Multinomial Probabilistic Modeling for Decompression Sickness using Gas Content and Bubble Volume Models

by

Amy Elizabeth King

Department of Mechanical Engineering and Materials Science
Duke University

Date:_____

Approved:

_____
Laurens Howle, Advisor

_____
Lawrence Virgin

_____
Brian Mann

_____
Thomas Witelski

Dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy in the Department of
Mechanical Engineering and Materials Science in the Graduate School of
Duke University

2020

ABSTRACT

Multinomial Probabilistic Modeling for Decompression Sickness using Gas Content and Bubble Volume Models

by

Amy Elizabeth King

Department of Mechanical Engineering and Materials Science
Duke University

Date:_____

Approved:

_____
Laurens Howle, Advisor

_____
Lawrence Virgin

_____
Brian Mann

_____
Thomas Witelski

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of
Mechanical Engineering and Materials Science in the Graduate School of
Duke University

2020

Abstract

Decompression sickness (DCS) is a condition resulting from the decrease in ambient pressure, both in hyperbaric and hypobaric environments. When ambient pressure is decreased, inert gas dissolved in the body's tissues can become supersaturated and form bubbles, the physiological precursor to DCS. The signs and symptoms of DCS range from mild joint pain or rash to serious cardiopulmonary or neurological dysfunction. DCS is treated with recompression therapy, in which a patient is recompressed in a hyperbaric chamber and then decompressed following a treatment schedule specific to their symptoms.

This work focuses on DCS arising from hyperbaric exposures, and specifically underwater diving. DCS is a risk faced by U.S. Navy divers during underwater missions. To gain insight into the level of risk posed by a particular time-depth profile, the Navy uses probabilistic decompression models. These models quantify the probability of DCS occurring using survival analysis and a gas content or bubble volume model to define risk. Current probabilistic models make a binomial prediction of the probability of DCS occurring and not occurring.

Risk is a two-faceted entity, in which both the probability of injury and severity of injury contribute to the level of risk posed by an activity. Many Department of

Defense organizations use two-dimensional risk assessment matrices to consider both the probability and severity of injury to manage operational risk. However, the U.S. Navy's probabilistic decompression models only predict the probability of injury and do not provide divers with any information regarding potential DCS symptom severity.

The goal of this work is to investigate the efficacy of a variety of multinomial probabilistic models which simultaneously predict the probability and severity of DCS injury. We first conducted an analysis on the BIG292 calibration dataset to uncover the source of the bimodal trend in DCS symptom onset times. We concluded that neither of the two peaks alone correspond to DCS symptoms or DCS resulting from a particular dive type, and rather are a product of dive trial medical surveillance protocol. Because the dataset's bimodal symptom onset time behavior is not related to the illness itself, it is not necessary for a probabilistic model to reproduce this trend.

Next, we developed 20 multinomial probabilistic models, testing the effectiveness of gas content versus bubble volume models to calculate risk, the justification of various gas content model parameters, the impact of using Type A/B versus Type I/II symptom severity splitting methods to define mild/serious DCS cases, and the influence of treating marginal DCS cases as separate, hierarchical events versus considering them non-events during model calibration.

The multinomial models presented in Chapters 3, 4, and 5 are able to accurately predict the incidence and severity of DCS as observed in the calibration data within 95% confidence. We find that trinomial models, which predict the probabilities of mild, serious, and no DCS, do perform better on the calibration dataset then their binomial counterparts. Multinomial models that predict the probabilities of marginal DCS are not able to accurately replicate the onset or distribution of marginal DCS cases in the calibration dataset, and the use of marginal cases during model calibration negatively affects the model's ability to accurately predict other types of DCS as well. The multinomial bubble volume models tested in this work were not able to achieve an optimal parameter set, and experienced model failure when predicting zero risk for a subset of dives.

We do not recommend the further use of multinomial models that predict the probability of marginal DCS separately from other symptom types. The multinomial bubble volume models should be reoptimized with a different algorithm and/or an alternative bubble nucleation criterion. To determine which multinomial probabilistic model presented here is optimal for U.S. Navy dive planning, all models should be evaluated on data that was not used for model calibration.

Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

Symbols

$D_{t,i}$	Compartmental bulk diffusion constant for inert gas
F_{N_2}	Fixed fraction of nitrogen gas
g_i or G_i	Compartmental gain for gas content (lowercase) or bubble volume (uppercase) models
k_i	Compartmental tissue half-time
LL	Log likelihood function
M_i	Compartmental elastic modulus
P_a	Arterial inert gas pressure
P_{amb}	Ambient pressure
P_b	Bubble pressure
P_{FVG}	Fixed venous gas partial pressure
P_{H_2O}	Water vapor partial pressure
P_{N_2}	Nitrogen gas partial pressure
$P_{T,i}$	Compartmental inert gas partial pressure
P_{XO}	Crossover pressure
PR	Pearson Residual

$\dot{Q}$	Flow rate
r	Risk function
R_{amb}	Rate of change of ambient pressure
$r_{b,i}$	Compartmental bubble radius
$r_{b,i}^0$	Compartmental bubble nucleonic radius
T	Right-censored time
T_1	Last known time diver was asymptomatic
T_2	First known time diver was experiencing DCS symptoms
Thr_i	Compartmental pressure threshold
$V_{b,i}$	Compartmental bubble volume
$V_{b,i}^0$	Compartmental bubble nucleonic volume
α_b	Ostwald solubility of nitrogen gas in blood
$\alpha_{t,i}$	Compartmental solubility of nitrogen gas in tissue
γ_i	Compartmental diffusivity fraction
λ_i	Compartmental sink
ξ_i	Compartmental total solubility
σ	Gas-liquid surface tension

Abbreviations

ATA	Atmospheres Absolute
3RUT	Three-Region Unstirred Tissue
BVM	Bubble Volume Model
CDF	Cumulative Density Function
DCIEM	Defence and Civil Institute of Environmental Medicine
DCS	Decompression Sickness
DMO	Diving Medical Officer
EE	Exponential-Exponential
fsw	Feet of seawater
INM	Institute of Naval Medicine
IRB	Institutional Review Board
LE	Linear-Exponential
NEDU	Navy Experimental Diving Unit
NMRC/NMRI	Naval Medical Research Center / Naval Medical Research Institute
NSMRL	Naval Submarine Medical Research Laboratory
ODE	Ordinary Differential Equation
ODF	Occurrence Density Function
PLB	Perfusion-Limited Base
PSI	Perceived Severity Index

1. Introduction

1.1 Brief History of Dive-Related Illnesses

The first recorded case of decompression sickness (DCS) dates back to 1670, when Sir Robert Boyle noted a bubble in the eye of a viper after being placed in a vacuum [1]. Through this experiment, he demonstrated bubble formation in living tissue following a reduction in ambient pressure. This was eight years after Boyle's law was published, which describes the inverse relationship between the pressure and volume of a gas. In the early nineteenth century, Dalton's gas law and Henry's law were published, both of which contribute to our modern understanding of DCS. Dalton's law of partial pressures states that the total pressure of a mixture of gases is equal to the sum of the partial pressures of each gas. Henry's law equates the amount of dissolved gas in a liquid to the partial pressure of the gas above the liquid.

In the 1840's, DCS was described in coal miners working in a pressurized mine in France, who experienced muscle cramps after leaving the pressurized environment [1, 2]. This phenomenon was described again in the early 1870's during the construction of the Brooklyn Bridge and the Eads Bridge. The large piers of these bridges were built using caissons, which were air-tight, pressurized chambers to keep mud and water from flooding the working environment. The construction of the east pier of the Eads Bridge exposed workers to 4.4 ATA, and after a 3-4 minute decompression from this environment, many workers experienced joint pain, numbness, paralysis, and there

were 12 recorded deaths [2]. After leaving the working environment, many laborers walked with a stooped posture, giving DCS its popular name of "the bends", derived from the similar fashionable posture of Victorian ladies called the Grecian bend [3].

In 1878, French physiologist Paul Bert published *La Pression barometrique* (Barometric Pressure), describing a series of experimental work on the physiological effects of hypobaric and hyperbaric exposures. He was able to conclude that DCS symptoms were caused by the formation of nitrogen bubbles in the blood and recommended DCS mitigation through slow decompression. Among Bert's findings was also central nervous system oxygen toxicity. He found that breathing oxygen at elevated pressures can be toxic, and symptoms of oxygen toxicity are sometimes called the "Paul Bert effect" [1, 2].

In the early twentieth century, though it was known that DCS arose from a decrease in pressure and the formation of nitrogen bubbles in the body, there were no standard methods for mitigating DCS risk. The Royal Navy commissioned Scottish physiologist J. S. Haldane to develop decompression schedules that could be used by Navy divers. The gas content model published by Haldane and coworkers [4] provided the foundation upon which all modern decompression models are based. Haldane's model calculated nitrogen gas content in five compartments (hypothetically representing tissues in the body), each with a unique gas absorption/elimination rate (also called a half-time). These half-times were chosen to be 5, 10, 20, 40, and 75 minutes.

Unlike previous researchers who proposed uniform decompression, Haldane's model recommended staged decompression, in which divers make stops during ascent to eliminate the absorbed inert gas. This decompression schedule allowed for a maximum reduction in absolute pressure by one-half, followed by a series of decompression stops to complete the ascent so as not to exceed this 2:1 ratio of air tension in the diver to the absolute pressure. The model was deterministic, meaning it predicted a binary outcome of DCS or no DCS based on the adherence to or violation of Haldane's "safe" decompression criteria [2, 4].

In the mid 1930's, the U.S. Navy observed occasional and sudden distress in divers surfacing from less than 33 feet of seawater (fsw). Thought to be aberrant cases of DCS, similar symptoms resulting from submarine escape training revealed pulmonary barotrauma. This lung injury results from breath-holding during ascent: as ambient pressure decreases, the air in the lungs will expand and, if the diver does not exhale, can rupture, releasing air bubbles into the pulmonary veins. These air bubbles, called Arterial Gas Embolisms, can travel to the heart and brain, causing the severe symptoms reported by the Navy [1].

Also in the mid 1930's, U.S. Navy Submarine Medical Officer Albert Behnke investigated the intoxicating effects of diving deeper than 150 feet. Behnke discovered that inert gases, such as nitrogen, can cause mental deterioration; this phenomenon is known as nitrogen narcosis. To avoid nitrogen narcosis, a series of experimental dives

were performed with helium-oxygen (heliox) breathing mixtures in 1937. In the 1960's, breathing mixtures of helium-nitrogen-oxygen (trimix) were investigated [1].

1.2 Physiology of DCS

This work focuses on hyperbaric exposures from underwater diving. At atmospheric pressure, the body's tissues are saturated with the partial pressure of inert gas (nitrogen) inspired. During decent to depth, the hydrostatic pressure of the water increases the ambient pressure experienced by a diver. For every 33 fsw (10 meters seawater) descended, the ambient pressure increases by one atmosphere (1 bar, 14.5 psig). As a diver descends, the partial pressure of inert gas in their blood and tissues, which was equilibrated with 1 ATA, is now lower than the ambient pressure. The diver breathes gas that is equilibrated to the new ambient pressure into the lungs, where it is absorbed into the blood and circulated to the tissues at this increased partial pressure. When a diver ascends back to the surface and ambient pressure decreases, the absorbed inert gas is now at an elevated partial pressure and can leave solution, forming gas bubbles [5].

The signs and symptoms of DCS can range from mild, such a skin rashes or joint pain, to more severe, including neurological and cardiopulmonary malfunction and even death. The U.S. Navy categorizes DCS into two severities: Type I and Type II [6]. Type I DCS includes joint pain, cutaneous symptoms, and lymphatic symptoms. Joint pain is the most common symptom of DCS, and typically occurs at the shoulder, elbow,

wrist, hand, knee, and ankle. Cutaneous symptoms include itching, skin rashes, and cutis marmorata (marbling). Lymphatic symptoms result from swelling and pain in the lymph nodes. Type II DCS includes all neurological, inner ear, and cardiopulmonary symptoms. Common neurological symptoms include numbness, paresthesia, muscle weakness, loss of motor skills, tremors, lightheadedness, amnesia, urinary malfunction, and paralysis. Inner ear DCS, which has been observed resulting from a switch from helium-oxygen mixtures to air during decompression, is marked by tinnitus, hearing loss, vertigo, nausea, and vomiting. Cardiopulmonary symptoms include coughing, chest pain, and dyspnea [6].

Marginal DCS is defined as symptoms associated with DCS that are experienced during or after a dive and are mild, then subsequently resolve spontaneously without recompression treatment [7, 8]. Examples of marginal DCS are pain in one joint lasting for less than 60 minutes, or pain in two joints lasting less than 30 minutes.

Howle *et al.* introduced an alternative approach to the classification of DCS symptom severity [9, 10]. Howle's system, called Perceived Severity Index (PSI), categorizes DCS symptoms into six indices. In order of increasing severity, these indices are constitutional or nonspecific (fatigue, dizziness, nausea), lymphatic or skin (itching, rash, marbling), pain (ache, cramps, joint pain), mild neurological (paresthesia, numbness, tingling), cardiopulmonary (cough, dyspnea, hemoptysis), and serious neurological (dysfunction of bladder, bowel, coordination, mood, vision, hearing).

Under this system, the Navy's traditional classifications indicate Type I includes constitutional, skin, and pain symptoms, while Type II includes mild neurological, cardiopulmonary, and serious neurological symptoms (see Table 1).

Howle *et al.* recommended an alternative severity categorization to Type I/II called Type A/B. Type A includes constitutional, lymphatic or skin, pain, and mild neurological symptoms, and Type B encompasses cardiopulmonary and serious neurological manifestations (Table 1).

Table 1: Distribution of PSI into Type I/II and Type A/B splitting.

PSI	Type I/II	Type A/B
6: Constitutional	Type I	Type A
5: Lymphatic or skin		
4: Pain		
3: Mild neurological	Type II	
2: Cardiopulmonary		Type B
1: Serious neurological		

The U.S. Navy Diving Manual outlines procedures for treatment of DCS based on symptom severity. These procedures, called treatment tables, outline recompression and subsequent decompression schedules and oxygen-breathing periods to be performed in a hyperbaric chamber [6].

1.3 Probabilistic Modeling of DCS

Haldane's decompression model was deterministic, as it predicted that DCS would not occur if the prescribed "safe" ascent criteria were followed, and DCS would occur if the ascent schedule was violated. This model has been improved and refined,

and the basic approach is still used today to generate staged decompression schedules. However, deterministic models are unable to describe the variability in symptoms observed in divers executing identical dive profiles. To address this limitation of deterministic models, Berghage *et al.* [11] and Weathersby *et al.* [12] proposed a probabilistic approach, in which each dive is assigned a non-zero probability of resulting in DCS. A distinct advantage of probabilistic models is model parameters can be estimated via maximum likelihood, allowing the model to be calibrated to a set of human dive data.

Using survival analysis to convert a risk function r into a probability [13], the probability of not experiencing DCS before T is

$$P(0) = e^{-\int_0^T r dt}. \tag{1.1}$$

The risk function is defined by a gas content model (Section 1.4 Gas Content Models) or bubble volume model (Section 1.5 Bubble Volume Models) and contains all the information about the gas kinetics and ambient pressure. The probability of DCS occurring before time T is

$$P(DCS) = 1 - e^{-\int_0^T r dt}. \tag{1.2}$$

Weathersby *et al.* improved upon this concept with the addition of onset time information [14]. Given a symptom onset time window, where T_1 is the last time a diver was definitely symptom-free, and T_2 is the first time the diver was definitely

experiencing symptoms, the combined probability of being symptom-free until T_1 and the onset of DCS occurring between T_1 and T_2 is

$$P(DCS) = P(0)_{0\to T_1} P(DCS)_{T_1\to T_2} = \left(e^{-\int_0^{T_1} r dt} \right) \left(1 - e^{-\int_{T_1}^{T_2} r dt} \right). \tag{1.3}$$

The parameters of a model used to generate the risk function r can be estimated with likelihood maximization. Given a set of calibration dive data containing the time-depth profiles and symptom histories, the agreement between the model's predictions and the true outcomes from the data can be quantified as the likelihood. Maximizing the likelihood will thus maximize the accuracy of the model's predictions. The likelihood function for the i^{th} dive is defined as

$$L_i = P(DCS_i)^{\delta_i} (1 - P(DCS_i))^{(1-\delta_i)} \tag{1.4}$$

where $\delta_i = 1$ if the i^{th} exposure resulted in DCS, and $\delta_i = 0$ if the i^{th} exposure did not result in DCS. The definition of $P(DCS)$ in Eq. (1.3) can be used in Eq. (1.4). The likelihoods of each dive exposure are multiplied. Because the probabilities are always less than 1, the likelihood parameter can become quite small, so the log likelihood is used to quantify the agreement between the model's predictions and the calibration data:

$$LL = \sum_{i=1}^{N} \ln\left[P\left(DCS_i\right)^{\delta_i} (1 - P(DCS))^{1-\delta_i} \right]. \tag{1.5}$$

To find the best set of model parameters, Eq. (1.5) is maximized.

1.4 Gas Content Models

In Chapters 3 and 4, the risk function is defined using three parallelly-perfused compartments (Figure 1), similar to those used by Boycott *et al.* in the Haldane model [4]. Each compartment is assumed to be well-stirred and has a unique half-time. Models with strictly exponential gas uptake and elimination are known as Exponential-Exponential (EE). Thalmann *et al.* introduced models with both linear and exponential gas kinetics, known as Linear-Exponential (LE), in which gas elimination kinetics are switched from exponential to linear after a specific level of supersaturation is achieved [15, 16]. This crossover pressure is intended to represent the level of supersaturation at which dissolved inert gas leaves solution and enters the gas phase. While the crossover pressure is exceeded, linear gas kinetics are used to simulate the transfer of the gas bubble into the solution.

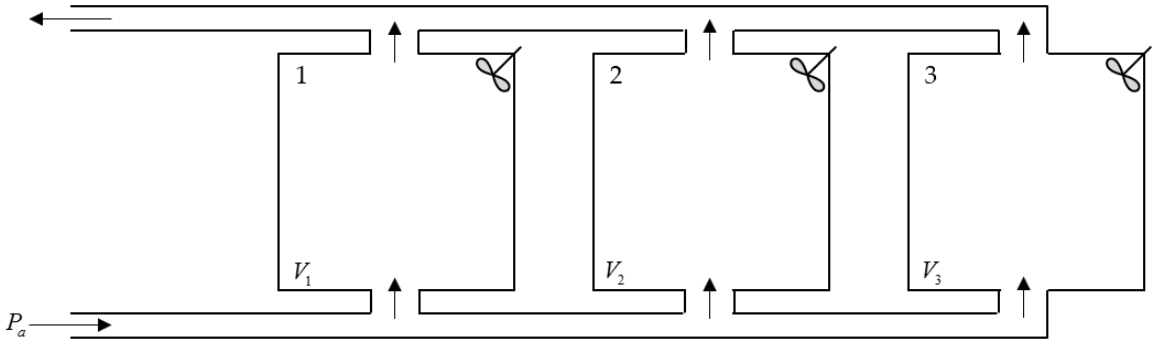

Figure 1: Three parallelly-perfused, well-stirred PLB compartments.

Considering a single, well-mixed perfused compartment, also called the Perfusion-Limited Base (PLB) model, the time rate of change of the inert gas tension in the compartment as dictated by mass conservation is

$$\frac{dP_{T,i}}{dt} = \frac{\dot{Q}}{V_i}(P_{T_{in}} - P_{T,i_{out}}) \tag{1.6}$$

where $P_{T,i}$ is the partial pressure of inert gas in the i^{th} compartment, t indicates time, $\dot{Q}$ is the flow rate, V_i is the volume of the compartment, $P_{T_{in}}$ is the inert gas wash-in, and $P_{T,i_{out}}$ is the inert gas wash-out. Assuming the compartment is well-stirred dictates that $P_{T,i} = P_{T,i_{out}}$. Note this first-order linear ordinary differential equation can be solved analytically. We will assume the input pressure is a linear ramp, i.e.

$$P_{T_{in}} = P_{T_{in}}^0 + R_{in}t \tag{1.7}$$

where $P_{T_{in}}^0$ is the input inert gas tension at $t = 0$ and R_{in} is the rate of increase of input pressure. Substituting Eq. (1.7) into Eq. (1.6), defining tissue rate $k = \frac{Q}{V}$, and solving Eq. (1.6) for P_T yields

$$P_T = \alpha e^{-kt} + R_{in}t + \beta \tag{1.8}$$

where α and β are defined as

$$\begin{aligned} \alpha &= P_T^0 - P_{T_{in}}^0 + k^{-1}R_{in}, \\ \beta &= P_{T_{in}}^0 - k^{-1}R_{in}. \end{aligned} \tag{1.9}$$

The dive trial data used in this dissertation is for air or nitrox diving, so nitrogen is the inert gas of interest. Given the dive's time-depth profile, absolute ambient pressure can be converted from feet sea water (fsw) depth as

$$1\ fswg\left(\frac{1atm}{33.08\ fswg}\right)+1atm\rightarrow ATA. \tag{1.10}$$

Assuming the diver is breathing a fixed fraction of nitrogen gas, the partial pressure of inspired nitrogen gas is calculated as

$$P_{N_2}=F_{N_2}\left(P_{amb}-P_{H_2O}\right) \tag{1.11}$$

where F_{N_2} is the fixed fraction of nitrogen gas (0.79 in air), P_{amb} is the absolute ambient pressure in ATA, and P_{H_2O} is the water vapor pressure (constant 0.0617 ATA). Eq. (1.11) can be used as input pressure $P_{T_{in}}$ in Eq. (1.6).

For each dive segment, the ambient pressure is

$$P_{amb}=P_{amb}^{0}+R_{amb}t \tag{1.12}$$

where P_{amb}^{0} is the ambient pressure at the start of the dive leg and R_{amb} is the rate of change in ambient pressure. The absolute ambient pressure calculated with Eq. (1.12) is used in Eq. (1.11) to determine the partial pressure of the inspired inert gas. From here, the tissue tension values, P_T, can be calculated at each time step.

The instantaneous risk in this gas-content model compartment is proportional to the level of inert gas supersaturation, given by

$$r_i(t)=\frac{P_{T,i}-P_{amb}-Thr_i+P_{FVG}}{P_{amb}} \tag{1.13}$$

where P_{FVG} is the constant fixed venous gas pressure (6.34 fsw), determined by summing the partial pressures of venous gases O_2 (2.00 fsw), CO_2 (2.30 fsw), and H_2O

vapor (2.04 fsw). Thr_i is an optimized threshold parameter that serves to delay risk accumulation until the supersaturation exceeds a critical threshold, and has been found to significantly improve a model's performance [15, 17]. The subscript i indicates the individual compartment. This instantaneous risk is multiplied by an optimized gain parameter, g_i, and summed over the entire dive for each of i compartments. Thus the form of Eq. (1.2) specific to EE gas kinetics is

$$P(DCS) = 1 - e^{-\sum_i g_i \int_0^T r_i(t)dt} \qquad r_i \geq 0 \tag{1.14}$$

and likewise when using the time window of symptom onset with Eq. (1.3).

The addition of linear gas kinetics adds an optimized crossover pressure, P_{XO}, above which compartmental gas kinetics transition from exponential to linear. Thalmann *et al.* [15] defined inert gas tension as

$$P_T = P_{amb} + A \tag{1.15}$$

where

$$A = P_{XO} - P_{FVG}. \tag{1.16}$$

Thalmann then defined the "inert gas burden", P'_T, such that

$$\begin{aligned} P_T &= P'_T \qquad when \;\; P'_T < P_{amb} + A \\ P_T &= P_{amb} + A \;\; when \;\; P'_T \geq P_{amb} + A. \end{aligned} \tag{1.17}$$

Eq. (1.6) can be altered to incorporate the inert gas burden as

$$\frac{dP_T'}{dt} = k\left(P_{T_{in}} - P_T\right). \tag{1.18}$$

As ambient pressure increases, gas uptake is always exponential. When ambient pressure decreases and the inert gas burden exceeds $P_{amb} + A$, inert gas elimination switches from exponential to linear kinetics, signifying the presence of a gaseous phase [15].

1.5 Bubble Volume Models

Thalmann's LE1 probabilistic model added linear gas washout to the EE1 gas content model to prolong inert gas washout, simulating gas transfer from a gas bubble to the tissue [15, 16]. To better describe the formation and growth of an inert gas bubble, which is the mechanism initiating DCS *in vivo*, Gerth and others introduced a variety of models in which the risk function is defined by the inert gas bubble volume [18, 19]. Rather than calculating risk proportional to a critical level of inert gas supersaturation, these models define risk relative to a critical bubble volume:

$$\begin{aligned} r(t) &= \sum_{i=1}^{n} G_i \left(V_{b,i}(t) - V_{b,i}^0\right) \\ V_{b,i}(t) - V_{b,i}^0 &\geq 0 \end{aligned} \tag{1.19}$$

where G_i is the compartmental gain, $V_{b,i}$ is the compartmental bubble volume, and $V_{b,i}^0$ is the compartmental nucleonic bubble volume.

In Chapter 5, the mathematical model used to calculate bubble volume is the Three-Region Unstirred Tissue, or 3RUT, model. The 3RUT model contains a gas bubble

surrounded by an unstirred diffusion region that is perfused by a well-stirred tissue compartment [20, 21]. A schematic of three parallelly-perfused 3RUT model compartments is depicted in Figure 2.

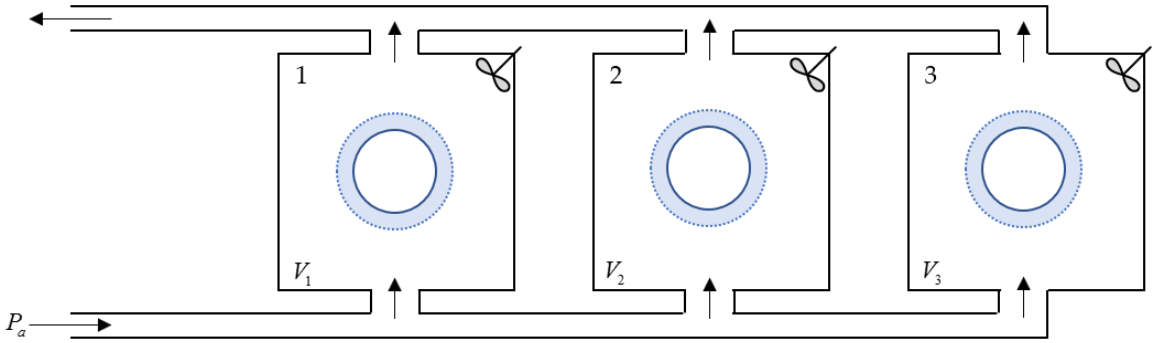

Figure 2: Three parallelly-perfused, well-stirred 3RUT compartments. Each compartment contains a bubble surrounded by an unmixed diffusive region, perfused by arterial blood flow.

Compartmental tissue tension is calculated according to Eq. (1.6). When the compartmental supersaturation exceeds ambient pressure according to Eq. (1.20),

$$P_{amb} < P_{T,i} + P_{FVG} - \frac{2\sigma}{r_{b,i}^0} \tag{1.20}$$

where σ is the gas-liquid surface tension, a bubble forms at the nucleonic radius, $r_{b,i}^0$ [18]. After bubble formation, the blood-tissue gas exchange is governed by

$$\frac{dP_{T,i}}{dt} = k_i \left(P_a - P_{T,i} \right) - \xi_i \frac{d\left(P_{b,i} V_{b,i} \right)}{dt} \tag{1.21}$$

where the arterial gas tension, P_a, is calculated with Eq. (1.11), $P_{b,i}$ is the compartmental bubble pressure, $V_{b,i}$ is the compartmental bubble volume, and ξ_i is the compartmental total solubility, defined as

$$\xi_i = \frac{1}{\alpha_{t,i} V_{t,i}}. \tag{1.22}$$

In Eq. (1.22), $\alpha_{t,i}$ is the compartmental solubility of nitrogen gas in tissue, and $V_{t,i}$ is the compartmental volume.

The bubble dynamics are calculated according to the desired mathematical model for bubble evolution. In this work, we will use the 3RUT model with a single bubble [20]. According to this model, the change in bubble radius is calculated as

$$\frac{dr_{b,i}}{dt} = \frac{\gamma_i \left(P_{T,i} - P_{b,i}\right)\left[\lambda_i + \frac{1}{r_{b,i}}\right] - \frac{r_i}{3}\frac{d(P_{amb})}{dt}}{\left(P_{amb} - P_{FVG}\right) + \frac{4\sigma}{3r_{b,i}} + \frac{8\pi}{3} M_i r_{b,i}^3} \tag{1.23}$$

where M_i is the compartmental elastic modulus, and the compartmental sink (λ_i) and diffusivity fraction (γ_i) are defined as

$$\begin{aligned} \lambda_i &= \left(\sqrt{\tau_i D_{t,i}}\right)^{-1}, \\ \gamma_i &= \alpha_{t,i} D_{t,i}. \end{aligned} \tag{1.24}$$

In Eq. (1.24), $D_{t,i}$ is the compartmental bulk diffusion constant for inert gas, and τ_i is calculated from Ostwald solubility of nitrogen gas in blood $\alpha_b = 1.410E-02$, the tissue-gas solubility $\alpha_{t,i}$, and the compartmental blood flow rate $\dot{Q}_i$ as

$$\tau_i = \frac{\alpha_{t,i}}{\alpha_b \dot{Q}_i}. \tag{1.25}$$

The volume of a spherical bubble is calculated as

$$V_{b,i} = \frac{4\pi}{3} r_{b,i}^3 \tag{1.26}$$

and this can be used to calculate the bubble pressure, $P_{b,i}$:

$$P_{b,i} = P_{amb} - P_{FVG} + \frac{2\sigma}{r_{b,i}} + V_{b,i} M_i. \tag{1.27}$$

The adjustable parameters in the 3RUT bubble volume probabilistic model are compartmental gain G_i, rate $k_i = \frac{1}{\tau_i}$, elastic modulus M_i, nucleonic bubble radius $r_{b,i}^0$, total solubility ξ_i, sink λ_i, diffusivity fraction γ_i, and the surface tension σ.

1.6 Dive Data

All probabilistic models presented in this work were calibrated against subsets of empirical human dive trial data published by Temple in two Naval Medical Research Center (NMRC) reports [7, 8]. This public data does not require institutional review board (IRB) approval. Temple's reports contain a compilation of dive profiles and symptom histories of air and other nitrogen-oxygen mixture dive exposures conducted by the U.S., U.K., and Canadian militaries from 1944-1997. These research trials were conducted in hyperbaric chambers, including both wet and dry diving, with medical officers monitoring divers to diagnose any onset of DCS. The data are presented with time-depth profiles for each dive, inspired gas mixtures, dive conditions (wet or dry), and DCS symptom descriptions and onset time windows if DCS occurred.

The first data subset used in this work, called BIG292, contains 3,322 exposures of bounce dives and saturation dives, with depths ranging from 20 to 604.2 fsw and durations ranging from 0.64 to 12,960 minutes. There is a total of 1,038 unique time-depth profiles in this dataset.

A bounce dive is the most common dive type, in which a diver descends to the maximum depth for a short time, followed by a direct ascent to the surface or an ascent with decompression stops. A decompression stop is an isobaric hold during ascent to allow for elimination of inert gas from the diver's body. There are two categories of bounce diving in the BIG292 dataset: single dives and repetitive/multilevel dives. A single dive consists of descent to the maximum depth, a period of time spent at maximum depth, followed by ascent to the surface with decompression stops if necessary. Repetitive/multilevel dives consist of more than one bounce dive with surface intervals between dives for inert gas elimination, or dives with significant periods spent at multiple depths. In contrast to bounce dives, a saturation dive is when the diver remains at depth for long enough that the partial pressure of inert gas in his/her tissues becomes completely equilibrated with the ambient pressure, followed by a single decompression to the surface [6]. See Table 2 for the breakdown of dive type in the BIG292 dataset.

Table 2: Dive types and DCS outcomes in the BIG292 dataset.

Dive Type	Exposures	Full DCS	Marginal DCS
Single Air	1005	53	13
Repetitive & Multilevel Air	565	34	15
Single Nonair	678	25	18
Repetitive & Multilevel Nonair	607	26	3
Saturation	467	52	61
Total	*3322*	*190*	*110*

As seen in Table 2, while only 3% of the total dive exposures in the BIG292 dataset resulted in marginal DCS, over half (55%) of these marginal events occurred from saturation diving. However, saturation diving constitutes only 14% of the total dataset.

The second data subset used in this work is called NMRI98 (also referred to as p97). This dataset contains the entirety of the BIG292 data and adds 1,013 dive exposures, totaling 4,335 exposures with 1,306 unique dive profiles. These additional dive exposures used increased oxygen content in the divers' inspired gas. There are two categories of these dives: in-water decompression with oxygen, and surface decompression with oxygen. During in-water oxygen decompression dives, the diver switches to a breathing mixture of high oxygen content or pure oxygen during decompression. Surface decompression with oxygen dives have the diver brought to the surface before completing an adequate decompression, placed in a hyperbaric chamber, recompressed, and then complete their decompression while breathing oxygen. The p97

dataset contains 224 cases of full DCS and 127 cases of marginal DCS. Table 3 lists the breakdown of DCS events by dive type in the p97 data.

Table 3: Dive types and DCS outcomes in the NMRI98 (p97) dataset.

Dive Type	Exposures	Full DCS	Marginal DCS
Single Air	1005	53	13
Repetitive & Multilevel Air	565	34	15
Single Nonair	678	25	18
Repetitive & Multilevel Nonair	607	26	3
Saturation	467	52	61
Surface Decompression with Oxygen	427	11	1
In-Water Oxygen Decompression	586	23	12
Total	*4335*	*224*	*127*

1.7 Multinomial Probabilistic Models

Probabilistic DCS models used by the U.S. Navy today predict only the probability of the occurrence of DCS, and do not provide any information about the severity of DCS symptoms. The severity of DCS injury ought to be considered along with the probability of DCS injury, as an activity with a high probability of a mild injury or a low probability of a serious injury could both be considered high risk [22]. Both the probability and severity of injury are used in a risk assessment matrix, in which each cell is assigned a level of risk based on the particular combination of intersecting probability and severity. Such risk assessment matrices are used in many industries and facets of the Department of Defense to manage risk, including the U.S. Army [23] and the U.S. Marine Corps [24].

The U.S. Navy has guidelines for the acceptable occurrence of DCS based on symptom severity. During a dive mission, the U.S. Navy will allow for no more than two cases of Type I DCS per 100 dives, or one case of Type II DCS per 1,000 dives [25]. The goal of this dissertation is to explore models for predicting both the probability of occurrence and severity of DCS, in order to provide the U.S. Navy dive planner with a more complete spectrum of risk information.

In Chapter 2, we interrogate the BIG292 empirical DCS calibration dataset, used to estimate probabilistic model parameters. The onset times of DCS symptoms observed in the data are bimodal, however probabilistic models used by the U.S. Navy to predict the probability of the occurrence of DCS predict a unimodal peak in DCS onset over time. If the data's bimodality in DCS onset is a true component of the illness caused by symptom severity or dive type, and not a product of experimental methods, probabilistic models would need to replicate this bimodality in DCS onset to provide a better fit to the data. The source of the bimodal DCS onset times in the BIG292 dataset is investigated in Chapter 2. This work is published in *Computers in Biology and Medicine* [26].

Previously, Howle *et al.* developed a trinomial gas content model, which simultaneously predicts the occurrence of serious, mild, and no DCS [10]. In Chapters 3, 4, and 5, we further that work by deriving and analyzing gas content and bubble volume models that make multinomial predictions. In Chapter 3, we present a trinomial

marginal gas content model that predicts the probabilities of full (serious and mild), marginal, and no DCS outcomes. This work is published in *Computers in Biology and Medicine* [27]. In Chapter 4, we develop a tetranomial gas content model that predicts the probabilities of serious, mild, marginal, and no DCS outcomes. Finally, in Chapter 5, we investigate a trinomial 3RUT bubble volume model that predicts the probabilities of mild, serious, and no DCS outcomes. Our conclusions and recommendations are presented in Chapter 6.

2. Bimodal Decompression Sickness Onset Times Are Not Related to Dive Type or Event Severity

This chapter is adapted from a publication in *Computers in Biology and Medicine*.

Secondary authors are Dr. F. Gregory Murphy and Dr. Laurens Howle.

[26] A.E. King, F.G. Murphy, L.E. Howle, *Bimodal Decompression Sickness Onset Times are Not Related to Dive Type or Event Severity*. Computers in Biology and Medicine, 91 (2017) 59-68. https://doi.org/10.1016/j.compbiomed.2017.10.010

2.1 Introduction

Decompression sickness (DCS) is a condition associated with depressurization of the body from underwater diving. During a dive, exposure to increased ambient pressure allows elevated partial pressures of inert gas in the lung to dissolve into the blood. When this blood circulates, the inert gas can diffuse into the body's tissues. During decompression and after surfacing from a dive, the excess inert gas is normally circulated back to the lungs to be exhaled. However, if the ambient pressure is reduced sufficiently far below the partial pressure of the dissolved gases, then gaseous bubbles may form in the blood and/or tissues. The signs and symptoms of DCS can include, but are not limited to joint pain, paresthesia, fatigue, abdominal pain, and paralysis [3]. DCS cases are typically categorized into either Type I (also called mild) or Type II (also called serious), in which Type I includes pain-only cases and Type II includes neurological and cardiopulmonary cases. In addition, DCS manifestations which subsequently spontaneously resolve without recompression treatment are categorized as marginal

DCS cases. Examples of marginal cases are pain in one joint lasting less than 60 minutes or pain in two joints lasting less than 30 minutes [7, 8].

Decompression modeling originated in the early 20th century when Boycott *et al.* introduced the theory that DCS was caused by the formation of bubbles in the body during decompression due to the elevated partial pressure of dissolved nitrogen gas in the body's tissues [4]. The model presented by Boycott and coworkers, later known as the Haldane model, was deterministic, as DCS could be avoided if a set of criteria were followed and was inevitable if those criteria were violated. However, deterministic modeling cannot account for the variation in DCS occurrence and symptoms present in divers executing identical dive profiles as recorded in empirical dive data [7, 8]. This variability in DCS outcome prompted the development of probabilistic models, introduced by Berghage *et al.* [11] and Weathersby *et al.* [12], which compute a non-zero probability of DCS occurrence for a given dive profile. Such probabilistic models used to predict the incidence and onset time of DCS rely on risk calculated from survival analysis [13] and either a gas content or bubble model. These models allow dive profiles to be created with a level of risk tailored to the diver's objective. An advantage of probabilistic modeling is that their parameters can be calibrated with empirical dive data via numerical optimization. Model parameters can be estimated to maximize the likelihood, which is a statistic that quantifies the agreement between the model and the corresponding experimental data. In addition, including the time of onset of DCS

symptoms from experimental data during optimization has been shown to improve a model's ability to describe the data [14]. To facilitate calibration of probabilistic DCS models with experimental dive data, Temple *et al.* published a compilation of dive profile and DCS manifestation descriptions corresponding to both air and nitrogen-oxygen human dive trials conducted by the United States, United Kingdom, and Canadian militaries between 1944 and 1997 [7, 8]. These research trials were conducted in hyperbaric chambers and include both wet and dry dives during which a medical officer monitored divers and determined the time of onset of DCS symptoms. Temple's report includes the bottom times, depths, and ascent rates which characterize each dive profile, and the corresponding dive conditions (wet or dry), inspired gas mixtures, DCS symptom descriptions and onset times, and references to the originating dive trial reports. The dive types performed during these research trials include single air, single non-air, repetitive and multilevel air, repetitive and multilevel non-air, air and oxygen decompression, saturation, sub-saturation, surface decompression with air, and surface decompression with oxygen. The calibration set known as the BIG292 standard DCS dataset is a subset of the data presented by Temple *et al.* that includes a portion of the single air, single non-air, repetitive and multilevel air, repetitive and multilevel non-air, and saturation dive types. This calibration set has been used in optimizing the parameters for a probabilistic model known as the LE1-USN93 model [28]. The LE1 model consists of three perfusion-limited parallel compartments, two with mono-

exponential gas uptake and elimination and one with mono-exponential uptake and linear elimination after a crossover tension is exceeded [15]. The BIG292 calibration dataset is analyzed in the present work.

An occurrence density function (ODF) describes the number of occurrences of a particular event per unit of time, and can be used to graphically assess the agreement between a model's estimations and observed DCS occurrences and onset times. These plots map time relative to the final surface interval on the abscissa and the number of DCS occurrences on the ordinate. A probabilistic model that most accurately predicts the onset time of DCS would generate an ODF which closely resembles that of empirical dive data. The ODF constructed with the BIG292 dive dataset is bimodal, peaking in DCS occurrences at both the completion of decompression and two hours following decompression. However, current probabilistic models, including the LE1-USN93 [28] and the BVM(3) [18], used to predict the onset time of DCS do not produce bimodal ODFs. The ODFs of the LE1-USN93 and BVM(3) models each contain only one peak, located after the completion of decompression. Simulating the bimodality of the empirical data would improve the fit of the model to the data, creating a better likelihood match.

Recently, Hada [29] investigated using inert gas input delay in a class of probabilistic pharmacokinetic models with perfusion coupled compartments [30] and perfusion-diffusion coupled compartments in an effort to align model onset time

predictions with the bimodal onset times found in the BIG292 data. Of the 11 delay-differential probabilistic pharmacokinetic models Hada optimized and analyzed, many showed an improvement in model fit with the addition of the single-parameter input delay but none showed enough improvement by the Akaike Information Criterion to justify adding input delay. Additionally, none of the models, when optimized on the BIG292 data, predicted bimodal ODFs. This finding motivated our present study to investigate bimodality of the BIG292 dive data. We wish to know if there is a feature, such as dive type, event severity, symptom type, or breathing gas, generates the two peaks in the ODF. If so, this might inform what model changes could lead to improved onset time prediction. If no feature can be identified, or if the bimodality is a result of some type of measurement bias, then attempts to reproduce bimodality in model prediction are unlikely to be successful or useful.

2.2 Methods

2.2.1 Calibration Dataset

The BIG292 standard DCS dataset from two Naval Medical Research Institute (NMRI) reports was used [7, 8] in this study. The BIG292 dataset, which is a subset of the dive data detailed in [7, 8], contains dive profiles from 3,322 exposures of air and nitrogen-oxygen diving conducted by the United States, United Kingdom, and Canadian militaries between 1944 and 1997. The BIG292 dataset includes single air, single non-air, repetitive and multilevel air, repetitive and multilevel non-air, and saturation dive

types, resulting in 190 DCS cases and 110 marginal DCS cases. Marginal DCS is defined as a case involving signs or symptoms associated with DCS that were deemed not serious enough to be treated with recompression and subsequently spontaneously resolved [7, 8]. In the BIG292 dataset, all DCS cases and 68 of the 110 marginal DCS cases are reported with symptom onset times T_1 and T_2, where T_1 is the last known time a diver was symptom free, and T_2 is the earliest time the diver was definitely experiencing symptoms. Following our previous work on the efficacy of using marginal DCS events in fitting probabilistic DCS models, we scored marginal cases as non-events when considering the BIG292 dataset in this work so that only full DCS events were analyzed [31, 32]. Because these dive trial data are de-identified and are freely available to the public in the form of two U.S. Government reports, IRB approval was not required for this retrospective study.

The 190 DCS cases in the BIG292 dataset can be further classified by perceived severity index (PSI) [9, 33]. As introduced by Howle *et al.*, the PSI scale is defined with the following six indices, in order of increasing severity: constitutional (fatigue, nausea, dizziness), skin bends (rash, itching, marbling), pain (aches, joint pain, stiffness), mild neurological (numbness, paresthesia), cardiopulmonary (dyspnea, cough, hemoptysis), and serious neurological (dysfunction of vision, hearing, bladder, bowel, coordination) [33]. Based on the DCS symptom descriptions in the two NMRI reports [7, 8], the 190 DCS cases were each assigned an index by Howle *et al.*, with 6 indicating constitutional

and 1 indicating serious neurological. If a DCS case fell into more than one of these categories, it was assigned an index corresponding to the highest severity present. Traditionally, DCS is categorized into Type I (mild) and Type II (serious), where Type I includes the PSI categories of constitutional, skin, and pain, and Type II includes mild neurological, cardiopulmonary, and serious neurological manifestations. An alternative approach to classifying DCS severity was proposed by Howle *et al.* [33], called Type A/B splitting. Type A (mild) includes the PSI categories of constitutional, skin, pain, and mild neurological, while Type B (serious) includes the cardiopulmonary and serious neurological PSI categories. In the BIG292 dataset, there are 152 cases of Type I DCS and 38 cases of Type II DCS. Applying Type A/B splitting, the BIG292 dataset contains 170 Type A and 20 Type B DCS cases. When exploring DCS symptom type as a potential source of the bimodal ODF in this work, both Type I/II and Type A/B splitting were applied to the BIG292 data. DCS cases corresponding to each individual PSI were also examined.

2.2.2 Computational Modeling

Many probabilistic DCS models are derived using the methods of survival analysis [13]. For these models, the probability of DCS is defined as

$$P\left(DCS\right)=1-e^{-\sum_i g_i \int r_i dt} \tag{2.1}$$

where $P(DCS)$ is the probability of a DCS event occurring, the index i counts over the risk-bearing model compartments, g_i is the i^{th} compartmental gain, r_i is the i^{th}

compartmental hazard function, and the definite integral containing the hazard function is evaluated from the beginning of the exposure to the right censoring time. It should be noted that the risk function of the form of Eq. (2.1) assumes a time-uniform event probability. Models differ by their definitions of compartmental hazard functions, r_i.

The event probability of DCS defined in Eq. (2.1) can be modified to include the time of symptom onset (event window), which was first introduced into the field of probabilistic DCS modeling by Weathersby *et al*. [14]. The BIG292 data define the DCS event window using times T_1, the last time the diver was known to be symptom-free and T_2, the first time the diver was known to be symptomatic. For a profile which results in DCS, the event probability can be expressed as

$$P(DCS) = P(0)_{0\to T_1}\, P(DCS)_{T_1\to T_2} = e^{-\sum_i g_i \int_0^{T_1} r_i dt}\left(1 - e^{-\sum_i g_i \int_{T_1}^{T_2} r_i dt}\right) \tag{2.2}$$

where $P(0)_{0\to T_1}$ is the probability that the diver remained asymptomatic from the beginning of the exposure until time T_1, and $P(DCS)_{T_1\to T_2}$ is the probability that the diver is bent during the event window $T_1 \to T_2$. The event windowing information is also used in constructing the ODF, as explained below.

Although recent work has considered the use of Bayesian inference in optimizing probabilistic DCS models [34], most models are optimized by maximizing the log likelihood function

$$LL = \sum_{n=1}^{N} \ln\left(P(DCS_n)^{\delta_n} (1 - P(DCS_n))^{(1-\delta_n)} \right) \tag{2.3}$$

in order to find the best set of model parameters [12]. In Eq. (2.3), n counts over the N exposures and $\delta_n = 1$ if the n^{th} exposure resulted in DCS and $\delta_n = 0$ otherwise. Typically, marginal DCS cases are assigned a fractional weight, $\delta_n = 0.1$, but as discussed above, we de-rate marginal DCS events to non-events [31, 32].

Occurrence density functions that display both the experimentally observed and model estimated DCS onset can be used to visualize a model's agreement with dive trial data. The ODF of observed DCS data is calculated using convolution with a top hat function [35]. Time is divided into one-hour bins, where a top hat function is turned on at the start of each bin and off at the end of each bin. To generate the model-predicted ODF, the probability of a DCS occurrence in each individual bin is calculated with Eq. (2.2) (with T_1 and T_2 equal to the bin's time interval bounds) for a particular exposure. The sum of the probabilities for all exposures in each bin is plotted against time relative to the final surface interval to create the occurrence density function [15, 18].

2.2.3 Data Partitioning

In order to determine the cause of the bimodal peaks in the BIG292 dive data ODF, many different schemes for partitioning the data were examined. DCS occurrence density functions were computed and plotted with the data partitioned by dive type, DCS event severity, DCS symptom type, institution conducting the dive trial, and

chronology of the dive trial data. Marginal DCS cases were not included in these ODFs, as they were scored as non-events; apart from one ODF which considers only marginal events. Each ODF was determined to be bimodal, unimodal, or ambiguous in shape by visual inspection of the number of peaks present. The BIG292 onset time bimodality could be attributed to dive type, event severity, symptom type, institution, or chronology if one of these methods of data partitioning resulted in unimodal ODFs. For reference, the onset time determination method was extracted from each dive report (Table 4) to determine post-dive medical protocol.

Table 4: DCS onset time measurement method of dive reports in the BIG292 dataset. The NEDU 1-99 / NMRC 99-01 and NSMRL 1200 followed a consistent method of examining divers immediately post-dive, 2 hours post-dive, and the day after the dive. All DCIEM reports followed a similar procedure of monitoring divers during the dive and for 1.5-3 hours post-dive. The NEDU reports 11-80, 1-84, and 8-85 do not explicitly document that divers were examined at 2 hours post-dive, although the NEDU 11-80 does indicate that divers were released from the hyperbaric chamber facility at this 2-hour mark.

Report Contributing to the BIG292 Dataset	DCS Onset Time Measurement Method
NEDU 11-80 [36]	Following decompression, divers remained at the dive chamber for 2 hours then were required to be within 30 min of facility for the next 4 hours. The dive chamber facility was prepared to provide DCS treatment for 24 hours post-dive.
NEDU 1-84 [37]	A US Navy Diving Medical Officer (DMO) examined all divers after surfacing.
NEDU 8-85 [38]	Divers could report symptoms of DCS at all times, including during the dive. All DCS determination was made by a DMO.

NEDU 1-99 / NMRC 99-01 [39]	Divers were examined by a DMO during surface intervals (for repetitive dives) and immediately after completing each dive. If no DCS symptoms were present, divers were re-examined 2 hours later. Divers were required to be in the presence of someone who could recognize DCS for the next 24 hours. The DMO then re-examined each diver the next morning after their dive.
NMRI 86-97 [40]	Divers were examined by a DMO upon completion of the dive and again after 2 hours following decompression. Divers could report symptoms for 18 hours post-dive and were interviewed by a DMO the morning after the dive.
NSMRL 1200 [41]	The DMO examined divers immediately post-dive, 2 hours post-dive, and 24 hours post-dive. Divers could report symptoms at any time.
DCIEM 80-R-32 [42]	Divers were monitored with Doppler Prechordal Bubble Detector, both at rest and after performing an exercise during the dive. Divers were monitored with this device pre-dive, at 15 min intervals during the dive, immediately post-dive, and periodically for at least 3 hours post-dive. Divers could report symptoms at any time. The decision to treat for DCS was not based on doppler results, however these results were used by the DMO to determine if diver symptoms required recompression.
DCIEM 81-R-02 [43]	Divers were monitored with Doppler Prechordal Bubble Detector, both at rest and after performing an exercise during the dive. Divers were monitored with this device pre-dive, at 15 min intervals during the dive, immediately post-dive, and periodically for at least 3 hours post-dive. Divers could report symptoms at any time.
DCIEM 82-R-38 [44]	Divers were monitored with Doppler Prechordal Bubble Detector, both at rest and after performing an exercise during the dive. Divers remained at rest for 90 min post-dive. Divers could report symptoms at any time.
DCIEM 84-R-72 [45]	Divers were monitored with the Doppler Bubble Detector before each dive, and at 30 min intervals for at least 2 hours post-dive while resting. If bubbles were detected, the diver remained under observation until bubbles diminished. Divers could report symptoms at any time. The DMO determined treatment based on symptoms, not bubble grades.

DCIEM 84-R-73 [46]	Divers were monitored with the Doppler Bubble Detector before each dive, and at 30 min intervals for at least 2 hours post-dive while resting. If bubbles were detected, the diver remained under observation until bubbles diminished. Divers could report symptoms at any time. The DMO determined treatment based on symptoms, not bubble grades.
DCIEM 85-R-18 [47]	Divers were monitored with the Doppler Bubble Detector before each dive, and at 30 min intervals for at least 2 hours post-dive while resting. If bubbles were detected, the diver remained under observation until bubbles diminished. Divers could report symptoms at any time. The DMO determined treatment based on symptoms, not bubble grades.

First, the BIG292 dive data files were categorized by dive type and breathing gas, which included single air, repetitive and multi-level air, single non-air, repetitive and multi-level non-air, and saturation (Table 5) [48]. The ODFs for a subset of the BIG292 data were examined using these dive type categories previously by Thalmann *et al*. [15], however they were not seeking to determine the cause of the bimodal ODF. Dive profiles were also partitioned by event severity. Both Type I/II and Type A/B splitting were used, in which Type I and Type A are considered mild DCS, while Type II and Type B are considered serious DCS. The distribution of PSI classifications into each severity splitting method is summarized in Table 6. Next, DCS data were separated according to their PSI, and the number of DCS occurrences with each PSI are listed in Table 7.

Table 5: Data partitioned by dive type. The data files in the BIG292 dataset were categorized by dive type and ODFs were generated for each grouping (Figure 7), excluding marginal DCS cases. All ODFs were bimodal except for that of the saturation dives, which was ambiguous in shape. All bimodal plots had the first peak at 0 hours and the second peak at 2 hours following decompression.

Dive Type	BIG292 Files	Number of DCS Occurrences (excluding marginal DCS cases)	ODF Shape
Single Air	EDU885A	53	Bimodal
	DC4W		
	SUBX87		
	NMRNSW		
	PASA		
	NSM6HR		
Repetitive and Multi-level Air	EDU885AR	34	Bimodal
	DC4WR		
	PARA		
	PAMLA		
Single Non-air	NMR8697	25	Bimodal
	EDU885M		
	EDU1180S		
Repetitive and Multi-level Non-air	EDU184	26	Bimodal
	PAMLAOD		
	PAMLAOS		
	EDU885S		
Saturation	ASATEDU	52	Ambiguous
	ASATNMR		
	ASATNSM		
	ASATARE		

Table 6: Data partitioned by event severity (Type I/II and Type A/B classifications) [10]. Each DCS case in the BIG292 dataset was assigned a PSI value based on the reported symptoms. These cases were then categorized into Type I or Type II, then Type A or Type B, by their PSI. Type I and Type A are considered mild cases of DCS, while Type II and Type B are serious cases. ODFs were generated for Type I, Type II, Type A, and Type B DCS cases, excluding marginal DCS cases (Figure 8). All plots were bimodal, with the first and second peaks occurring at 0 and 2 hours after the final decompression, respectively.

<table>
<tr><th>Perceived Severity Index</th><th>Type I/II</th><th>Type A/B</th></tr>
<tr><td rowspan="2">Constitutional
Skin
Pain
Mild Neurological
Cardiopulmonary
Serious Neurological</td><td>Type I
152 DCS Occurrences
Bimodal</td><td rowspan="2">Type A
170 DCS Occurrences
Bimodal</td></tr>
<tr><td>Type II
38 DCS Occurrences
Bimodal</td><td>Type B
20 DCS Occurrences
Bimodal</td></tr>
</table>

Table 7: Data partitioned by symptom type. Each DCS case in the BIG292 dataset was assigned a PSI value based on the reported symptoms, and ODFs were generated for each symptom type. The ODFs for pain and serious neurological symptoms were bimodal. The first and second peaks occurred at 0 hours and 2 hours after the final surface interval, respectively. ODFs for other symptom types did not have enough exposures to produce non-ambiguously-shaped plots.

Symptom Type	Number of DCS Occurrences (excluding marginal DCS cases)	ODF Shape Based on Symptom Type
Constitutional	1	Insufficient data
Skin	1	Insufficient data
Pain	149	Bimodal
Mild Neuro	18	Insufficient data
Cardio	2	Insufficient data
Serious Neuro	18	Bimodal

The BIG292 dive data were then partitioned by originating dive report, followed with grouping by the institution that conducted each dive trial [7, 8]. The institutions included in the BIG292 dataset are the Navy Experimental Diving Unit (NEDU), Naval Medical Research Institute (NMRI), Naval Submarine Medical Research Laboratory (NSMRL), Defense and Civil Institute of Environmental Medicine (DCIEM), and Institute of Naval Medicine (INM) (Table 8). Finally, the data were split by the year the dive trials were conducted and organized chronologically by end date (Table 9). Some dive trials spanned many years, so the corresponding data files were divided based on the date the data were collected. This chronological list was grouped into year ranges each containing similar quantities of DCS occurrences. These year ranges, which can be found in Table 9, are 1978-1983, 1984, 1985-1987, and 1988-1992, which include 44, 57, 43, and 46 DCS events respectively. There were two exceptions to this chronological method of data partitioning. First, though the dive trials contained in the DC4W data file spanned 1979-1986, all eight DCS events occurred during 1978-1983, so this data file was included in that chronology range. Second, the ASATNMR data file was composed of dive trials during June-August 1986, and then another during July 1988. The single DCS occurrence corresponding to these 50 dives occurred during the July 1988 dive trial, so the data from this file was included in the 1988-1992 range. ODFs were generated for each dive report, for all dive data generated by each institution, and for each of the year ranges outlined above.

Table 8: Data partitioned by institution. The data files in the BIG292 dataset were categorized by the institution conducting the dive trial. ODFs were generated separately for each dive report, and then for all the reports published by each institution. All marginal DCS cases were excluded. The first column indicates the institution (Navy Experimental Diving Unit, Naval Medicine Research Center, Naval Submarine Medical Research Laboratory, Defense and Civil Institute of Environmental Medicine, Institute of Naval Medicine) that conducted the dive trials. The second column includes each dive report published by that institution, and the third column indicates the dive files in the BIG292 dataset that were described by each report. Columns four and five show the number of DCS occurrences in each dive file and the shape of the ODF for those exposures, respectively. The sixth column indicates the ODF shape for all the dive data from each institution. The NEDU 11-80, NEDU 1-84, NMRI 88-06, NSMRL 1200, and INM dives have too few DCS events to produce meaningful results. The saturation dives produce an ODF with ambiguous shape (Figure 7). All bimodal plots have a first peak at 0 hours and a second peak at 2 hours after the final decompression. All ambiguously shaped plots have a peak at 0 hours after the final decompression, however the T1 and T2 times documented in these reports are consistent with those that produce bimodal plots.

Institution	Report	Files	Number of DCS Occurrences (excluding marginal DCS cases)	ODF Shape Based on Report	ODF Shape Based on Institution
NEDU	NEDU 11-80 [36]	EDU1180S	10	Insufficient data	Bimodal
	NEDU 1-84 [37]	EDU184	11	Insufficient data	
	NEDU 8-85 [38]	EDU885A	49	Bimodal	
		EDU885AR			
		EDU885M			
		EDU885S			
	NEDU 1-99 [39]	PASA	36	Bimodal	
		PAMLA			
		PARA			
		PAMLAOS			
		PAMLAOD			
NMRC	NMRI 86-97	NMR8697	11	Bimodal	Bimodal

	[40]				
	NMRI Protocol 88-06 (no report published)	NMRNSW	5	Insufficient data	
	NMRC 99-02 [7, 8]	ASATEDU	52	Ambiguous	
		ASATNMR			
		ASATNSM			
		ASATARE			
	NMRC 99-01 [39]	PASA	36	Bimodal	
		PAMLA			
		PARA			
		PAMLAOS			
		PAMLAOD			
NSMRL	NSMRL 1200 [41]	NSM6HR	3	Insufficient data	Ambiguous
DCIEM	Several [42-47]	DC4W	11	Bimodal	Bimodal
		DC4WR			
INM	No report published	SUBX87	2	Insufficient data	Ambiguous

Table 9: Data partitioned by dive trial chronology [7, 8]. Each dive data file was ordered (and split if necessary) based on dive trial end date. Though the DC4W data file spanned 1978-1986, all eight DCS events occurred during 1978-1983, so this data file has been included in that chronology range. The ASATNMR data file is composed of dive trials during June-August 1986, and then another during July 1988. The 1 DCS occurrence corresponding to these 50 dives occurred during the July 1988 dive trial, so this data file was included in the 1988-1992 range. The ODFs for 1984, 1985-1987, and 1988-1992 are bimodal, with the first peak at 0 hours and the second peak at 2 hours following decompression. The ODF for dive trials completed by 1978-1983 is unimodal, with a peak at 0 hours after the final decompression (Figure 9).

Dive Trial End Date Range	Dive File	Range Dives Conducted During	Number of Exposures	Number of DCS Occurrences (excluding marginal DCS cases)	ODF Shape Based on Age

1978-1983	EDU1180S	1977-1978	120	10	44 DCS Occurrences Unimodal
	ASATNSM Profiles 14-17	Mar 1977-Feb 1979	23	1	
	DC4W	1978-1983	244	8	
	ASATEDU Profiles 1-5	1979	10	4	
	ASATEDU Profiles 6-9	1979	10	3	
	ASATEDU Profiles 10-12	1979	10	1	
	EDU184	Jul 1980-Aug 1980	239	11	
	ASATEDU Profiles 13-14	1981	10	1	
	ASATEDU Profiles 15-17	1981	11	1	
	ASATNSM Profiles 18-24	May 1979-Feb 1981	12	3	
	ASATEDU Profile 18	1982	10	0	
	ASATEDU Profiles 19-21	1983	10	1	
	ASATEDU Profiles 22-23	1983	10	0	
1984	ASATNSM Profiles 25-28	Sept 1982-Jan 1984	16	1	57 DCS Occurrences Bimodal
	ASATNSM Profiles 1-13	Nov 1983-May 1984	33	4	
	DC4WR	Feb 1984	12	3	
	EDU885A	Aug 1984-	483	30	

		Dec 1984			
	EDU885AR	Aug 1984-Dec 1984	182	11	
	EDU885M	Nov 1984	81	4	
	EDU885S	Dec 1984	94	4	
	ASATEDU Profiles 24-26	1984	10	0	
1985-1987	NMR8697	April 1983-Dec 1985	477	11	43 DCS Occurrences Bimodal
	ASATNSM Profiles 29-41	Sept 1984-Sept 1986	31	9	
	ASATARE	1984-1986	165	20	
	ASATEDU Profiles 27-28	1986	5	1	
	ASATNSM Profiles 42-45	Oct 1986-Jun 1987	17	0	
	ASATEDU Profile 29	April 1987	9	0	
	SUBX87	13-24 July 1987	58	2	
1988-1992	ASATNMR	Jun-Aug 1986, July 1988	50	1	46 DCS Occurrences Bimodal
	ASATEDU Profiles 30-31	1988	7	0	
	ASATEDU Profile 32	1988	8	1	
	NMRNSW	May 1988-Jan 1989	91	5	
	NSM6HR	1989, 1991	57	3	
	PAMLA 1	Feb 1991-Jun 1991, Jul 1991-Jan 1992	236	13	

	PASA	Mar 1991-Jun 1991	72	5	
	PARA	Jun 1991-Jan 1992	135	7	
	PAMLAOS	Jun 1991-Jan 1992	140	5	
	PAMLAOD	Jun 1991-Jan 1992	134	6	

2.3 Results

The DCS onset time measurement methods used in each dive report contributing to the BIG292 dataset are documented in Table 4. The NSMRL 1200 [41], NEDU 1-99/NMRC 99-01 [39], NMRI 86-97 [40] technical reports all followed the same onset time determination procedure: divers were examined by a U.S. Navy Diving Medical Officer (DMO) during surface intervals, immediately following the final decompression, 2 hours post-dive, and the day after the dive. All technical reports by the DCIEM [42-47] used the following onset time determination procedure: divers were monitored with a Doppler Prechordal Bubble Detector, both at rest and while exercising during the dive, and then at 30-minute intervals for at least 1.5-3 hours post-dive. If a diver's Doppler bubble score was still elevated at this time, he/she remained under medical surveillance until this score decreased. A diver's account of symptoms was used to determine the onset time of DCS by a medical officer.

The occurrence density function for the entire BIG292 dataset, excluding marginal DCS cases, is shown in Figure 3. The time scale of the ODF ranges from mid-dive (indicated by negative values for time) to post-dive (indicated by positive values

for time) with 0 hours representing the completion of decompression. The first peak of this bimodal plot occurs at 0 hours after decompression and the second peak occurs at 2 hours after decompression. A trace of the computational model LE1-USN93 parameterized without marginal DCS events [32] is also represented in Figure 3. Note that the LE1-USN93 ODF is not bimodal, with a peak at 0 hours after the final decompression.

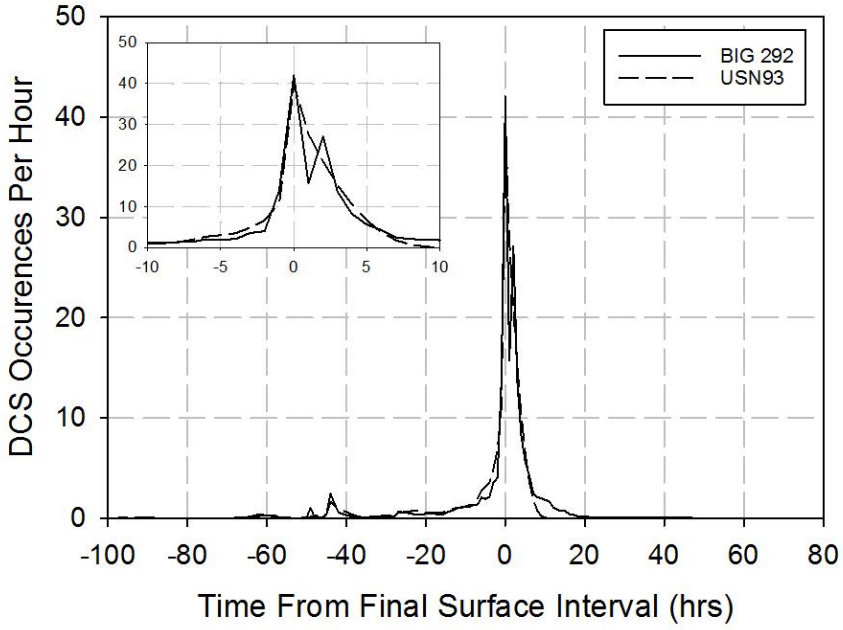

Figure 3: Occurrence density function of the BIG292 dataset and predictions from the LE1-USN93 computational model, excluding marginal DCS cases [32]. The ODF is a plot of the number of DCS occurrences per hour relative to the final surface time. Time less than zero indicates that the onset of DCS occurred before the completion of the dive. This plot is bimodal, with one peak at 0 hours and a second peak at 2 hours after the final surface time. The bimodality is not dependent on the presence or absence of marginal DCS cases. The dashed line shows the prediction from the LE1-USN93 computational model (recalibrated without marginal DCS events [32]). The ODF produced by this model is not bimodal. The magnified view of the ODF in the top left more clearly illustrates the bimodal behavior of the onset of DCS symptoms in dive trial data.

The ODF of marginal DCS cases is plotted in Figure 4. This plot contains a trace for all 110 marginal DCS cases (both with and without onset times), and a trace for the 68 marginal DCS cases with onset times. Both ODFs are unimodal with both peaks occurring at 1 hour after the final decompression. This unimodal marginal DCS data is excluded from all further ODF analysis in this work.

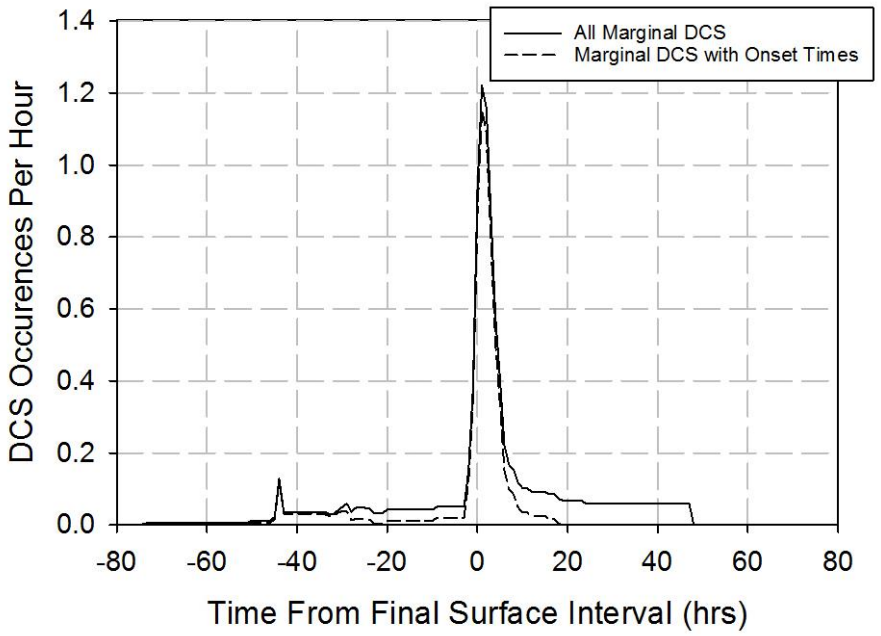

Figure 4: Occurrence density functions for marginal DCS cases. The solid line trace includes data for all 110 marginal DCS cases in the BIG 292 dataset, and the dashed line trace plots only marginal DCS cases with recorded onset times (68 cases). Both ODFs are unimodal, with both peaks at 1 hour after the final surface time.

DCS onset times T_1 and T_2 are ranked and plotted in Figure 5. DCS occurrences were sorted first by T_1, indicated by the leftmost solid line. DCS occurrences that share the same T_1 were then sorted by T_2, as indicated by the rightmost line. The gray shaded region indicates the time between T_1 and T_2. This plot serves to graphically display the

disparity in the symptom event window duration. The DCS events contributing to the first and second peaks are indicated in the figure. Despite this disparity in event window, there is no obvious indication in the figure of what causes the bimodal peak.

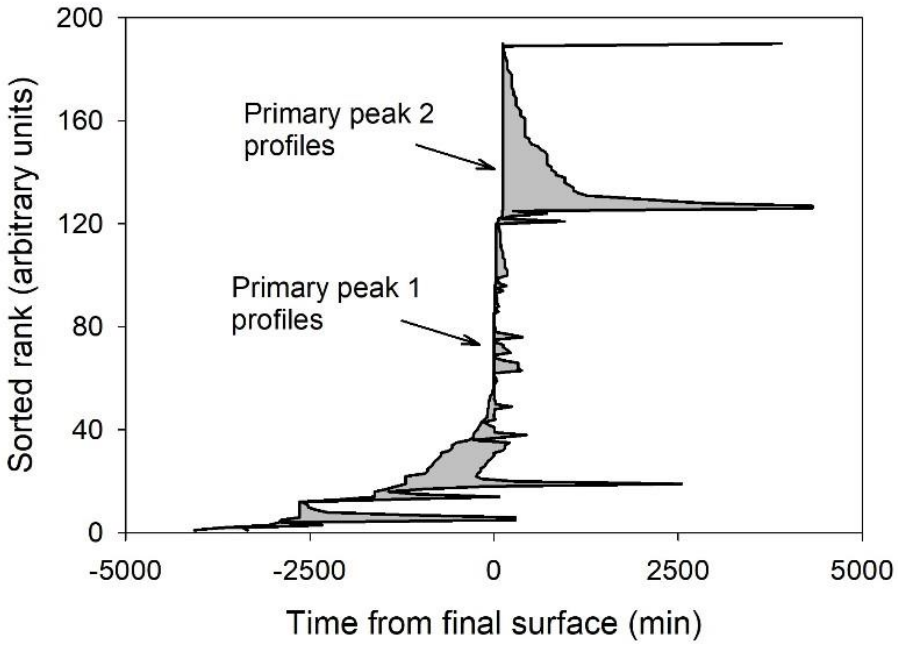

Figure 5: DCS onset time T_1 and T_2 ranking. T_1 is the last known time a diver was symptom free, and T_2 is the earliest time the diver was definitely experiencing symptoms. DCS occurrences were sorted first by T_1, indicated by the leftmost line. DCS occurrences that share the same T_1 were then sorted by T_2, as indicated by the rightmost line. The shaded region shows the timespan between T_1 and T_2. This plot serves to graphically display the disparity in the symptom event window size. The profiles primarily contributing to the first and second ODF peaks are indicated, as the two long vertical segments correspond to 0 and 2 hours. DCS events with T_1=0 hours generally correspond to short event windows, while those with T_2=2 hours are consistent with delayed DCS onset (as indicated by the large shaded region).

Histograms displaying the distributions of T_1 and T_2 can be found in Figure 6. The leftmost and rightmost bars in both histograms include all onset times before 6 hours prior to surfacing and after 6 hours post-dive respectively. The frequency of T_1

times peaks at 2 hours post-dive, with only one case having T_1 exceed 2 hours (by 1 minute). The frequency of T_2 times peaks immediately after surfacing and again after 2 hours post-dive, then a significant portion of T_2 times are after 6 hours post-dive.

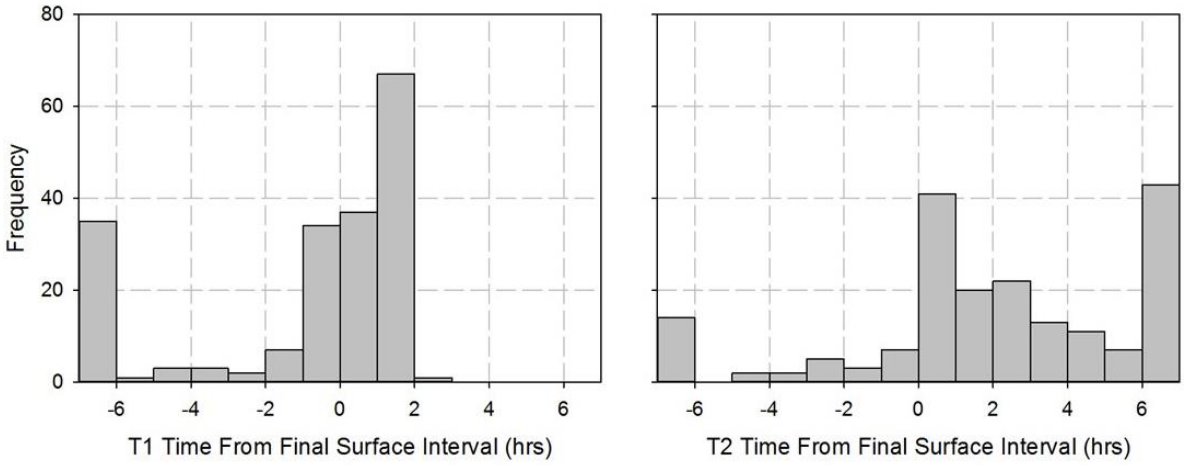

Figure 6: Distributions of T_1 and T_2 times of the DCS events in the BIG292 dataset. Time less than zero indicates that the onset of DCS occurred before the completion of the dive. The leftmost and rightmost bars include all onset times prior to -6 hours and after 6 hours relative to surfacing respectively. The most frequent T_1 time is 2 hours post-dive (as indicated by the peak frequency in the (1,2] bin range), with the only T_1 time exceeding 2 hours being 2 hours and 1 minute. Many of the T_1 times prior to -6 hours correspond to saturation diving. The peak in T_2 times during (0,1] hour post-dive relates to DCS events with rapid onset and correspond to T_1 times during decompression. The significant quantity of T_2 times after 6 hours post-dive indicates delayed symptom onset; many of these late T_2's correspond to T_1 at 2 hours.

2.3.1 Dive Type

DCS occurrence density functions were generated for each of the five dive type and breathing gas combinations included in the BIG292 dive dataset, excluding marginal DCS outcomes (Figure 7). The individual data files corresponding to each dive type and the shapes of the resulting ODF (bimodal, unimodal, or ambiguous) are reported in

Table 5. The single air, single non-air, repetitive and multi-level air, and repetitive and multi-level non-air ODFs have a bimodal shape. The first peak of all these plots occurs at 0 hours and the second peak occurs at 2 hours after the final decompression. Though these dive types differ in numbers of DCS occurrences, the timing of the peaks is consistent across dive type. The saturation dive ODF is ambiguous in shape, with a peak at 0 hours after decompression and a sharp decline in DCS occurrence after 2 hours. Unlike the bounce dives, saturation diving displays a substantial frequency of

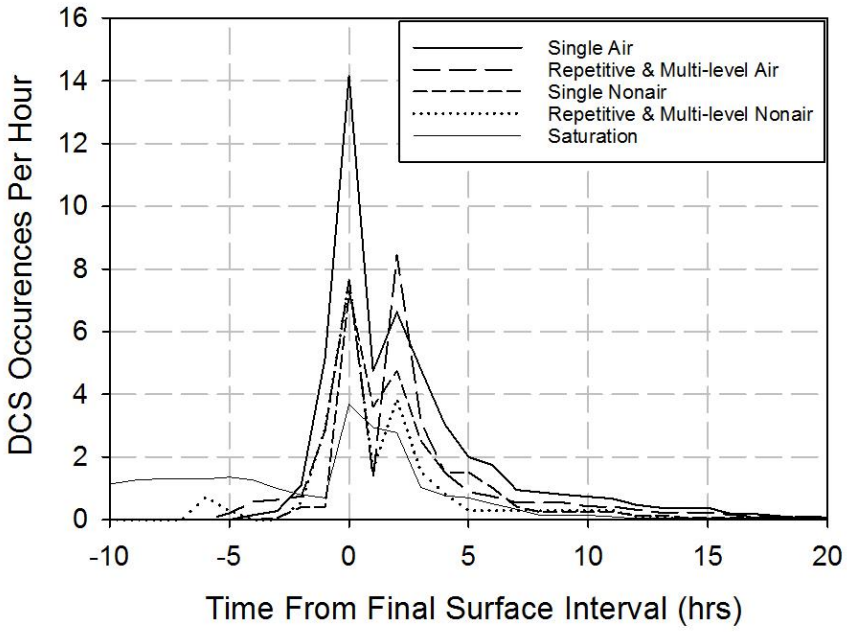

Figure 7: Occurrence density functions for the BIG292 dataset partitioned by dive type. See Table 5 for the categorization of the BIG292 data files. The ODFs for all dive types except saturation dives are bimodal, with the first peak at 0 hours and the second peak at 2 hours after the final decompression. Unlike the bounce dives, saturation diving displays a substantial frequency of DCS onset during decompression (prior to surfacing). However, this dive type difference does not affect the bimodality of the BIG292 ODF.

DCS onset during decompression (prior to surfacing). However, this dive type difference does not affect the bimodality of the BIG292 ODF.

2.3.2 Event Severity

Separate ODFs were generated for mild DCS and serious DCS cases, using both Type I/II and Type A/B data splitting and excluding marginal DCS cases. The distribution of PSI classifications into each splitting method is reported in Table 6, along with the number of DCS events in each mild or severe categorization and the shapes of the resulting ODFs. All plots are bimodal, with the first peak at 0 hours and the second peak at 2 hours after the final surface interval (Figure 8).

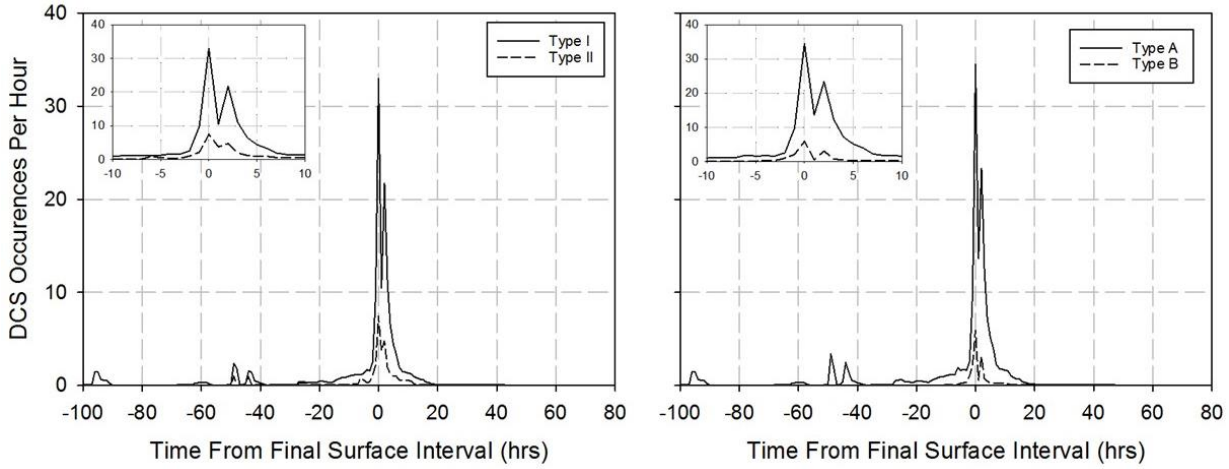

Figure 8: Occurrence density functions for the BIG292 dataset partitioned by event severity (Type I/II and Type A/B classifications). See Table 6 for categorization by PSI. All ODFs are bimodal, with the first peak at 0 hours and the second peak at 2 hours after the final decompression. The magnified views of each ODF in the top left corners more clearly illustrate the bimodal behavior of DCS symptom onset time.

2.3.3 Symptom Type

The shape of the ODF for each PSI is reported in Table 7. The ODFs for pain and serious neurological symptoms are bimodal, with the first peak at 0 hours and the second peak at 2 hours after the final decompression. The ODFs for constitutional, skin, mild neurological, and cardiopulmonary symptoms are ambiguously shaped, likely due to the low number of DCS events with these symptom types in the BIG292 data.

2.3.4 Institution

ODFs were generated for each individual dive report contained in the BIG292 dataset, then ODFs were created for all the dive reports published by each institution (Table 8). The NEDU 11-80 [36], NEDU 1-84 [37], and NSMRL 1200 [41] technical reports each generated ambiguously shaped ODFs, likely due to the low number of DCS events in each report. The NEDU 8-85 [38], NEDU 1-99/NMRC 99-01 [39], NMRI 86-97 [40], and all DCIEM technical reports (DCIEM 80-R-32 [42], DCIEM 81-R-02 [43], DCIEM 82-R-38 [44], DCIEM 84-R-72 [45], DCIEM 84-R-73 [46], DCIEM 85-R-18 [47]) each yielded bimodal ODFs. All bimodal ODFs had a first peak at 0 hours and a second peak at 2 hours after final decompression. All ambiguously shaped ODFs had a peak at 0 hours after the final decompression. However, by inspection of the T_1 and T_2 times corresponding to the ambiguous ODFs, these DCS event windows are consistent with those generating bimodal ODFs. Regardless of onset time measurement procedure, ODFs for all reports were either bimodal (with the first peak at 0 hours and second peak

at 2 hours after the final decompression) or ambiguously shaped; the latter likely due to insufficient quantity of DCS events documented by that particular dive report.

Saturation dive data were not partitioned by report or institution. The four saturation dive files (ASATARE, ASATNSM, ASATEDU, ASATNMR) contain data from 30 reports and dive test plans with publication spanning 1979-1992 [7, 8]. The removal of saturation dive data does not affect the bimodality of the ODFs of other dive types as many resulting DCS events occurred prior to surfacing (Figure 7), so further partitioning by dive report was not warranted.

2.3.5 Chronology of Dive Trials

ODFs were generated for dive trials completed during 1978-1983, 1984, 1985-1987, and 1988-1992 (Table 9). These temporal groupings were selected to each include similar quantities of DCS events. The ODFs for 1984, 1985-1987, and 1988-1992 are bimodal, with the first peak at 0 hours and the second peak at 2 hours following decompression. The ODF for dive trials completed during 1978-1983 is unimodal, with a peak at 0 hours after the final decompression (Figure 9).

2.4 Discussion

The goal of this work was to determine the cause of the bimodality in the BIG292 ODF to enable replication of this bimodal behavior in probabilistic models of DCS. The source of the bimodality would be revealed if a particular data partitioning method resulted in a set of unimodal ODFs, some peaking at 0 hours and others at 2 hours

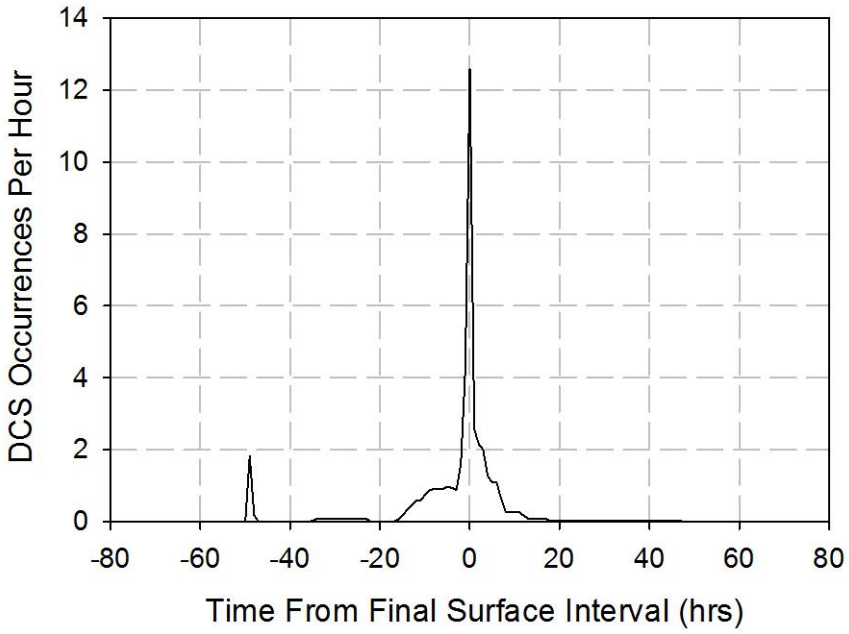

Figure 9: Occurrence density function for dive trials completed during 1978-1983 in the BIG292 dataset. See Table 9 for details about dive data files included in this time range. A total of 719 dives were conducted during this time, resulting in 44 DCS occurrences. Neglecting the two cases of DCS that occurred nearly 50 hours before the time of final surfacing, the ODF is unimodal, with a peak corresponding to the time of final surfacing.

following decompression. All but one iteration of data partitioning resulted in a bimodal or ambiguously shaped ODF; the latter likely due to an insufficient quantity of DCS events in that group.

Partitioning the data by dive type, DCS severity, DCS symptom type, or institution did not provide any insight into the source of the bimodal peak. The ODFs for single air, single non-air, repetitive and multi-level air, and repetitive and multi-level non-air dives were all bimodal. The ODF for saturation dives had only one peak, which occurred at 0 hours after decompression, but followed a bimodal shape, with a sharp

decline in DCS occurrences after 2 hours post-dive. No dive type could be identified as the source of either of the bimodal peaks. When analyzing Type I/II and Type A/B data splitting, the ODFs for all mild and severe categorizations were bimodal. Further, ODFs for each PSI with a sufficient number of DCS occurrences were bimodal. Therefore, the second peak in the BIG292 bimodal ODF cannot be attributed to a discrepancy in DCS onset time based on DCS symptom severity or type. Similarly, splitting the dive data based on the institution that conducted the dive trials, and further by each report published, generated either bimodal or ambiguously shaped ODFs; the latter likely due to an insufficient quantity of DCS occurrences in that dataset to produce a typical ODF. Thus, slight institutional differences in dive trial protocol are not responsible for the bimodality of the BIG292 ODF.

DCS occurrence data from dive trials completed during 1978-1983 produced a unimodal ODF. In all dive trial reports pertaining to the BIG292 dataset except the NEDU technical reports 11-80, 1-84, and 8-85, a strict onset time measurement protocol was established, in which all divers were examined by a DMO immediately after surfacing, and then again after 2 hours. NEDU technical reports 11-80, 1-84, and 8-85 do not explicitly document that divers were examined at 2 hours post-dive, although the NEDU 11-80 does indicate that divers were released from the hyperbaric chamber facility at this 2-hour mark. Thus, dive data produced by the NEDU during the 1978-1983 trial end date range may not have used the above onset time determination

protocol. In addition, one reviewer pointed out that the decompression schedules tested during this early date range were riskier, thus the onset of DCS events in these trials occurred during decompression or shortly thereafter. Many of the T_2 times reported in these trials are within 2 hours of decompression, and divers were under medical surveillance during this time. Thus, the corresponding onset time windows are smaller and the ODF is unimodal. As decompression schedules improved, divers were able to surface without incident but experienced DCS onset after surfacing. In these later trials, divers were monitored by a DMO for 2 hours (NEDU, NMRC, NSMRL) or 1.5-3 hours (DCIEM) before being released if no symptoms developed. For many DCS cases with onset after this window of post-dive medical surveillance, the T_1 time is 2 hours, which is when close medical examination of the diver ended (and corresponds to the second ODF peak). This bias towards a T_1 time of 2 hours with delayed-onset DCS is apparent in Figure 6, as the T_1 histogram contains a peak in the bin range (1,2]. Only one DCS case in the BIG292 dataset has a T_1 time greater than 2 hours, and this onset time is 2 hours and 1 minute. In Figure 5, DCS events with T_1=0 hours generally correspond to short event windows, while those with T_2=2 hours arise from delayed DCS onset (as indicated by the large shaded region). This is consistent with our conclusion that the T_1 for delayed DCS onset cases was biased towards T_1=2 hours. These long windows are a result of termination of continuous medical surveillance at 2 hours post-dive, thus the

last known time a diver was definitely asymptomatic was set as T_1=2 hours if symptoms developed after this time.

The bimodal shape of the BIG292 ODF is likely due to the medical surveillance protocol used in determining the onset time of DCS symptoms. It may be possible to eliminate this bimodality in future dive trial data if medical examination of divers is scheduled more frequently in the 24 hours following decompression. If divers were to be examined more often between 2 hours and 24 hours post-dive, there would be a lower number of T_1=2 hours cases, as T_1 would now reflect these later examination times. This could aid in shortening the event window, as T_1 and T_2 times would become more accurate.

Although computational replication of the bimodal shape of the BIG292 ODF would increase the likelihood, implying a better fit of the model to the data, doing so would be meaningless because there is not a distinct source (differences in dive type or DCS event severity) of the second peak. The single peak produced by current models of DCS which merges the bimodal peaks of the data in the ODF is very likely the correct solution. The ODF is still however a valuable tool for assessing the performance of models which predict the onset time of DCS. ODFs provide a qualitative metric to confirm that the onset time is being correctly predicted by the underlying model.

2.5 Conclusions

We investigated the potential sources of the bimodal shape of the BIG292 ODF with the goal of identifying features in the DCS cases that could potentially lead to an improvement in DCS model prediction. We found that bimodal shape of ODFs of DCS occurrences in empirical dive data were not related to dive types, severity of DCS symptoms, or symptom type. The DCS onset time determination protocol used by each institution that contributed human dive trial data to the U.S. Navy collection that became the BIG292 collection was reviewed. The common protocol between all institutions involved examination by a medical officer immediately after decompression and again in approximately 2 hours. The unimodal ODF corresponding to dive data from 1978-1983 is likely due to the higher risk decompression schedules that were tested during these trials, thus DCS onset occurred during decompression or shortly thereafter. In later trials with less risky decompression schedules (post-1984), DCS onset tended to occur after the 2-hour window of close medical surveillance. We conclude that the bimodality of the BIG292 ODF is likely due to a combination of delayed DCS onset in post-1984 trials, as decompression schedules became safer, and the protocol for determining DCS symptom onset time, in which divers were released from medical officer surveillance if no symptoms developed around 2 hours post-dive. When the time of symptom onset information is used in optimizing probabilistic DCS models, the log likelihood function implicitly aligns the model's ODF prediction with that of the

calibration data. Thus, if a model were able to replicate the data's bimodal ODF, the model would generate a greater log likelihood than a model which generated a unimodal ODF. This would lead one to conclude that the model replicating the bimodal ODF were the better of the two models. However, replicating the bimodal ODF shape in probabilistic models would likely be meaningless; as the ODF bimodality is not caused by dive type, DCS event severity, or symptom type but is rather related to the DCS onset time measurement protocol.

2.6 Acknowledgements

We would like to thank Dr. David Doolette of the Navy Experimental Diving Unit for his helpful discussion related to this work. We would also like to thank the five anonymous reviewers for their many helpful comments and suggestions. This work was supported by Naval Sea Systems Command (NAVSEA-00C - http://www.navsea.navy.mil/) under contracts N00024-13-C-4104 and N00024-17-C4317 and the Office of Naval Research under grant N000141310063. BelleQuant Engineering, PLLC provided computational resources. Neither the funding agency nor the commercial entity played any role in designing this study, data collection and analysis, decision to publish, interpreting the results, or writing the manuscript.

3. Trinomial Decompression Sickness Model using Full, Marginal, and Non-Event Outcomes

This chapter is adapted from a publication in *Computers in Biology and Medicine*.

Secondary authors are Nicholas Andriano and Dr. Laurens Howle.

[27] A.E. King, N.R. Andriano, L.E. Howle, *Trinomial Decompression Sickness Model using Full, Marginal, and Non-Event Outcomes*. Computers in Biology and Medicine, 118 (2020). https://doi.org/10.1016/j.compbiomed.2020.103640

3.1 Introduction

Decompression sickness (DCS) is a condition associated with reductions in ambient pressure during underwater diving, aviation, or other situations exposing humans to hyperbaric or hypobaric environments. Focusing here on hyperbaric exposures during deep-sea diving, inspired inert gas can dissolve into the blood at elevated pressures due to the hydrostatic pressure of water at depth. This dissolved gas is then circulated to the capillaries during diving, where it can diffuse into the tissues. Upon ascent, depressurization of the body can lead to supersaturation of the blood and tissues, causing the inert gas to leave solution and form bubbles, the putative initiating cause of DCS. The signs and symptoms of DCS can range from mild, such as joint pain or rash, to more severe, such as paralysis and death [3].

The Haldane decompression algorithm is credited as the first model used to generate decompression schedules to mitigate the occurrence of DCS [4]. This model used stage decompression to optimize the rate of inert gas washout by placing decompression stops at depths with substantial difference between the ambient pressure

and the inert gas tension in the body. This early model was deterministic and binary in outcome. Deterministic models predict the definite occurrence of DCS if the "safe" ascent criteria are violated, and likewise that the diver will not experience DCS if they adhere to the prescribed decompression schedule. Deterministic models do not allow for any variation in DCS outcome. However, this variation has been observed, as a single dive profile can result in DCS for some divers, and no DCS for others. In addition, a single individual repeating the same dive profile will not experience DCS every time [7, 8].

To address this variation in DCS outcome, Weathersby *et al.* [12] and Berghage *et al.* [11] introduced probabilistic DCS modeling, in which all dive profiles are assigned a non-zero probability of DCS. Probabilistic models use survival analysis to quantify risk, and allow for model parameters to be calibrated with empirical dive data [13]. However, neither deterministic nor probabilistic models used by the U.S. Navy to generate dive schedules address DCS symptom severity. Symptom severity is taken into account during U.S. Navy dive missions, as the Navy has established limits on the number of allowable cases of DCS for a given dive that vary based on symptom severity. The U.S. Navy categorizes symptom severity as Type I DCS (mild, pain-only) and Type II DCS (serious, cardiopulmonary or neurological), and will allow for a slightly higher probability of Type I than Type II DCS during dive planning [25]. Developing models to

predict both the probability and severity of DCS outcomes would provide dive planners with a more complete spectrum of risk information.

Previously, Howle *et al*. [10] explored multinomial probabilistic models in which the probabilities of multiple separate events are calculated simultaneously during model calibration. Howle derived and tested a trinomial severity model, which predicted probabilities of serious DCS, mild DCS, and no DCS. A third classification of DCS symptoms is called marginal DCS, in which symptoms associated with DCS last for only a short duration and resolve spontaneously without recompression treatment [7, 8]. For example, pain in one joint lasting for less than 60 minutes, or pain in two joints lasting less than 30 minutes, would be considered marginal DCS. Howle's trinomial severity model considered marginal DCS as non-events. The observed cases of serious and mild DCS were treated hierarchically, as a diagnosis of serious DCS would take precedence over and mask mild DCS if both types of symptoms were present. This model used a scaling factor to differentiate between serious and mild DCS during parameter calibration with empirical dive data. Two approaches to symptom severity classification were tested: the traditional Type I and Type II, and a novel classification system called Type A and Type B [9, 10]. Howle's Type A/B classification scale is defined by the Perceived Severity Index (PSI) [9, 10]. The PSI scale consists of six indices, listed here in order of increasing severity: constitutional or nonspecific (fatigue, dizziness, nausea), skin (rash, marbling), pain (joint pain, spasm, stiffness), mild neurological (numbness,

paresthesia), cardiopulmonary (cough, dyspnea), and serious neurological (dysfunction of bladder, coordination, mental status, vision). As a single DCS case could exhibit multiple symptom types, the PSI index is assigned corresponding to the most severe symptom present. In the traditional symptom classification system, Type I DCS would encompass constitutional, skin, and pain symptoms, and Type II DCS would include mild neurological, cardiopulmonary, and serious neurological manifestations. Howle *et al*. [10] presented an alternative approach, in which mild neurological symptoms are considered mild DCS rather than serious DCS. This system is called Type A/B splitting, in which Type A DCS includes constitutional, skin, pain, and mild neurological symptoms, and Type B DCS consists of cardiopulmonary and serious neurological cases. The categorization of DCS cases into mild and serious have important implications in U.S. Navy dive planning, as the Navy allows for a lower predicted probability of serious DCS than mild DCS for a given dive profile.

In some early probabilistic decompression models, marginal events were weighted as one-half of a full DCS event during model calibration. When U.S. Navy medical officers expressed a lower level of concern for marginal DCS, the weighting of marginal DCS was reduced to one-tenth of a full DCS event [16]. Further research on the efficacy of including marginal DCS as fractionally weighted events during probabilistic model fitting indicated that these events could skew the model's performance [31, 49].

In this chapter, we explore an alternative trinomial model, called the trinomial marginal model, in which the three states are full DCS, marginal DCS, and no DCS. Full DCS includes both mild and serious DCS, and marginal DCS is treated as a separate, fully-weighted event during model calibration. Full and marginal DCS are differentiated with a scaling factor during model calibration. The hierarchical nature of observed full and marginal DCS is reconciled with the competing probabilities used in the model. This trinomial marginal model is compared with its binomial counterpart and Howle's previous trinomial severity model.

3.2 Methods

3.2.1 Calibration Dataset

The BIG292 standard DCS dataset, available from two Naval Medical Research Institute (NMRI) reports [7, 8], was used in the fitting of all models in this work. No IRB approval was required for this work, as all the data used was de-identified and made publicly available in official Government reports. The dive trials described in this dataset were conducted by the U.S., U.K., and Canadian militaries from 1944-1997. The BIG292 data is a subset of the human dive trial data presented in the NMRI reports [7, 8], comprised of dive profiles and symptom histories for 3,322 air and N2-O2 exposures, including single bounce dives, repetitive and multilevel dives, and saturation dives conducted in hyperbaric chambers. These dive profiles range from 20 to 604.2 feet of

seawater (fsw) in depth and 0.64 to 12,960 minutes in duration. Of the 3,322 dive exposures, there were 190 cases of DCS, and 110 cases of marginal DCS.

In the BIG292 dataset, symptom onset time information is provided for all full DCS cases and 68/110 marginal DCS cases. This onset time information is given as times T1 and T2, where T1 is the last known time a diver was symptom-free, and T2 is the first known time the diver was symptomatic. Standard procedures have been established for determining this onset time window [50]. These symptom onset times can be used in calculating the probability of DCS and have been shown to improve a model's fit to the empirical dive data [14]; although this onset time window may be biased by the dive trial's medical surveillance protocol [26].

Based on the symptom histories given in the NMRI reports, the 190 full DCS cases can be further categorized into the U.S. Navy's traditional classification of severity, Type I (mild) and Type II (serious) [6], and Howle's alternative classification system of Type A (mild) and Type B (serious) [9, 10]. In the BIG292 dataset, there are 152 cases of Type I DCS and 38 cases of Type II DCS. Type A/B splitting, applied to the BIG292 dataset, yields 170 cases of Type A DCS (constitutional, skin, pain, and mild neurological) and 20 cases of Type B DCS (cardiopulmonary and serious neurological) [10]. The trinomial marginal model presented in this work groups mild and serious DCS into one category: full DCS.

3.2.2 DCS Models

Probabilistic models generally use parallel perfused compartments measuring gas content or bubble volume and survival analysis to calculate the probability of the occurrence of DCS [13]. The basic building block of these models is the Perfusion Limited Base Model (PLB), which is a single, well-stirred, perfused compartment. A set of three parallel PLB compartments, each with different parameter values, known as the EE1 model, has been studied and used extensively by the U.S. Navy. Our previous work investigated pharmacokinetic gas content models, which have coupled compartments that allow for gas transfer between compartments [30, 51]. These pharmacokinetic models perform well on specific dive types but do not outperform the U.S. Navy's model on the entire dataset, so they are not used here.

This chapter explores three variants of exponential-exponential (EE) and three variants of linear-exponential (LE) decompression models. All six models consist of three parallel, well-mixed compartments, each with a unique half-time for the uptake and elimination of inert gas, and have a baseline of six adjustable parameters (three tissue half-times and three gain parameters). The EE models use exponential gas kinetics in both gas uptake and elimination. Three variants of EE models were explored: EE1, EE1nt, and EE1 Full. The EE1 model includes a pressure threshold parameter in the third (slowest) compartment, which serves to modify risk accumulation. The EE1nt

model, "nt" indicating no threshold, does not contain any threshold parameters. The EE1 Full model uses a threshold parameter in all three compartments.

The LE models allow for a switch between exponential and linear gas kinetics during inert gas washout [15, 16]. Three variants of LE models were examined: LE1, LE1nt, and LE1 Full. The LE1 model builds upon the EE1 with the addition of a pressure crossover parameter in the second (intermediate) compartment, which dictates the pressure above which linear gas kinetics is used for gas wash-out rather than exponential wash-out. The LE1 model also contains the pressure threshold parameter in the third compartment. The LE1nt model does not contain any threshold parameter, but does include the aforementioned pressure crossover parameter in the second compartment. The LE1 Full model includes pressure crossover and threshold parameters in all three compartments, and thus has the highest number of degrees of freedom of all six models. A detailed derivation of these models can be found elsewhere [17].

3.2.3 Trinomial Marginal Model

Previous works by Weathersby *et al*. [12] and Thalmann *et al*. [15] modeled the probability of DCS with a binomial outcome: the probability of DCS and the probability of no DCS. The probability of DCS for a given dive is defined as

$$P(DCS) = 1 - e^{-\vec{g} \cdot \vec{R}} \tag{3.1}$$

where $\vec{g}$ and $\vec{R}$ are vectors of the compartmental gain and hazard respectively. The hazard function is derived from survival analysis and contains all information about gas kinetics; see [13, 17] for a thorough derivation. It follows that the probability of no DCS is

$$P(0) = 1 - P(DCS) = e^{-\vec{g}\bullet\vec{R}}. \tag{3.2}$$

The time of symptom onset information can be included in the hazard function and has been shown to improve model performance when used during model calibration [14]. This onset time information is used to create a joint probability of surviving DCS-free until T_1 and experiencing DCS between T_1 and T_2. This joint probability can be calculated as

$$P(DCS) = P(0)_{0\to T_1} P(DCS)_{T_1\to T_2} = e^{-\vec{g}\bullet\int_0^{T_1}\vec{r}dt}\left(1 - e^{-\vec{g}\bullet\int_{T_1}^{T_2}\vec{r}dt}\right). \tag{3.3}$$

The trinomial marginal model explored in this work uses a scale factor, a, to differentiate the probability of a full DCS event from a marginal DCS event, as shown in Eq. (3.4).

$$\begin{aligned} P_f^c &= 1 - e^{-a(\vec{g}\bullet\vec{R})} \\ P_n^c &= 1 - e^{-\vec{g}\bullet\vec{R}} \end{aligned} \tag{3.4}$$

Subscripts f and n correspond to full DCS and marginal DCS respectively, and the superscript c indicates that these are competitive probabilities. The term competitive

probability means P_f^c and P_n^c are the independent probabilities of full and marginal DCS occurring, respectively. The assignment of the scaling factor to the probability of full DCS is arbitrary. All mathematical notation used in this work is kept consistent with the previous work by Howle *et al*. [10] to aid in model comparison.

3.2.4 Competitive and Hierarchical Probabilities

The determination of a DCS outcome is dictated by the most severe symptom present. For example, a diver with symptoms associated with full DCS may also be experiencing marginal DCS, but the single recorded outcome will be full DCS. To reconcile this hierarchical system of classification of observed DCS with the competitive probabilities (Eq. (3.4)) used in modeling, Howle *et al.*'s [10] definition of hierarchical probabilities is applied here. The hierarchical probability of a less severe event occurring is defined as the joint probability of the competitive probability of the less severe event occurring and the competitive probability of the more severe event not occurring. In the trinomial marginal model, this means that the probability of a diver experiencing marginal DCS is multiplied by the probability of that diver not experiencing full DCS. Using Howle's compact notation, let

$$e^{-\vec{g}\cdot\vec{R}} = \xi \tag{3.5}$$

and

$$e^{-a(\vec{g}\cdot\vec{R})} = \xi^a. \tag{3.6}$$

The competitive probabilities in Eq. (3.4) can be converted to hierarchical probabilities as

$$P_f^h = P_f^c = 1 - \xi^a$$
$$P_n^h = P_n^c (1 - P_f^c) = \xi^a - \xi^{a+1} \tag{3.7}$$

where the superscript h indicates hierarchical probabilities. The probability of no DCS is thus

$$P_{0,tri\_marg}^h = 1 - P_f^h - P_n^h = \xi^{a+1} \tag{3.8}$$

as the law of total probability dictates that the sum of the probabilities of all events is equal to 1. It is important to note that the DCS outcomes recorded in empirical dive data correspond to hierarchical probabilities, rather than competitive probabilities.

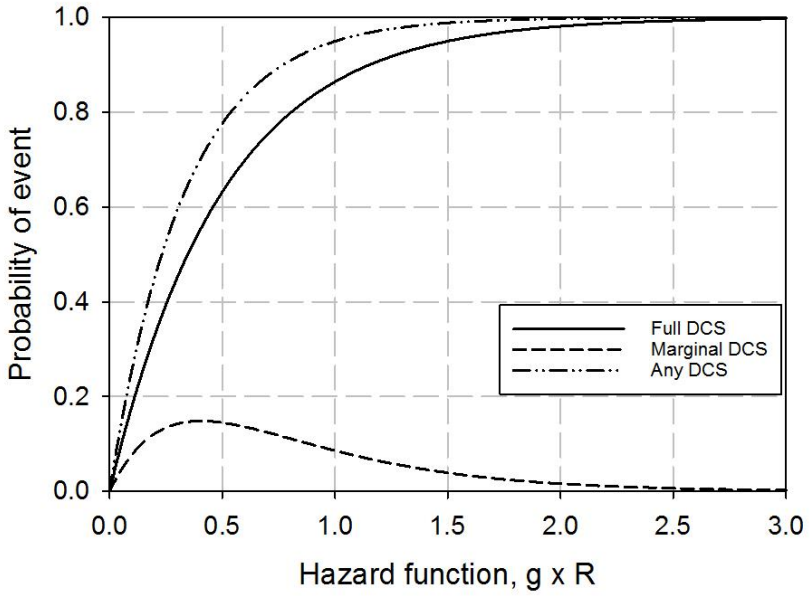

Figure 10: Probabilities of full and marginal DCS events with increasing value of the hazard function in the hierarchical model. The probability of observing marginal DCS decreases as the probability of full DCS increases, illustrating that full DCS masks marginal DCS. An arbitrary scale factor of $a = 2$ was used to generate these illustrative results.

Figure 10 plots the hierarchical probabilities for a single tissue compartment defined in Eqs. (3.7) and (3.8). A scale factor $a = 2$ was used to generate this plot, as there are roughly twice as many observed cases of full DCS as marginal DCS in the dataset. With increasing hazard function, the probability of marginal DCS increases then decreases. This decrease is due to the masking of marginal DCS by full DCS, as a diver becomes less likely to experience only marginal DCS with increasing risk.

3.2.5 Multinomial Likelihood Functions

To determine the model parameters which yield the best representation of the dive data, Weathersby *et al*. [12] applied the principle of maximum likelihood to decompression modeling. Bayesian estimation has also been explored for model fitting, and while it can provide additional information about the parameters' uncertainties, it has a much higher computational demand than likelihood maximization [34]. In this work, we use maximum likelihood parameter estimation, which uses repeated trials to select parameters that maximize the agreement between observed data and the model's predictions.

For a binomial model, the log likelihood function is defined as

$$LL_2 = \sum_{i=1}^{N} \ln\left[\left(1 - P(DCS)_i\right)^{1-\delta} \left(P(DCS)_i\right)^{\delta} \right] \tag{3.9}$$

where $P_{D,i}$ is the calculated probability of DCS occurring on dive profile i. For each dive, the exponent δ is set to 1 if DCS was observed, and 0 if no DCS was observed.

The natural log of this quantity is summed over all N dives in the dataset. Eq. (3.9) can be extended to our trinomial marginal model as

$$LL_3 = \sum_{i=1}^{N} \ln\left[\left(1 - P_{f,i}^h - P_{n,i}^h\right)^{1-\mu-\sigma} \left(P_{f,i}^h\right)^{\mu} \left(P_{n,i}^h\right)^{\sigma}\right] \tag{3.10}$$

where the exponents are set as $\mu = \sigma = 0$ for no DCS, $\mu = 1, \sigma = 0$ for full DCS, and $\mu = 0, \sigma = 1$ for marginal DCS. Thus, full DCS and marginal DCS are considered separate, fully-weighted hierarchical events.

3.2.6 DCS Model Optimization and Statistical Methods

Model parameters are determined through maximization of Eq. (3.10). A detailed explanation of the optimization used in this study can be found in [17]. This optimization can be quite computationally expensive, as some parameters are nearly collinear, resulting in an ill-conditioned Hessian matrix which slows convergence. To mitigate this computational demand, Howle analytically derived an exact solution for the optimal gain given the other parameters [52]. This decreases the number of parameters to be optimized, alleviating some of the computational intensity. A derivation of the exact gain calculation for a three-state model can be found in our previous work [10].

Each of the six model variants was optimized for 256 random initial guesses of the model parameter set. The log likelihood difference test was used for model comparison, as each model has a different number of adjustable parameters and cannot

be compared directly [13]. To categorize model improvement, a value of $p < 0.05$ was considered significant, and a value of $p < 0.01$ was considered highly significant. 95% confidence intervals on the predicted parameters were estimated according to the method presented in Ref. [13] by Dr. Gerth. In this method, The Hessian matrix was approximated by perturbing each parameter. The covariance matrix was calculated as the negative inverse Hessian, and the estimated parameter standard errors were taken as the diagonal components of this covariance matrix. SigmaPlot v14 was used to calculate and plot the 95% confidence limits and 95% prediction limits for the fit of the model to the data [53].

When comparing the success of each model, the log likelihoods cannot be compared directly when the models differ in degrees of freedom. The log likelihood difference test can be used to compare model fit [13], shown in Eq. (3.11)

$$\Delta LL_{ij} = \chi^2 = -2\left(LL_i - LL_j\right). \tag{3.11}$$

The log likelihood difference values between all models can be compared against the critical Chi-squared (χ^2) value for significance ($p < 0.05$) and the critical Chi-squared value for high significance ($p < 0.01$) based on the number of additional degrees of freedom from one model to the other. This log likelihood difference test can be used to determine if the addition of a parameter to the model with more degrees of freedom provides statistically significant improvement in model performance.

3.3 Results

The optimized trinomial marginal model variants were examined to determine the best performing model using the log likelihood difference test. The fit of this model to the empirical dive data was explored with the model's predicted vs. observed probabilities of DCS and the cumulative density functions for DCS events predicted by the model. The trinomial marginal model's predicted probabilities were also compared with the binomial model.

3.3.1 Trinomial Marginal Model Comparison

The optimal model parameter values and log likelihood values for all model variants can be found in Table 10, with the 95% confidence intervals for the LE1 model.

Table 10: Parameter values for the EE1 and LE1 model variants. 95% confidence intervals are provided for the LE1 model parameters, which yielded the best model fit to the data.

	EE1nt	EE1	EE1 Full	LE1nt	LE1	LE1 Full
$1/k_1$ (min)	1.340	1.092	1.932	1.570	3.509 ± 0.2165	5.426
$1/k_2$ (min)	227.7	231.1	566.8	228.2	63.79 ± 17.44	541.6
$1/k_3$ (min)	570.4	617.1	217.6	590.7	548.3 ± 45.09	146.4
g_1 (min^{-1})	1.465E-03	2.167E-03	1.227E-03	1.378E-03	4.962E-04 ± 3.002E-04	2.028E-04
g_2 (min^{-1})	2.545E-04	2.738E-04	7.830E-04	2.445E-04	6.303E-05 ± 1.661E-05	6.302E-04
g_3 (min^{-1})	1.770E-04	4.014E-04	2.408E-04	1.525E-04	7.327E-04 ± 1.570E-04	1.764E-04
PXO_1 (fsw)	∞	∞	∞	∞	∞	1.8787
PXO_2 (fsw)	∞	∞	∞	3.526E-01	7.467E-02 ± 1.065E-02	4.5414
PXO_3 (fsw)	∞	∞	∞	∞	∞	7.8777
Thr_1 (fsw)	0	0	1.298E-01	0	0	1.082E-02
Thr_2 (fsw)	0	0	1.635E-01	0	0	1.125E-01
Thr_3 (fsw)	0	1.230E-01	-2.701E-02	0	1.203E-01 ± 1.749E-02	-6.318E-02
a	1.609	1.762	1.547	1.732	1.596 ± 3.134E-03	1.452
LL_3	-1516.66	-1492.45	-1488.93	-1514.70	-1485.39	-1492.40

Table 11: Log likelihood difference comparison between trinomial marginal models determined by Chi-squared distributions for the respective degree of freedom differences. The degrees of freedom of each model are listed in parentheses. Each value in the table is the log likelihood difference comparison between the row and column models. A bold value indicates the model listed in the column offers significant improvement ($p < 0.05$) over the model listed in the row. A bold and underlined value indicates the column model demonstrates highly significant improvement ($p < 0.01$) over the row model.

	EE1nt (7)	EE1 (8)	LE1nt (8)	LE1 (9)	EE1 Full (10)	LE1 Full (13)
EE1nt (7)	-	**<u>48.413</u>**	**3.912**	**<u>62.524</u>**	**<u>55.462</u>**	**<u>48.520</u>**
EE1 (8)		-	-44.502	**<u>14.111</u>**	**7.049**	0.107
LE1nt (8)			-	**<u>58.612</u>**	**<u>51.550</u>**	**<u>44.609</u>**
LE1 (9)				-	-7.062	-14.004
EE1 Full (10)					-	-6.942
LE1 Full (13)						-

The log likelihood difference comparisons between each of the six models can be found in Table 11, located in the corresponding row-column intersections. The log likelihood difference values reading down each column are the comparisons between that column's model and those with fewer degrees of freedom, and reading left to right across a row are the comparisons between that row's model and those with more degrees of freedom. The LE1 model provides significant improvement over all models with fewer degrees of freedom, and no models with more degrees of freedom than the LE1 offer significant improvement. The addition of a pressure threshold parameter is justified, as the EE1 outperforms the EE1nt, and the LE1 outperforms the LE1nt. The LE1 model provides significant improvement over the EE1 model, justifying the addition of a single crossover pressure. The models with more degrees of freedom than

the LE1 (EE1 full and LE1 full) do not offer any significant improvement. From these log likelihood difference comparisons, we can conclude that the LE1 variant is the best performing trinomial marginal model.

The observed probability of DCS is plotted against the predicted probability of DCS in Figure 11 for the LE1 trinomial marginal model. This plot was generated by sorting all dive exposures in the dataset by the model's predicted probability of no DCS from smallest to largest. Exposures were then grouped into bins containing equal numbers of full (or marginal) events. For each bin, the predicted PDCS was calculated as the sum of the per-exposure probabilities of full (or marginal) DCS events in the bin divided by the actual number of observed cases of full (or marginal) DCS. The observed PDCS was calculated as the number of observed full (or marginal) DCS cases in the bin divided by the total number of exposures in the bin. Ten bins were used for full DCS (19 observed cases per bin) and five bins were used for marginal DCS (22 observed cases per bin). The full (triangles, left panel) and marginal (circles, right panel) data points were plotted with 95% confidence and 95% prediction bands, as represented by the long-dashed and solid lines respectively. The dotted lines represent the linear fit of the observed to predicted probability of DCS, with $r^2_{marginal} = 0.80$ and $r^2_{full} = 0.64$. A model that perfectly fits the dive data would have observed PDCS equal to predicted PDCS, thus all points would fall on the line of identity. It is interesting to note that the

marginal DCS data points are more linear than full DCS, but the full DCS regression line is closer to the line of identity.

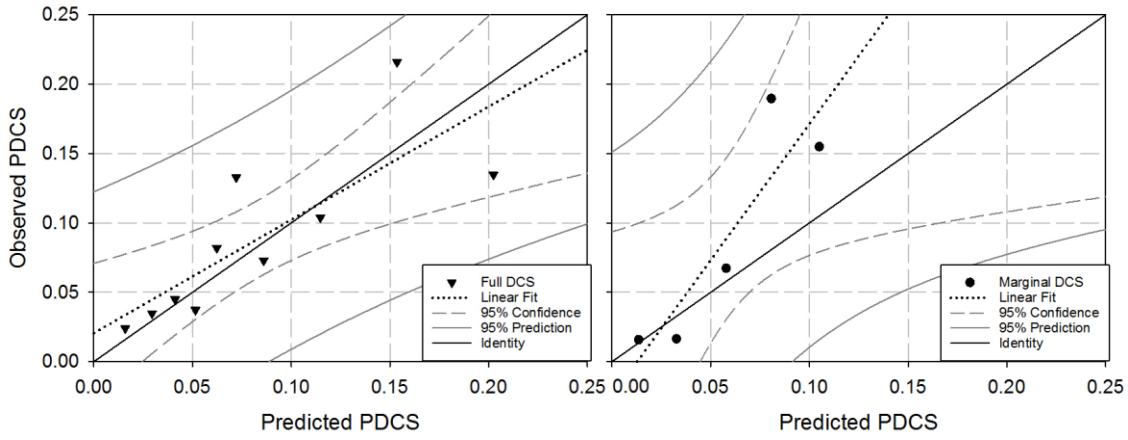

Figure 11: Trinomial marginal LE1-predicted probabilities of DCS versus observed probabilities of full (left) and marginal (right) DCS. These data points were generated by separating the marginal (full) DCS exposures into five (ten) bins, each containing 22 (19) observed marginal (full) DCS cases. The predicted and observed probabilities of DCS were then calculated for each bin. A larger bin size was used for the marginal DCS cases to best reflect the probability of observed DCS, where some dive profiles resulted in multiple observed marginal DCS cases. These probabilities are plotted with a linear fit (dotted lines, $r^2_{marginal} = 0.80$, $r^2_{full} = 0.64$) and the 95% confidence (long dashed line) and 95% prediction (solid line) bands.

3.3.2 Predictions on Data

Table 12 shows the number of observed and marginal DCS cases in the BIG292 dataset, along with the trinomial marginal LE1 model's predictions for full and marginal DCS. The 95% confidence bands are given for the total number of predicted marginal and full DCS cases. The dive data is categorized by dive type: single air, repetitive and

Table 12: DCS occurrences and trinomial marginal model predictions for the BIG292 dataset.

		Observed DCS			LE1 Trinomial Marginal Predicted DCS		
	Exposures	Full	Marginal	Total	Full	Marginal	Total
Single Air							
EDU885A	483	30	0	30	22.07	13.23	35.3
DC4W	244	8	4	12	4.59	2.81	7.4
SUBX87	58	2	0	2	0.16	0.1	0.26
NMRNSW	91	5	5	10	4.33	2.59	6.92
PASA	72	5	2	7	2.08	1.27	3.35
NSM6HR	57	3	2	5	3.46	2.05	5.51
Rep&Mult Air							
EDU885AR	182	11	0	11	9.4	5.59	14.99
DC4WR	12	3	0	3	0.74	0.44	1.18
PARA	135	7	3	10	7.8	4.62	12.42
PAMLA	236	13	12	25	15.7	9.24	24.94
Single Non-air							
NMR8697	477	11	18	29	12.27	7.48	19.75
EDU885M	81	4	0	4	2.42	1.48	3.9
EDU1180S	120	10	0	10	5.66	3.37	9.03
Rep&Mult Non-air							
EDU184	239	11	0	11	11.36	6.79	18.15
PAMLAOD	134	6	0	6	6.61	3.98	10.59
PAMLAOS	140	5	3	8	4.77	2.89	7.66
EDU885S	94	4	0	4	2.9	1.77	4.67
Saturation							
ASATEDU	120	13	27	40	16.53	9.06	25.59
ASATNMR	50	1	0	1	4.62	2.66	7.28
ASATNSM	132	18	21	39	25.17	13.17	38.34
ASATARE	165	20	13	33	20.08	11.22	31.3
Totals	*3322*	*190*	*110*	*300*	*182.72 ± 25.9*	*105.81 ± 15.0*	*288.53 ± 29.9*

multilevel air, single non-air, repetitive and multilevel non-air, and saturation dives. The model's predictions for total number of full DCS and marginal DCS cases do match the observed data within their 95% confidence intervals.

3.3.3 Binomial to Trinomial Probability Shift

The influence of marginal DCS events treated as separate, hierarchical events is illustrated in Figure 12, which shows the shift in the predicted probability of full DCS events between the trinomial marginal LE1 and binomial LE1 models. The parameters used for each model to generate Figure 12 can be found in Table 13. Marginal DCS cases are weighted as non-events in the binomial LE1 model, following previous work on the efficacy of marginal DCS events in model fitting [31, 49]. For this comparison of the trinomial marginal to binomial LE1 models, all dives that did not result in full DCS are categorized as no DCS (including marginals). In Figure 12, the trinomial marginal LE1 model predicted probabilities of full DCS for all observed full DCS cases are plotted against that of the binomial LE1 model (black circles). Likewise, for dive profiles that did not result in DCS, the trinomial marginal LE1 sum of the model predicted probabilities for no DCS and marginal DCS are plotted against the corresponding binomial LE1 predicted probabilities for no DCS. Observed marginal DCS cases are delineated from no DCS cases with the white square markers.

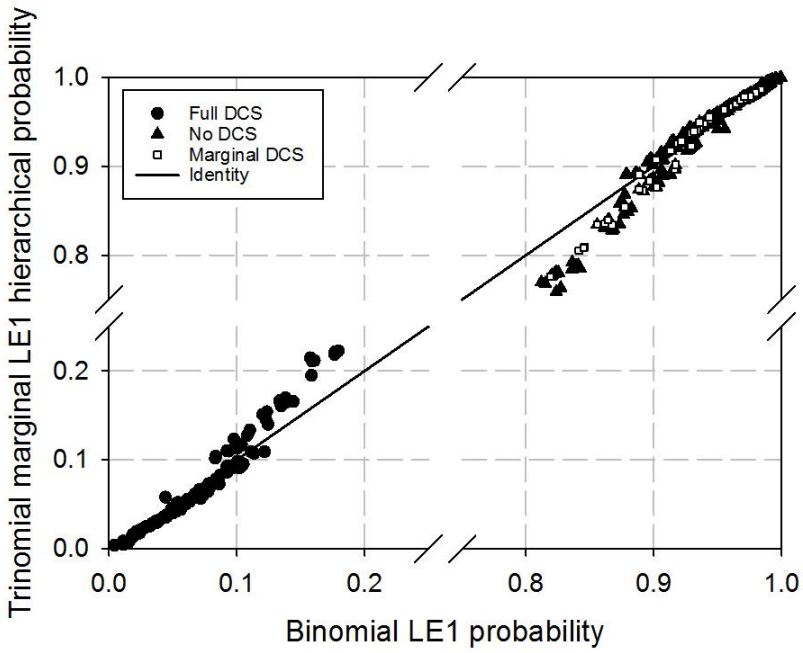

Figure 12: Binomial to trinomial probability shift plot. For the dives that resulted in DCS, the binomial LE1 predicted probability is compared to the predicted full DCS probability from the trinomial marginal LE1 model with solid black circles. For dive exposures that did not result in full DCS, the predicted probabilities of no DCS are shown with black triangles and white squares. The trinomial marginal model's plotted probabilities for no DCS are the sum of the predicted marginal and predicted no DCS probabilities. Exposures with observed marginal DCS are delineated with white square markers.

The line of identity in Figure 12 can be used to compare the predictions of both models. A point that falls below the line of identity indicates the trinomial marginal model has a lower predicted event probability than the binomial model. Conversely, a point above the line of identity indicates the trinomial marginal model has a higher predicted event probability than the binomial model. It is thus evident that for some dives, the trinomial marginal model predicts a lower probability of no DCS and a higher probability of full DCS than the binomial model.

Table 13: Optimized parameters used for the Binomial LE1 and Trinomial Marginal LE1 models in Figure 12.

	Binomial LE1	Trinomial Marginal LE1
$1/k_1$ (min)	2.350 ± 0.09401	3.509 ± 0.2165
$1/k_2$ (min)	63.02 ± 39.49	63.79 ± 17.44
$1/k_3$ (min)	504.4 ± 55.85	548.3 ± 45.09
g_1 (min^{-1})	1.985E-03 ± 1.200E-3	4.962E-04 ± 3.002E-04
g_2 (min^{-1})	1.069E-04 ± 4.523E-05	6.303E-05 ± 1.661E-05
g_3 (min^{-1})	9.665E-04 ± 2.658E-04	7.327E-04 ± 1.570E-04
PXO_1 (fsw)	∞	∞
PXO_2 (fsw)	7.236E-02 ± 2.052E-02	7.467E-02 ± 1.065E-02
PXO_3 (fsw)	∞	∞
Thr_1 (fsw)	0	0
Thr_2 (fsw)	0	0
Thr_3 (fsw)	9.963E-02 ± 1.459E-02	1.203E-01 ± 1.749E-02
a	-	1.596 ± 3.134E-03
LL_2	-963.308	-
LL_3	-	-1485.39

The linear regression line fitted to the full DCS predicted probabilities has a slope of 1.320 ($r^2 = 0.9627$) and the linear fit for the no DCS predicted probabilities has a slope of 1.226 ($r^2 = 0.9584$). This indicates there is slightly more scatter in the no DCS probabilities. These slopes can be used to approximate the probability shift from the binomial to the trinomial marginal model, i.e. $P_{tri,\,full} \approx 1.320 P_{bin,\,full}$.

3.3.4 Cumulative Density Function

The cumulative density functions plotted in Figure 13 can be used to visually compare DCS onset times for both the dive data and the model's predictions. For the

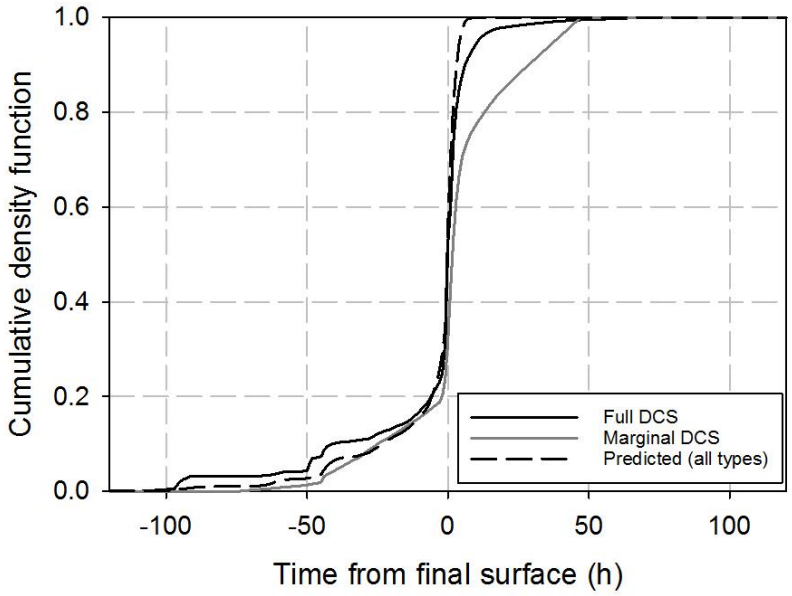

Figure 13: Trinomial marginal cumulative density function. Full DCS cases (black, solid curve) and marginal DCS cases (gray, solid curve) are shown for the empirical dive data. The cumulative density function for the predicted full and marginal DCS cases fall on the same curve (black, dashed).

predicted cases, the hierarchical probability model simply scales event probabilities to distinguish between marginal and full DCS. Thus the predicted cumulative density functions for both marginal and full DCS fall on the same curve (dashed black line). The observed full (solid black line) and marginal (solid gray line) DCS cumulative probabilities differ from approximately 100 hours prior to surfacing to 15 hours prior to surfacing, and from 0 to 48 hours after surfacing. The cumulative probability of observed marginal DCS cases is always less than that of full DCS cases. Prior to surfacing, the model under predicts the cumulative probability of full DCS, and after

surfacing, the model over predicts the cumulative probability of both full and marginal DCS. The model predicts that DCS events occur sooner after surfacing than seen in the data.

Figure 4 illustrates the disparity in the onset time information for full DCS versus marginal DCS in the data. While all full DCS cases are reported with the symptom onset time window T_1 and T_2, only 68 of the 110 marginal cases have this information. Thus, the T_2, or the first known time the diver was experiencing symptoms, for these cases without onset times was set to the right-censored time (24 or 48 hours after surfacing). The impact of this assignment of T_2 is evident in the marginal DCS cumulative density function, as the cumulative probability lags behind full DCS until 48 hours post-dive. The trinomial marginal model does not replicate this trend, and over-predicts both full and marginal DCS after the final surface interval.

3.4 Discussion

The trinomial marginal model explored in this work is a continuation of the multinomial modeling developed by Howle *et al*. [10]. The conversion between competitive probabilities, used by the model, to hierarchical probabilities, observed in the data, was adapted for full and marginal DCS, and a multinomial log likelihood function was used for model calibration. The analysis of the trinomial marginal model was conducted in accordance with our previous work [10] to expedite comparison.

Of the six model variants optimized in this work, the log likelihood difference test indicated the LE1 provided the best fit to the BIG292 dataset, as it performed significantly better than models with fewer degrees of freedom, and models with more degrees of freedom performed worse than the LE1. This justified using both the pressure crossover parameter in the second model compartment and the pressure threshold parameter in the third model compartment.

Although the LE1 trinomial marginal model is able to predict, within 95% confidence intervals, the number of observed full and marginal DCS cases, the distribution of these predictions among the dataset is not directly aligned with the observed cases. The model predicts full DCS more accurately than marginal DCS. It has been shown that the inclusion of marginal DCS in model calibration can impair a model's performance [31, 49]. Marginal DCS events make up only 3.3% of the BIG292 dataset, and 55% of these marginal events resulted from saturation diving. Because saturation diving only constitutes 14% of the BIG292 dataset, the frequency of observed marginal DCS causes undue risk to be associated with this dive type during model calibration.

The weighting of marginal DCS events in binomial models has historically been subject to debate and change. The original weighting of 0.5 was somewhat arbitrary, and was subsequently decreased to 0.1. The comparison of our tristate model, in which marginal DCS is considered a fully weighted event separate from full DCS, with

Howle's trinomial severity model [10], in which marginal events are classified as non-events, can provide insight into the additional impact of marginal events on model calibration. Howle demonstrated that his trinomial LE1nt severity model, which simultaneously predicted the probabilities of serious, mild, and no DCS, outperformed the binomial LE1nt model. We cannot compare our trinomial marginal LE1 model directly with the binomial LE1 model, as the treatment of marginal events differs between the two. When comparing the distribution of each of Howle's and our model's predictions of the data (see Table 12), it is interesting to investigate both models' success with saturation diving. In the dive data, a total of 52 full and 61 marginal DCS cases were observed resulting from saturation diving. The trinomial marginal model predicts 66.4 full DCS and 36.11 marginal DCS events. Howle's trinomial severity model predicted a total of 40.9 full DCS events (mild + serious), which is more accurate than the trinomial marginal model's corresponding prediction of full DCS. The trinomial marginal model only attributed roughly a third of the total predicted marginal cases to saturation diving, the rest distributed approximately uniformly among other dive types, resulting in the over-prediction of the occurrence of marginal DCS for single air, and repetitive and multilevel air and nonair diving.

The binomial to trinomial probability shift plot in Howle's work [10] indicated that both models' probabilities of no DCS were nearly identical, and his trinomial severity model predicted lower probabilities for full DCS, especially those with serious

symptoms, when compared with the binomial model. Howle's optimal model parameters for the trinomial severity model were found to be similar to the binomial model, generating the tight correlation between the binomial and trinomial severity models' predicted probabilities. The slopes of the linear fit lines for mild (0.894) and serious (0.109) DCS are nearly identical to the ratios of mild and serious DCS events to the total number of full DCS cases in the data, i.e. for mild, 170/190 = 0.895, and for serious, 20/190 = 0.105. Our binomial to trinomial marginal probability shift plot (Figure 12) has more scatter, and over-predicts the probability of full DCS when compared with the binomial counterpart. This illuminates one effect that the inclusion of marginal DCS cases has on model fitting, as the optimal parameters for the trinomial marginal are quite different from the binomial and trinomial severity models. In Figure 12, there is no trend in the marginal DCS data points that distinguishes them from no DCS events.

The cumulative density function for the BIG292 dataset and the model's predictions (Figure 13) indicates that the model under-predicts the probability of full DCS prior to surfacing, and over-predicts both full and marginal DCS after surfacing. The model is able to replicate the trend of full DCS onset, but does not predict the late onset of marginal DCS symptoms. As symptom onset time information is missing for almost 40% of the marginal cases, the assigned onset windows are quite large, ranging from one to five days. For comparison, the reported onset time windows for the other 68 marginal events range from 15 minutes to 30 hours (1.25 days). Because these reported

onset time windows are imprecise, the trinomial marginal model's inability to predict late onset for marginal DCS does not necessarily indicate poor performance.

3.5 Conclusions

The trinomial marginal model presented in this work is an augmentation of the traditional binomial probabilistic decompression model. This model predicts the probabilities of the occurrence of full, marginal, and no DCS simultaneously. The LE1 trinomial marginal model provided the best fit to the data, justifying the addition of both the pressure threshold and crossover pressure parameters. Analysis of the LE1 trinomial marginal model indicated there is room for improvement. The model's predicted probabilities of DCS do not align directly with observed probabilities of DCS. The lack of recorded symptom onset times for many of the marginal DCS cases and subsequent right censoring is unfavorable during the optimization process, as these large event windows make model parameters more difficult to refine than cases with recorded onset times.

Howle's trinomial severity model [10] and our trinomial marginal model treat marginal DCS cases differently; as Howle's model considers marginal DCS as a non-event, while we weight marginals as fully separate hierarchical events from full DCS. Because of this differential treatment, the two models cannot be directly compared. For future work, it is possible to derive a four-state (tetranomial) model, that simultaneously predicts mild, serious, marginal, and no DCS outcomes. This model could be compared

directly against the trinomial marginal model, with the eventual goal of determining which multinomial probabilistic model performs best, and whether these models provide a significant improvement over the binomial probabilistic model.

3.6 Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1644868 and by the U.S. Navy, Naval Sea Systems Command under contracts #N00024-13-C-4104 and #N00024-17-C-4317. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the U.S. Navy. Computational resources were provided by BelleQuant Engineering, PLLC. Neither of the funding agencies nor the commercial entity played any role in designing this study, data collection and analysis, decision to publish, interpreting the results, or writing the manuscript.

4. Tetranomial Decompression Sickness Model using Serious, Mild, Marginal, and Non-Event Outcomes

4.1 Introduction

Decompression sickness (DCS) is a condition resulting from a reduction in ambient pressure. This can occur during hyperbaric exposures, such as ascent from a deep-sea dive, and hypobaric exposures, such as ascent to altitude. When ambient pressure is reduced, inert gas which had been inspired, circulated, and dissolved into the body's blood and tissues at the previous elevated pressure can leave solution, forming bubbles and causing DCS. Signs and symptoms of DCS can range from mild skin rashes and joint pain to serious neurological and cardiological malfunction, and even death [3]. Marginal DCS is defined as symptoms typically associated with DCS that are mild and resolve spontaneously without recompression treatment, such as pain in one joint lasting for less than 60 minutes or pain in two joints lasting less than 30 minutes [7, 8]. Focusing in this work on hyperbaric exposures, DCS is of particular concern for U.S. Navy diver planners, as onset of symptoms can result in premature termination of undersea missions.

The first known decompression model to mitigate the risk of DCS was created by Boycott *et al*. [4] in the early twentieth century, known as the Haldane Model. The Haldane model generated decompression schedules using stage decompression to control the rate of inert gas washout from the body during ascent. This model was deterministic, meaning it predicted that DCS would absolutely occur if the proposed

"safe" ascent criteria were violated, and would not occur if these criteria were followed. While this early model did reduce the prevalence of DCS, some divers who complied with the prescribed "safe" decompression schedules still experienced DCS.

Probabilistic decompression modeling was introduced by Weathersby *et al*. [12] and Berghage *et al*. [11] to simulate the variation in DCS onset and severity experienced by divers executing the same dive profile as seen in empirical dive data [7, 8]. Probabilistic decompression algorithms use either gas content or bubble models and survival analysis to generate a probability of DCS for each dive profile [13]. A significant advantage of probabilistic modeling over deterministic modeling is that model parameters can be calibrated with empirical dive data. Probabilistic models used today to predict the probability of the occurrence of DCS do not provide any information about symptom severity. DCS severity predictions would be advantageous as they would allow safety analysis to be conducted on military diving operations.

Both the probability of DCS occurrence and symptom severity are of high concern to the U.S. Navy when planning undersea missions. When planning dives, the U.S. Navy has previously stated that a 2.0% risk of Type I (mild) DCS and a 0.1% risk of Type II (serious) DCS is acceptable [25]. Additionally, U.S. Navy Dive Medical Officers have indicated a low level of concern for marginal DCS [16]. DCS symptom onset can result in premature termination of U.S. Navy diving missions. The addition of the proposed multi-state probabilistic decompression model that predicts both the

occurrence and severity of DCS to dive planning would allow dive supervisors to tailor undersea missions to the acceptable level of risk for the divers.

Howle *et al*. [10] introduced a multinomial probabilistic decompression model, which simultaneously predicted the probability of three outcomes for a given dive profile: mild DCS, serious DCS, and no DCS. Howle tested two classifications of DCS cases as mild and serious based on the symptom histories published in the data set used for model calibration [7, 8], one in accordance with current U.S. Navy severity definitions [6] and one novel approach [9, 10]. Howle's trinomial model considered marginal DCS as non-events following previous research on the effectiveness of marginal events in probabilistic model calibration [31, 49]. This trinomial model was compared with a binomial model (predicting full DCS and no DCS outcomes), and it was concluded that the trinomial model provided statistically significant improvement over the binomial model in its ability to fit empirical dive data.

In the previous chapter, we modified Howle's trinomial model by analyzing the multi-state outcome of full DCS, marginal DCS, and no DCS [27]. Historically, marginal DCS events have been included in probabilistic decompression models as fractionally weighted during model calibration. Originally, marginal events were assigned a weighting of 0.5, indicating they were half as important as a full DCS event during model fitting. This weighting was later reduced to 0.1 when U.S. Navy Medical Officers indicated a low level of concern for marginal DCS, to ensure that marginal DCS cases

did not cause undo risk to be associated with particular dives during model calibration [16]. Further research on the impact of fractionally weighted marginal events in probabilistic model fitting has indicated that fractionally weighted marginal DCS events may hinder a model's performance [31, 49]. To address this issue, we developed the aforementioned trinomial marginal model, which considered marginal DCS to be a fully-weighted hierarchical outcome separate from full DCS. This model could not be compared directly with Howle's trinomial model, which classified marginal DCS as nonevents, though we found the inclusion of marginal events in this fashion may have skewed the distribution of predictions on the data. In this chapter, we continue the investigation of multinomial probabilistic modeling by optimizing a tetranomial model with mild DCS, serious DCS, marginal DCS, and no DCS outcomes.

4.2 Methods

4.2.1 Calibration Data

The model presented in this study was calibrated with the BIG292 standard DCS data set, which is a subset of data presented in two Naval Medical Research Institute (NMRI) reports [7, 8]. The BIG292 data contains 3,322 exposures of air and nitrogen-oxygen diving conducted by the United States, United Kingdom, and Canadian militaries from 1944-1997. This data set includes the dive profile, dive conditions (wet or dry), inspired gas, and DCS outcome and symptom history for each exposure. The BIG292 data set contains a total of 190 DCS cases and 110 marginal DCS cases resulting

from single air, single non-air, repetitive and multilevel air, repetitive and multilevel non-air, and saturation dive types. Marginal DCS is defined as signs or symptoms associated with DCS that persist for a short duration and spontaneously resolve without recompression treatment [7, 8]. The dive data used in this study are de-identified and available to the public in the form of two U.S. government reports, and no IRB approval was required for the present study.

If DCS occurs, the onset time window of DCS symptoms can be characterized by times T_1 and T_2, where T_1 is the last known time a diver was asymptomatic and T_2 is the first known time a diver was definitely experiencing DCS symptoms [16]. In the BIG292 data set, all 190 full DCS cases and 68 of the 110 marginal DCS cases are reported with symptom onset times T_1 and T_2. These symptom onset times can be used in probabilistic DCS modeling to improve model fitting [14]. In our previous work, we found the onset time window provided by T_1 and T_2 are not related to DCS symptom severity, and may actually be biased by the medical surveillance protocol of each dive trial [26].

4.2.2 DCS Event Severity

DCS cases are categorized into Type I (also called mild or pain-only) or Type II (also called serious or neurological) [6, 25]. A novel method of categorizing DCS cases was proposed by Howle *et al.*, in which the 190 full DCS cases in our calibration data set are classified by perceived severity index (PSI) [9, 10]. These indices for describing DCS symptoms, in order of least to most severe, are: constitutional/nonspecific (dizziness,

fatigue, nausea), lymphatic/skin (itching, rash, marbling), pain (ache, joint pain, spasm), mild neurological (paresthesia, numbness, tingling), cardiopulmonary (hemoptysis, dyspnea, cough), and serious neurological (dysfunction of bladder, coordination, mental status).

The dive data published in the two NMRI reports [7, 8] included symptom descriptions for each case of full and marginal DCS, so Howle *et al*. assigned each case a severity index 1-6 [10]. If a case exhibited symptoms corresponding to more than one severity category, the most severe index present was selected.

The traditional categorization of Type I DCS corresponds to constitutional, skin, and pain manifestations, while mild neurological, cardiopulmonary, and serious neurological cases are considered Type II DCS [6]. Howle *et al*. [10] proposed Type A/B splitting, in which Type A DCS includes constitutional, skin, pain, and mild neurological symptoms, while Type B DCS corresponds to cardiopulmonary and serious neurological. The number of DCS occurrences in the BIG292 data set corresponding to each PSI and classified by both Type I/II and Type A/B splitting are summarized in Table 1 (in Chapter 1).

4.2.3 DCS Models

Probabilistic DCS models use survival analysis with a gas content or bubble volume model to quantify the risk of DCS occurrence for a given dive profile [13]. Our tetranomial probabilistic models extended the exponential-exponential (EE) and linear-

exponential (LE) gas content models described by Thalmann [15], which consist of three stirred, parallel perfused gas compartments. The models in this chapter were added to our previously developed DCS modeling and optimization system, described in Refs. [17, 52].

Twelve probabilistic decompression model variants were tested to determine which model parameters were statistically justified for the tetranomial model. The base model was the EE1 model, which consists of three well-mixed, parallel-perfused compartments. Each compartment exhibits exponential gas kinetics and has a unique half-time. The slowest compartment has a pressure threshold parameter, allowing for greater gas supersaturation before risk accumulation. We also tested two additional variants of this EE1 model – one without any threshold parameter (EE1nt), and one with threshold parameters in all three compartments (EE1 full).

Next, we tested the LE1 model, which augments the EE1 model by allowing for a switch between exponential and linear gas kinetics at an optimized crossover pressure in the middle compartment [15, 16]. The two variants of that LE1 model were one without any threshold parameter (LE1nt), and one with both threshold and crossover pressure parameters in all three compartments (LE1 full). A detailed derivation of these models can be found in Ref. [17]. A summary of the free parameters in each model variant can be found in Table 15.

These three EE1 and three LE1 models were tested with Type I/Type II splitting, and again with Type A/Type B splitting, totaling 12 model variants.

4.2.4 Tetranomial Model

The binomial probability of DCS occurrence, as defined by Weathersby *et al*. [12], is

$$P(DCS) = 1 - e^{-\vec{g}\cdot\vec{R}} \tag{4.1}$$

and the probability of DCS not occurring is dictated by the law of total probability as

$$P(0) = 1 - P(DCS) = e^{-\vec{g}\cdot\vec{R}}, \tag{4.2}$$

where $\vec{g}$ is a vector of each compartment's gain and $\vec{R}$ is a vector containing each compartment's risk information. The risk function is derived from survival analysis and quantifies the gas kinetics in each compartment; a detailed derivation can be found in Refs. [13, 17].

It has been shown that including the DCS symptom onset time information in Eq. (4.1) can improve a model's performance, and is done by calculating the joint probability of surviving DCS-free until T_1 and experiencing DCS during the onset time window T_1-T_2 [14]. This joint probability can be written as

$$\begin{aligned} P(DCS) &= P(0)_{0\to T_1} P(DCS)_{T_1\to T_2} \\ &= \left(e^{-\vec{g}\cdot\int_0^{T_1}\vec{r}dt} \right)\left(1 - e^{-\vec{g}\cdot\int_0^{T_1}\vec{r}dt} \right). \end{aligned} \tag{4.3}$$

These equations can be extended to the proposed tetranomial model, in which the probabilities of serious, mild, marginal, and no DCS are all calculated simultaneously. Competitive probabilities, meaning probabilities for each event independent of any other event occurring, are derived from Eq. (4.1) using fitted scale factors a and b to differentiate between DCS severity:

$$\begin{aligned} P_s^c &= 1 - e^{-a(\vec{g}\cdot\vec{R})} \\ P_m^c &= 1 - e^{-\vec{g}\cdot\vec{R}} \\ P_n^c &= 1 - e^{-b(\vec{g}\cdot\vec{R})}. \end{aligned} \tag{4.4}$$

In Eq. (4.4), P_s^c, P_m^c, and P_n^c are the competitive probabilities of serious, mild, and marginal DCS respectively.

4.2.5 Competitive and Hierarchical Probabilities

Observed cases of DCS are categorized hierarchically. For example, the diagnosis of serious DCS would take precedence over mild and marginal DCS if mild and/or marginal DCS symptoms were present, and mild DCS takes precedence over marginal DCS. The calculated DCS probabilities in Eq. (4.4) are defined competitively, and consequently must be converted to hierarchical probabilities to accurately reflect the diagnoses in our dive data. These hierarchical probabilities, labeled with a superscript h, can be calculated from competitive probabilities as the joint probability of the event's independent probability and the probability that the more severe event(s) does not occur:

$$
\begin{aligned}
P_s^h &= P_s^c \\
P_m^h &= P_m^c \left(1 - P_s^c\right) \\
P_n^h &= P_n^c \left(1 - P_s^c\right)\left(1 - P_m^c\right) \\
P_{0,tet}^h &= 1 - P_s^h - P_m^h - P_n^h
\end{aligned}
\tag{4.5}
$$

The sum of the probabilities of all events is equal to 1 by the law of total probability. For comparison with Ref. [10] and Chapter 3, we can rewrite Eq. (4.5) in Howle's compact notation, where a quantity ξ is defined as

$$
e^{-\vec{g}\bullet\vec{R}} \tag{4.6}
$$

and

$$
\begin{aligned}
e^{-a(\vec{g}\bullet\vec{R})} &= \xi^a \\
e^{-b(\vec{g}\bullet\vec{R})} &= \xi^b .
\end{aligned}
\tag{4.7}
$$

Eqs. (4.6) and (4.7) can be substituted into the hierarchical probabilities defined in Eq. (4.5):

$$
\begin{aligned}
P_s^h &= 1 - \xi^a \\
P_m^h &= \xi^a - \xi^{a+1} \\
P_n^h &= \xi^{a+1} - \xi^{a+b+1} \\
P_{0,tet}^h &= \xi^{a+b+1}
\end{aligned}
\tag{4.8}
$$

The hierarchical probabilities of serious, mild, and marginal DCS are plotted with increasing hazard function for a single compartment in Figure 14. The probability of serious DCS increases with increasing hazard function, while the probabilities of mild and marginal DCS increase and then decrease. This plot illustrates the masking of less severe DCS events by more severe DCS, i.e. a diver diagnosed with serious DCS may

have also been experiencing mild DCS symptoms. We hypothesize that as the risk function increases, it is more likely that the diver will develop serious DCS symptoms and thus more likely to be diagnosed with serious DCS and less likely to be diagnosed with mild or marginal DCS.

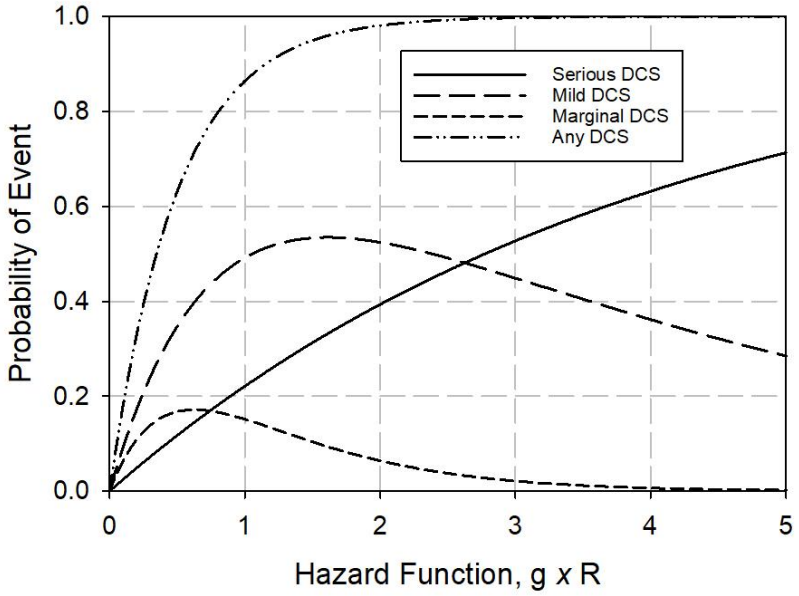

Figure 14: Probabilities of serious, mild, and marginal DCS events with increasing hazard function in the hierarchical model. The masking of mild DCS by serious DCS, and marginal DCS by mild and serious DCS, is illustrated by the decreasing probabilities of mild and marginal DCS events with increasing hazard function. Arbitrary scale factors of $a = 0.25$ and $b = 0.75$ were used to generate this plot.

4.2.6 Multinomial Likelihood Functions

Probabilistic DCS models are advantageous in their capacity to be calibrated with empirical dive data. To determine optimal model parameters, Weathersby *et al*. [12] used maximization of the likelihood function. Other optimization methods, such as

Bayes optimization, have also been used to estimate probabilistic DCS model parameters [34]. Although Bayesian optimization can provide a clearer picture of estimated parameters' uncertainties, it has a high computational cost, so maximum likelihood optimization is used in the present work.

For a binomial model predicting the probabilities of full and no DCS, the log likelihood function is

$$LL_2 = \sum_{i=1}^{N} \ln\left[P(DCS)_i^{\delta} \left(1 - P(DCS)_i^{1-\delta}\right)\right], \tag{4.9}$$

where $P(DCS)_i$ is the probability of DCS occurring for the i^{th} of N total dives, calculated with Eq. (4.1) or (4.3). The exponent δ signals the observed outcome of the i^{th} dive, where $\delta = 1$ if DCS occurred, and $\delta = 0$ if DCS did not occur. This function is optimized to maximize the model's fit to the data.

For our tetranomial model, the hierarchical probabilities defined in Eq. (4.5) can be used in a multinomial log likelihood function to calculate the fit of the model to the calibration data set:

$$LL_4 = \sum_{i=1}^{N} \ln\left[\left(P_{s,i}^h\right)^{\sigma} \left(P_{m,i}^h\right)^{\mu} \left(P_{n,i}^h\right)^{\nu} \left(1 - P_{s,i}^h - P_{m,i}^h - P_{n,i}^h\right)^{1-\sigma-\mu-\nu} \right], \tag{4.10}$$

where index i counts over each dive exposure and the observed dive outcome is expressed with

$$\begin{aligned}
&\sigma = 1, \mu = \nu = 0 \quad \text{for serious DCS} \\
&\mu = 1, \sigma = \nu = 0 \quad \text{for mild DCS} \\
&\nu = 1, \sigma = \mu = 0 \quad \text{for marginal DCS} \\
&\sigma = \mu = \nu = 0 \quad \text{for no DCS.}
\end{aligned} \tag{4.11}$$

The model is optimized with serious, mild, and marginal DCS treated as separate, hierarchical events distinguished by scaling factors.

We can collapse the tetranomial log likelihood function in Eq. (4.10) to an equivalent trinomial marginal log likelihood function by combining the probabilities of serious and mild DCS to represent full DCS as

$$LL_{43} = \sum_{i=1}^{N} \ln\left[\left(P_{s,i}^{h} + P_{m,i}^{h} \right)^{\sigma+\mu} \left(P_{n,i}^{h} \right)^{\nu} \left(1 - P_{s,i}^{h} - P_{m,i}^{h} - P_{n,i}^{h} \right)^{1-\sigma-\mu-\nu} \right]. \tag{4.12}$$

We will use this deflated log likelihood to compare the tetranomial model in this work with the trinomial marginal model in the previous chapter [27].

4.2.7 DCS Model Optimization and Statistical Methods

The optimal parameters for the tetranomial model were determined with maximization of the tetranomial log likelihood function in Eq. (4.10). A thorough description of the maximization technique used herein can be found in Ref. [17]. The optimization of Eq. (4.10) is computationally expensive because some model parameters are nearly collinear. To reduce the number of optimized parameters, Howle previous derived an analytical solution for the optimal compartmental gain values given the rest of the parameter set [52], which can be extended to these multinomial models [10].

All 12 model variants were optimized from 1024 random initial guesses, and the parameter set yielding the maximum log likelihood was chosen for each model. Because these model variants differ in the number of adjustable parameters, their log likelihoods cannot be compared directly, so the log likelihood difference test was used, defined in [13] as

$$\Delta LL_{ij} = \chi^2 = -2\left(LL_i - LL_j\right), \tag{4.13}$$

Where LL_i and LL_j are the log likelihoods of the models being compared. The log likelihood difference comparison value, ΔLL_{ij}, for each model pair can be compared against the Chi-squared distribution value for significant ($p < 0.05$) or highly significant ($p < 0.01$) improvement based on the number of additional degrees of freedom from one model to the other.

The 95% confidence intervals on the optimized parameters were calculated according to Ref. [13]. In this method, the covariance matrix was taken as the negative inverse of the approximate Hessian, and the estimated parameter standard errors were the diagonal components of this covariance matrix.

SigmaPlot v14 [53] was used to generate and plot the 95% confidence limits and 95% prediction limits on the models' fits to the data set.

4.3 Results

In the subsections below, all 12 optimized model variants are compared and the best performing model chosen using the log likelihood difference test. The best model's

predictions on the dive data set are examined, along with the cumulative density function for predicted cases and predicted vs. observed probabilities of DCS. The tetranomial model is then compared with the trinomial and trinomial marginal models from our previous work.

4.3.1 Tetranomial Model Comparison

For each of the 12 model variants, the parameter sets yielding the best log likelihood were chosen for comparison. The log likelihoods of each splitting type (I/II and A/B) model pair can be compared directly, and for all six pairs, the A/B models performed better than the corresponding I/II models (Table 14). The optimal parameter

Table 14: Maximum log likelihood for each of the 12 tetranomial models optimized from 1024 random initial guesses. Each of the six models (EE1, EE1nt, EE1 Full, LE1, LE1nt, and LE1 Full) was tested with both Type I/II and Type A/B DCS severity splitting. For each of these six models, the log likelihoods of Type I/II and Type A/B splitting can be compared directly to determine which splitting method yields the best model performance.

Model	# DOF	LL	Severity Splitting Type Winner
EE1 NT I/II	8	-1612.30041	EE1 NT A/B
EE1 NT A/B	8	-1581.05407	
EE1 I/II	9	-1589.44908	EE1 A/B
EE1 A/B	9	-1560.50726	
LE1 NT I/II	9	-1609.74076	LE1 NT A/B
LE1 NT A/B	9	-1578.65117	
LE1 I/II	10	-1583.42341	LE1 A/B
LE1 A/B	10	-1549.5327	
EE1 Full I/II	11	-1588.09186	EE1 Full A/B
EE1 Full A/B	11	-1559.70088	
LE1 Full I/II	14	-1586.30942	LE1 Full A/B
LE1 Full A/B	14	-1562.72073	

sets for these six A/B splitting models (EE1, EE1nt, EE1 Full, LE1, LE1nt, and LE1 Full) can be found in Table 15, along with the 95% confidence intervals for the LE1 model.

Table 15: Optimal model parameters for all EE1 and LE1 model variants. All model variants in the above table used Type A/B splitting. 95% confidence intervals are given for the LE1 model parameters, which provided the best fit to the BIG292 data set.

	EE1nt	EE1	LE1nt	LE1	EE1 Full	LE1 Full
$1/k_1$ (min)	2.295	4.957	3.001	3.496 ± 0.1510	7.802	1.585
$1/k_2$ (min)	245.9	267.6	211.5	63.83 ± 22.86	496.6	578.2
$1/k_3$ (min)	619.6	619.2	607.8	548.1 ± 42.89	149.4	151.0
g_1 (min^{-1})	1.212E-03	3.767E-04	7.665E-04	7.138E-04 ± 4.337E-04	3.743E-04	1.085E-03
g_2 (min^{-1})	4.222E-04	4.719E-04	3.458E-04	9.036E-05 ± 2.603E-05	1.110E-03	5.991E-04
g_3 (min^{-1})	2.032E-04	1.363E-03	2.369E-04	1.049E-03 ± 2.129E-04	1.226E-04	3.461E-04
PXO_1 (fsw)	∞	∞	∞	∞	∞	2.429
PXO_2 (fsw)	∞	∞	0.2897	0.07471 ± 0.01127	∞	4.821
PXO_3 (fsw)	∞	∞	∞	∞	∞	2.708
Thr_1 (fsw)	0	0	0	0	0.07158	0.1220
Thr_2 (fsw)	0	0	0	0	0.1127	0.08404
Thr_3 (fsw)	0	0.2185	0	0.1202 ± 0.01134	-0.06614	-0.02619
a	0.1134	0.1087	0.1124	0.1127 ± 0.01552	0.1235	0.1250
b	0.6756	0.6489	0.6869	0.6981 ± 0.01142	0.7173	0.6000
P(N)	106.83	100.99	107.1	105.83 ± 12.57	102.46	93.11
P(M)	167.41	167.88	165.31	163.2 ± 20.78	153.45	165.27
P(S)	19.72	19.23	19.31	19.31 ± 3.566	19.87	21.6
LL_4	-1581.05	-1560.51	-1578.65	-1549.53	-1559.70	-1562.72

The comparisons between the six A/B model variants, which differ in the number of degrees of freedom, were performed with the log likelihood difference test. These log likelihood difference test values (ΔLL_{ij}) can be found in Table 16 for all A/B splitting models, and the Chi-squared distribution values for one to six additional degrees of freedom are in Table 17. In Table 16, the number of adjustable parameters for each

model is listed in parenthesis. The log likelihood difference value between each model pair is listed in the corresponding row-column intersection. Reading down a column compares that column's model to models with less degrees of freedom, and reading across a row compares that row's model with models having more degrees of freedom. A bold value indicates the column model provides significant improvement ($p < 0.05$) over the row's model, and a bold and underlined value indicates the column model provides highly significant improvement ($p < 0.01$) over the row's model. We can see that the use of one pressure threshold parameter is justified, as the EE1 and LE1 provided highly significant improvement over the EE1nt and LE1nt respectively. The crossover pressure parameter enabled the LE1 model to perform significantly better than the EE1 model. However, the EE1 Full and LE1 Full models did offer significant

Table 16: Log likelihood difference comparison for all tetranomial models using Type A/B splitting. Each model's number of adjustable parameters is listed in parenthesis. The log likelihood difference value between any two of the six models is located in the corresponding row-column intersection. Values in bold indicate the column model provides significant improvement (p < 0.05) over the row model, and bold and underlined values indicate the column model provides highly significant improvement (p < 0.01) over the row model.

	EE1nt (8)	EE1 (9)	LE1nt (9)	LE1 (10)	EE1 Full (11)	LE1 Full (14)
EE1nt (8)	-	**<u>41.094</u>**	**4.806**	**<u>63.043</u>**	**<u>42.706</u>**	**<u>36.667</u>**
EE1 (9)		-	-36.288	**<u>21.949</u>**	1.613	-4.427
LE1nt (9)			-	**<u>58.237</u>**	**<u>37.901</u>**	**<u>31.861</u>**
LE1 (10)				-	-20.336	-26.376
EE1 Full (11)					-	-6.040
LE1 Full (14)						-

improvement over the EE1 and LE1 models respectively, so the addition of threshold and crossover pressure parameters to all compartments is not justified. We can conclude that the LE1 model provided the best fit to our data, as the LE1 model provided highly significant improvement over the EE1, EE1nt, and LE1nt models. Models with more adjustable parameters than the LE1 (the EE1 Full and LE1 Full) did not offer any improvement. Therefore, all following discussion will pertain to the LE1 model with Type A/B splitting.

Table 17: Chi-squared distribution values for 0.95 ($p < 0.05$) and 0.99 ($p < 0.01$) based on the number of additional degrees of freedom.

Δ DOF	$p<0.05$	$p<0.01$
1	3.841	6.635
2	5.991	9.210
3	7.815	11.345
4	9.488	13.277
5	11.070	15.086
6	12.592	16.812

The observed probabilities of DCS in the data set and the LE1 tetranomial model's predicted probabilities of DCS are plotted in Figure 15 for marginal (black diamonds, top right), mild (gray circles, top left), serious (white triangles, bottom) DCS. This plot was generated by first sorting the model's per-dive exposure predictions by the probability of no DCS. These per-dive predictions were then placed in bins with equal numbers of observed serious, mild, or marginal DCS cases. For this plot, we used 10

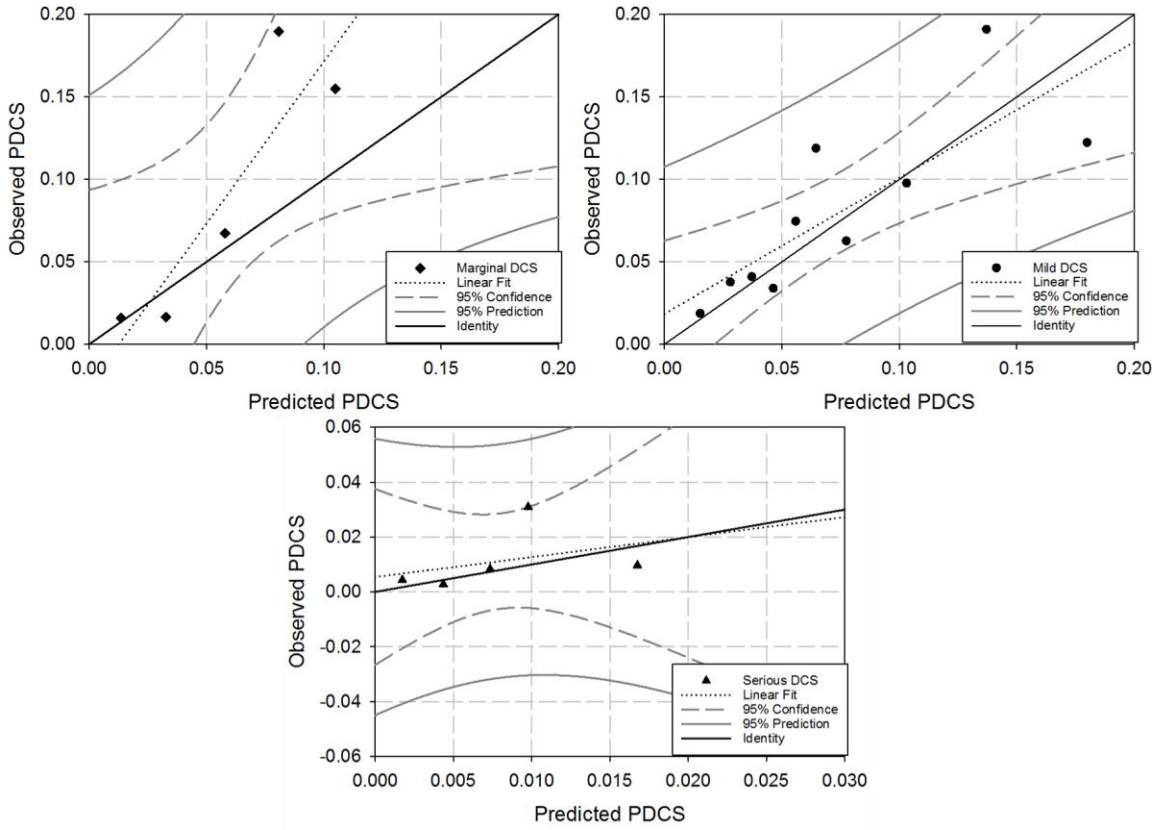

Figure 15: Tetranomial LE1 observed probabilities of DCS vs. predicted probabilities of DCS. To generate these data points, the predicted probabilities of marginal (top left), mild (top right), and serious (bottom) DCS for each dive were sorted by the predicted probability of no event, from smallest to largest. These predicted probabilities were then separated into five, five, and ten bins with equal numbers of serious, mild, or marginal DCS events respectively. The predicted and observed probabilities of DCS were calculated and plotted for each bin. These probabilities are plotted with a linear fit ($r^2_{marginal} = 0.80$, $r^2_{mild} = 0.65$, and $r^2_{serious} = 0.14$) and the 95% confidence and 95% prediction bands

bins of 17 mild DCS outcomes each, 5 bins of 4 serious DCS outcomes, and 5 bins of 22 marginal DCS outcomes. The predicted probabilities of DCS were calculated as the sum of the model's per-dive exposure predictions for that DCS severity divided by the total number of exposures in the bin. The observed probabilities of DCS were calculated as

the number of observed DCS outcomes in the bin divided by the total number of dive exposures in the bin. The linear fits for the serious, mild, and marginal DCS data points were plotted ($r^2_{serious} = 0.14$, $r^2_{mild} = 0.65$, and $r^2_{marginal} = 0.80$), along with the 95% confidence and 95% prediction bands. The line of identity was also plotted (black line). If a model's predictions were perfectly aligned with the data set, all points in this plot would fall on the line of identity. Like the trinomial model in the previous chapter [27], the marginal DCS data points are less scattered than that of serious and mild DCS. The mild DCS model predictions align the closest with the data set, as the mild DCS linear fit line aligns closer to the line of identity than that for serious or marginal DCS.

4.3.2 Predictions on Data

The tetranomial LE1 model's predicted DCS outcomes and the observed DCS cases in the data can be found in (Table 18). The dive data is separated by dive type, which includes single air, single non-air, repetitive and multilevel air, repetitive and multilevel non-air, and saturation diving. The 95% confidence intervals are listed for the model's total predictions of serious, mild, marginal, and any DCS. These predictions do match the observed number of cases within 95% confidence. From Table 18, we can see that the model underpredicts mild, serious, and marginal DCS occurrence for single air diving.

Table 18: DCS occurrences and tetranomial model predictions for the BIG292 data set.

		Observed DCS				LE1 AB Tetranomial Predicted DCS			
	Exposures	Mild	Serious	Marginal	Total	Mild	Serious	Marginal	Total
Single Air									
EDU885A	483	27	3	0	30	19.77	2.29	13.24	35.3
DC4W	244	7	1	4	12	4.11	0.47	2.81	7.39
SUBX87	58	0	2	0	2	0.14	0.02	0.10	0.26
NMRNSW	91	4	1	5	10	3.88	0.45	2.59	6.92
PASA	72	4	1	2	7	1.87	0.21	1.27	3.35
NSM6HR	57	3	0	2	5	3.10	0.36	2.05	5.51
Rep&Mult Air									
EDU885AR	182	11	0	0	11	8.42	0.98	5.60	15
DC4WR	12	3	0	0	3	0.66	0.08	0.44	1.18
PARA	135	6	1	3	10	6.98	0.81	4.62	12.41
PAMLA	236	9	4	12	25	14.05	1.64	9.25	24.94
Single Nonair									
NMR8697	477	9	2	18	29	11.00	1.26	7.48	19.74
EDU885M	81	4	0	0	4	2.17	0.25	1.48	3.9
EDU1180S	120	9	1	0	10	5.07	0.59	3.38	9.04
Rep&Mult Nonair									
EDU184	239	11	0	0	11	10.17	1.18	6.79	18.14
PAMLAOD	134	5	1	0	6	5.92	0.68	3.98	10.58
PAMLAOS	140	5	0	3	8	4.28	0.49	2.89	7.66
EDU885S	94	4	0	0	4	2.60	0.30	1.77	4.67
Saturation									
ASATEDU	120	11	2	27	40	14.70	1.80	9.05	25.55
ASATNMR	50	1	0	0	1	4.12	0.49	2.66	7.27
ASATNSM	132	18	0	21	39	22.31	2.80	13.16	38.27
ASATARE	165	19	1	13	33	17.88	2.16	11.22	31.26
Totals	*3322*	*170*	*20*	*110*	*300*	*163.2 ± 20.78*	*19.31 ± 3.566*	*105.83 ± 12.57*	*288.34 ± 24.5*

The imbalance in the distribution of marginal DCS events in the data is evident when considering saturation diving. More than half of the marginal DCS events (55%) in the data set occur from saturation diving, though the entire data set is comprised of only 3% marginal events and 14% saturation dives. In Table 18, we can see that the tetranomial model does not reproduce this skew in observed marginal DCS cases, as the model predicts only 33% of marginal cases occurring from saturation diving.

4.3.3 Tetranomial Model vs. Trinomial Marginal Model

The model parameters used in the trinomial marginal LE1 model along with model performance analysis can be found in Chapter 3 [27]. We can calculate the tetranomial model's equivalent trinomial marginal log likelihood using Eq. (4.12). For the optimized tetranomial LE1 model parameter set, $LL_{43} = -1485.4$, which is nearly identical to the optimal trinomial marginal LE1 log likelihood found in Chapter 3 [27]. This indicates that the performance of the tetranomial model is on par with the trinomial marginal model when using the BIG292 data set. This is likely because both models optimized to nearly identical parameter sets.

The shift in predicted dive exposure probabilities between the trinomial marginal and tetranomial models is plotted in Figure 16. In Figure 16, the trinomial marginal and tetranomial models' predicted probabilities for full DCS are plotted for all full observed DCS cases (gray circles), predicted probabilities of marginal DCS for observed marginal DCS cases (white diamonds), and predicted probabilities of no DCS for observed no

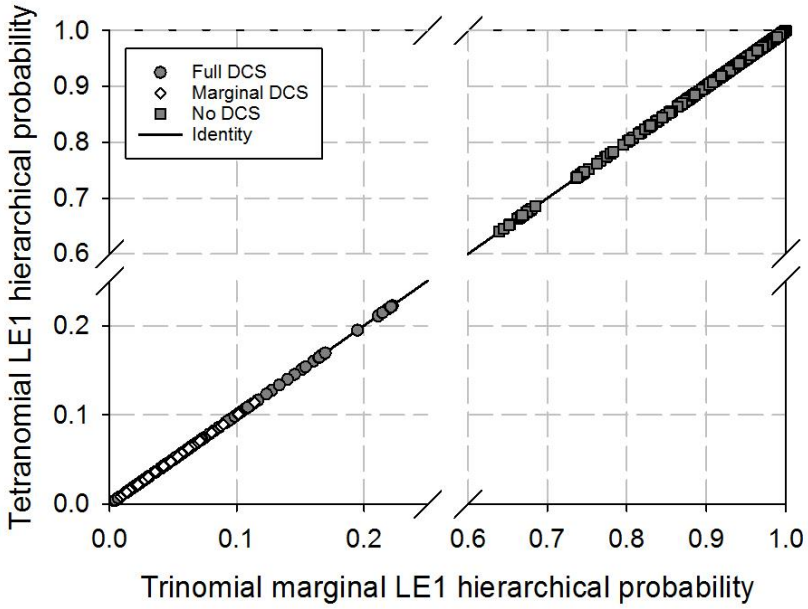

Figure 16: Trinomial marginal to tetranomial probability shift. For dives that resulted in full DCS, the sums of the tetranomial predicted probabilities of serious and mild DCS are plotted against the trinomial marginal predicted probability of full DCS (gray circles). For dive exposures that resulted in marginal DCS and no DCS, the tetranomial model predicted probabilities of marginal (white diamonds) and no DCS (gray squares) respectively are compared with that of the trinomial marginal model.

DCS cases (gray squares). All these data points fall close to the line of identity, indicating that these models make nearly identical predictions on the data set. The slope of the linear fit to the full DCS data points is 0.9978 ($r^2 > 0.9999$), for marginal DCS data points is 1.000 ($r^2 = 1.000$), and for no DCS data points is 0.9985 ($r^2 > 0.9999$). These slopes approximate the probability shift between the two models, i.e. $P_{tet,\,full} \approx P_{tri\_m,\,full}$ and $P_{tet,marg} \approx P_{tri\_m,marg}$. The tetranomial and trinomial marginal models' agreement in hierarchical probabilities for each DCS cases is a result of both models optimizing to

nearly identical parameter sets, so one model does not offer significant performance improvement over the other on this data set.

4.3.4 Tetranomial Model vs. Trinomial Model

The shift in predicted dive exposure probabilities between the trinomial and tetranomial models is plotted in Figure 17. The model parameters used in this trinomial LE1nt model can be found in Ref. [10]. Both models use DCS Type A/B splitting (see Table 2). In Figure 17, the trinomial and tetranomial models' predicted probabilities of mild DCS for dive exposures that resulted in mild DCS are plotted with gray circles, and

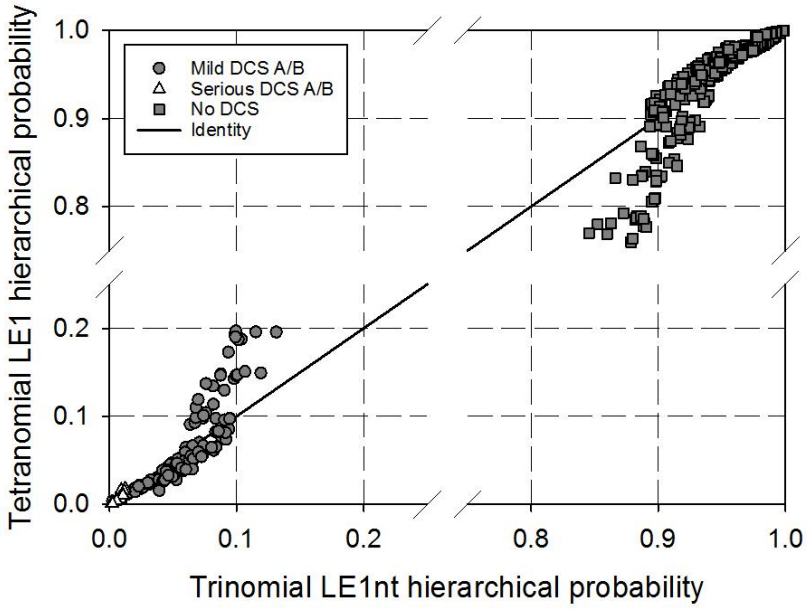

Figure 17: Trinomial to tetranomial probability shift. For dives that resulted in mild or serious DCS, the tetranomial model predicted probabilities of mild or serious DCS are plotted against that of the trinomial model. For dives that resulted in no DCS (including marginal DCS), the sum of the predicted probabilities of marginal and no DCS for the tetranomial model are compared with the trinomial model's predicted probabilities of no DCS.

likewise for serious DCS in white triangles. The trinomial model's predicted probabilities of no DCS and the tetranomial model's predicted probabilities of no- and marginal DCS for dive exposures that did not result in full DCS are plotted with gray squares. The mild DCS and serious DCS data points that fall above the line of identity indicate the tetranomial model predicted a greater probability of occurrence of DCS for those exposures than the trinomial model, and the no DCS points that fall below the line of identity indicate the tetranomial model predicted a lower probability of no DCS for those exposures compared with the trinomial model.

The slope of the linear fit to the mild DCS data points is 1.720 ($r^2 = 0.7358$) and the slope of the linear fit to the serious DCS data points is 1.228 ($r^2 = 0.7250$). The line of linear fit to the no DCS points has a slope of 1.352 ($r^2 = 0.7333$). All three sets of data points have similar amounts of scatter, as all have an r^2 value of approximately 0.73. Using these slopes to approximate the trinomial to tetranomial probability shift, $P_{tet,mild} \approx 1.720 P_{tri,mild}$ and $P_{tet,ser} \approx 1.228 P_{tri,ser}$. Thus, the tetranomial model predicts a greater probability of mild/serious DCS for some mild/serious DCS cases and a lower probability of no DCS for some no DCS cases when compared with the trinomial model.

4.3.5 Cumulative Density Function

Cumulative density functions (CDF) can be used to visually inspect the DCS symptom onset time agreement between a model's predictions and empirical data. A probabilistic DCS model that performs well on the dataset would produce a CDF that

closely replicates that of the data. An in-depth analysis of the BIG292 data set density function was performed in our previous work [26], and it is important to note that the DCS symptom onset times reported in the data may have been biased by the medical surveillance protocol.

The cumulative density functions for the mild, serious, and marginal DCS BIG292 data are plotted in Figure 18 as the solid black curve, solid gray curve, and dashed gray curve respectively. The dashed black line represents the cumulative density function for the tetranomial model's predictions of all DCS types, as scaling factors are used by the model to delineate these severities and thus their cumulative density functions fall on the same curve.

The tetranomial model's predicted CDF indicates the model most severely over-predicts serious DCS prior to surfacing, then over-predicts the onset of all severities of DCS immediately after surfacing. The tetranomial model's onset time predictions are closely aligned with the marginal DCS cases' onset times until approximately 7 hours prior to surfacing. After surfacing, the marginal DCS data CDF lags behind the mild and serious DCS curves, as the 42 of 110 marginal DCS cases reported without onset times were assigned T_2, or the first known time the diver was experiencing symptoms, at the studies' right-censored times (24 or 48 hours). Because the onset time windows for these 42 marginal cases are imprecise, the tetranomial model's predicted CDF's inability to replicate late onset for marginal cases may not indicate an issue with the model.

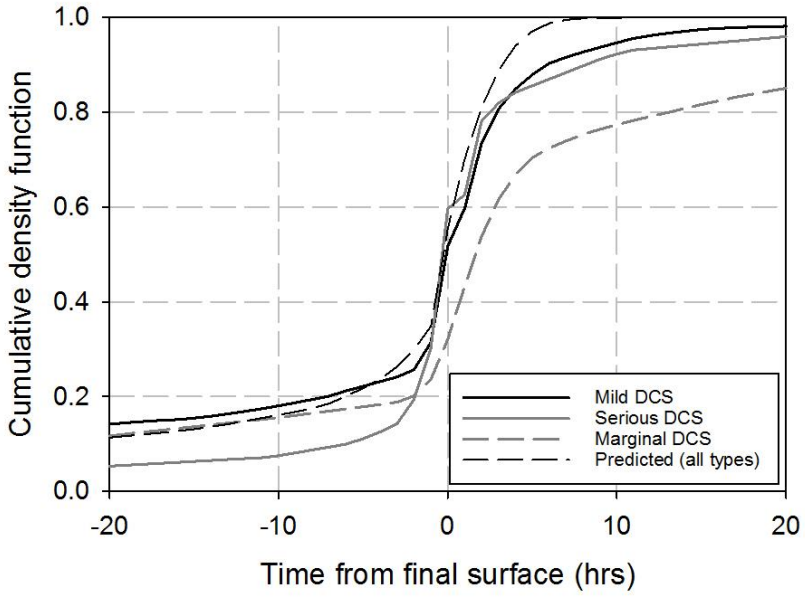

Figure 18: Tetranomial cumulative density function. Mild DCS (black, solid curve), serious DCS (solid, gray curve), and marginal DCS (dashed, gray curve) cases are shown for the BIG292 dive data set. The cumulative density functions for predicted mild, serious, and marginal DCS fall on the same curve (black, dashed).

4.3.6 Pearson Residual

The Pearson Residual group statistic was calculated for each dive type in the BIG292 data set according to Ref. [13], i.e.

$$PR_j = \frac{\left(obs_j - pred_j\right)^2}{pred_j\left(1 - \frac{pred_j}{N_j}\right)}, \tag{4.14}$$

where subscript j indicates the data group, obs_j is the number of observed events in the group, $pred_j$ is the number of model-predicted events for that group, and N_j is the

total number of exposures in group j. The sum of the Pearson Residuals for all j groups is equal to the Chi-squared statistic:

$$\chi^2 = \sum_{j=1}^{J} PR_j. \tag{4.15}$$

In this statistical analysis, the null hypothesis is that the model-predicted incidence of DCS is equal to the incidence of DCS observed in the BIG292 dataset. A high χ^2 value (and corresponding low p-value) indicates that the model's predictions are not consistent with the observed occurrence of DCS in the data.

The chi-squared values calculated from the Pearson Residual of each dive type according to Eqs. (4.14) and (4.15) for the binomial and trinomial LE1nt models in Ref. [10], the trinomial marginal LE1 model in Chapter 3 [27], and the tetranomial LE1 model presented in this chapter can be found in Table 19. It is evident from Table 19 that the

Table 19: Pearson Residual group statistic (χ^2) and corresponding p-value calculated for each model's predictions of DCS incidence. A high χ^2 value (and corresponding low p-value) indicates that the model's predictions are not consistent with the observed occurrence of DCS in the data.

	Number of DOF	Pearson Residual Full DCS	Pearson Residual Mild DCS	Pearson Residual Serious DCS	Pearson Residual Marginal DCS
Binomial LE1nt [10]	7	8.465 p=0.294			
Trinomial LE1nt [10]	8		8.421 p=0.393	4.527 p=0.807	
Trinomial Marginal LE1 [27]	9	12.270 p=0.199			36.568 p=0.000031
Tetranomial LE1	10		7.597 p=0.668	9.246 p=0.509	36.612 p=0.000066

trinomial marginal model's and the tetranomial model's predictions of marginal DCS do not align with the observed incidence of marginal DCS, because these groups have a high χ^2 value (and corresponding low p-value).

4.4 Discussion

The tetranomial model presented here serves as a continuation of the trinomial LE1nt model published by Howle *et al*. [10] and the trinomial marginal LE1 model explored in Chapter 3. All model formulation and analyses were conducted in accordance with those works. In this Discussion section, we will compare all three models.

In this chapter, we optimized six tetranomial model variants: EE1, EE1nt, EE1 Full, LE1, LE1nt, and LE1 Full. These model variants were tested with both Type I/II and Type A/B splitting, and the Type A/B splitting models outperformed all their corresponding Type I/II splitting models. The log likelihood difference test was used to determine that the LE1 model, with a pressure crossover parameter in the second compartment and a pressure threshold parameter in the third compartment, provided the best fit to the BIG292 data set.

The tetranomial LE1 model predicted the distribution and onset of mild DCS cases better than that of serious and marginal DCS. In Figure 15, the linear fit line for the mild DCS data is closest to the line of identity, and in Figure 18, the predicted CDF is follows closest to the mild DCS curve when compared with serious and marginal DCS.

These figures also illustrate that the model is least accurate in predicting both the distribution of marginal DCS cases within the data set and their onset times. These graphical results are verified in Table 19, as the Pearson Residual Chi-squared value is lowest for mild DCS, followed closely by serious DCS. Howle's trinomial model does not follow this trend, and predicts serious DCS more accurately than mild DCS (Table 19). All three models' CDFs indicate they are able to accurately predict the onset of serious and mild DCS around the time of surfacing (Figure 13, Figure 18, Ref. [10]).

The high Pearson Residual Chi-squared value for marginal DCS indicates that both the trinomial marginal and tetranomial models' predictions are not aligned with the incidence of marginal DCS in the BIG292 data set. The distribution of marginal DCS cases in the BIG292 data set is skewed towards saturation diving, as 55% of the BIG292 marginal cases occur from saturation diving, and saturation diving only constitutes 14% of the total data. Both the trinomial marginal and the tetranomial LE1 models are unable to reproduce this skew, and only predict 34% of marginal DCS cases occurring from saturation diving. In addition, the marginal cases with right-censored T_2 times may not accurately reflect the true symptom onset times. Neither the trinomial marginal nor the tetranomial models predict the onset time delay created by this right-censoring (Figure 13, Figure 18). This may not indicate an inherent flaw in these models' ability to predict marginal DCS, rather points to an issue with potentially inaccurate data.

When comparing this tetranomial LE1 model with the trinomial marginal LE1 model in Figure 16, all data points fall close to the line of identity. Both models make nearly identical predictions on the data set. In Table 18, the sums of the tetranomial model's mild DCS and serious DCS predictions for each dive type are nearly equivalent to the trinomial marginal model's predictions for full DCS (Table 12). Both models optimized to nearly identical parameter sets. When using the tetranomial model's equivalent trinomial marginal log likelihood to compare these two models, no clear winner emerges.

The optimal tetranomial model parameter set is quite different from the trinomial model's optimal parameters in Ref. [10], which considers marginal DCS events as non-events. In Figure 17, the tetranomial model predicts a higher probability of mild and serious DCS than the trinomial model for some mild and serious DCS cases, and a lower probability of no DCS than the trinomial model for some no DCS cases. The increase in scatter of these data points when compared with Figure 16 illustrates the difference in optimal parameter sets which alters each models' predictions. It could be argued that the tetranomial model would generate more conservative "safe" ascent criteria than the trinomial model, as the tetranomial model predicts increased probabilities of DCS and decreased probabilities of no DCS than the trinomial model.

When the trinomial model was compared with a binomial model in Ref. [10], the probability shift plot showed a similar trend as Figure 16 and both optimal parameter

sets were nearly identical. However, the trinomial model's equivalent binomial log likelihood indicated the trinomial model performed highly significantly better than the binomial model on the BIG292 data set.

4.5 Conclusions

The tetranomial model explored in this work simultaneously predicts the hierarchical probabilities of serious, mild, marginal, and no DCS. The derivation of these hierarchical probabilities and the multinomial log likelihood function used during model calibration are extensions of the previous Howle *et al*. publication [10].

Both the trinomial marginal model in Chapter 3 and tetranomial model presented here are unable to accurately replicate the occurrence of marginal DCS events observed in the BIG292 dataset. These marginal DCS events may hinder model fit during calibration. The trinomial LE1nt model in Ref. [10] demonstrated highly significant improvement over the binomial LE1nt model, both considering marginal DCS as non-events. Using the Pearson's χ^2 statistic as a metric, we find that the trinomial LE1nt model's predictions are most closely aligned with the incidence of observed DCS in the data. We therefore recommend the use of the trinomial LE1nt model from Ref. [10] with the event categories of serious, mild, and no-DCS, Type A/B severity splitting, and marginal events scored as non-events. This trinomial probabilistic model can be used to generate dive schedules specific to symptom severity, to better tailor dive missions to the acceptable level of risk for the divers.

4.6 Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1644868 and by the U.S. Navy, Naval Sea Systems Command under contracts #N00024-13-C-4104 and #N00024-17-C-4317. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the U.S. Navy. Computational resources were provided by BelleQuant Engineering, PLLC. Neither funding agency nor the commercial entity played any role in designing this study, data collection and analysis, decision to publish, interpreting the results, or writing the manuscript.

5. Three-Region Unstirred Tissue Bubble Volume Trinomial Decompression Sickness Model using Serious, Mild, and Non-Event Outcomes

5.1 Introduction

Decompression sickness (DCS) is a condition resulting from decreasing ambient pressure, and can arise in both hyperbaric and hypobaric environments. This work focuses on hyperbaric exposures, and in particular underwater diving. As a diver makes their descent, ambient pressure increases. The diver breathes gases equilibrated with ambient pressure, and that inspired gas is absorbed into the blood and circulated to the tissues. When the diver makes their ascent, ambient pressure decreases, and the gas inspired at elevated pressure can become supersaturated. This can cause gas to leave solution and form bubbles, which is the underlying mechanism perpetrating DCS symptoms [3].

The U.S. Navy has three classifications for DCS symptoms: Type I (mild), Type II (serious), and marginal. The signs and symptoms of Type I DCS include joint pain, cutis marmorata, and paresthesia. Type II DCS is characterized by more severe cardiopulmonary and neurological symptoms, such as hemoptysis, dyspnea, dysfunction of coordination, vision, and hearing, and loss of consciousness. Type I and Type II DCS can be treated with hyperbaric chambers, in which the patient is recompressed, then decompressed according to the appropriate treatment schedule [6]. Marginal DCS is symptoms associated with DCS that persist for only a short period of

time and resolve spontaneously without recompression treatment. An example of marginal DCS is pain in one joint lasting for less than 60 minutes, or pain in two joints lasting less than 30 minutes [7, 8]. Though DCS is treatable, if symptom onset occurs mid-dive, it can diminish a diver's ability to safely complete decompression. DCS poses a great risk to military divers and can cause premature termination of diving missions.

In the early 20th century, Scottish physiologist J. S. Haldane was commissioned by the U. K. Royal Navy to create diving ascent schedules for mitigating the risk of DCS. Haldane and coworkers published a five-compartment gas content model [4], referred to now as the Haldane model, which simulated the uptake and elimination of nitrogen gas in the body during diving. Each of the five compartments, intended to represent the body's tissues, had a unique prescribed rate of inert gas exchange with the blood, hypothetically demonstrating variation in the degree of blood perfusion. Haldane proposed that if, during ascent, the partial pressure of inert gas in each compartment did not exceed twice the ambient pressure, a diver could avoid DCS. Decompression tables were generated based on this principle, in which one or more decompression stops were assigned during ascent to allow for inert gas equilibration.

The Haldane model was deterministic, meaning it purported that a diver was "safe" if they followed the appropriate decompression schedule, and would experience DCS if they violated the schedule. However, in empirical dive data, there is variation in DCS symptoms (or lack thereof) among divers performing identical dive profiles [7, 8].

In response to this observed variation, Weathersby *et al.* [12] and Berghage *et al.* [11] developed probabilistic models for DCS. Such models use a gas content model (like the Haldane model) or a bubble volume model in conjunction with survival analysis to report a probability of experiencing DCS for a given dive. Probabilistic models have several adjustable parameters which can be calibrated with empirical dive data. Neither deterministic nor probabilistic DCS models provide information regarding DCS symptom severity.

The gas content models used currently by the U.S. Navy consist of three parallelly-perfused compartments, each with a different rate of gas uptake and elimination (referred to as a tissue half-time). One model that has been particularly successful in its ability to predict empirical dive data, in particular DCS cases with late onset, is known as the LE1 model [15, 16]. This model improved upon previous gas content models by implementing asymmetrical gas kinetics, in which inert gas elimination occurs slower than gas uptake, representing the influence of gas bubble dynamics. This can result in a longer duration of inert gas supersaturation and thus compartmental risk accumulation, enabling the model to predict occurrence of DCS long after the final surface interval as observed in some DCS cases.

Following the success of the LE1 model, whose linear gas washout was inspired by bubble behavior *in vivo*, Tikuisis and coworkers developed the first probabilistic model to predict DCS using bubble physics [54, 55]. Their goal was to create a

predictive model more closely aligned with the physiological mechanism prompting DCS, so that the model could perform successfully on dives outside the calibration dataset.

Though the LE1 model does perform well on empirical dive data, it does not allow for an increase in DCS risk after decompression and subsequent delay in DCS onset as observed in decompression from sea level to altitude (hypobaric exposures). With the twofold goal of creating a model with flexible gas kinetics to model hyperbaric and hypobaric decompression data, and to model the mechanism prompting DCS *in vivo*, Dr. Gerth and others [18] developed a bubble volume probabilistic model based on Van Liew *et al.*'s theoretical two-region (2R) model of bubble behavior in tissues [56]. Dr. Gerth's probabilistic model, called BVM3, consisted of three parallelly-perfused gas compartments, each capable of nucleating a bubble according to the 2R model. BVM3 was calibrated with empirical dive data and performed on-par with the LE1, though it had almost 20 additional adjustable parameters. When tested on a validation dataset containing high-oxygen exposures, it was found that the bubble volume model successfully accumulated higher DCS risk and better predicted DCS incidence during high-oxygen decompression [18]. Because this model could predict increasing DCS risk after decompression is completed, it could be applied to both hyperbaric and hypobaric exposures, and even flying after diving. However, without a more expansive training dataset, the bubble model's number of parameters was not statistically warranted, and

these parameters did not reflect true physiologic values. It was concluded that this bubble volume model, though inspired by DCS physiology, was only a mathematical descriptor of the processes involved in the illness [18].

Some of the limitations faced by Dr. Gerth's BVM3 model from Ref. [18] were derived from unphysical assumptions of the 2R model [57]. A true physiological model of bubble behavior in extravascular tissue with blood flow is complex, requiring partial differential equations with many parameters. Simpler models of bubble dynamics in the context of DCS were derived and tested by Srinivasan and others, with the goal of reducing the number of model parameters and enabling the practical implementation of ordinary differential equations [57-60]. A summary of these ODE models of bubble dynamics can be found in Ref. [20]. Of these models presented, the Three-Region Unstirred Tissue model, or 3RUT model, is the favorite because its underlying assumptions are most aligned with the true physiology and it provides the greatest behavioral flexibility.

The 3RUT model consists of a spherical gas bubble surrounded by an unstirred diffusive region in a finite-volume, well-stirred tissue perfused with arterial blood. A detailed derivation of the 3RUT model can be found in Ref. [59], and the 3RUT model accommodating multiple bubbles in each compartment, or 3RUT-MB, can be found in Ref. [58] and Appendix A of Ref. [21].

Recently, the 3RUT-MB model has been augmented with exercise data and used to predict altitude DCS [21]. However, none of the bubble volume models used by the U.S. Navy predict any information about DCS symptom severity. When conducting a Naval diving mission, dive medical officers will allow for a higher frequency of Type I (mild) DCS than Type II (serious) DCS occurrence before needing to terminate the mission [25]. Both the probability of injury and the severity of the injury contribute to the risk of an activity. To incorporate both probability and severity factors in probabilistic modeling for DCS, Howle *et al*. created a trinomial gas content model that simultaneously predicts the probability of mild, serious, and no DCS events [10]. In Chapters 3 and 4, we continue this work by developing trinomial marginal and tetranomial probabilistic models with risk dependent on gas supersaturation in tissue [27, 61]. In this chapter, we use the 3RUT model to develop a trinomial probabilistic bubble volume model for DCS, predicting the occurrence of serious, mild, and no DCS.

5.2 Methods

5.2.1 3RUT Model

The 3RUT model is a formulation for quantifying bubble dynamics in tissue, developed by Dr. Srinivasan and others [58, 59]. The three regions defined in this model are a spherical gas bubble, an unstirred diffusive region surrounding the bubble, and a well-stirred finite-volume tissue perfused by arterial blood.

The volume of the gas bubble, V_b, can be calculated as the volume of a sphere:

$$V_{b,i}(t) = \frac{4\pi}{3} r_{b,i}^3(t), \tag{5.1}$$

where $r_{b,i}$ is the time-dependent bubble's radius in the i^{th} compartment. This radius is calculated from ambient pressure, P_{amb}, compartmental tissue tension, $P_{T,i}$, and bubble pressure, $P_{b,i}$, as

$$\frac{dr_{b,i}}{dt} = \frac{\gamma_i \left(P_{T,i} - P_{b,i}\right)\left[\lambda_i + \frac{1}{r_{b,i}}\right] - \frac{r_{b,i}^3}{3}\frac{d(P_{amb})}{dt}}{\left(P_{amb} - P_{FVG}\right) + \frac{4\sigma}{3r_{b,i}} + \frac{8\pi}{3} M_i r_{b,i}^3}, \tag{5.2}$$

where γ_i is the compartmental diffusivity fraction, λ_i is the compartmental sink term, σ is the bubble surface tension, and M_i is the compartmental bubble's elastic modulus. The bubble pressure in each compartment, $P_{b,i}$, is calculated from ambient pressure as

$$P_{b,i} = P_{amb} - P_{FVG} + \frac{2\sigma}{r_{b,i}} + V_{b,i} M_i. \tag{5.3}$$

The rate of change of tissue tension in a bubble-baring compartment is calculated as

$$\frac{dP_{T,i}}{dt} = k_i \left(P_a - P_{T,i}\right) - \xi_i \frac{d\left(P_{b,i} V_{b,i}\right)}{dt}, \tag{5.4}$$

where P_a is the arterial partial pressure of inert gas, ξ_i is the compartmental total solubility, and k_i is the tissue half-time.

Note that Eq. (5.2) is undefined if the bubble's radius shrinks to zero. To avoid this mathematical instability, the 3RUT model dictates a bubble nucleation radius, $r_{b,i}^0$.

Before a bubble nucleates, the compartmental tissue tensions are calculated according to the EE1 model (i.e. Eq. (5.4) without the right-most term). Bubble nucleation is determined by tissue saturation. The models presented in this chapter moderated bubble nucleation according to the bubble formation criterion presented by Dr. Gerth and others in Ref. [18], in which a bubble forms if

$$P_{amb} < P_{T,i} + P_{FVG} - \frac{2\sigma}{r_{b,i}^0}. \tag{5.5}$$

After bubble nucleation, Eqs. (5.2)-(5.4) are used to model the bubble's growth. A complete derivation of Eqs. (5.2)-(5.4) can be found in Ref. [59].

5.2.2 Probabilistic Model

The bubble volume model presented in this chapter consists of three parallelly-perfused bubble-baring compartments, with bubble behavior following the 3RUT model. The risk of a given dive is dependent on the relative increase in bubble volume:

$$\begin{aligned} r(t) &= \sum_{i=1}^{3} g_i \left(V_{b,i}(t) - V_{b,i}^0 \right) \\ V_{b,i}(t) - V_{b,i}^0 &\geq 0. \end{aligned} \tag{5.6}$$

Because the bubble never shrinks below the nucleonic size, risk is always greater than or equal to zero.

The probability of the occurrence of DCS can be calculated from this risk function using survival analysis [13]:

$$P(DCS) = 1 - e^{-\int_0^T r(t)dt}. \tag{5.7}$$

In this formulation, DCS is treated as a binomial process. The probability of not experiencing DCS is defined as

$$P(0) = 1 - P(DCS) = e^{-\int_0^T r(t)dt}. \tag{5.8}$$

Weathersby *et al*. [14] found that the inclusion of DCS onset times when estimating model parameters enabled probabilistic models to provide a better fit to calibration data. DCS symptom onset times are reported in empirical dive data as a window from T_1 to T_2, in which time T_1 was the last time a diver was not experiencing DCS symptoms, and T_2 was the first time a diver was noted to be experiencing DCS symptoms [14]. Weathersby augmented Eq. (5.7) by calculating the probability of DCS from the joint probability of DCS not occurring until T_1, and DCS occurring between T_1 and T_2, as follows:

$$\begin{aligned} P(DCS) &= P(0)_{0 \to T_1} P(DCS)_{T_1 \to T_2} \\ &= \left(e^{-\int_0^{T_1} r(t)dt} \right) \left(1 - e^{-\int_{T_1}^{T_2} r(t)dt} \right). \end{aligned} \tag{5.9}$$

The 3RUT model's adjustable parameters are compartmental gains, tissue half-times, diffusivity fractions, sinks, total solubilities, elastic moduli, bubble nucleonic radii, and the surface tension. These parameters can be estimated using maximum likelihood optimization, first applied to probabilistic models for DCS by Weathersby *et al.* [12]. The likelihood function, shown in Eq. (5.10), quantifies the agreement between the model's predictions and the observed prevalence of DCS in the calibration dataset.

$$LL = \sum_{i=1}^{N} \ln\left[P\left(DCS_i\right)^{\delta_i} \left(1 - P(DCS_i)\right)^{1-\delta_i} \right] \tag{5.10}$$

In Eq. (5.10), subscript i counts over the N dives in the calibration dataset. For each dive, the probability of the occurrence of DCS is calculated according to Eq. (5.7), and the exponent δ_i is set to 1 if DCS was observed, and 0 if DCS was not observed. To calibrate a probabilistic model, model parameters are adjusted to maximize Eq. (5.10). The calibration dataset used in this work is discussed in Section 5.2.4.

5.2.3 Trinomial Formulation

The primary objective of this chapter is to derive a bubble volume model that simultaneously predicts the probabilities of the occurrence of serious, mild, and no DCS. The binomial formulation of Eqs. (5.7)-(5.10) must be altered to accommodate multiple probabilities. In accordance with our previous work on multinomial probabilistic modeling [10, 27, 61], the probabilities of serious DCS (denoted with subscript s) and mild DCS (denoted with subscript m) can be distinguished mathematically by a scale factor, a.

$$\begin{aligned} P_s^c &= 1 - e^{-a\left(\int r(t)dt\right)} \\ P_m^c &= 1 - e^{-\left(\int r(t)dt\right)} \end{aligned} \tag{5.11}$$

In Eq. (5.11), the superscript c indicates these probabilities are competitive probabilities. A competitive probability is the independent probability of an event happening, whether or not any other event has occurred. In empirical dive data, DCS

severity is diagnosed hierarchically. If a diver is experiencing both mild DCS and serious DCS symptoms, their diagnosis will be serious DCS. To properly represent observed DCS cases in data, the competitive probabilities of serious and mild DCS in Eq. (5.11) must be converted to hierarchical probabilities. In the hierarchical model, the probability of mild DCS occurrence is the joint probability of the independent probability of mild DCS and the probability of serious DCS not occurring, as shown in Eq. (5.12).

$$\begin{aligned} P_s^h &= P_s^c = 1 - e^{-a\left(\int r(t)dt\right)} \\ P_m^h &= P_m^c\left(1 - P_s^c\right) = \left(1 - e^{-\int r(t)dt}\right)e^{-a\left(\int r(t)dt\right)} \end{aligned} \tag{5.12}$$

The superscript h denotes hierarchical probabilities. The masking of mild DCS by serious DCS in this hierarchical formulation is illustrated in Figure 19.

The probability of no event occurring is

$$P_{0,tri}^h = 1 - P_s^h - P_m^h = e^{-(a+1)\int r(t)dt}. \tag{5.13}$$

Note the law of total probability dictates that the sum of the probabilities of all events must equal 1, and indeed $P_s^h + P_m^h + P_{0,tri}^h = 1$.

The log likelihood function in Eq. (5.10) must be altered to accommodate the trinomial model, as follows:

$$LL_3 = \sum_{i=1}^{N} \ln\left[\left(P_{s,i}^h\right)^{\sigma}\left(P_{m,i}^h\right)^{\mu}\left(1 - P_{s,i}^h - P_{m,i}^h\right)^{\mathbf{1}-\sigma-\mu}\right]. \tag{5.14}$$

The exponents σ and μ are modulated to reflect observed occurrences of DCS in the calibration dataset, where $\sigma = 1$ if serious DCS occurred, $\mu = 1$ if mild DCS occurred, and $\sigma = \mu = 0$ if no DCS occurred.

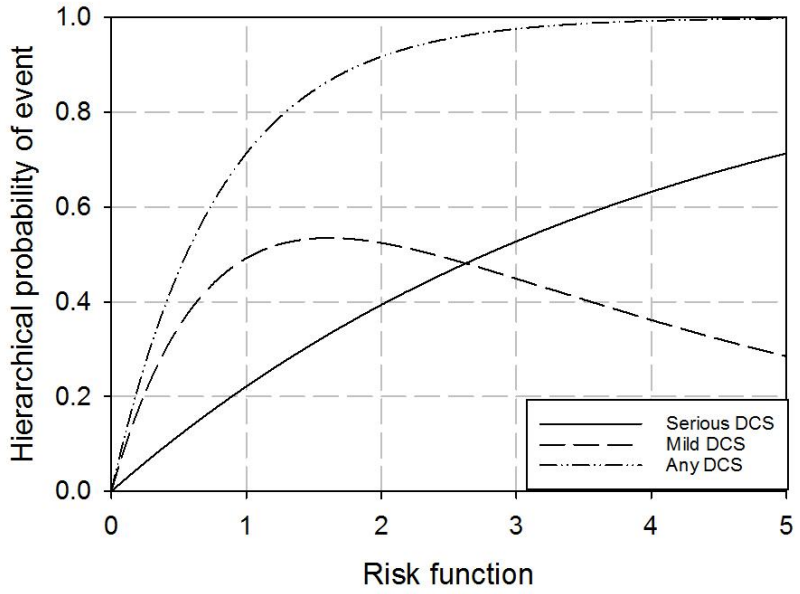

Figure 19: Hierarchical probabilities of serious DCS, mild DCS, and any DCS event occurring with increasing risk function. The hierarchical probabilities were calculated with Eq. (5.11), using a scale factor of $a = 0.25$. As the risk function increases, the probability of mild DCS increases then decreases, illustrating the masking of mild DCS by serious DCS.

To enable direct comparison between the binomial and trinomial bubble volume models presented in this chapter, we can reduce Eq. (5.14) to the equivalent binomial log likelihood as

$$LL_{32} = \ln \sum_{i=1}^{N} \left[\left(P_{s,i}^{h} + P_{m,i}^{h} \right)^{\sigma+\mu} \left(1 - P_{s,i}^{h} - P_{m,i}^{h} \right)^{1-\sigma-\mu} \right], \tag{5.15}$$

since the sum of the probabilities of serious and mild DCS provides the trinomial model's prediction for the occurrence of full DCS.

Following our previous research on the impact of marginal DCS events on model calibration, all binomial and trinomial models presented in this chapter categorize marginal events as non-events [31, 49].

5.2.4 Calibration Dataset

Model parameters were calibrated with the p97 DCS dataset, which is a subset of data published by Temple *et al*. in Refs. [7, 8]. This dataset consists of 4,335 exposures of human dives conducted by the U.S., U.K. and Canadian militaries between 1944 and 1997. These dive trials used hyperbaric chambers to control ambient pressure, and diving occurred in both wet and dry conditions. Seven different types of dives were tested in these trials: single dives with air and with nitrox, repetitive dives with air and nitrox, saturation dives, surface oxygen decompression dives, and in-water oxygen decompression dives. Of these 4,335 dive exposures, there were 224 observed cases of full DCS and 127 cases of marginal DCS.

Temple's reports contain symptom histories for all the observed cases of full and marginal DCS, along with the onset time window of DCS symptoms all full DCS events and some marginal DCS events. Howle and coworkers used the reported symptom histories to apply their Perceived Severity Index (PSI) system of severity classification to this dataset [10]. The six PSI indices, in order of least to most severe, are: constitutional:

dizziness, fatigue, nausea; lymphatic or skin: rash, marbling, itching; pain: joint pain, ache, cramps; mild neurological: paresthesia, tingling, numbness; cardiopulmonary: hemoptysis, cough, dyspnea; serious neurological: dysfunction of coordination, bladder, bowel, mental status, vision, consciousness [10]. Howle assigned each case of full DCS in the p97 dataset a PSI according to the most severe symptom present. See Table 20 for the breakdown of PSI in the p97 dataset, with two symptom severity classification systems: Type I and Type II [6], and Type A and Type B [10].

Table 20: Distribution of Type I/II and Type A/B severities in the p97 dataset.

<table>
<tr><th>PSI</th><th>Occurrences in p97 Dataset</th><th>Type I/II</th><th>Type A/B</th></tr>
<tr><td>Constitutional</td><td>2</td><td rowspan="2">Type I
177 DCS Occurrences</td><td rowspan="3">Type A
199 DCS Occurrences</td></tr>
<tr><td>Lymphatic or skin</td><td>1</td></tr>
<tr><td>Pain</td><td>174</td><td rowspan="3">Type II
47 DCS Occurrences</td></tr>
<tr><td>Mild Neurological</td><td>22</td><td rowspan="2">Type B
25 DCS Occurrences</td></tr>
<tr><td>Cardiopulmonary</td><td>3</td></tr>
<tr><td>Serious neurological</td><td>22</td><td></td></tr>
</table>

5.2.5 Optimization and Statistical Methods

Three 3RUT bubble volume models were optimized in this chapter: a binomial model, and two trinomial models predicting the probabilities of mild, serious, and no DCS outcomes. One trinomial model used Type I/II severity splitting, and the other used Type A/B splitting. These models were optimized by maximizing the log likelihood (binomial: Eq. (5.10), trinomial: Eq. (5.14)) with random initial guesses against the p97 dataset in our previously developed optimization system [17]. The equivalent

binomial log likelihood was calculated for the trinomial models according to Eq. (5.15) to compare trinomial model performance with the binomial model. SigmaPlot v14 was used to generate and plot 95% prediction and confidence limits for the models' fits to the calibration dataset [53].

The Pearson Residual group statistic was used to quantify model performance on each dive type and on the dataset as a whole. In accordance with Ref. [13], the Pearson Residual is calculated as

$$PR_j = \frac{\left(obs_j - pred_j\right)^2}{pred_j\left(1 - \frac{pred_j}{N_j}\right)}, \tag{5.16}$$

where j counts over each data group, obs_j is the number of observed events in the j^{th} group, $pred_j$ is the number of model-predicted events in the j^{th} group, and N_j is the total number of exposures in the j^{th} group. Summing the Pearson Residuals for all data groups gives the Chi-squared statistic,

$$\chi^2 = \sum_{j=1}^{J} PR_j. \tag{5.17}$$

A high χ^2 value indicates the model's predictions are not aligned with the observed events in the dataset.

5.3 Results

5.3.1 Trinomial Model Comparison

The best solutions achieved by our optimization process were close but not at maximum log likelihoods. One of the limitations of model parameter estimation via log likelihood maximization is the potential for model failure when a parameter is adjusted such that the model predicts diminishing probability of DCS on a dive exposure with observed DCS [17].

The parameter sets yielding best model fits to the p97 data for the binomial and both trinomial models can be found in Table 21, along with the binomial model's log likelihood (LL_2, Eq. (5.10)), the trinomial models' log likelihoods (LL_3, Eq. (5.14)), and the trinomial models' equivalent binomial log likelihoods (LL_{32}, Eq. (5.15)). The trinomial models' log likelihoods can be compared directly, and the trinomial model with Type A/B severity splitting achieved the best fit to the calibration dataset. This trinomial A/B model also outperformed the binomial model, as its equivalent binomial log likelihood (-1258.69) is slightly higher than the binomial model's (-1260.34). The trinomial I/II model did not provide a better fit to the data than the binomial model.

The observed probability of DCS in the p97 dataset was plotted against the trinomial A/B model's predicted probabilities of DCS in Figure 20. These plots were generated by sorting the model's three-state predictions for each dive exposure by $P(0)$, and grouping the observed cases of mild and serious DCS into 10 and 5 bins of 20 and 5

cases of DCS, respectively. The observed probability of DCS was calculated as the number of DCS events in each bin divided by dive exposures divided in that bin. The model's predicted probability of DCS was calculated as the sum of the model's

Table 21: Optimized parameters used for the Binomial, Trinomial A/B, and Trinomial I/II bubble volume models.

	Binomial	Trinomial AB	Trinomial I/II
$1/k_1$ (min)	3.936E+04	2.265E+03	1.755E+04
$1/k_2$ (min)	3.100E+02	3.106E+02	2.988E+02
$1/k_3$ (min)	2.146E+06	2.513E+06	2.439E+06
G_1	4.275E+01	1.274E-01	2.115E+01
G_2	1.200E+02	1.043E+02	9.903E+01
G_2	5.470E+03	4.981E+03	4.006E+03
ξ_1 (cm^{-3})	1.342E+01	1.202E+01	2.151E+01
ξ_2 (cm^{-3})	4.471E+02	3.775E+02	4.507E+02
ξ_3 (cm^{-3})	2.975E+00	7.467E+00	7.652E+00
γ_1 (cm^2min^{-1})	5.912E-03	5.444E-03	5.386E-03
γ_2 (cm^2min^{-1})	6.592E-04	5.031E-04	4.620E-04
γ_3 (cm^2min^{-1})	4.734E+00	6.632E+00	6.725E+00
λ_1 (cm^{-1})	6.675E-03	6.744E-03	8.543E-03
λ_2 (cm^{-1})	2.993E-02	2.887E-02	3.536E-02
λ_3 (cm^{-1})	1.116E-03	8.139E-04	7.987E-04
σ ($atm \cdot cm^{-1}$)	2.466E-06	2.300E-06	2.486E-06
r_1^0 (cm)	2.617E-03	2.204E-03	2.865E-03
r_2^0 (cm)	1.272E-03	1.616E-03	1.739E-03
r_3^0 (cm)	1.460E-03	3.443E-03	3.593E-03
M_1 ($atm \cdot cm^{-3}$)	9.195E+05	9.044E+05	8.580E+05
M_2 ($atm \cdot cm^{-3}$)	2.264E+05	1.805E+05	1.951E+05
M_3 ($atm \cdot cm^{-3}$)	2.483E+05	2.086E+05	2.202E+05
a	--	9.995E-02	2.486E-01
LL_2	-1260.34	--	--
LL_3	--	-1332.75	-1373.46
LL_{32}	--	-1258.69	-1262.99

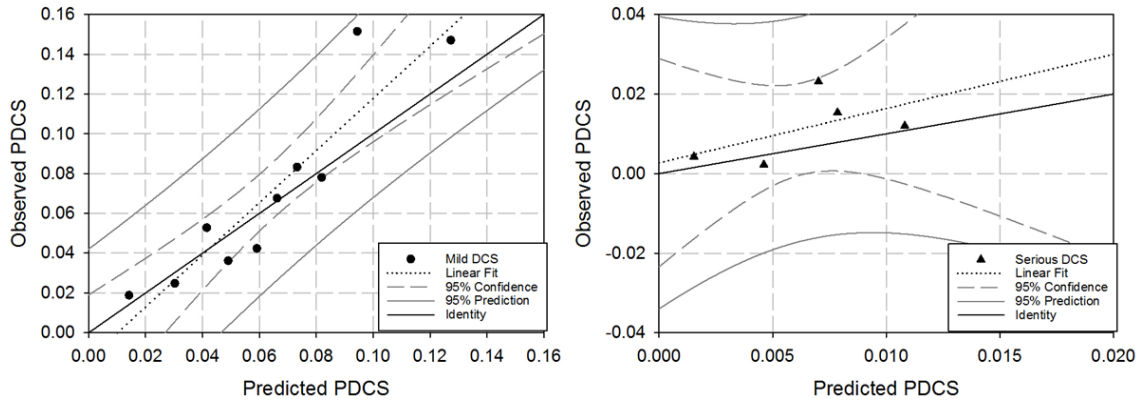

Figure 20: Trinomial A/B predicted probabilities versus observed probabilities of mild (left) and serious (right) DCS. These data points were generated by sorting the model's predictions of the probabilities of mild and serious events into bins with equal numbers of observed cases. 10 bins were used for the 199 mild DCS cases, and 5 bins were used for the 25 serious DCS cases in the p97 dataset. The probabilities are plotted with a linear fit (dotted black lines, $r^2_{mild} = 0.85$, $r^2_{serious} = 0.32$) and 95% confidence (dashed gray lines) and 95% prediction (solid gray lines) limits.

predictions for exposures in that bin divided by the number of exposures in the bin. The linear fit lines of the data points are plotted (dotted, black), along with the 95% confidence (dashed, gray) and 95% prediction (solid, gray) limits and the line of identity (solid, black). If a model was able to perfectly fit the data, all data points in Figure 20 would fall on the line of identity. The regression lines in this figure can help us determine how well the model's predictions are aligned with the observed probabilities of DCS. The serious DCS data points are less linear ($r^2_{serious} = 0.32$) than the mild DCS data points ($r^2_{mild} = 0.85$), indicating the model is better at predicting the occurrence of mild DCS than serious DCS in the p97 dataset.

5.3.2 Predictions on Data

The binomial and trinomial A/B models' predictions for each data file in the p97 dataset are given in Table 22. Both models underpredict the total number of observed DCS cases for single air, saturation, and in-water oxygen decompression diving.

Both models failed to predict any DCS from the SUBX87 data file. The data in this file is derived from open water submarine escape trials performed in Norway in July 1987, consisting of 58 dives with 2 serious DCS cases. The submarine escape dives saw a maximum depth of 602.4 fsw (19.2 ATA) and averaged around 2 minutes in duration. These short, deep dives are intended to simulate exiting from a disabled submarine. The bubble volume models tested in this chapter were unable to predict any amount of risk for these dives, meaning no compartment saw the nucleation of a bubble.

In Table 23, the models' predictions on each dive type are summarized, along with the Pearson χ^2 value for the binomial model's predictions of full DCS and the trinomial A/B model's predictions for mild and serious DCS. The null hypothesis tested here is that the model's predictions are a true reflection of the data, and a low χ^2 value supports this hypothesis. These Pearson χ^2 values further indicate that the trinomial A/B model is a better predictor of mild than serious DCS.

Table 22: DCS occurrences and trinomial AB model predictions for the p97 dataset.

		Observed DCS			Predicted DCS			
					Binomial	Trinomial AB		
Dive Type	Exposures	Mild	Serious	Total	Total	Mild	Serious	Total
Single Air								
EDU885A	483	27	3	30	24.2	22.1	2.3	24.4
DC4W	244	7	1	8	3.5	3.3	0.3	3.6
SUBX87	58	0	2	2	0	0	0	0
NMRNSW	91	4	1	5	6.6	5.5	0.6	6.1
PASA	72	4	1	5	3.3	2.8	0.3	3.1
NSM6HR	57	3	0	3	4.7	4.1	0.4	4.5
Rep&Mult Air								
EDU885AR	182	11	0	11	9.3	8.1	0.8	8.9
DC4WR	12	3	0	3	0.7	0.6	0.1	0.7
PARA	135	6	1	7	10.4	8.3	0.9	9.2
PAMLA	236	9	4	13	17.8	16.4	1.7	18.1
Single Nonair								
NMR8697	477	9	2	11	19.6	15.9	1.6	17.5
EDU885M	81	4	0	4	3.5	3.1	0.3	3.4
EDU1180S	120	9	1	10	6.2	6.2	0.6	6.8
Rep&Mult Nonair								
EDU184	239	11	0	11	13.4	12.1	1.3	13.4
PAMLAOD	134	5	1	6	9	8.2	0.9	9.1
PAMLAOS	140	5	0	5	6.8	6.3	0.6	6.9
EDU885S	94	4	0	4	4.7	4.2	0.40	4.6
Saturation								
ASATEDU	120	11	2	13	9.8	9.2	1.0	10.2
ASATNMR	50	1	0	1	3.2	3	0.3	3.3
ASATNSM	132	18	0	18	17.8	14.0	1.5	15.5
ASATARE	165	19	1	20	15.8	14.6	1.5	16.1
Air + O_2 Decompression								
NMR94EOD	284	14	3	17	14.8	13.2	1.4	14.6
DC8AOD	256	2	1	3	5.1	4.7	0.5	5.2
DC8AOW	46	3	0	3	1.0	1.0	0.1	1.1
Surface Decompression with O_2								

DC8ASUR	358	9	1	10	11.8	9.0	0.9	9.9
DCSUREP	69	1	0	1	2.8	2.2	0.2	2.4
Totals	*4335*	*199*	*25*	*224*	*225.8*	*198.1*	*20.5*	*218.6*

Table 23: Binomial and trinomial A/B predictions on each p97 dive type with the Pearson Chi-Squared statistic.

		Observed DCS			Predicted DCS Trinomial A/B		Binomial
Dive Type	Exposures	Mild	Serious	Total	Mild	Serious	Total
Single Air	1005	45	8	53	37.8	3.9	42.3
Rep&Mult Air	565	29	5	34	33.4	3.5	38.2
Single Nonair	678	22	3	25	25.2	2.5	29.3
Rep&Mult Nonair	607	25	1	26	30.8	3.2	33.9
Saturation	467	49	3	52	40.8	4.3	46.6
Air + O_2	586	19	4	23	18.9	2	20.9
Decompression Surface	427	10	1	11	11.2	1.1	14.6
Decompression with O_2							
Total	*4335*	*199*	*25*	*224*	*198.1*	*20.5*	*225.8*
Pearson χ^2					*5.552*	*9.001*	*7.763*
					p=0.9977	*p=0.9999*	*p=0.9959*

5.3.3 Binomial to Trinomial Probability Shift

In Figure 21, model-predicted probabilities of DCS and no DCS are plotted for the trinomial models against the binomial model. For all dive exposures that resulted in mild or serious DCS, the trinomial models' predictions of the probabilities of mild or serious DCS are plotted against the binomial model's predictions of the probabilities of full DCS. For dives that did not result in DCS, the trinomial models' predicted probabilities of no DCS are plotted against that of the binomial model. The trinomial

A/B model's predictions are delineated with gray markers, and the trinomial I/II model predictions are shown with white markers.

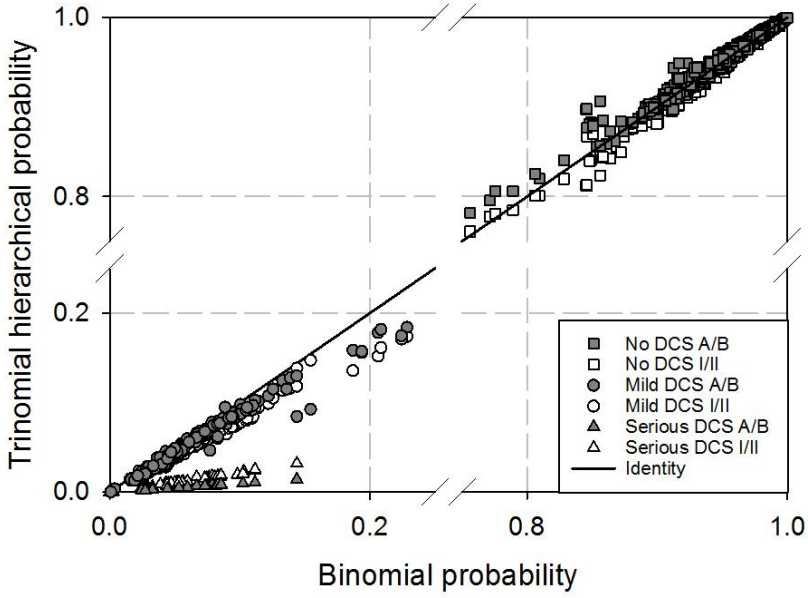

Figure 21: Trinomial to binomial probability shift plot. For dive exposures that resulted in mild or serious DCS, each trinomial models' predictions are plotted against the corresponding binomial model's predictions for that dive. Likewise for dives that did not result in DCS, the trinomial models' predictions are plotted against the binomial model's predictions of no DCS. Gray data points are the shift between trinomial A/B and binomial probability, and white data points display the trinomial I/II to binomial probability shift. Mild DCS events are indicated in circles, serious DCS as triangles, and no DCS as square markers.

If the trinomial and binomial models made the same predictions, all data points in Figure 21 would fall on the line of identity. Any point below the line of identity indicates the binomial model made a higher prediction for the probability of DCS or no DCS for that dive, and any point above the line of identity means a trinomial model predicted a higher probability of DCS or no DCS for that dive.

By fitting a linear regression line to each set of data points, we can quantitively assess the shift in probability between the trinomial and binomial models. The linear regression slopes and corresponding r^2 values for each set of data points in Figure 21 can be found in Table 24. For dives that resulted in DCS, both trinomial models predict a lower probability of the occurrence of DCS than the binomial model. The trinomial A/B model makes a slightly higher prediction of the probability of mild DCS and a slightly lower prediction of the probability of serious DCS than the trinomial I/II model for dives with DCS. The data points corresponding to dives with serious DCS outcomes have the least amount of scatter. The trinomial models' predictions of the probabilities of no DCS events shift in opposite ways; the trinomial A/B model predicts higher probabilities than the binomial model, and the trinomial I/II model predicts lower probabilities than the binomial model.

Table 24: Slope of Linear Regression and r^2 value for each probability shift in Figure 21.

Figure 21 Probability Shift	Slope of Linear Regression	r^2
Trinomial A/B Mild DCS to Binomial	0.8331	0.9610
Trinomial I/II Mild DCS to Binomial	0.7876	0.9755
Trinomial A/B Serious DCS to Binomial	0.0987	0.9912
Trinomial I/II Serious DCS to Binomial	0.2221	0.9887
Trinomial A/B No DCS to Binomial	0.9509	0.9714
Trinomial I/II No DCS to Binomial	1.0366	0.9835

5.3.4 Cumulative Density Function

The cumulative density functions (CDFs) for the calibration dataset and the trinomial A/B model's predictions are plotted in Figure 22. The CDF helps us to visualize how a model's DCS onset time predictions align with symptom onset in the data. The CDFs for the trinomial model's predictions of mild and serious DCS are identical because the model distinguishes between severities mathematically with a scaling factor, so when the CDFs are computed, that scaling factor divides out. Figure 22 shows the model's predicted CDF for all DCS types (dashed line, black) and the CDFs

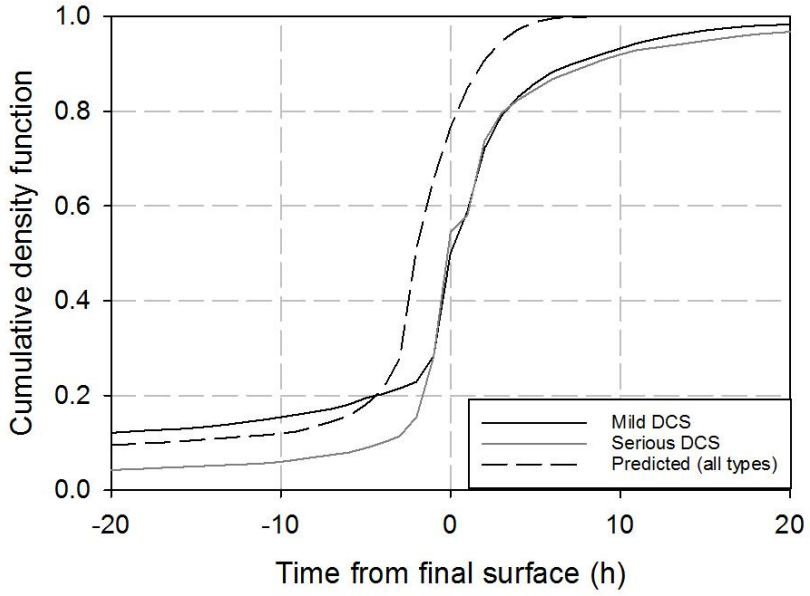

Figure 22: Trinomial A/B cumulative density function. The cumulative density functions for mild DCS and serious DCS in the p97 dataset are the black curve and the gray curve, respectively. The model's predicted cumulative density function for all DCS events is plotted with the dashed black line. Because the model distinguishes between mild and serious DCS with a scaling factor, both predicted cumulative density functions fall on the same curve.

for the observed cases of mild DCS (solid line, black) and serious DCS (solid line, gray) in the p97 dataset. Time on the x-axis is relative to the time of final surfacing, negative time indicating decompression has not yet been completed. The model's predicted CDF under-predicts the cumulative probability of mild DCS and over-predicts the cumulative probability of serious DCS until 4 hours prior to surfacing, after which the model over-predicts the cumulative probability of all types of DCS. Thus the model is predicting that more DCS events occur prior to surfacing than is observed in the data.

5.4 Discussion

In this Chapter, we optimized a binomial and two trinomial 3RUT bubble volume models to the p97 dataset. The trinomial A/B model provides a better fit to the calibration dataset than the binomial bubble volume model. From Figure 20 and the Pearson χ^2 values in Table 23, we can conclude that the trinomial A/B model predicts the occurrence of mild DCS in the dataset more accurately than that of serious DCS. This is similar to our findings in Chapter 4, in which the tetranomial A/B LE1 gas content model was a better predictor of mild DCS than serious or marginal DCS.

The trinomial A/B model is particularly proficient at predicting the occurrence of mild DCS resulting from the high oxygen content dives that augment the BIG292 dataset to create the p97 dataset. When comparing the success of this model on the various dive types with the models presented in Chapters 3 and 4, all multinomial models presented in this dissertation make significant under-predictions on the occurrence of full (mild

and serious) DCS resulting from single air dives. The trinomial marginal LE1 gas content model from Chapter 3 and the tetranomial A/B LE1 gas content model from Chapter 4 both over-predict the occurrence of full (mild and serious) DCS from saturation diving, a trend thought to be related to the concentration of marginal DCS events resulting from saturation diving. The trinomial A/B bubble volume model presented in this Chapter, which considered marginal DCS cases as non-events, underpredicts the occurrence of mild DCS from saturation diving, and in general is more accurate than the preceding models for predictions on saturation diving.

The results presented in this Chapter can also be compared with those published by Howle *et al*. in Ref [10], who optimized binomial and trinomial (mild, serious, and no DCS) gas content models. Unlike our trinomial bubble volume models, Howle's trinomial gas content models optimized to nearly identical parameter sets as the binomial counterpart. This resulted in their trinomial models making nearly identical predictions on the dataset as the binomial model. While both Howle's trinomial A/B and trinomial I/II gas content models outperformed the binomial model, only our trinomial A/B bubble volume model provided a better fit to the data than the binomial model. Howle's trinomial A/B gas content model's equivalent binomial log likelihood saw much more substantial improvement over the binomial model than our trinomial A/B model over our binomial model.

The trinomial models presented in this chapter were unable to predict any risk for submarine escape dives (see Table 22). Bubble volume model risk is derived from bubble nucleation, and our models did not nucleate a bubble on these short, deep profiles. Bubble nucleation occurred if Eq. (5.5) was satisfied. There are other options for bubble formation criteria, such as including the elastic modulus in Eq. (5.5), or modeling nucleation as a balance between tissue supersaturation and surface energy [62]. Perhaps one of these alternatives would allow our models to generate DCS risk from submarine escape dive data.

Although bubble volume models have the advantages of more realistically simulating the physiological precursor to DCS and more accurately predicting DCS for dives with higher oxygen content when compared with gas content models, bubble volume models are vulnerable to overfitting. With 23 adjustable parameters calibrated against 4,335 dive exposures, our trinomial models may be overfit to the p97 dataset. To investigate the possibility of overfitting, our trinomial A/B model should be tested on dive data outside the calibration dataset.

5.5 Conclusions

The trinomial bubble volume models presented in this chapter serve to augment the binomial bubble volume models currently used by the U.S. Navy to estimate the risk of diving missions. By providing the U.S. Navy divers with both the probability and

severity of DCS occurring, our trinomial model presents divers with a more complete set of risk information.

The trinomial model with Type A/B severity splitting provided a better fit to the calibration dataset than the binomial model and the trinomial model with Type I/II splitting. This trinomial A/B model was most accurate in predicting the occurrence of mild DCS in the dataset, and was able to successfully predict risk for dives using oxygen during in-water decompression and surface decompression. However, the trinomial A/B model was not able to assign risk to submarine escape dives, likely because the short nature of these exposures did not allow for adequate tissue supersaturation required to trigger bubble nucleation in the model. This issue may be addressed by implementing a different condition for bubble nucleation, or by adjusting the optimization algorithm to achieve a more optimal parameter set.

Because the trinomial A/B model has 23 adjustable parameters and was calibrated on a dataset with only 4,335 dive exposures, it is possible that this model was overfit to the data. To test for overfitting, the model should be evaluated on dive data from outside the p97 dataset. In addition, it would be worthwhile to investigate the critiques presented in Ref. [62] of this formulation of bubble volume model, and potentially implement some of the suggested improvements.

5.6 Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1644868 and by the U.S. Navy, Naval Sea Systems Command under contracts #N00024-13-C-4104 and #N00024-17-C-4317. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the U.S. Navy. Computational resources were provided by BelleQuant Engineering, PLLC. Neither of the funding agencies nor the commercial entity played any role in designing this study, data collection and analysis, decision to publish, interpreting the results, or writing the manuscript.

6. Conclusions

There are two factors that contribute the risk of an activity: the probability of injury, and the severity of injury. An activity's risk rises with the increase in the probability of injury and with the increase in the severity of injury. Risk analysis can be conducted with a risk assessment matrix, which visualizes harm probability against harm severity. Risk assessment matrices are used by many facets of the U.S. Department of Defense to manage operational risk.

To mitigate the risk of DCS for diving missions, the U.S. Navy currently uses binomial probabilistic models, which predict the probability of DCS occurring and not occurring for a given time-depth dive profile. These models provide Navy divers with DCS incidence and onset time predictions, but do not predict any information regarding the severity of potential injury. With the goal of providing U.S. Navy divers with a more complete risk prediction for dive missions, we first interrogated the BIG292 calibration dataset for evidence of DCS severity-dependent onset time information, then developed a variety of multinomial gas content and bubble volume probabilistic models that predict the probability of no DCS, serious DCS, mild DCS, and/or marginal DCS.

Probabilistic decompression models have adjustable parameters that can be calibrated to an empirical dive dataset, such as the BIG292 dataset. This data was recorded with the symptom onset time window when DCS occurred. The occurrence density function, which visualizes the frequency in DCS onset over time relative to the

completion of decompression, shows two peaks in DCS onset in the BIG292 dataset. The probabilistic models used by the U.S. Navy do not predict this temporal bimodality in symptom onset. After investigating the source of the bimodal behavior, we concluded that the DCS symptom onset time windows recorded in the BIG292 dataset may have been biased by dive trial protocol, which dictated that divers underwent medical exams immediately after completing decompression and again two hours later. The two peaks in reported symptom onset coincide with these two medical exams, likely because medical officers made diagnoses during these scheduled exams. The bimodal trend in symptom onset time is present in the occurrence density functions generated by subsets of the data for each classification of DCS symptom severity and dive type. Probabilistic models' predictions of a single peak in DCS symptom onset over time do not indicate a defect with the models, but rather a bias in the data, and it is thus not necessary to develop a model that replicates this bimodal trend.

Probabilistic decompression models can be modified to predict both the probability and severity of DCS injury. Previously, Howle *et al*. developed a trinomial gas content model that predicts the probability of mild, serious, and no DCS. Expanding on that work, we developed and optimized 20 multinomial models in this dissertation: six trinomial marginal gas content models, 12 tetranomial gas content models, and two trinomial bubble volume models. Of all multinomial gas content models tested, the LE1 formulation provided the best fit to the calibration dataset. We found that multinomial

models that predict mild and serious DCS performed best with Type A/B severity splitting when compared with Type I/II splitting. A summary of the best performing multinomial models' Pearson χ^2 and p-values can be found in Table 25.

Table 25: Pearson χ^2 and p-values for all multinomial models. The null hypothesis tested here is that model predictions are identical to observations in the data. A high χ^2 (and low p-value) suggests rejection of the null hypothesis, meaning the model's predictions are not consistent with observation.

Multinomial Model	Number of DOF	Pearson χ^2 Full DCS	Pearson χ^2 Mild DCS	Pearson χ^2 Serious DCS	Pearson χ^2 Marginal DCS
Binomial LE1nt Gas Content [10]	7	8.465 p=0.294			
Trinomial A/B LE1nt Gas Content [10]	8		8.421 p=0.393	4.527 p=0.807	
Trinomial Marginal LE1 Gas Content (Chapter 3)	9	12.270 p=0.199			36.568 p=0.000031
Tetranomial A/B LE1 Gas Content (Chapter 4)	10		7.597 p=0.668	9.246 p=0.509	36.612 p=0.000066
Binomial 3RUT Bubble Volume (Chapter 5)	22	7.736 p=0.998			
Trinomial A/B 3RUT Bubble Volume (Chapter 5)	23		5.552 p=0.999	9.007 p=0.996	

The trinomial marginal LE1 A/B model and the tetranomial LE1 A/B model, which both considered marginal DCS to be a hierarchical event, did not accurately

predict the occurrence, onset, or distribution of marginal DCS cases in the calibration dataset. The inclusion of marginal DCS data during model fitting may have hindered those models' ability to predict the probabilities of other types of DCS as well. Both the trinomial marginal LE1 model and the tetranomial A/B LE1 model over-predict the occurrence of full DCS from saturation diving, which is the dive type with the largest frequency of marginal DCS cases in the data. The trinomial marginal LE1 model is not able to predict the probability of full DCS as well as the binomial LE1nt or binomial 3RUT models, which consider marginal DCS a non-event. The probability shift between the tetranomial A/B and trinomial marginal LE1 models indicates both models make nearly identical predictions on the dataset. Many of the marginal DCS cases in the calibration dataset were not reported with symptom onset times, so the onset times windows were assigned to be between the divers' last medical checks and the trials' right-censored times. These large event windows hinder model optimization, as symptom onset times are used to refine model parameters. We do not recommend further development of multinomial models that predict marginal DCS separately, and rather suggest reassigning marginal DCS cases to non-events in the calibration dataset.

The bubble volume models presented in Chapter 5 yielded the highest p-values in Table 25, indicating those models' predictions are most closely aligned with the calibration dataset, even though they did not achieve a maximum log likelihood during model fitting. These models were not able to accumulate any risk for submarine escape

dives, which are short, deep dives intended to simulate the decompression burden from exiting a disabled submarine. An additional issue with these bubble volume models is they may have more adjustable parameters than warranted by the number of dive exposures in the p97 dataset. Future work should include testing all multinomial models presented in this work on dive data from outside the calibration datasets. Poor performance on new dive data would indicate the model is overfit to the calibration dataset and cannot be extrapolated to dive profiles that are different from those used during model fitting.

References

[1] E.P. Kindwall, A Short History of Diving and Diving Medicine, in: A.A. Bove, J. Davis (Eds.) Diving Medicine. W. B. Saunders, Philadelphia, 1990, pp. 1-8.

[2] D.H. Elliott, J. Vorosmarti, An Outline History of Diving Physiology and Medicine, in: A.O. Brubakk, T.S. Neuman (Eds.) Bennett and Elliott's Physiology and Medicine of Diving. Saunders, Great Britain, 2003, pp. 4-16.

[3] R.E. Moon, R.D. Vann, P.B. Bennett, The Physiology of Decompression Illness. Scientific American, 273(2): 54-61, 1995.

[4] A.E. Boycott, G.C.C. Damant, J.S. Haldane, The Prevention of Compressed-Air Illness. Journal of Hygiene, 8(03): 342-443, 1908.

[5] R.D. Vann, Mechanisms and Risks of Decompression, in: A.A. Bove, J.C. Davis (Eds.) Diving Medicine. W.B. Saunders, Philadelphia, 1990, pp. 29-45.

[6] U.S. Department of Navy, U.S. Navy Diving Manual. Revision 7, Naval Sea Systems Command, Washington, D.C., 2016.

[7] D.J. Temple, R. Ball, P.K. Weathersby, E.C. Parker, S.S. Survanshi, The Dive Profiles and Manifestations of Decompression Sickness Cases After Air and Nitrogen-Oxygen Dives. Volume I: Data Set Summaries, Manifestation Descriptions, and Key Files. NMRC 99-02(Vol. I), Naval Medical Research Center, Bethesda, MD, 1999.

[8] D.J. Temple, R. Ball, P.K. Weathersby, E.C. Parker, S.S. Survanshi, The Dive Profiles and Manifestations of Decompression Sickness Cases After Air and Nitrogen-Oxygen Dives. Volume II: Complete Profiles and Graphic Representations for DCS Events. NMRC 99-02(Vol II.), Naval Medical Research Center, Bethesda, MD, 1999.

[9] R.D. Vann, P.J. Denoble, L.E. Howle, P.W. Weber, J.J. Freiberger, C.F. Pieper, Resolution and Severity in Decompression Illness. Aviation, Space, and Environmental Medicine, 80(5): 466-471, 2009.

[10] L.E. Howle, P.W. Weber, E.A. Hada, R.D. Vann, P.J. Denoble, The Probability and Severity of Decompression Sickness. PLOS ONE, 12(3), 2017.

[11] T.E. Berghage, J.M. Woolley, L.J. Keating, The Probabilistic Nature of Decompression Sickness. Undersea Biomedical Research, 1(2): 189-196, 1974.

[12] P.K. Weathersby, L.D. Homer, E.T. Flynn, On the Likelihood of Decompression Sickness. Journal of Applied Physiology: Respiration Enviromental Exercise Physiology, 57(3): 815-825, 1984.

[13] W.A. Gerth, Overview of Survival Functions and Methodology, in: P.K. Weathersby, W.A. Gerth (Eds.) Fifty-first Workshop of the Undersea and Hyperbaric Medical Society, Undersea and Hyperbaric Medical Society, Seattle, WA, 2002, pp. 1-48.

[14] P.K. Weathersby, S.S. Survanshi, L.D. Homer, E. Parker, E.D. Thalmann, Predicting the Time of Occurrence of Decompression Sickness. Journal of Applied Physiology, 72(4): 1541-1548, 1992.

[15] E.D. Thalmann, E.C. Parker, S.S. Survanshi, P.K. Weathersby, Improved Probabilistic Decompression Model Risk Predictions Using Linear-Exponential Kinetics. Undersea and Hyperbaric Medical Society, 24(4): 255-274, 1997.

[16] E.C. Parker, S.S. Survanshi, P.K. Weathersby, E.D. Thalmann, Statistically Based Decompression Tables VIII: Linear-Exponential Kinetics. NMRI 92-73, Naval Medical Research Institute, Bethesda, MD, 1992.

[17] L.E. Howle, P.W. Weber, R.D. Vann, A Computationally Advantageous System for Fitting Probabilistic Decompression Models to Empirical Data. Computers in Biology and Medicine, 39(12): 1117-1129, 2009.

[18] W.A. Gerth, R.D. Vann, Probabilistic Gas and Bubble Dynamics Models of Decompression Sickness Occurrence in Air and Nitrogen-Oxygen Diving. Undersea and Hyperbaric Medical Society, 24(4): 275-292, 1997.

[19] W.A. Gerth, R.D. Vann, Development of Iso-DCS Risk Air and Nitrox Decompression Tables Using Statistical Bubble Dynamics Models. Duke University Medical Center, Durham, NC, 1996.

[20] R.S. Srinivasan, W.A. Gerth, Mathematical Models of Diffusion-Limited Gas Bubble Evolution in Perfused Tissue. NEDU TR 13-05, Navy Experimental Diving Unit, Panama City, FL, 2013.

[21] W.A. Gerth, R.S. Srinivasan, F.G. Murphy, K.A. Gault, A Probabilistic Model of Altitude Decompression Sickness Based on the 3RUT-MB Model of Gas Bubble Evolution in Perfused Tissue. NEDU TR 18-01, Navy Experimental Diving Unit, Panama City, FL, 2018.

[22] W.W. Lowrance, Of Acceptable Risk: Science and the Determination of Safety, William Kaufmann, Inc., Los Altos, CA, 1976.

[23] U.S. Department of the Army, Combat Health Support in Stability Operations and Support Operations. Appendix C: Risk Assessment. Field Manual 8-42, Washington, D.C., 1997.

[24] Operational Risk Management (ORM), OPNAV Instruction 3500.39A Marine Corps Order 3500.27A.

[25] H.D. Van Liew, E.T. Flynn, Decompression Tables and Dive-Outcome Data: Graphical Analysis. Undersea & Hyperbaric Med., 32(4): 184-198, 2005.

[26] A.E. King, F.G. Murphy, L.E. Howle, Bimodal Decompression Sickness Onset Times are Not Related to Dive Type or Event Severity. Computers in Biology and Medicine, 91: 59-68, 2017.

[27] A.E. King, N.R. Andriano, L.E. Howle, Trinomial Decompression Sickness Model using Full, Marginal, and Non-Event Outcomes. Computers in Biology and Medicine, 118, 2020.

[28] E.C. Parker, S.S. Survanshi, P.K. Weathersby, Improving on a "Good" Model, in: P.K. Weathersby, W.A. Gerth (Eds.) Fifty-first Workshop of the Undersea and Hyperbaric Medical Societ, Undersea and Hyperbaric Medical Society, Seattle, WA, 2002, pp. 137-141.

[29] E.A. Hada, On the advancement of probabilistic models of decompression sickness (Doctoral Dissertation) (Doctoral Dissertation). Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, 2016.

[30] F.G. Murphy, E.A. Hada, D.J. Doolette, L.E. Howle, Probabilistic Pharmacokinetic Models of Decompression Sickness in Humans, Part 1: Coupled Perfusion-Limited Compartments. Computers in Biology and Medicine, 86: 55-64, 2017.

[31] L.E. Howle, P.W. Weber, R.D. Vann, M.C. Campbell, Marginal DCS Events: Their Relation to Decompression and Use in DCS Models. Journal of Applied Physiology, 107(5): 1539-1547, 2009.

[32] F.G. Murphy, The impact of weighting marginal DCS events as non-events, pharmacokinetic gas content models, and optimal decompression schedule calculation (Doctoral Dissertation) Mechanical Engineering and Materials Science, Duke University, Durham, NC, 2017.

[33] L.E. Howle, P.W. Weber, E.A. Hada, R.D. Vann, P.J. Denoble, The probability and severity of decompression sickness. PLOS ONE, 12(3): e0172665, 2017.

[34] L.E. Howle, P.W. Weber, J.M. Nichols, Bayesian Approach to Decompression Sickness Model Parameter Estimation. Computers in Biology and Medicine, 82: 3-11, 2017.

[35] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, 2nd ed. Cambridge University Press, Cambridge, UK, 1992.

[36] E.D. Thalmann, I.P. Buckingham, W.H. Spaur, Testing of Decompression Algorithms for Use in the U. S. Navy Underwater Decompression Computer. Phase I. NEDU TR 11-80, Navy Experimental Diving Unit, Panama City, FL, 1980.

[37] E.D. Thalmann, Phase II Testing of Decompression Algorithms for Use in the U.S. Navy Underwater Decompression Computer. NEDU TR 1-84, Navy Experimental Diving Unit, Panama City, FL, 1984.

[38] E.D. Thalmann, Air-N_2O_2 Decompression Computer Algorithm Development. NEDU TR 8-85, Navy Experimental Diving Unit, Panama City, FL, 1986.

[39] E.D. Thalmann, P.C. Kelleher, S.S. Survanshi, E.C. Parker, P.K. Weathersby, Statistically Based Decompression Tables XI: Manned Validation of the LE Probabilistic Model for Air and Nitrogen-Oxygen Diving. NMRI 99-01/NEDU TR 1-99, Joint Technical Report: Naval Medical Research Center and Navy Experimental Diving Unit, Bethesda, MD, 1999.

[40] P.K. Weathersby, B.L. Hart, E.T. Flynn, W.F. Walker, Human Decompression Trial in Nitrogen-Oxygen Diving. NMRI 86-97, Naval Medical Research Institute, Bethesda, MD, 1986.

[41] C.L. Shake, P.K. Weathersby, D. Wray, E.C. Parker, Decompression Sickness Resulting from Long Shallow Air Dives. NSMRL 1200, Naval Submarine Medical Research Laboratory, 1996.

[42] R.Y. Nishi, K.E. Kisman, I.P. Buckingham, B.C. Eatock, G. Masurel, XDC-2 Digital Decompression Computer: Assessment of Decompression Profiles by Ultrasonic Monitoring, Phase I: 36-54 msw. DCIEM 80-R-32, Defence and Civil Institute of Enviromental Medicine, Downsview, ON, 1980.

[43] R.Y. Nishi, B.C. Eatock, I.P. Buckingham, G. Masurel, XDC-2 Digital Decompression Computer: Assessment of Decompression Profiles by Ultrasonic Monitoring, Phase II: 30-75 msw. DCIEM 81-R-02, Defence and Civil Institute of Enviromental Medicine, Downsview, ON, 1981.

[44] R.Y. Nishi, B.C. Eatock, I.P. Buckingham, B.A. Ridgewell, Assessment of Decompression Profiles by Ultrasonic Monitoring, Phase III: No-Decompression Dives. DCIEM 82-R-38, Defence and Civil Institute of Environmental Medicine, Downsview, ON, 1982.

[45] G.R. Lauckner, R.Y. Nishi, B.C. Eatock, Evaluation of the DCIEM 1983 Model for Compressed Air Diving (Series A-F). DCIEM 84-R-72, Defence and Civil Institute of Environmental Medicine, Downsview, ON, 1984.

[46] G.R. Lauckner, R.Y. Nishi, B.C. Eatock, Evaluation of the DCIEM 1983 Model for Compressed Air Diving (Series G-K). DCIEM 84-R-73, Defence and Civil Institute of Environmental Medicine, Downsview, ON, 1984.

[47] G.R. Lauckner, R.Y. Nishi, B.C. Eatock, Evaluation of the DCIEM 1983 Model for Compressed Air Diving (Series L-Q). DCIEM 85-R-18, Defence and Civil Institute of Environmental Medicine, Downsview, ON, 1985.

[48] E.C. Parker, S.S. Survanshi, P.B. Massell, P.K. Weathersby, Probabilistic Models of the Role of Oxygen in Human Decompression Sickness. Journal of Applied Physiology, 84(3): 1096-1102, 1998.

[49] F.G. Murphy, A.J. Swingler, W.A. Gerth, L.E. Howle, Iso-Risk Air No Decompression Limits After Scoring Marginal Decompression Sickness Cases as Non-Events. Computers in Biology and Medicine, 92: 110-117, 2018.

[50] P.K. Weathersby, S.S. Survanshi, R.Y. Nishi, E.D. Thalmann, Statistically Based Decompression Tables VII: Selection and Treatment of Primary Air and N2O2 Data. NSMRL 1182/NMRI92-85, Joint Technical Report: Naval Submarine Medical Research Laboratory and Naval Medical Research Institute, Bethesda, MD, 1992.

[51] F.G. Murphy, E.A. Hada, D.J. Doolette, L.E. Howle, Probabilistic Pharmacokinetic Models of Decompression Sickness in Humans, Part 2: Coupled Perfusion-Diffusion Models. Computers in Biology and Medicine, 92: 90-97, 2018.

[52] L.E. Howle, Analytic Gain in Probabilistic Decompression Sickness Models. Computers in Biology and Medicine, 43: 1739-1747, 2013.

[53] Systat Software Inc, 2107 North First Street, Suite 360. San Jose, CA 95131.

[54] P. Tikuisis, K.A. Gault, R.Y. Nishi, Prediction of Decompression Illness using Bubble Models. Undersea & Hyperbaric Med., 21(2): 129-143, 1994.

[55] P. Tikuisis, R.Y. Nishi, Role of Oxygen in a Bubble Model for Predicting Decompression Illness. DCIEM 94-04, Defence and Civil Institute of Environmental Medicine, Ontario, Canada, 1994.

[56] H.D. Van Liew, M.P. Hlastala, Influence of Bubble Size and Blood Perfusion on Absorption of Gas Bubbles in Tissues. Respiration Physiology, 7: 111-121, 1969.

[57] R.S. Srinivasan, W.A. Gerth, M.R. Powell, Mathematical Models of Diffusion-Limited Gas Bubble Dynamics in Tissue. Journal of Applied Physiology, 86: 732-741, 1999.

[58] R.S. Srinivasan, W.A. Gerth, M.R. Powell, Mathematical Model of Diffusion-Limited Evolution of Multiple Gas Bubbles in Tissue. Annals of Biomedical Engineering, 31: 471-481, 2003.

[59] R.S. Srinivasan, W.A. Gerth, M.R. Powell, Mathematical Model of Diffusion-Limited Gas Bubble Dynamics in Unstirred Tissue with Finite Volume. Annals of Biomedical Engineering, 30: 232-246, 2002.

[60] R.S. Srinivasan, W.A. Gerth, M.R. Powell, A Mathematical Model of Diffusion-Limited Gas Bubble Dynamics in Tissue with Varying Diffusion Region Thickness. Respiration Physiology, 123: 153-164, 2000.

[61] A.E. King, L.E. Howle, Tetranomial Decompression Sickness Model using Serious, Mild, Marginal, and Non-Event Outcomes. In submission, 2020.

[62] C. Walsh, E. Stride, U. Cheema, N. Ovenden, A Combined Three-Dimensional In Vitro-In Silico Approach to Modelling Bubble Dynamics in Decompression Sickness. Journal of The Royal Society Interface, 14(137), 2017.

Biography

Amy King earned a B.S. in Mechanical Engineering from the University of Massachusetts Amherst in 2016, and a M.S. in Mechanical Engineering from Duke University in 2019. While attending Duke University, she received the James B. Duke Fellowship (2016) and the National Science Foundation Graduate Research Fellowship (2018).

Publications:

A.E. King, F.G. Murphy, L.E. Howle, *Bimodal Decompression Sickness Onset Times are Not Related to Dive Type or Event Severity*. Computers in Biology and Medicine, 91 (2017) 59-68.

A.E. King, N.R. Andriano, L.E. Howle, *Trinomial Decompression Sickness Model using Full, Marginal, and Non-Event Outcomes*. Computers in Biology and Medicine, 118 (2020).

A.E. King, L.E. Howle, *Tetranomial Decompression Sickness using Serious, Mild, Marginal, and Non-Event Outcomes*. In submission.

A.E. King, L.E. Howle, *Three-Region Unstirred Tissue Bubble Volume Trinomial Decompression Sickness Model using Serious, Mild, and Non-Event Outcomes.* In preparation.

Reports:

Howle L.E., Freiberger J.J., Diehl S., Natoli M., King A.E., Vann R.D., *Transfer of Duke Dive Trial Data to U.S. Navy – Vol. 1 MK15 UBA Dive Trials – Rev. A.* Report to Deep Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 1049 pages.

Howle L.E., Freiberger J.J., Diehl S., Natoli M., King A.E., Vann R.D., *Transfer to Duke Dive Trial Data to U.S. Navy – Vol. 2 – Fife Deep Air Dive Trials – Rev. A.* Report to Deep Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 14 pages.

Howle L.E., Freiberger J.J., Diehl S., Natoli M., King A.E., Vann R.D., *Transfer to Duke Dive Trial Data to U.S. Navy – Vol. 3 – SIO2 Dive Trials – Rev. A.* Report to Deep

Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 44 pages.

Howle L.E., Freiberger J.J., Diehl S., Natoli M., King A.E., Vann R.D., *Transfer to Duke Dive Trial Data to U.S. Navy – Vol. 4 – NOAA Dive Trials – Rev. A.* Report to Deep Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 29 pages.

Howle L.E., Freiberger J.J., Diehl S., Natoli M., King A.E., Vann R.D., *Transfer to Duke Dive Trial Data to U.S. Navy – Vol. 5 – Ascent Rate Trials – Rev. A.* Report to Deep Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 38 pages.

Howle L.E., Dunford R., Denoble P., King A.E., Vann R.D., *Transfer to Duke Dive Trial Data to U.S. Navy – Vol. 6 – Project Dive Exploration – Rev. A.* Report to Deep Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 8675 pages.

Howle L.E., King A.E., Vann R.D., *Transfer to Duke Dive Trial Data to U.S. Navy – Vol. 7 – NMRC-99 augmented with Perceived Severity Index – Rev. A.* Report to Deep Submergence Biomedical R&D Program, Naval Sea Systems Command, 00CM, U.S. Navy. 2019. 19 pages.

Abstracts & Presentations:

King A.E., Murphy F.G., Howle L.E., *Bimodal decompression sickness onset times are not related to dive type or event severity*, Proceedings of the 2017 Undersea and Hyperbaric Medical Society Meeting. Naples, FL. 2017.

Lewis D.S., King A.E., Murphy F.G., Kim C.Y., Nelson R.C., Howle L.E., *Hemodialysis Catheter Integrity during Mechanical Power Injection of Iodinated Contrast Medium*, Proceedings of the 2017 Radiological Society of North America 2017, Chicago, IL. 2017

King A.E., Howle L.E., Vann R.D., *Tetranomial Probabilistic Modeling of Serious, Mild, Marginal, and No Decompression Sickness*, Proceedings of the 2019 Undersea and Hyperbaric Medical Society Meeting. Rio Grande, PR. 2019.