_2 (fswa)	Rats	DCS Rats	Incidence	95% Limits
90	95.9	50	0	0.0	0 - 0.071
100	103.8	30	0	0.0	0 - 0.116
120	119.6	120	5	0.042	0.013 - 0.095
150	143.3	60	7	0.117	0.048 - 0.226
175	163.1	60	25	0.424	0.281 - 0.543
200	182.8	60	40	0.678	0.526 - 0.779
225	202.6	66	58	0.879	0.775 - 0.946
250	222.3	60	60	1.00	0.940 - 1.00
300	261.8	10	10	1.00	0.589 - 1.00

* Two rats (one at 175 and one at 225 fsw exposure) had ambiguous symptoms after decompression. These rats were ignored in calculation of raw incidence and binomial limits, and were assigned outcomes of 0.5 in likelihood calculations.

1157481

TABLE 3.

Likelihood Results for Models and Rat Decompression on Air

	Model <Eq.> with Parameters ± SE*		Log Likelihood
1.	Multistep (noncontinuous) response; 9 averages		-145.52
	Pressure Difference <9>, Hill Response <7>		
2.	R50 = 134.5 ± 2.3	n = 8.78 ± 0.76	-150.30
3.	R50 = 177 ± 18	n = 1.00	-285.68
	Pressure Difference <9>, Risk Function <8>		
4.	$k = 9.8 \pm 8.1 \times 10^{-5}$	n = 6.18 ± 0.55	-145.91
5.	k = 0.152 ± 0.012	n = 1.00	-270.26
	Pressure Difference <9>, Normal Distribution		
6.	Mean = 169.6 ± 2.2 SD = 25.9 ± 2.0		-148.13

* Parameters are given with error estimates of 1 SE calculated as asymptotic errors in the likelihood. If no error is shown, the parameter was fixed at the value shown.

1157482

TABLE 4.

Human One-step Decompression Breathing He-O_2

Depth (fswg)	PIO2 (ATA)	P1 (fswa)	P2 (fswg)	Outcome			Source
				Safe	Marg.	DCS	
33	.45	49.9	0	6	0	0	RNPL
33	.22	57.3	0	21	0	0	RNPL
79	.45	95.8	33	3	0	0	RNPL
79	.22	103.2	33	20	1	4	RNPL
85	.45	101.8	33	3	0	0	RNPL
85	.22	109.2	33	1	1	1	RNPL
135	.22	159.2	79	3	0	0	RNPL
144	.22	168.2	79	3	0	0	RNPL
154	.22	178.2	79	29	0	3	RNPL
167	.22	191.2	79	0	0	1	RNPL
190	.22	214.2	154	3	0	0	RNPL
210	.22	234.2	154	3	0	0	RNPL
226	.22	250.2	154	11	0	2	RNPL
240	.22	264.2	154	9	0	0	RNPL
246	.22	270.2	154	0	3	1	RNPL
226	.22	250.2	167	0	0	3	RNPL
302	.22	326.2	226	3	0	0	RNPL
328	.22	352.2	226	5	1	0	RNPL
328	.40	346.2	312	30	0	0	RNPL
354	.22	378.2	226	3	0	0	RNPL
328	.22	352.2	272	3	0	0	RNPL
328	.22	352.2	276	9	0	0	RNPL
436	.22	462.2	397	3	0	0	RNPL

1157483

TABLE 4.

Human One-step Decompression Breathing He-O_2 (con.)

449	.22	473.2	410	3	0	0	RNPL
449	.22	473.2	426	3	0	0	RNPL
492	.22	516.2	476	3	0	0	RNPL
590	.22	614.2	574	3	0	0	RNPL
689	.22	713.2	673	15	0	0	RNPL
225	.40	243.3	150	15	0	0	EDU
230	.40	248.3	150	5	0	0	EDU
400	.40	408.2	300	10	1	0	EDU
410	.40	418.2	300	5	0	0	EDU
750	.375	769.0	600	9	0	1	EDU
1000	.375	1019.0	850	5	0	0	EDU
1000	.375	1019.0	832	0	0	1	EDU
1000	.375	1019.0	820	15	0	0	EDU
1000	.375	1019.0	810	5	0	0	EDU
1000	.375	1019.0	800	0	0	1	EDU
1208	.375	1227.0	1020	6	0	0	EDU
1400	.375	1419.0	1180	6	0	0	EDU
1400	.375	1419.0	1166	6	0	0	EDU

1157484

TABLE 5.

Likelihood Results for Models and Human Decompression on Helium-Oxygen*

Model <Eq.> with Parameters ± SE†				Log Likelihood
No Pressure effect <5>				
1.	c = 0.069 ± 0.015			-78.11
Pressure Difference <9>, Hill Response <7>				
2.	R50 = 843 ± 194	n = 1.0		-78.23
Pressure Ratio <11>, Hill Response <7>				
3.	R50 = 4.16 ± .99	Pt = 0.0	n = 1.0	-72.85
4.	R50 = 2.12 ± .90	Pt = 24.4 ± 9.6	n = 1.0	-69.32
Pressure Ratio <11>, Risk Function <8>				
5.	k = .226 ± .051	Pt = 0.0	n = 1.0	-72.94
6.	k = .44 ± .17	Pt = 24.4 ± 9.3	n = 1.0	-69.37
Pressure Difference <9>, Hill Response <7>, R50 as polynomial <12>				
7.	a = 176 ± 196		c = 0.0	-71.47
	b = 3.6 ± 1.5		n = 1.0	
8.	a = 307 ± 306		c = $4.5 \pm 5.9 \times 10^{-3}$	-70.86
	b = 1.5 ± 3.2		n = 1.0	

* Marginal results treated as X = 0.5.

† Parameters are given with error estimates of 1 SE calculated as asymptotic errors in the likelihood. If no error is shown, the parameter was fixed at the value shown.

1157485

TABLE 6.

Effect of Marginal Cases on Human DCS Modeling

	Marginal = Safe	Marginal = 0.5	Marginal = DCS
Model 5. Pressure ratio, Risk function, all n = 1.0			
k	0.35 ± 0.15	0.44 ± 0.17	0.52 ± 0.18
Pt	24.3 ± 9.6	24.4 ± 9.3	24.5 ± 8.2
LL	-61.99	-69.37	-76.14
P1 (5%)	62.0	61.3	60.8
Model 7. Pressure difference, Hill Eq., R50 dependence on P2, n = 1.0			
a	310 ± 383	307 ± 306	305 ± 232
b	2.5 ± 4.1	1.5 ± 3.2	0.8 ± 2.5
c	$3.4 \pm 6.7 \times 10^{-3}$	$4.5 \pm 5.9 \times 10^{-3}$	$5.1 \pm 5.2 \times 10^{-3}$
LL	-63.24	-70.86	-77.75
P1 (5%)	52.1	50.2	48.9

1157486

FIGURE LEGENDS

Figure 1. Log likelihood function for parameter c in Eq. 5; the estimate of underlying probability in an experiment producing three cases of bends out of 10 dives (solid line), or 30 cases of bends out of 100 dives (broken line). To allow both curves to been seen together, the latter has been raised by adding 50.0 to the LL.

Figure 2. Log likelihood function for a simple dose response model, Eq. 6, in human air decompression experiments from Table 1. The abscissa is the parameter value for 50% bends incidence.

Figure 3. Sudden decompression of rats from P_{N2} (alveolar N_2 pressure = P1) to an ambient pressure of 1 ATA plotted on the abscissa. The ordinate is the fraction of rats exhibiting serious symptoms of DCS (squares), with error bars corresponding to 95% confidence limits on the binomial observations. Dashed lines correspond to models 2 and 3 of Table 3.

Figure 4. Human single-step decompression on helium. Pressure reduction (P1-P2) is plotted against the final pressure (P2). Each open figure represents the condition used in one or more trials tabulated in Table 4. The lines are estimates of conditions producing a 5% incidence of DCS calculated from best-fit parameters of models 6 and 8 in Table 5.

Figure 5. Conditions of human one-step decompression experiments breathing helium and oxygen. Ordinate is probability of DCS, abscissa is P1 (tissue He pressure). Triangles are raw incidence from the P2 = 154 fswg subset of the data in Table 4; error bars are 95% binomial limits on the raw data. Two dose-response curves are plotted (entries 6 and 8 from Table 5).

Figure 6. Human single-step decompression on helium. Pressure reduction (P1-P2) is plotted against the final pressures (P2) using the same coordinates as in Fig. 4. The two shaded areas are approximate 1 SE error bounds on 5% and

1157487

10% estimated incidences of bends; incidence lines calculated from best-fit parameters of Model 8 in Table 5. The solid line is the currently allowable USN limit at an assumed inspired oxygen level of 0.4 ATA.

1157488