A simple probabilistic model for standard air dives that is focused on total decompression time.

H. D. VAN LIEW, E. T. FLYNN

Navy Experimental Diving Unit, Panama City, FL 32407

Van Liew, Flynn ET. A simple probabilistic model for standard air dives that is focused on total decompression time. Undersea Hyperb Med 2005; 32(4):199-213. A statistical fit of an algorithm to "calibration data" gives parameter values for a "probabilistic decompression model." Our objective is to prepare a simple model that will estimate risk of decompression sickness (DCS) in air dives. We develop a logistic regression model using calibration data from carefully controlled experimental dives recorded in the U.S. Navy Decompression Database. We exclude saturation dives, which can have very long decompression times. For most depths, our model's prescriptions for 2% probability of DCS avoid the experimental DCS cases without mandating excessive time at decompression stops. Our model indicates that the long decompression times prescribed by some previous probabilistic models are not necessary. Our model cannot be used operationally because it cannot calculate depths and times at decompression stops; however, there is general concurrence between our model and prescriptions of a deterministic model known as the VVal-18 Algorithm; this supports the adoption of the VVal-18 Algorithm for operational use on decompression dives.

INTRODUCTION

A decompression table tells a diver how to minimize the chance of decompression sickness (DCS). For short dives, direct ascent to the surface is possible. For longer dives, the diver must spend times prescribed by the table at decompression stops that are part way toward the surface. The sum of times at decompression stops plus the time required to travel to the surface is the "total decompression time" (TDT).

Decompression tables can be generated from "probabilistic" or "deterministic" models. Probabilistic models apply statistical techniques to information about whether or not a certain dive profile resulted in DCS (1). Probabilistic models imply that the response to decompression stress is graded, so that even ostensibly safe dives carry a small but finite risk, in contrast to "deterministic models," which are formulated from intuition and assumptions about risk of DCS and imply that divers are either safe or unsafe. To produce a probabilistic model, assumptions about risk are incorporated into an algorithm; parameter values of the algorithm are estimated by statistical analysis of the dive-outcome data, sometimes called "calibration" or "training" data.

Standard air dives can be defined as dives to depths of 25 to 190 feet of sea water, gauge (fswg; 1 fsw = 3.063 kPa; 33.08 fswg = 2 atmospheres absolute), inclusive, with bottom times of 720 min or less, where bottom time is defined as the elapsed time from leaving the surface to leaving the bottom depth. A recently published probabilistic model (2), which we will call the "NMRI'98 Model" (Model 2), evaluated at 2.2% probability of DCS *(Pdcs),* prescribes TDTs that are far longer than TDTs of experimental standard air dives abstracted from carefully documented U.S. Navy dive trials (3). The long TDTs also conflict with a deterministic model, the "VVal-18 Algorithm"

(4,5). The NMRI'98 calibration data includes saturation dives (dives longer than a day), which can have TDTs that are substantially longer than TDTs for standard air dives. An earlier probabilistic model, which we will call the "NMRI'93 Model," also used saturation dives and also prescribes long times at decompression stops for reasonable DCS risks around 2% (6).

The objectives for the present work were: a) to test whether or not a probabilistic model that is based only on standard air dives gives TDTs for standard air diving that are as long as those found with previous probabilistic models (2,6), b) to test the hypothesis that including saturation dives may cause a probabilistic decompression model to prescribe excessively long TDTs, c) to assess whether the VVal-18 Algorithm (4,5) gives appropriate decompression prescriptions, and d) to develop a convenient way to estimate DCS risk for air dives.

METHODS

Calibration dataset.

The U.S. Navy Decompression Database (7,8) consists of a number of computer files, each representing a particular decompression study. We used single-level, non-repetitive, nitrogen-oxygen experimental dives documented in the U.S. Navy Decompression Database to generate a dataset for calibration of our "StandAir" Model (see Table 1; the date specifies the particular version of the file that we used and the Cases Pred and % Pred columns preview predictions from the statistical fit covered in the Results section). Each of the 19 source files contains a series of sequential entries that provide information about persons who followed a particular dive profile. Each profile entry bears a summary heading that describes the profile and its outcome, followed by a

TABLE 1. STANDAIR CALIBRATION DATASET: DETAILS OF SOURCE FILES

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Source file	Date	Entries	Person -dives	Cases Obs	Cases Pred	% Obs	% Pred
1	DC4D	10/9/97	247	775	19	17.4	2.5%	2.2%
2	DC4W	12/21/93	141	240	8	6.7	3.3%	2.8%
3	EDU1157	9/23/97	27	46	15	17.7	32.6%	38.5%
4	EDU1351NL	12/3/96	43	143	2	3.9	1.4%	2.7%
5	EDU159AVL*	9/30/97	5	11	5	1.1	45.5%	10.0%
6	EDU545	11/20/97	42	94	18	10.8	19.1%	11.5%
7	EDU557	5/29/97	135	568	27	51.6	4.8%	9.1%
8	EDU849LT2	5/5/97	74	141	26	15.8	18.4%	11.2%
9	EDU849S2	6/27/97	35	60	13	13	21.7%	21.7%
10	EDU885A	12/20/93	82	483	30	24.1	6.2%	5.0%
11	EDUAS45	1/15/98	10	14	3	2.3	21.4%	16.4%
12	NMR8697	1/29/91	229	477	11	13.9	2.3%	2.9%
13	NMR97NOD	8/19/97	9	103	3	2.6	2.9%	2.5%
14	NMRNSW	1/29/91	43	86	5	4.5	5.8%	5.2%
15	NSM6HR	12/20/93	19	57	3	2.3	5.3%	4.0%
16	PASA	5/26/92	26	72	5	2.5	6.9%	3.5%
17	RNPL52BL	7/20/95	27	192	1	6.9	0.5%	3.6%
18	RNPL57L	7/21/95	50	50	9	3.2	18.0%	6.4%
19	RNPLX50	9/19/97	16	57	5	7.6	8.8%	13.3%
			—	—	—	—	———	———
	Totals		1,260	3,669	208	207.9 + 57.9**	5.67%***	5.67%****

* File EDU159AVL is a modification of file EDU159A based on a re-reading of the report.
** 95% confidence interval
*** = total CasesObs/total Person-dives
**** = total CasesPred/total Person-dives

series of lines showing the depth/time nodes of the profile.

To prepare our calibration dataset, we carefully studied the details of the time and depth profiles of each entry, using the heading of the entry as a guide. When the recorded information in a profile indicated that there was a small deviation from a square-wave exposure to depth, we made appropriate corrections so that the corrected depth/bottom time pattern corresponded approximately to square-wave behavior. All profiles of EDU885A needed correction for lags at 7 fswg at the beginning of the dives. For details of our methods for adjusting the data, see our other publication (3).

We consider "marginal" cases, which do not require treatment, to be non-cases. Of the 3,669 person-dives in the StandAir calibration dataset, 45% are no-stop dives, or nearly so, on the basis that their TDT is less than 1.5 times that expected for an ascent rate of 30 fsw/min; average depth for these is 92 fswg. To increase the size of the dataset, we augmented the air-breathing dives with 515 no-stop dives (11 DCS cases) for which the breathing gas at depth was a nitrogen-oxygen mixture with inspired oxygen fraction (FO2) different from that of air (0.21). For these we calculated an "equivalent air depth" (EAD) in fswg to replace the actual depth (D) in fswg:

$$1)\quad EAD = 1.266 \cdot (1 - FO_2) \cdot (D + 33) - 33$$

In the calibration dataset, bottom times are 5 to 720 min inclusive, average 98 min; depths (or equivalent air depths) are 28 to 300 fswg inclusive, average 118 fswg; and TDTs are 0.6 to 1,445 min inclusive, average 43 min. Many ascent rates are 60 fsw/min, whereas the rate currently prescribed for the U.S. Navy is 30 fsw/min (9). Figure 1 illustrates the distributions of variables. A statistical model is most reliable in regions that have the most data points; the reliability of the StandAir Model will be low for bottom times greater than 79 min and for dives with TDTs longer than 20 min. Fortunately the bulk of the calibration data is in the low-incidence region, which is of interest for operational diving: for 80% of the person-dives, incidence of DCS estimated by the StandAir Model is below 3%; for 63%, below 2%; and for 47%, below 1%.

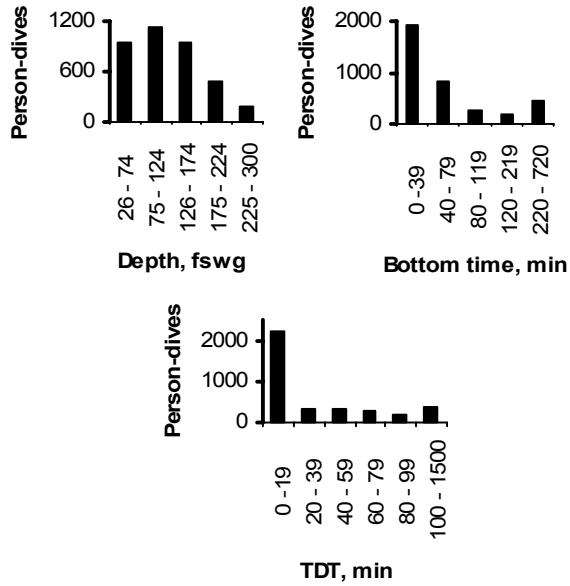

Fig. 1. Histograms of the calibration data for the StandAir Model.

Statistical analysis:

We use the logistic regression paradigm to characterize probability of DCS *(pdcs)*: *LOGIT* is defined as the logarithm of an odds ratio: *LOGIT+In(Pdcs(1-Pdcs)):*

$$2)\quad Pdcs = 1 / (1 - \exp(-LOGIT))$$

The NONLIN module of a commercial statistical program, SYSTAT, allows us to use a complex formular for the *LOGIT* term. Equation 3 is the particular *LOGIT* function we use for our StandAir Model. Symbol D represents dive depth in fswg; *T* is bottom time

in min; *TDT* is total decompression time in min; and *a, b, c, d, f,* and *g* are parameters estimated by the logistic regression process.

3) $LOGIT = a + b \cdot (D - c) \cdot [1 - exp(-D \cdot T^f)]/(TDT - g)$

From collective experience, we presume that for a given depth and bottom time, TDT is by far the most important determinant of DCS risk. The basis for Equation 3 is the obvious statement that for dives with depths and bottom times outside the region of "no-stop" diving, DCS can be avoided by time spent at decompression stops. In no-stop diving there is a hyperbolic relationship between dive depth and time that can be spent at the depth; this may mean that DCS can result from an excessive "dose" of depth multiplied by time ($D \cdot T$). We consider that TDT serves as an "antidote" to a dose of ($D \cdot T$) that would be intolerable for a no-stop dive. Thus, for a given ($D \cdot T$) product, probability of DCS is low if TDT is large. We are not able to make any statement about whether deep decompression stops are more or less effective than the relatively shallow stops used by the U.S. Navy because all our experimental data are shallow-stop dive profiles.

Figure 2 shows the TDT vs. bottom time traces at several depths for a decompression table generated by the deterministic VVal-18 Algorithm (4,5). All tables we examined in our earlier work (3) showed the same pattern: a fan of lines seems to radiate from a point on the zero bottom-time axis at a negative TDT.

We used the fan idea and a combination of insight with trial and error to devise Equation 3. A simple equation for the reciprocal of the slope of one of the projection lines in Figure 2 is $Y_1 = T/(TDT - g)$, where Y_1 is defined by the equation and *g* stands for the intersection of the line with the TDT axis at zero bottom time. The fan shape could be accommodated by

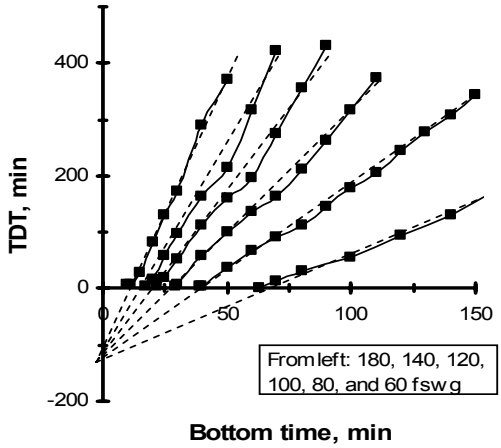

Fig 2. Bottom time/TDT combinations according to the VVal-18 Algorithm for several depths. Dashed projections lines are drawn by eye.

adding depth to the formula $Y_2 = D \cdot T/(TDT - g)$. We found empirically that subtracting a constant from depth improved the fit of the data: $Y_3 = (D - c) \cdot T/(TDT - g)$. To account for the expectation that TDT levels off as tissues become saturated with inert gas, we changed the bottom-time term: $Y_4 = (D - c) \cdot [1 - exp(-d \cdot T)]/(TDT - g)$, where *d* regulates the rate of leveling off. Finally, we raised bottom time to a power to allow the possibility that the leveling off may not correspond to a single exponential function: $Y_5 = (D - c) \cdot [1 - exp(-d \cdot T^f)]/(TDT - g)$. The Y_5 formulation gives Equation 3.

For maximum likelihood estimation of parameters, the NONLIN module of the SYSTAT program uses a *LOSS* function, an alternative to least squares of residuals. The log-likelihood (*LL*) is specified as the negative of the *LOSS* value. Equation 4 gives the *LOSS* statement for a single entry of dive-outcome data in the dataset; the *LOSS* function would be the sum of all *LOSS* statements for all entries of the dataset.

$$4) \quad LOSS = - DIVES \cdot [DCS \cdot ln (ESTIMATE) + (1 - DCS) \cdot ln (1 - ESTIMATE)]$$

In Equation 4, *DIVES* is the number of person-dives in the group that makes up the entry, *DIVES* multiplied by *DCS* is the number of subjects in the group having DCS, *ln* is natural logarithm, *ESTIMATE* is the estimated *Pdcs*, and *DIVES* multiplied by *(1 – DCS)* is the number of subjects in the group without DCS. In our case, the *DCS* variable is either 1 or 0 because divers who dived together and who did not contract DCS (up to 86 divers in a profile) are listed together in the same profile entry in the database (DCS = 0), but divers who contracted DCS (DCS = 1) are listed separately. Parameter estimation minimizes the positive *LOSS* function, which is the same as maximizing the negative *LL,* by a Quasi-Newton technique that uses numeric estimates of the first and second derivatives of the *LOSS* function to seek a minimum. In addition to estimating parameters, the NONLIN module computes asymptotic standard errors (*ASE*) and the asymptotic correlation matrix by estimating the Hessian matrix after iterations have stopped.

Once the parameters of Equation 3 are found by statistical analysis, *Pdcs* can be tabulated for *D, T,* and *TDT* of each entry in the calibration dataset. To generate a table that relates *TDT* to depth and bottom time at constant *Pdcs*, we find *TDT* for particular combinations of *D*, *T*, and *Pdcs* values; a computer program calculates *Pdcs* over ranges of *D, T*, and *TDT* by solving Equation 3 for a run of depths and, within each depth, for a run of bottom times, and within each bottom time, for a run of *TDTs*. The program writes the *D, T, TDT*, *Pdcs*, and confidence interval to a file when the *Pdcs* just exceeds the predetermined target; the program then continues to the next bottom time and eventually to the next depth.

Three kinds of 95% confidence intervals are of interest: a) The binomial theorem calculates confidence intervals for the "true" incidence of DCS in the entire population from a limited sample of person-dives. b) The SYSTAT program calculates confidence intervals for estimated parameters. c) We calculate confidence intervals for the probability estimates of the model as outlined by Ku (10) from the parameters, *ASE* values, and the asymptotic correlation matrix.

To evaluate our completed model, we use the grouped dive-outcome dataset developed in another communication (3); depths of experimental dive trials are rounded to the nearest 10 fswg and bottom times are rounded to the nearest 5 min. Traces of a table's instructions on bottom-time/TDT plots are compared with the locations of circles (DCS-free profiles) and triangles (profiles that gave rise to DCS).

Combination Model.

We test the hypothesis that in calibrating a model for standard air dives, inclusion of saturation dives causes the model to prescribe unnecessarily long decompression times. For this test, we deliberately add 240 air-breathing saturation dives to the calibration dataset used with the StandAir Model to produce a "Combination" Model. Among the saturation dives, only 78 are for dives with long TDTs, from 1,200 to 2,162 min with average of 1,656 min. The rest of the saturation dives are for shallower depths that allow much more rapid ascents, with average TDT of 5.7 min. Thus only a little more than 2% of the dives in the Combination dataset have long TDTs. Sources of the saturation dives are ASATFR85, ASATNMR, ASATNSM, EDU545, and NMR9209 (7,8). Methods for preparing the data are the same as given above for our StandAir Model and we use the StandAir Model's *LOGIT* function, Equation 3.

A Likelihood Ratio test (*LR* test) allows

a decision about whether groups of data, when combined, are described as well as when two separate models describe the respective groups (11-13). For use with the *LR* test, we developed a third model, called the OnlySat Model, using Equation 3 calibrated only with the saturation dives from the Combination dataset; LL was –64.29. The parameters of the OnlySat Model are not shown here. The *LR* test indicates that the separate models, OnlySat and StandAir, provide a better fit to the data than the Combination Model does. Nevertheless, we continued development of the Combination Model to test our hypothesis about using saturation dives to calibrate a model for standard air dives. The conclusion from the *LR* test applies only to our models. To find out whether combination of non-saturation and saturation dives is advisable for other models, comparable *LR* tests should be carried out. Indeed, Hays and coworkers (11) reported that for certain models, *LR* tests showed that saturation and non-saturation dive data should not be combined, but in the same publication the authors did combine the two kinds of data for another group of models.

RESULTS

Table 2 lists the parameters estimated for the StandAir and Combination models by fitting Equation 3 to their respective calibration datasets; the correlation matrix for the StandAir Model is at the bottom of the table. For both models, the differences between the *LL* value and the *LL* for the null model, in which *Pdcs* equals the DCS incidence in the dataset, have high significance by a likelihood ratio test. The high ratios of Param/ASE indicate that the parameters are tightly estimated. The correlation matrix for the StandAir Model shows high correlation between parameters *c* and *d*. As seen in Table 1, the StandAir Model predicts the 208 cases in the StandAir calibration data. We have not discovered any reason for the disagreement between predictions and observations for some of the files (see

TABLE 2. PARAMETERS FOR THE MODELS

StandAir Model *LL* = -689.144, Null Model *LL* = -798.98, incidence = 5.67%

Parameter	Estimate	*ASE**	Param/*ASE*	Wald 95% Confidence Interval
a	-6.022169	0.277405	-21.7	-6.57 – -5.48
b	86.596315	18.887942	4.58	49.5 – 123.6
c	25.091718	2.038656	12.31	21.09 – 29.09
d	0.002929	0.000832	3.52	0.0013 – 0.0046
f	0.918547	0.041705	22.0	0.837 – 1.000
g	-170.304442	21.500126	-7.92	-212 – -128

Combination Model *LL* = -763.395, Null Model *LL* = -880.21, incidence = 5.94%

Parameter	Estimate	*ASE**	Param/*ASE*	Wald 95% Confidence Interval
a	-5.792136	0.234149	-24.8	-6.25 – -5.33
b	134.799017	15.251709	8.80	105 – 165
c	18.135384	1.477299	12.15	15.2 – 21.0
d	0.004029	0.000732	5.42	0.00256 – 0.0055
f	0.782505	0.024983	30.57	0.733 – 0.832
g	-284.304696	31.050956	-8.94	-345 – -223

Correlation matrix, StandAir Model

	a	*b*	*c*	*d*	*f*	*g*
a	1.000					
b	0.132158	1.000				
c	0.427921	0.417632	1.000			
d	-0.569450	-0.541532	-0.92333	1.000		
f	0.330143	-0.453320	0.503751	-0.458797	1.000	
g	0.074836	-0.075632	0.739574	-0.580370	0.590638	1.000

*ASE = asymptotic standard error

EDU159AVL, EDU557, EDU849LT2, and RNPL57L). To examine the possibility that data from recent dive trials might differ from earlier work, we sorted the calibration data into two groups, "new" and "old;" we found that the StandAir Model fit both sets equally well.

There are many sources of variability in the calibration dataset, such as changes in dive practice over the years, differences between the dives in different test settings (e.g., wet vs. dry dives, exercising vs. sedentary subjects, warm vs. cool temperature, acclimatized vs. non-acclimatized divers, etc), and differences in DCS criteria with different investigators. When estimated *Pdcs* is 2%, 95% confidence intervals for the StandAir Model show that the true *Pdcs* could actually be somewhere between 1.1 and 2.9%. For all *Pdcs* values of interest, the average width of the confidence intervals for the StandAir Model is about 60% as large as the *Pdcs*. Note that the computation of confidence intervals by the Ku method (10) does not account for number of data points, so for practical purposes, the uncertainties in data-poor regions, such as regions having long bottom times and long TDTs, are probably greater than the estimates given by the Ku calculation.

Predictions versus observations.

Figure 3 compares estimates from the StandAir Model with observed outcomes. To produce the graphs, we sorted the entire dataset by a particular variable and then divided the sorted set into 7 bins containing approximately equal numbers of person-dives. For optimizing parameters, it would be preferable to have the DCS cases spread equitably with regard to depth, bottom time, and TDT, but Figure 3 shows that the bins have unequal numbers of DCS cases.

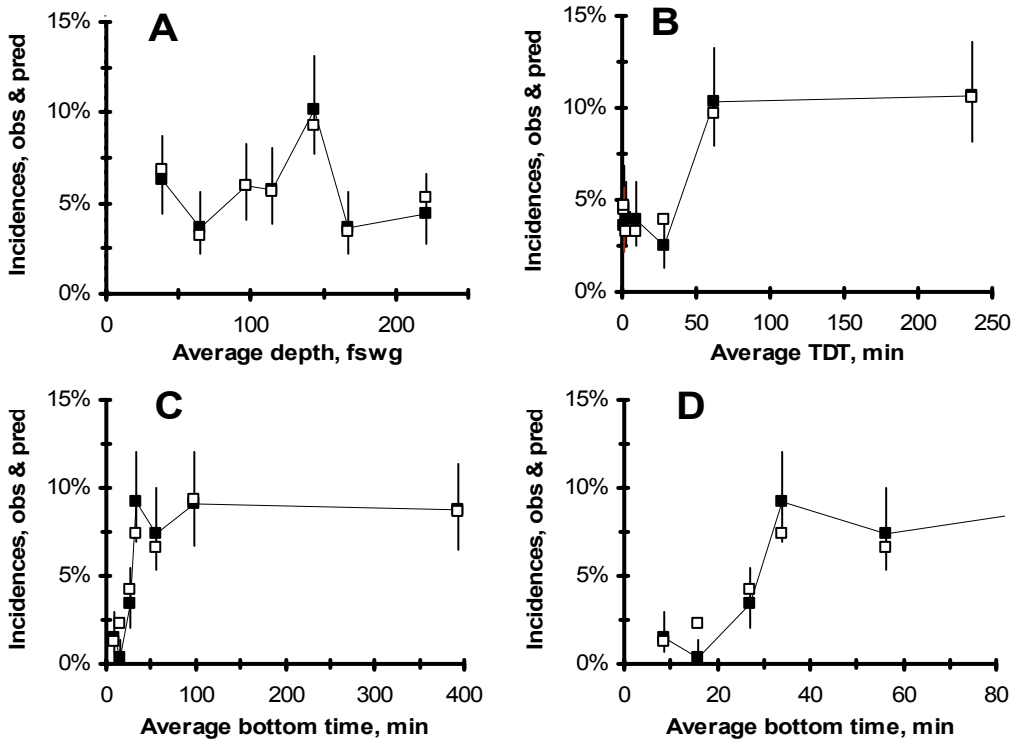

Fig. 3. Incidences as functions of a variable for subdivisions of the data containing equal numbers of person-dives. Filled rectangles = observations, with trace to lead the eye; open rectangles = predictions; some of the symbols superimpose on each other. Vertical line segments show 95% confidence intervals, calculated by the binomial theorem, around observed incidences. A: incidences versus depth. B: same format, but for TDT. C: same format, for bottom time. D: short bottom times.

The incidence is high for long TDTs (Figure 3B) and long bottom times (Figure 3C). In all cases, predicted values follow the observed values closely. The enlargement of the short bottom-time range in Figure 3D shows the only case in which the predicted incidence is outside the confidence interval for the observed incidence. Chi-square values for the plots in Figure 3 are 3.2 for depth, 11.0 for bottom time, and 8.0 for TDT. The chance value is 12.6 for 95% confidence with 6 degrees of freedom, so none of the differences are greater than those expected by chance.

To obtain the values for the X-axis in Figure 4A, we calculated *Pdcs* for each row in the StandAir calibration data, sorted according to estimated *Pdcs*, and then divided the sorted data into bins having approximately the same number of person-dives. We then added up the numbers of DCS cases predicted by the StandAir Model for the dives in each bin. We obtained values for the Y-axis by counting the number of DCS cases observed among the dives in each bin.

The data points in Figure 4A approximate a straight line but are slightly above the dotted identity line for most of its extent and below it at the extreme left. A least-squares trend line (see equation) is essentially superimposed on the identity line. Except for the point farthest to the right, the horizontal lines for uncertainty in predicted incidence are shorter than the vertical lines for uncertainty in observed incidence. The chi-square value for the DCS cases represented by the points in Figure 4A is 7.5, and the chance value is 9.5 for 95% confidence and 4 degrees of freedom, so the difference between observed and predicted DCS cases is not greater than expected by chance. The chi-square value for the cases represented by the three data points in Figure 4B is 7.3, and the chance value is 6.0 for 95% confidence and 2 degrees of freedom.

In the operationally important low-incidence region below 3% *Pdcs*, the model predicts higher incidence than is actually observed (Figure 4B). The heavy curve shows an equation for the three points that was

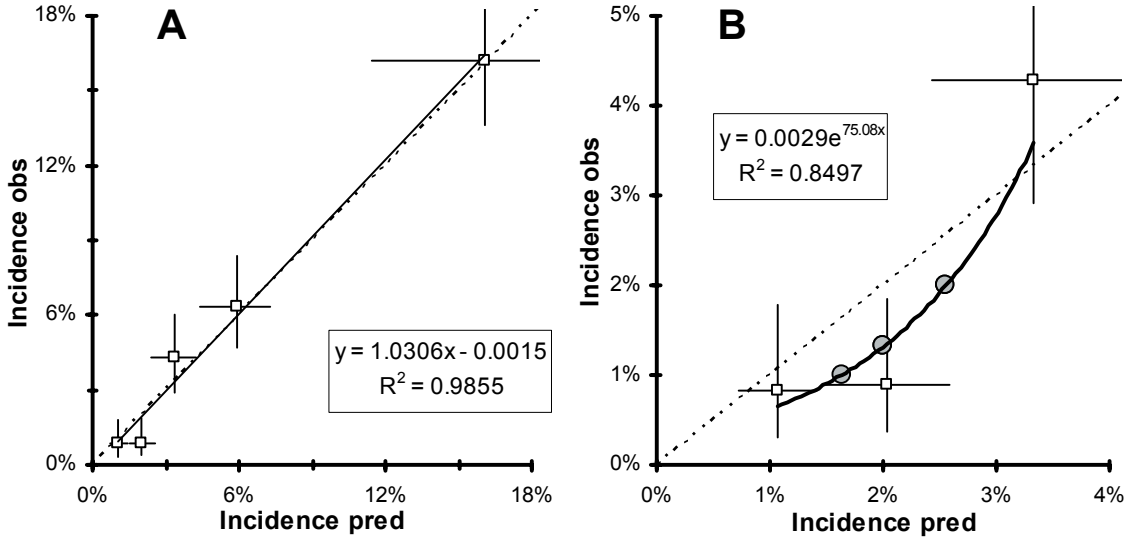

Fig. 4. A: observed incidence of DCS is plotted against incidence predicted by the StandAir Model. Dotted 45_o line = identity; solid line is the least-squares line through all the points; R_2 and the equation for the line are in the box; vertical and horizontal lines show 95% confidence intervals around observed and predicted incidences. B: low-incidence region of panel A; heavy curve is a least-squares trend curve through three data points (see box). See text for explanation of gray circles.

obtained by a least-squares process; the equation of the curve (in the box of Figure 4B) gives observed incidence as a function of predicted incidence. It can be seen that at observed incidence of 1.0%, the predicted incidence is 1.6% (lower gray circle); at observed incidence of 1.3%, predicted incidence is 2.0% (middle gray circle); and at observed incidence of 2.0%, the predicted incidence is 2.6% (upper gray circle). The discrepancy between prediction and observation is less for 3% incidence. For operational dives with desired incidence around 1% or 2%, our StandAir Model apparently overestimates the risk by about 0.6%. In other work (14), we describe a method for adjusting or "fine tuning" probabilistic models in the low-incidence region to make them as practical as possible for operational use.

Prediction versus observation for grouped test dives.

In Figure 5, symbols are from our previous work (3); they show grouped data from trials that resulted in DCS incidence at depths between 145 and 154 fswg, rounded to 150 fswg; bottom times are rounded to the nearest 5 min. Black triangle symbols represent a group for which it can be said, with 95% confidence, that "true" incidence is 5% or higher. For gray triangles, it can be said with 95% confidence that the DCS incidence in the group is 2% or higher but it cannot be said that incidence is higher than 5%. With white triangles, DCS occurred in the group, but confidence statements about 2% or 5% incidence are not possible, either because of insufficient numbers of dives or because of innocuous dives. Small circles are for groups in which DCS did not occur.

According to Figure 5, the StandAir Model instructs divers to spend enough time at decompression stops to avoid DCS; all the table traces avoid all the triangles with one exception, a white triangle at 15-min bottom time. In Figure 5A, the trace for the NMRI'98 Model lies far above the triangles except when TDT is very low, above the traces for the Val-18 table, and above the trace for 2% *Pdcs* by the StandAir Model. According to the trend curve in Figure 4B, the dotted *Pdcs* trace for 1.6% in Figure 5B matches observed incidence of 1.0%, the dotted *Pdcs* trace for 2.6% matches observed incidence of 2.0%, and the *Pdcs* trace for 3.0% approximately matches the observed incidence. Thus the trend curve of Figure

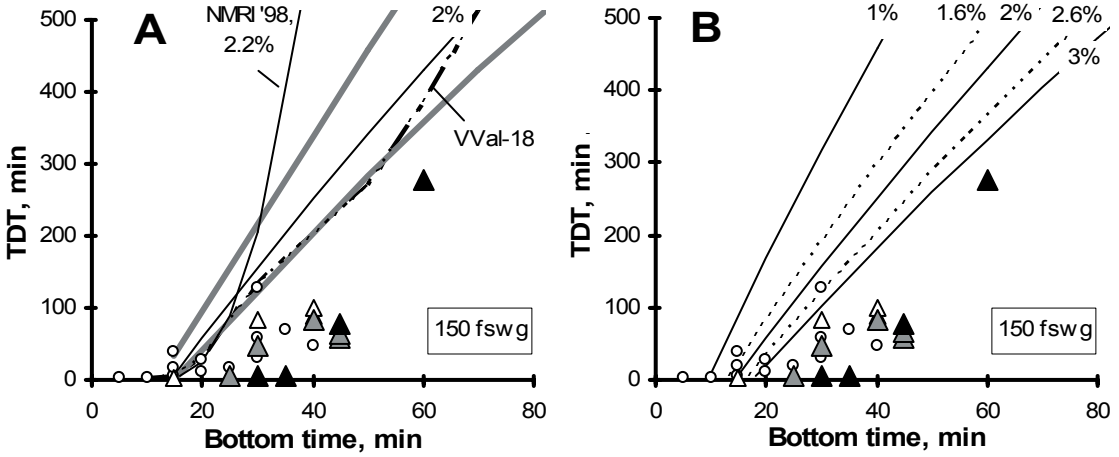

Fig. 5. Table traces and dive outcomes; see text for explanation of symbols. A: trace for 2% Pdcs according to the StandAir Model with heavy gray traces showing 95% confidence interval for the 2% Pdcs trace; labeled traces show prescriptions by the NMRI'98 and VVal-18 Models. B: StandAir Model traces for five Pdcs values.

4B suggests that the TDTs prescribed by the isopleths for 1% and 2% *Pdcs* in Figure 5 may be longer than necessary to give the stated risk of DCS; the same can be said for Figure 6.

The pattern of the relationship between DCS cases and *Pdcs* isopleths seen at 150 fswg in Figure 5 is repeated over the whole range of depths for standard air diving. Figure 6 shows six of the possible 16 graphs for depths from 40 to 190 fswg at 10 fswg intervals. In Figure 6 and in graphs for other depths that have few DCS cases (not shown), the positions of the table traces and the grouped dive-outcome data invite several contentions. a) The traces for 1, 2, and 3% *Pdcs* span a wide range of TDT for any given bottom time and the VVal-18 traces tend to be close to the StandAir traces for 2% and 3% *Pdcs*. b) Several of the graphs have regions with circles to the left of the 2% *Pdcs* trace and triangles to the right, but these regions occur only at low TDTs (see 100 fswg in Figure 6 for example). For a satisfying evaluation, we would want circles to the left and triangles to the right of the table traces at high as well as low TDTs in Figure 6, but there are no instances of such a configuration. c) When the dive-outcome triangles are in regions with relatively high TDT, they are not far below the table traces for 3% *Pdcs* (all panels in Figure 6). This suggests that the prescribed TDTs for these dives are not too long. d) Triangles above the table traces would strongly suggest that the prescribed TDTs are too short. Except for profiles with very low TDTs, no dive-outcome points, neither circles nor triangles, are above the table traces. Therefore, this evaluation does not yield a definite answer to the question of whether the StandAir Model prescribes TDTs that are too short. e) A gray or black triangle near a table trace indicates a serious risk of DCS, whereas white triangles could be due to chance. In Figure 6, only white triangles are near the table traces in the graphs for 60 and 80 fswg. A few black and gray triangles are near the bottom ends of the table traces; we consider dives with very low TDTs and no-stop dives in another report (15).

For an additional evaluation, we calculated the *Pdcs* by the StandAir Model for every point in the grouped data that make

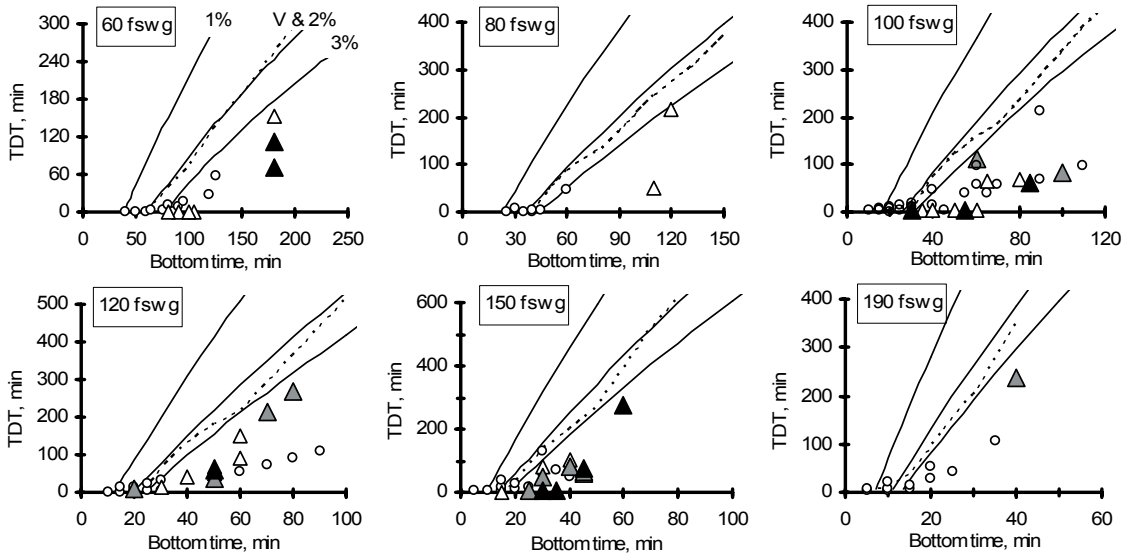

Fig. 6. Selected depths that have DCS cases near table traces for the StandAir Model; dive-outcome symbols are as in Figure 5. Solid traces are for Pdcs of 1% (left), 2% (middle), and 3% (right). Dotted traces are for VVal-18 (V).

up the evaluation dataset (3), and then plotted the difference between observed and expected incidence against each of the variables. These graphs (not shown here) did not reveal any areas of particularly bad fit with regard to depth, bottom time, or TDT.

DISCUSSION

Skepticism and probabilistic models.

Underlying every decompression model is a judgment or belief about how DCS is related to the variables — depth, bottom time, TDT, and any others that the model entails. With probabilistic modeling, this belief is embodied in the algorithm used to fit the calibration data. Previous probabilistic decompression models used elements of traditional knowledge about the etiology of DCS (1,2,6,16-19). This includes the ideas that the body is composed of a number of "tissue compartments" that exchange inert gas at various rates, that the supersaturation which a particular tissue can tolerate without developing DCS is limited, and in some models, that diffusive exchange of gas bubbles is an important variable (16,17). Recent models use time of DCS onset to attempt to refine risk estimates more closely (18,19). Some variables — e.g., bubbles detected in venous blood and the time at which symptoms occur — concern the course of the DCS disease process rather than the cause of the process, so use of such variables in fitting equations may not help the model to distinguish between risky and safe dives.

There is good reason to be skeptical about the prescriptions of a probabilistic model. The framework of relationships between variables that is established by the statistical process is the essential element of a probabilistic model, but the framework may be flawed. In fitting all regions as well as possible, the fit may be poor in a region of particular interest, such as the operationally important region of low *Pdcs*. Regions that are particularly vulnerable to error are at the edges of the range of the model, where prescriptions tend to be extrapolations from the bulk of the calibration data. To become aware of problem areas, we advocate using observed data retrospectively, as in Figures 5 and 6.

The Combination Model.

Comparison of observed vs. predicted DCS incidence in the variables of the calibration dataset (analogous to Figure 3) and of observed incidence with *Pdcs* (analogous to Figure 4) for the Combination Model show that the differences for explanatory variables, or for the model as a whole, are greater than expected by chance, whereas for the StandAir Model the differences are no greater than expected by chance. Figure 7 illustrates how dose-response curves for the StandAir and Combination Models cross over each other. In mid-range of Figure 7A, the curve for the StandAir Model is steeper than the curve for the Combination Model, so the curves cross at about 46 min.

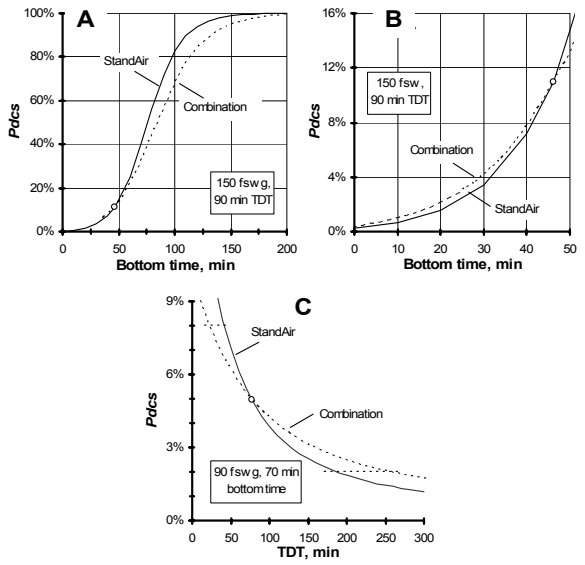

Fig. 7. Sample dose-response curves for the Combination and StandAir Models. Panels B and C are scaled to emphasize the low-risk region. Circles show crossover points.

The upper parts of the curves are extrapolations beyond the bulk of the data; above 40% *Pdcs*, there are only 39 person-dives in the calibration datasets. In the low-risk region, between 0.5 and 3% *Pdcs*, the TDTs prescribed by the Combination Model are longer than TDTs by the StandAir Model, but the crossovers change the situation for hazardous dives. Thus in Figure 7C, the Combination Model prescribes longer TDT than the StandAir Model for 2% *Pdcs* (lower dotted horizontal line segment), but the upper dotted horizontal line segment shows that for 8% *Pdcs*, the Combination Model requires shorter TDT than the StandAir Model; TDT is less than half as long.

We conclude that our hypothesis about saturation dives is valid: a few saturation dives with long TDTs can degrade the fit of a model to be used for standard air dives. We infer that the saturation dives act as outlying points that bias the model. For the sample depths shown in Figure 8 and for all other depths in the range of standard air diving, the Combination Model prescribes longer TDTs than does the StandAir Model. For operational dives meant to incur reasonable risk, such as *Pdcs* of 1% or 2%, excessively long TDTs for ordinary non-saturation dives would be detrimental in terms of time and safety. Depths of operational dives are determined by the job requirement. Time spent at decompression stops would interfere with useful work and would prolong the diver's exposure to risks inherent in being underwater.

The long TDTs of our Combination Model lead us to suspect that inclusion of saturation dives in the calibration data is the major cause of the long TDTs of the NMRI'98 Model (2) and the NMRI'93 Model (6), but the problem may, of course, have other explanations. The issue could be clarified by re-estimating the parameters of the NMRI'98 and NMRI'93 Models with calibration datasets that exclude saturation dives.

As a general principle it would be wise to calibrate a model for a given situation with data that is limited to that particular situation, excluding dive trials that are outside the region of interest. However, exclusion of outlying data or other questionable data decreases the size of the calibration dataset, with associated loss of precision. To understand the relation between size of the dataset and precision, we generated a false Combination dataset by duplicating each row so the false dataset had

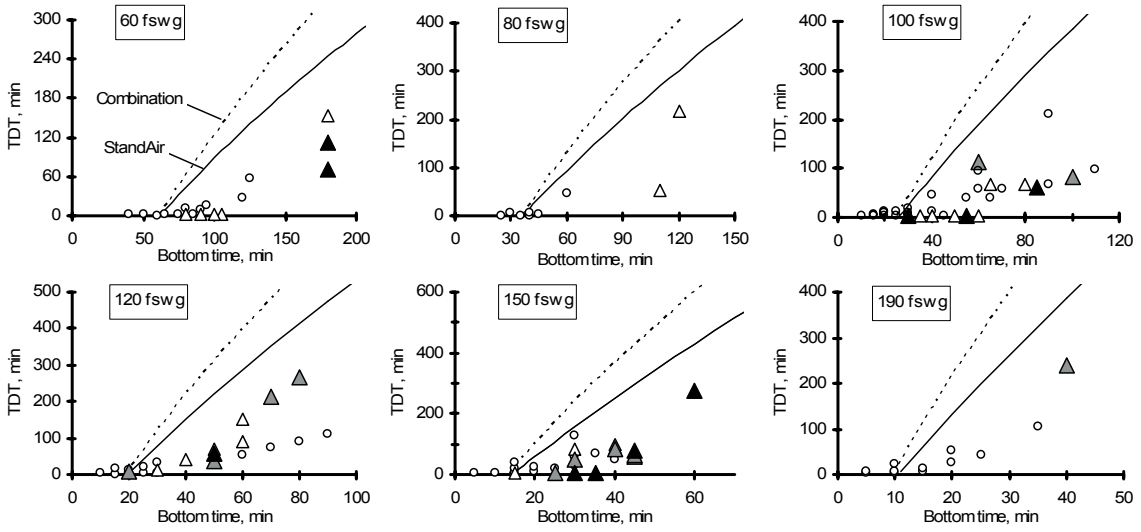

Fig. 8. Table traces for 2% Pdcs for the Combination Model (dashed) and the StandAir Model (solid); same depths as in Figure 6; dive-outcome symbols are as in Figures 5 and 6.

twice as many person-dives, but no change in balance. The false dataset yielded the same estimated parameters and correlation matrix as the original dataset. The negative *LL* value for the false dataset was double that found with the original set and the *ASE* and the confidence intervals of the parameters were 70% of their original values.

Operational diving.

Our graphs and results of chi-square analyses show that the StandAir Model is a good fit to the calibration dataset. In line with the Naval Sea Systems Command consensus that more than 2 cases of treatable DCS per 100 dives for non-emergency diving are highly undesirable (personal communication, C. A. Murray; 2000), we find no cause for concern about the StandAir Model's prescriptions for deep dives that have appreciable TDTs; the DCS cases for these are below the 2% table traces in Figure 6, evidence that the StandAir Model is effective in prescribing TDTs that minimize DCS. Apparently the long TDTs prescribed by NMRI'98 and similar models (2,6) are not warranted, but this conclusion is weakened by lack of dives in the high-TDT region, and by the large uncertainties in TDT for a certain *Pdcs*, as illustrated by confidence intervals in Figure 5A. In the absence of data in a depth/bottom time/TDT region, the model's estimates are extrapolations and interpolations from data-rich regions; thus, the StandAir Model's prescriptions for high TDT are projections from low-TDT regions.

For dives shallower than 60 fswg, the calibration dataset has 641 person-dives with 35 DCS cases, but average TDT for these is only 3.5 min. This means that the StandAir Model's TDT values for depths shallower than 60 fswg are extrapolations from deeper dives that have appreciable TDTs. For shallow dives, especially for 40 fswg, we therefore believe that the StandAir Model prescriptions for TDTs are questionable.

Conventional decompression tables generated both by deterministic models (4,5,9,20-22) and by probabilistic models (1,2,6,16-18) account for the inert gas partial pressures in a series of tissue compartments during depth/time maneuvers in a dive profile. A decompression stop is prescribed whenever a compartment is at risk of developing DCS. Our simple StandAir Model, generated from Equation 3, does not track tissue compartments during the dive, so it cannot prescribe decompression stop depths and times and is therefore not appropriate for operational use, although it is appropriate for estimation of risk. However, TDT values prescribed by the deterministic VVal-18 Algorithm and by our StandAir Model for 2% *Pdcs* are similar for most depths (see Figure 6), an indication that the VVal-18 Algorithm is acceptable for operational use for most depths. The TDTs for the VVal-18 table tend to be shorter than the TDTs that give 2% *Pdcs* by the StandAir Model; the ratio of TDTs (VVal-18/StandAir) averages 0.87 ± 0.22 (SD). We can use the StandAir Model to generate probability values for the depth/bottom-time/TDT profiles prescribed by the VVal-18 Algorithm with *Pdcs* close to 2% (average = 2.3 ± 0.3% (SD)). According to the trend curve in Figure 4B, *Pdcs* of 2.3% is associated with observed incidence of 1.6%.

CONCLUSIONS

a) Our probabilistic StandAir Model, generated from nonsaturation dives only, appears to prescribe TDT times that are acceptable for standard air diving to most depths. Questionable model prescriptions are at the two extremes of the depth continuum: there are few dives having appreciable time at decompression stops at the shallowest depths, and cases of DCS near the StandAir Model prescriptions for dives with short TDTs suggest

that the model may be too liberal for deep no-stop dives and for deep dives having short TDTs. b) If we are correct in our assertion that the long times at decompression stops mandated by some other probabilistic models are not warranted, it follows that the estimates of *Pdcs* by the other models are not accurate. c) Estimates using the StandAir Model parameters indicate that the VVal-18 Algorithm gives *Pdcs* near 2% and therefore is acceptable for operational use to most depths.

ACKNOWLEDGEMENT

This work was supported in part by the Deep Submergence Biomedical Development Program, Naval Sea Systems Command, Task Numbers 63713N S0099 01A 99-04 and 63713N S0099 01A 01-07. The opinions and assertions contained herein are the private ones of the authors and are not to be construed as official or as reflecting the views of the Navy Experimental Diving Unit, or the U.S. Navy.

REFERENCES

1. Weathersby PK, Homer LD, Flynn ET. On the likelihood of decompression sickness. *J Appl Physiol* 1984; 57:815–824.
2. Parker EC, Survanshi SS, Massell PB, Weathersby PK. Probabilistic models of the role of oxygen in human decompression sickness. *J Appl Physiol* 1998; 84:1096–1102.
3. Van Liew HD, Flynn ET. Decompression tables and dive-outcome data: graphical analysis. *Undersea Hyperb Med* 2005; (32):187-198.
4. Thalmann ED. Suitability of the USN MK 15 (VVAL-18) decompression algorithm for air diving. Final Report on U.S. Navy Contract N0463A-96-M-7036. Durham, NC: Duke University, 1997.
5. Thalmann ED. Phase II testing of decompression algorithms for use in the U.S. Navy Underwater Decompression Computer, NEDU TR 1-84. Panama City, FL: Navy Experimental Diving Unit, 1984.
6. Survanshi SS, Parker EC, Thalmann ED, and Weathersby PK. Statistically based decompression tables XII, Volume I: repetitive decompression tables for air and constant 0.7 ata PO_2 in N_2 using a probabilistic model. NMRI Technical Report 97-36. Bethesda MD, 1997.
7. Temple DJ, Ball R, Weathersby PK, Parker EC, Survanshi SS. The dive profiles and manifestations of decompression sickness cases after air and nitrogen-oxygen dives, Volumes I and II, NMRC Technical Report 99-02. Bethesda MD, 1999.
8. Weathersby PK, Survanshi SS, Nishi RY, Thalmann ED. Statistically based decompression tables VII: selection and treatment of primary air and N_2O_2 data. Joint Report, U.S. Naval Submarine Medical Research Laboratory, rep. 1182, Groton CT; NMRI Technical Report 92-85, Bethesda MD, 1992.
9. Commander, Naval Sea Systems Command, U.S. Navy Diving Manual, Revision 4:SS521-AG-PRO-010. Arlington, VA: Naval Sea Systems Command, 1999.
10. Ku HH. Notes on the use of propagation of error formulas. *J Res Natl Bur Stand. C. Eng Instr* 1966; 70C:263–273.
11. Hays JR, Hart BL, Weathersby PK, Survanshi SS, Homer LD, Flynn ET. Statistically based decompression tables IV: extension to air and N_2O_2 saturation diving. U.S. Naval Medical Research Institute, rep. 86-51, Bethesda MD, l986.
12. Weathersby PK, Survanshi SS, Nishi RY. Relative decompression risk of dry and wet chamber air dives. *Undersea Biomed Res* 1990; 17:333-352.
13. Ball R, Lehner CE, Parker EC. Predicting risk of decompression sickness in humans from outcomes in sheep. *J Appl Physiol* 1999; 86:1920-1929.
14. Van Liew HD, Flynn ET. Use of dive-outcome information three times in preparation of a probabilistic decompression model. *Undersea Hyperb Med* 2002; 29:111-112.
15. Van Liew HD, Flynn ET. Probability of decompression sickness in no-stop air diving and subsaturation diving. *Undersea Hyperb Med* (In press).
16. Gerth WA, Vann RD. Probabilistic gas and bubble dynamics models of decompression sickness occurrence in air and nitrogen-oxygen diving. *Undersea Hyperb Med* 1997; 24:275–292.
17. Tikuisis P, Gault KA, Nishi RY. Prediction of decompression illness using bubble models. *Undersea Hyperb Med* 1994; 21:129–143.
18. Weathersby PK, Survanshi SS, Homer LD, Parker E, Thalmann ED. Predicting the time of occurrence of decompression sickness. *J Appl Physiol* 1992; 72:1541–1548.
19. Conkin J, Kumar V, Powell MP, Foster PP, Waligora JM. A probabilistic model of hypobaric decompression sickness based on 66 chamber tests. *Aviat Space Environ Med* 1996; 67:176–183.
20. Gernhardt ML. Development and evaluation of a

decompression stress index based on tissue bubble dynamics, Ph.D. dissertation. Philadelphia, PA: University of Pennsylvania, 1991.

21. Lambertsen CJ. IFEM /Sub Sea International Inc. Air Diving Tables. Institute for Environmental Medicine, Document NO-S-MA-DTA01, Revision A, Philadelphia PA, 1995.

22. Nishi RY, Hobson BA, Lauckner GR. DCIEM/ Canadian Forces air decompression tables and procedures. Downsview, ONT, Canada: DCIEM, 1986.