Undersea Biomedical Research, Vol. 12, No. 3, September 1985

Statistical aspects of the design and testing of decompression tables

L. D. HOMER and P. K. WEATHERSBY

Naval Medical Research Institute, Bethesda, MD 20814

Homer LD, Weathersby PK. Statistical aspects of the design and testing of decompression tables. Undersea Biomed Res 1985; 12(3):239–249.—Before a decompression procedure is recommended for general use it is subjected to a limited number of human trial dives. Based on the trial, one attempts to reject unsafe procedures but accept those with a low incidence of decompression sickness (DCS). Binomial confidence regions are often so broad that even after 40 dives it may be impossible to distinguish between the possibility that the table being tested has a 0.6% risk of DCS and the possibility that it has a 17% risk. Our proposed alternative is to select some rule (e.g., one or more cases of DCS in 10 dives) for rejecting tables and to calculate the probabilities of accepting tables as a function of the probability of DCS. With such calculations we conclude that (a) generally one cannot reduce the risk of adopting unsafe tables without increasing the risk of rejecting safe ones unless one chooses to increase the number of test dives; (b) truncated sequential designs could reduce the number of dives required for testing by 15 to 20%; and (c) rules similar to the ones tested will always have a zone of indifference. Tables with a probability of DCS in this zone will be accepted or rejected with nearly equal frequency even if tested with hundreds of dives. The use of models describing the probability of DCS as a function of dive parameters should allow us to combine information from dives previously analyzed separately and perhaps to improve our selection of new tables to be tested.

binomial distribution
mathematical model
sequential design
power curves

A safe decompression schedule is one in which decompression sickness (DCS) rarely occurs, whereas an unsafe decompression schedule is one in which it occurs more often than is deemed reasonable. When testing new decompression procedures one tries to retain procedures with a low incidence of DCS, discard procedures leading to a high incidence of DCS, and to do so with as few trial decompressions as possible. Historically, decompression trials seem to have been conducted with little guidance from statistical theory.

Suppose we observe 2 cases of DCS in 40 trials of a new decompression procedure. Is this procedure safe? Unfortunately, experience has taught us that if we were to repeat our 40 trials we might encounter no cases of DCS or one, two, or more cases. Before we can judge the meaning of 2 out of 40 we must be able to assess the variability

we may encounter. Presumably, each time a decompression is undertaken there is a fixed but unknown "true" probability that DCS will occur. This probability will be different for different decompression schedules but we will assume it is the same for all individuals and the same from day to day for a given decompression schedule. [We use the terms "decompression table" and "decompression schedule" nearly interchangeably to describe the decompression procedure to be tested. To be precise, our remarks apply to a single specified dive and decompression (a "schedule") and require the additional restrictive assumption of *equal* unknown incidence to apply to a related set of procedures (a "table")]. These assumptions, which are consistent with a large population study of human altitude decompression (1), allow the variability to be predicted using the binomial distribution (2). Tables predict that if we observe 2 cases of DCS out of 40 trials, the 95% confidence region for the true incidence of DCS ranges from 0.6 to 17%. An incidence of 0.6% would be as safe as some procedures already in use, but a 17% incidence is altogether unsatisfactory. We conclude that this procedure might be safe, but then again it might not. We do not believe calculation and inspection of such confidence regions is the best way to address our problem. One deficiency is that the region depicted is a two-tail region, implying that we are just as fearful of approving a schedule with a DCS incidence less than 0.6% as we are of approving one with a 17% incidence. Another deficiency is that we still don't know whether to call this procedure safe or unsafe. Instead we will suggest approaches that recognize the two kinds of mistakes we may make when declaring a table to be safe. One mistake is to reject a safe schedule, the other is to retain a poor one. If we recognize that we may err by accepting unsafe tables or err by rejecting safe ones, we may develop alternative ways of looking at our problem that recognize the two kinds of mistakes to be made and permit us to calculate the probability of making such mistakes. Thus armed, we may hope to improve our rules for accepting new tables.

Selection rules for a fixed number of dives

If we adopt a rule for rejecting a schedule, we may calculate, using the binomial distribution, the probability of rejecting decompression schedules with that rule under different assumptions about the true incidence of DCS. For example, let us suppose we adopt the policy of running 10 trials and rejecting as unsafe any decompression procedure leading to one or more cases of DCS. If the true frequency of decompression sickness using the table in question is P, then the probability of observing k cases in N trials is, according to the binomial distribution,

$$B(k \mid N) = \{N!/(k!(N-k)!)\}\, P^k\, (1-P)^{N-k} \qquad (1)$$

The factorial notation, $N!$, means $(1)(2) \ldots (N-2)(N-1)(N)$, and 0! is defined as one. For example, 3! is 6, 4! is 24, and so on. Thus,

$$B(1 \mid 10) = 10!/(9!1!)\, P^1\, (1-P)^9 = 10\, P(1-P)^9$$

Now our rule asserts we will only accept the table if we have no cases of DCS in 10 trials. The probability of this outcome is $B(0{:}10)$, which is equal to $(1-P)^{10}$. The upper curve of Fig. 1 shows the probability that no cases of DCS will occur in 10 trials when the true probability is between 0 and 0.2. Such plots are termed power curves in statistics. Since the table will be accepted if there are no cases of DCS in

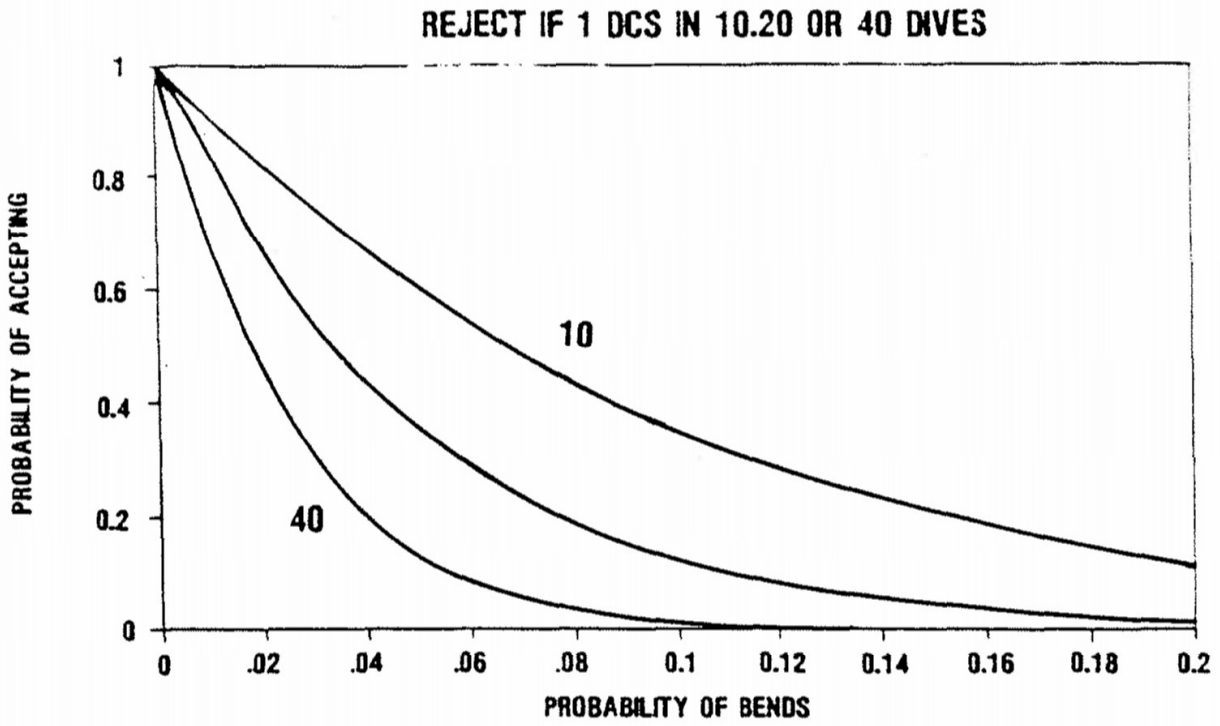

Fig. 1. Probability of accepting a decompression schedule against the underlying "true" incidence of DCS under several fixed trial rules. The 3 cases plotted correspond to strategies of accepting the schedule if no cases of DCS occur in trials of 10, 20, or 40 successive dives.

10 trials, the curve is equivalent to the probability of accepting the table. If the true incidence is low, the procedure works quite well. If the incidence of DCS were 0.6%, our rule would accept the table over 94% of the time. For bad tables it does not perform well. A decompression table with a DCS frequency of 13% would not always lead to one or more cases of bends in 10 trial dives. About 25% of the time such a table would lead to no bends at all in 10 dives and thus would have a probability of 0.25 of being accepted. As Fig. 1 shows, we could improve our rejection of bad tables by adopting a 1-out-of-20 rule and improve it still further with a 1-out-of-40 rule. A 1-out-of-40 rule would reject a table with 13% DCS incidence over 99% of the time. Unfortunately, it would also reject a table with a 0.6% incidence over 20% of the time. By adopting a 1-out-of-40 rule instead of a 2-out-of-40 rule, we can improve our rejection of poor tables only by increasing the risk of rejecting a good table.

Can we improve the balance between accepting good tables and rejecting poor ones by adopting a 2-out-of-40 rule? If we observe two or more cases of DCS in 40 trials we will reject the proposed table, otherwise we will accept it. Our acceptance probability, A, will then be the sum of $B(0 \mid 40)$ and $B(1 \mid 40)$, since zero or one case of DCS is the only acceptable outcome. Using our binomial formula we would get

$$A = B(0 \mid 40) + B(1 \mid 40) = (1\text{-}p)^{40} + 40\, P^1\, (1\text{-}P)^{39}$$

Then if the true incidence of DCS is 0.6%, our rejection rule would retain the schedule 97.5% of the time and reject it 2.5%. On the other hand, if the true incidence were 13%, our 2-out-of-40 rule will reject the schedule 97.5% of the time and retain it 2.5% of the time. As shown in Fig. 2, if the true incidence lies between these values, rejection and retention are more evenly balanced. For example, if the incidence of DCS is 5%, then our rule will reject the schedule 60% of the time and retain it 40%

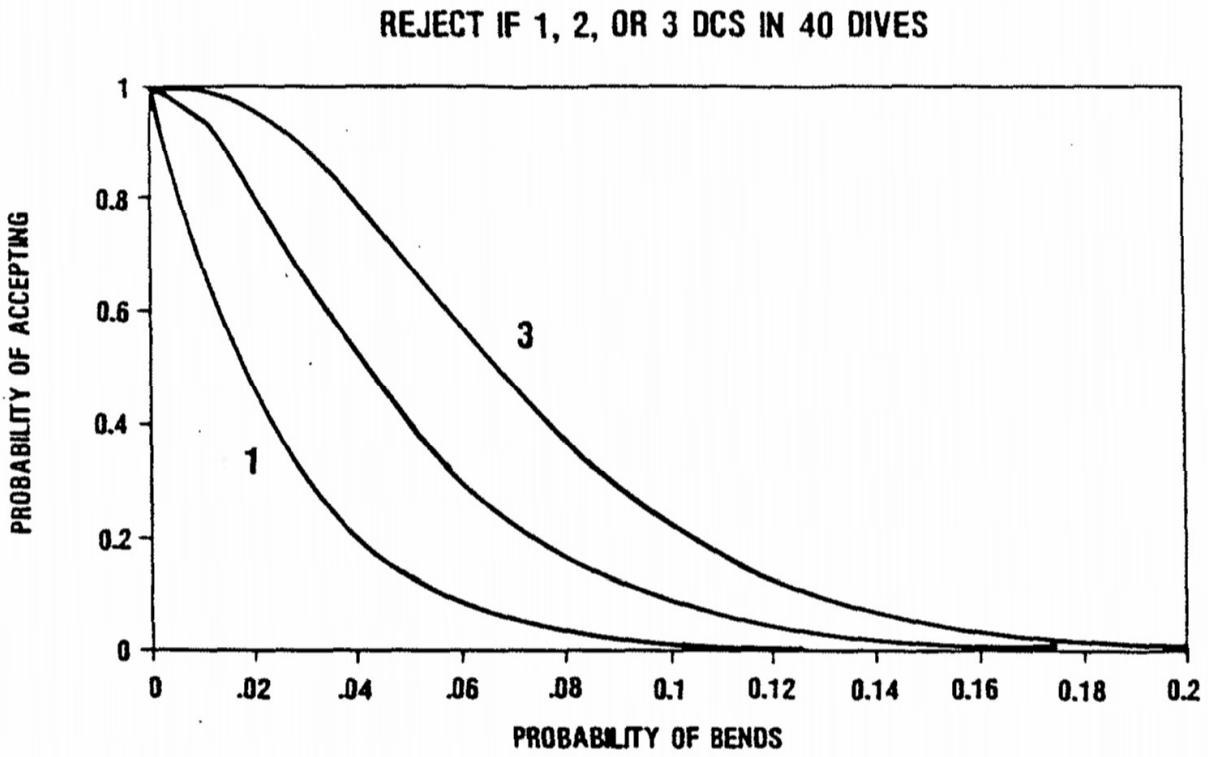

Fig. 2. Probability of accepting a decompression schedule against the underlying "true" incidence of DCS under several different fixed trial rules. The cases of rejecting tables for 1, 2, or 3 cases of DCS out of 40 repeated dives are shown.

of the time. The 1-out-of-40 rule gives better rejection of poor tables but only by increasing the chance of rejecting good tables. Figure 2 also shows the expected acceptance frequencies for a 3-out-of-40 rule. This does a fine job of accepting good tables, but the possibility of accepting poor tables is clearly greater than for the 1-out-of-40 or 2-out-of-40 rules. It is a general characteristic of any selection rule for diving schedules that for a given, fixed number of test dives improved rejection of poor schedules can only be purchased at the price of increasing the possibility of rejecting good schedules.

To improve our selectivity it is necessary to increase the number of test dives. Figure 3 shows the expected acceptance frequencies for rejection rules 2/40, 5/100, and 15/300. For these three rules the proportion of cases of DCS leading to rejection of the table is the same. The slope of the 2/40 curve is gradual, whereas that of the 15/300 is flatter and higher in the range of good tables, drops more precipitously in some intermediate range, and is lower and flatter for poor tables. The 15/300 curve is thus more selective in discriminating more sharply between good tables and poor tables. In intermediate ranges near a DCS frequency of 5%, each of these rules will accept the table about 50% of the time. It is important to understand that any reasonable selection rule will have such a zone of indifference; that is, there will be a range of DCS frequencies for which the rule will accept the tables about as often as it rejects them. As we increase the number of trials this zone of indifference becomes narrower.

Sequential trials

An altered strategy can improve the selectivity without altering the average number of trial dives. If there are 2 cases of DCS in our first 2 dives it is not necessary to

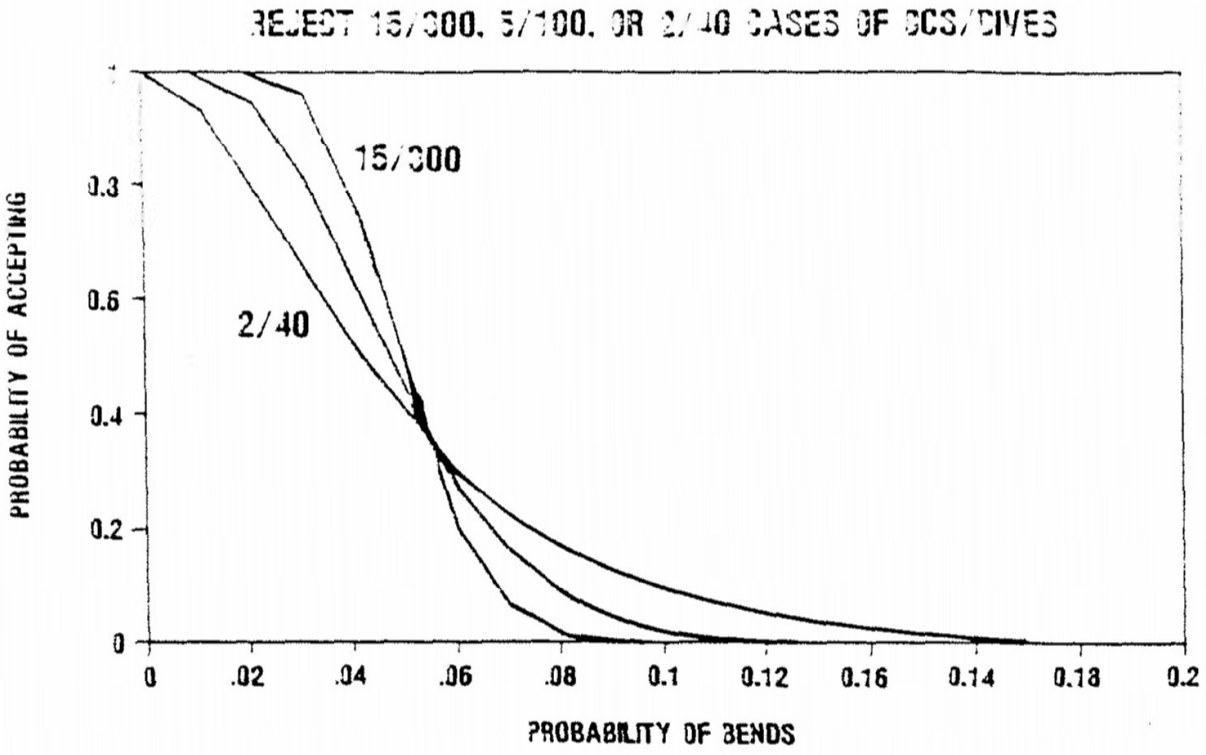

Fig. 3. Probability of accepting a decompression schedule against the underlying "true" incidence of DCS for different length trials all limited by a raw incidence of 5%. Rules displayed are for 2 cases out of 40, 5 out of 100, and 15 out of 300.

complete the remaining 38 test dives to reject the decompression schedule with a 2-out-of-40 rule. If we are willing to extend our trial series when we do not find 2 cases of DCS in the first 40 dives, we can improve our selectivity while maintaining an average near 40 dives per trial. Precisely how to devise such strategies is described in books on sequential analysis or sequential trials (3).

Figure 4 shows an example of such a rule. The starting point is the lower left corner where no dives have been performed and no case of DCS encountered. After each dive, one plots the number of cases of DCS against the total number of dives. If the plot reaches or crosses the upper sloping line we reject the table and stop testing. If the plot crosses or reaches the lower line we accept the table. Notice that we can accept a table after only 28 uneventful dives (where the acceptance line leaves the 0 cases axis), and that we must have more than 2 cases of DCS to reject if the second case occurs after dive 20 (where the rejection line rises above 2 cases of DCS). Nothing is specified about stopping after a certain number of dives. Formulas are also available to calculate an average number required (3). The scheme shown would stop after an average of about 30 to 35 dives for good tables and after an average of 17 or fewer dives for poor tables. In between, however, we will often exceed 40 dives. For a DCS frequency of 3% we may expect an average of 44 dives, with most, but not all, sequences ending before 100 dives. The formulas given by Wald (3) for the lines of the selection rule require the tester to select a bends incidence $p0$ that is low enough so that we would not wish to reject a table with a DCS incidence as low as $p0$ more than α proportion of the time. Also, the tester must select an incidence

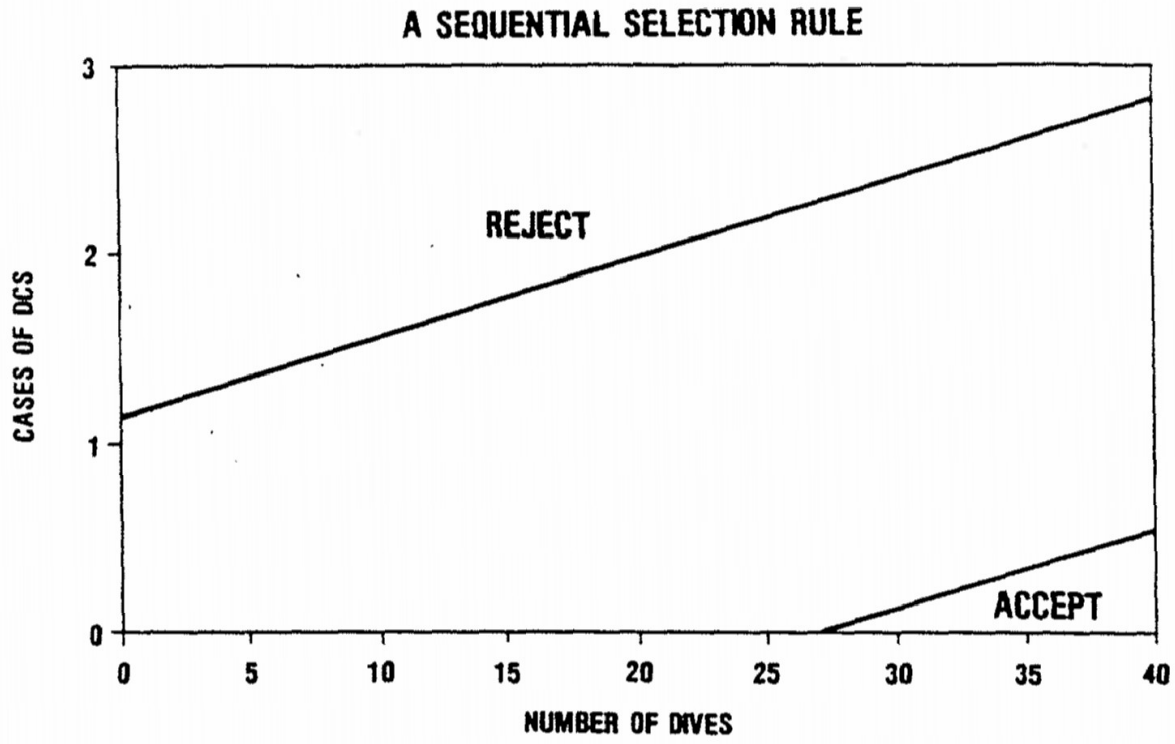

Fig. 4. An open sequential design of a decompression trial. After each trial dive, one moves a step to the right and upward one unit only if a case of DCS occurs. The trial ends when either the upper or lower bound is crossed.

$p1$ so high that we wish to accept such an incidence no more than β proportion of the time. Then the intercept of the lower line is,

$$h0 = \log[\beta/(1-\alpha)]/[\log(p1/p0) - \log((1-p1)/(1-p0))]. \quad (2)$$

The slope is given by,

$$s = \log[(1-p0)/(1-p1)]/[\log(p1/p0) - \log((1-p1)/(1-p0))]. \quad (3)$$

The intercept of the top line is,

$$h1 = \log[(1-\beta)/\alpha]/(\log(p1/p0) - \log((1-p1)/(1-p0))]. \quad (4)$$

The slope of the upper line is the same as that for the lower line. For the rule of Fig. 4 we selected $p0 = 0.006$, $\alpha = 0.025$, $p1 = 0.13$, $\beta = 0.025$. These values lead to $h0 = 1.14$, $h1 = 1.14$, $s = 0.042$. The performance of this sequential rule is compared with that of the 2/40 and 5/100 rules in Fig. 5. As may be seen, it runs about the same risk of accepting poor tables as the 2/40 rule, but is almost as good as the 5/100 rule in accepting tables with DCS frequencies below 1%. Unfortunately, the price to be paid is that one would have to accept occasional, very long test runs when tables with intermediate DCS frequencies are encountered. This is likely to be unacceptable to most groups with the responsibility of testing tables.

Closed sequential trials

An alternative is to take a rule like the one of Fig. 4 and modify it by stipulating that the trials will terminate after some maximum number of dives even if the boundaries have not been crossed, and in that case the decision to accept or reject the table

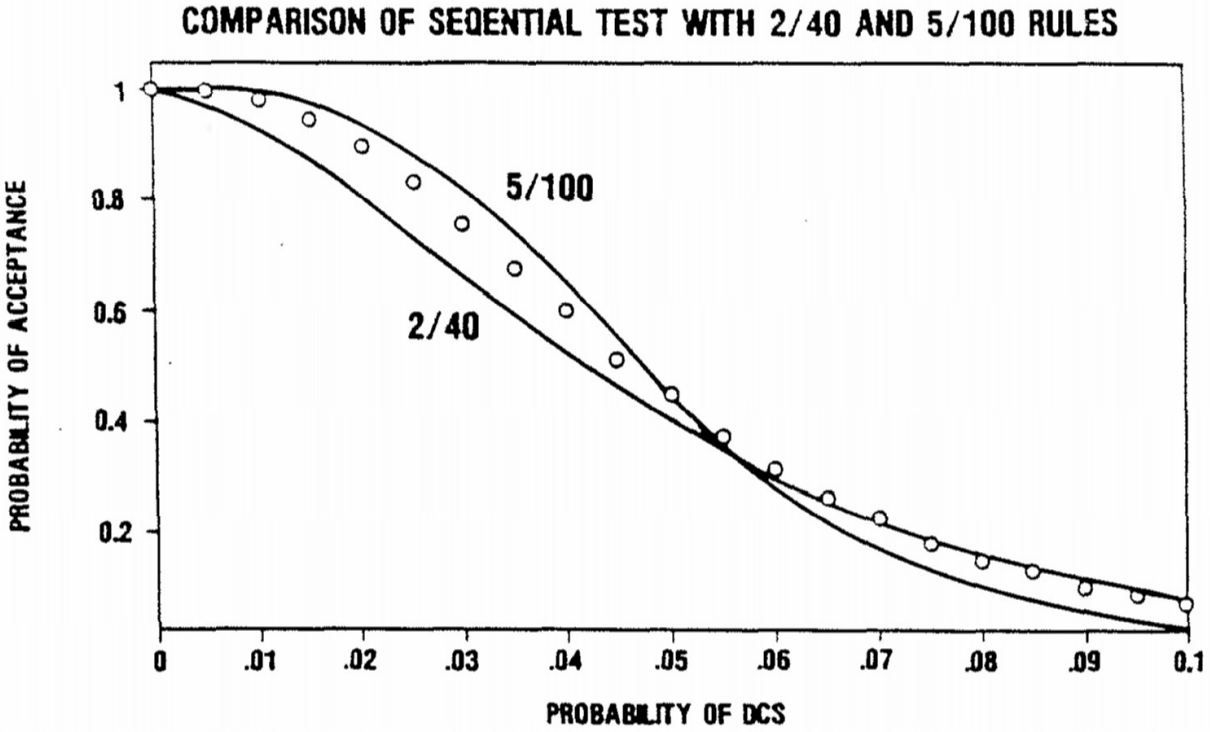

Fig. 5. Comparison of the probability of accepting a decompression schedule against the underlying "true" incidence of DCS for several possible acceptance rules. The lines of a fixed 2/40 or 5/100 trial are the same as Fig. 3, while the *dots* correspond to the open sequential trial of Fig. 4.

will be made on the basis of some accessory rule. For example, we might limit the total number of dives to 40 and, if the boundaries have not been crossed after 40 dives, accept the table if the last plotted point is closer to the acceptance boundary and reject the table if the last plotted point is closer to the rejection boundary. The boundaries for this new truncated rule are shown in Fig. 6. Notice that 2 cases of DCS lead to rejection, 0 cases in 28 trials to acceptance, and 1 out of 40 to acceptance. The upper boundary now stops at 2 because even if 3 is not reached 2/40 is closer to the upper boundary than to the lower boundary. We have evaluated the performance of this truncated sequential design by using a random number generator to simulate the outcome of thousands of dives at each of many DCS frequencies between 0 and 0.2. Figure 7 compares the performance of this truncated sequential rule to that of a simple 2/40 rule and to the untruncated design. As may be seen, the performance of this rule is about the same as that of the 2/40 rule in separating good and bad diving tables.

The truncated design, however, represents a definite improvement over the 2/40 rule when viewed from the standpoint of the number of trials to be run. Figure 8 shows a plot of the average expected number of dives needed to reach a decision for DCS frequencies from 0 to 0.2. The open sequential design requires longer trials in the midrange and occasional, very long trials to attain the performance displayed in Fig. 7. The 2/40 rule is similar to the other 2 rules in the expected length of trials for poorer tables, but because it lacks the lower boundary shown in Fig. 6, it is relatively inefficient in deciding if we are dealing with a good table. Accordingly, the truncated sequential design appears in several respects to be an attractive compromise between an open sequential design and a threshold rejection rule.

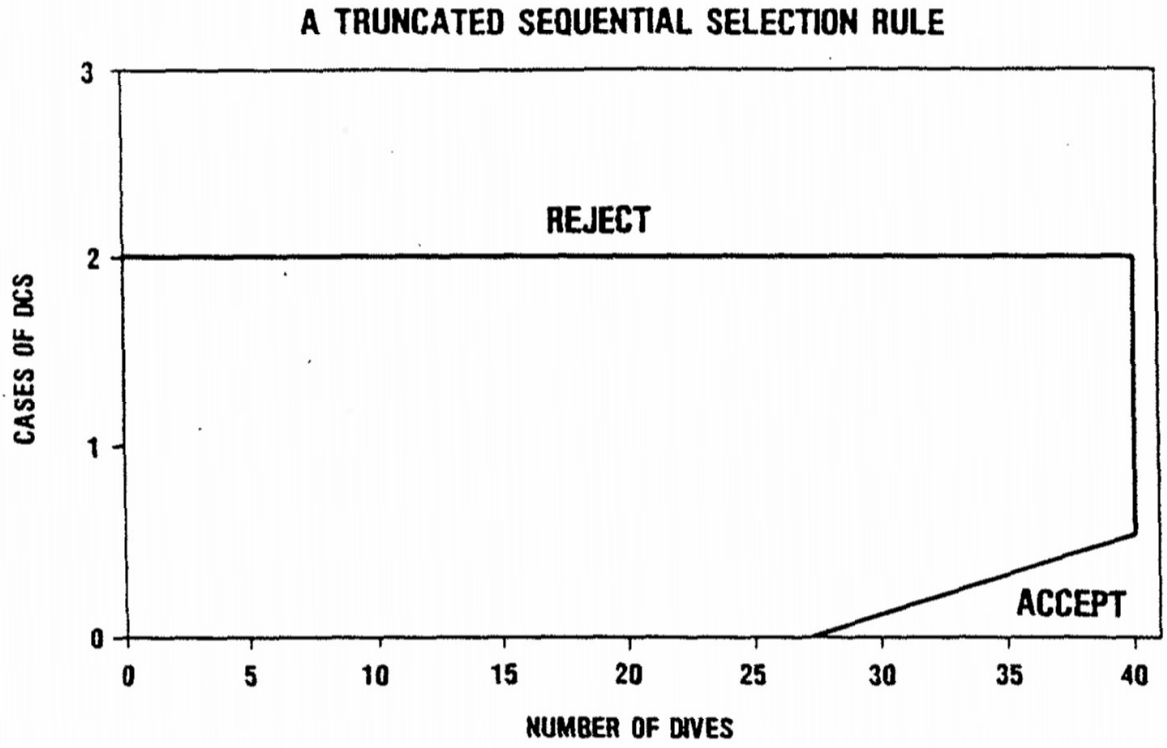

Fig. 6. A closed sequential design of a decompression trial. Rules follow those of Fig. 4, except the trial is stopped after the 40th dive even if *upper* and *lower lines* have not been crossed.

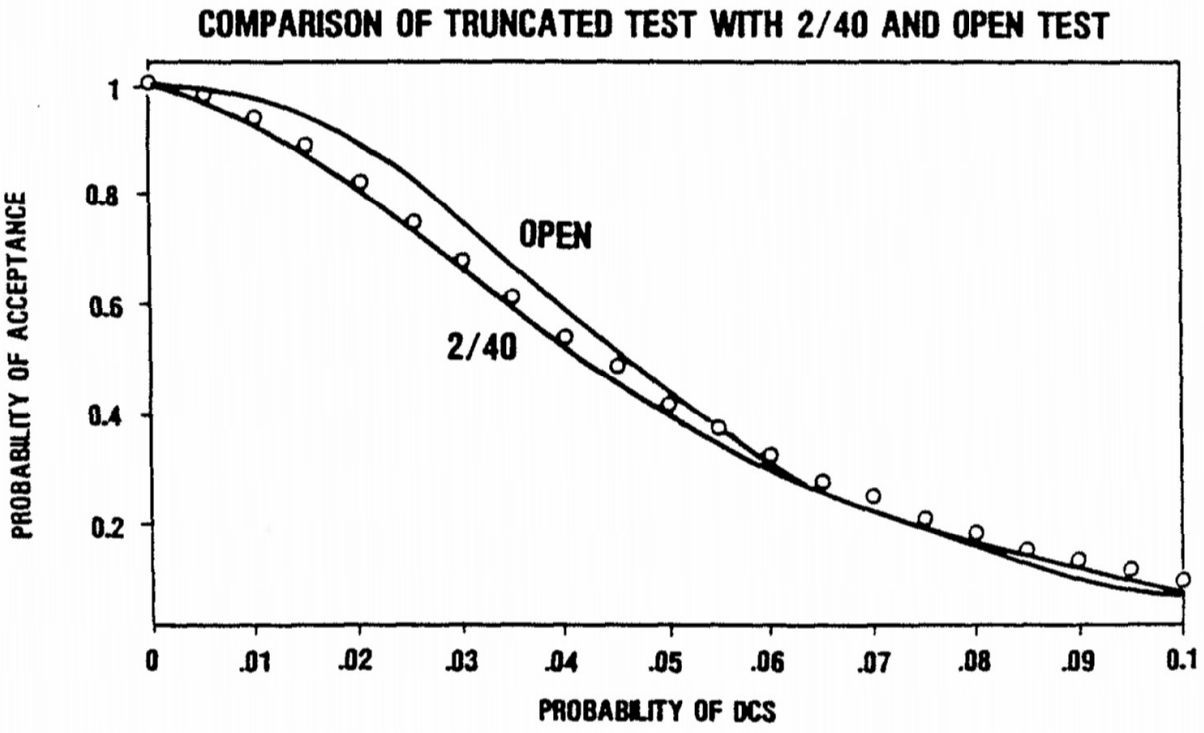

Fig. 7. Comparison of the probability of accepting a decompression schedule according to the truncated sequential rule of Fig. 6 *(dots)* with the fixed 2/40 rule and with the *open* sequential design of Fig. 4.

Models of decompression

Another approach is to assess decompression schedules within the framework of a mathematical model. The incidence of DCS is represented as a function of the time-depth history of the dive and some reasonably small set of unknown parameters.

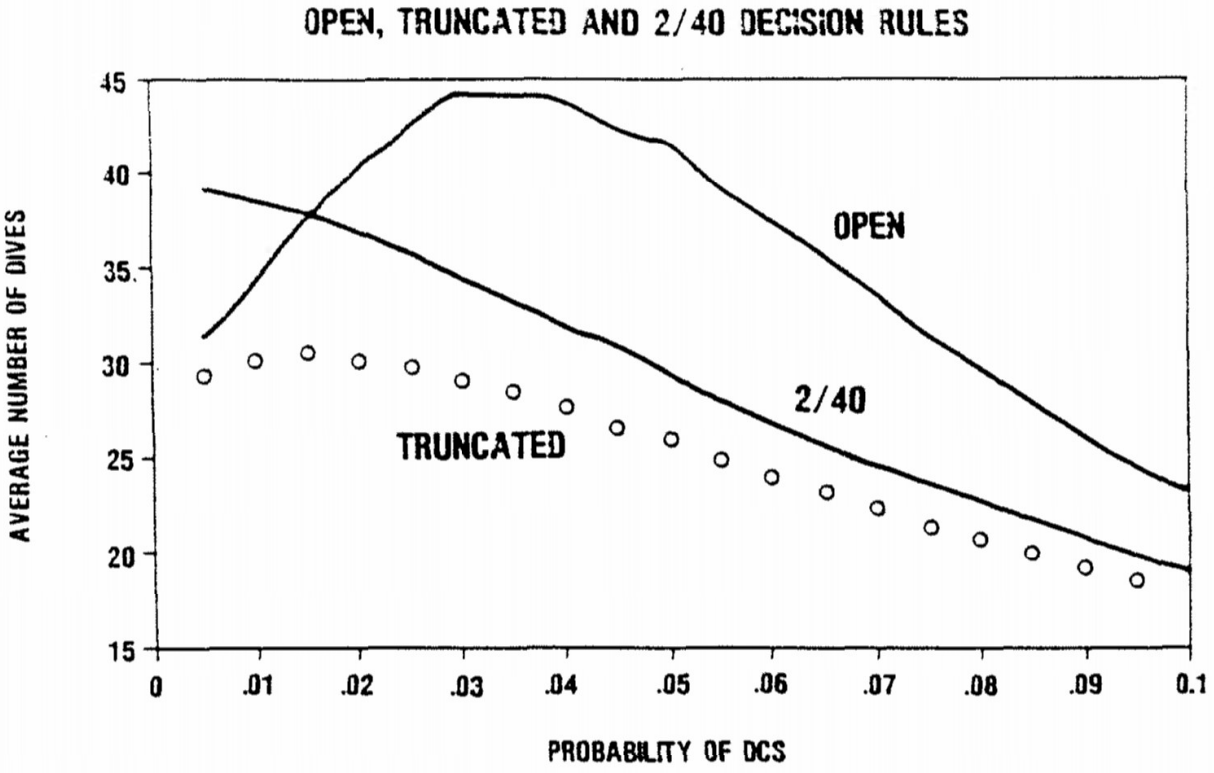

Fig. 8. Expected average number of repeated trials against the underlying "true" probability of DCS of a decompression schedule. The *dots* follow from the truncated design of Fig. 6. The other *lines* correspond to the closed sequential trial of Fig. 4 and the 2/40 rule.

Then the experience of many dives can be pooled to give estimates of fewer parameters than if we treat each schedule as a separate problem. If a sufficiently general model were developed, it could be used to generate tables or to monitor decompression procedures with complicated diving profiles. Testing would then consist largely in refinement of the parameter estimates and evaluation of the need for improving the model. No doubt the same data sets could, and would, be used to compare different models based on competing theories of decompression sickness.

One exceedingly simple empirical model that appears promising describes the probability, P, that DCS has occurred before time, t, as

$$P = 1 - \exp(-c_o \int^t [P_T - P_a]/P_a \, dx) \qquad (5)$$

P_T and P_a represent tissue inert gas pressure and ambient pressure, respectively, and c is a proportionality constant. The model is a simple way to assert that both depth and time increase the probability of DCS, but it invokes no other mechanisms. Both are estimated from data using the principle of maximum likelihood (4, 5). We have calculated P_T using a single exponential model with a time constant that must also be estimated. For example, in comparing the model to the 568 trial dives used to achieve the current United States Navy standard air tables (6), our best estimates of C range from 0.5 to $5 \times 10^{-3} \cdot \text{min}^{-1}$ (95% confidence region). The uncertainty in the half time places it between 20 and 460 min with a mean of 240 min. Despite large confidence regions for the parameters, the model does better in predicting bends than using a single DCS incidence ($P < 0.001$). The ability to separate a number of different dives according to their DCS incidence is shown in Fig. 9. The model was used to group

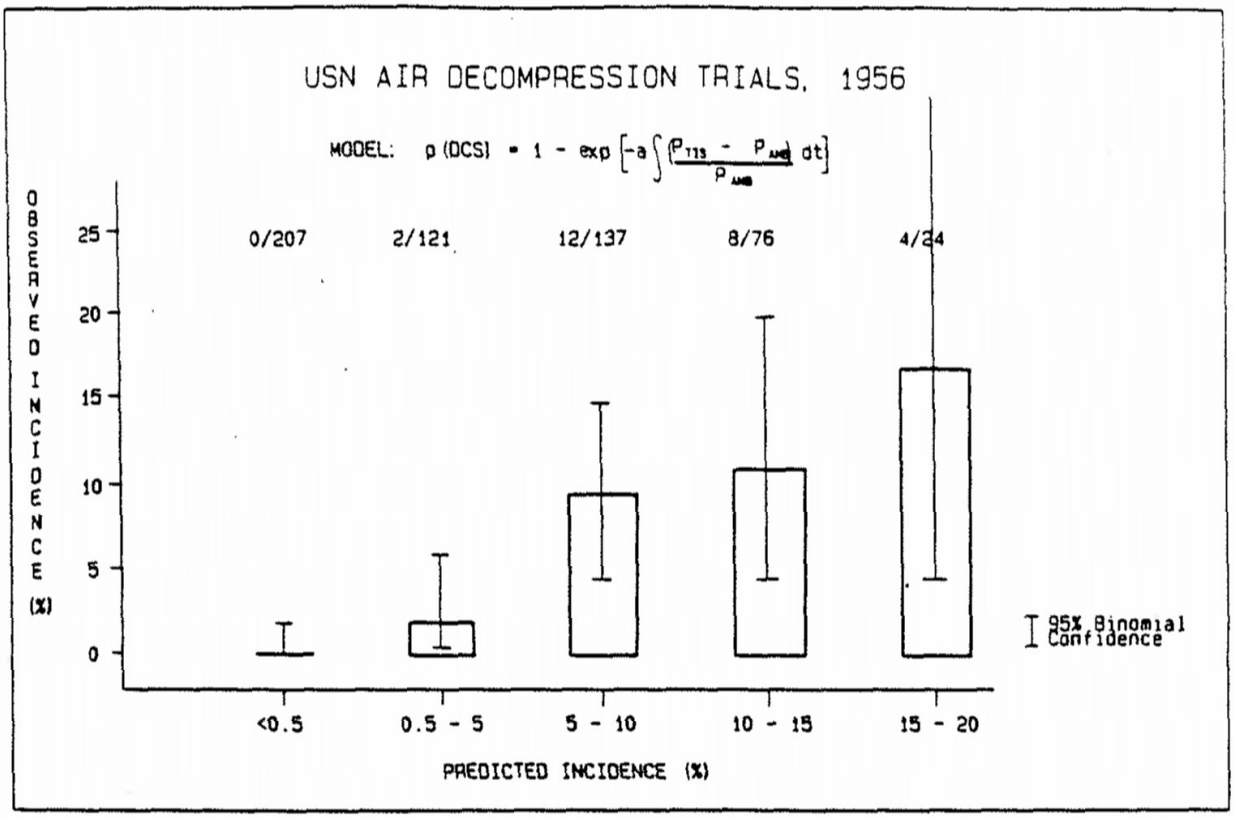

Fig. 9. Observed vs. model prediction of DCS in a large series of decompression trials with different schedules. The model was fit by maximum likelihood to a trial of 565 dives with 88 schedules that resulted in 26 cases of DCS.

the divers into 5 categories from low risk to high risk dives and the actual frequencies were then calcualted for each of these groupings.

DISCUSSION

Decompression tables, regardless of their basis of calculation, have always been tested to some extent before being accepted or rejected. The rejection rules have ranged from 1 case in 4 trials (6) to 2 or 3 cases in 30 trials (7), but the consequences of such rules have not been carefully examined. Trials are expensive in financial and human terms, so an efficient design of the trial is advantageous. To our knowledge, sequential designs have never been actually used, although mention of them appears in a recent British report (8).

Untruncated sequential designs are probably impractical to consider in assessing diving schedules. On the other hand, no one wishes to continue testing tables that result in a high proportion of bends in early trials, even though one may have planned a large number of dives. A sensible alternative is to select a strategy for rejecting tables based on any combination of past practices and intuitive insights agreeable to the participants. If a statistician is consulted, formulas for sequential designs might also be used. The performance of such strategies may then be assessed, as we have done, by using a random number generator to simulate the risk of decompression. Such assessments should take into account the fact that a dive is usually made with several divers rather than one at a time, as has been assumed in our calculations.

This research was supported by the Naval Medical Research and Development Command, Research Work Unit Nos. M0099901A.0002 and M0099901A.0005. The opinions and assertions contained herein are the private ones of the writers and are not to be construed as official or as reflecting the views of the Navy Department or the naval service at large.—*Manuscript received for publication July 1984; revision received November 1984.*

REFERENCES

1. Gray JS, Wigodsky HS, Magland RL, Green HL. Studies on altitude decompression sickness. IV. Attempts to avoid decompression sickness by selection of resistant personnel. J Aviat Med 1947; 34:88–95.
2. Feller W. An introduction to probability theory and its applications, vol 1, 2nd ed. New York: John Wiley & Sons, 1950.
3. Wald A. Sequential analysis. New York: John Wiley & Sons, 1947.
4. Kendall MG, Stuart A. The advanced theory of statistics, vol 1, 2nd ed. New York: Hafner Publishing Co., 1963.
5. Weathersby PK, Homer LD, Flynn ET. On the likelihood of decompression sickness. J Appl Physiol 1984; 57:815–825.
6. Des Granges M. Standard air decompression table. Washington, DC: U.S. Navy Experimental Diving Unit Rep 5-57, 1957.
7. Thalmann ED, Buckingham IP, Spaur WH. Testing of decompression algorithms for use in the U.S. Navy underwater decompression computer. Phase I. Panama City, FL: U.S. Navy Experimental Diving Unit Rep 11-80, 1980.
8. Shields TG. Retrial at sea of 70 and 80 metre 15 minute trimix decompression tables. AMTE(E) R83-401. London: HMSO, 1982.