Review article

Severity scoring in decompression illness

Simon Mitchell

Key words
Decompression sickness, decompression illness, severity, treatment sequelae, morbidity, review article

Abstract

(Mitchell S. Severity scoring in decompression illness. *SPUMS J*. 2005; 35: 199-205.)
Scoring systems for disease severity may be designed to predict outcome (prognostication), to facilitate clinical decision making or to stratify patients into subgroups of comparable severity to enable comparisons of therapeutic strategies. This review considers systems that have been proposed for scoring severity of decompression illness (DCI). Most attempts to score severity of DCI have focused on prognostication. One system has been designed to be used to stratify DCI severity in clinical trials, and to provide a linear index of severity that may be used to quantify recovery.

Statement on terminology

This paper adopts the editorial policy described by Francis and Mitchell with respect to the terminology of dysbaric disease.[1] Specifically, the term 'decompression sickness' (DCS) will explicitly refer to symptomatic evolution of bubbles from dissolved inert gas in blood and/or tissues, and the term 'arterial gas embolism' (AGE) will imply the introduction of bubbles to the pulmonary veins as a result of pulmonary barotrauma. The term 'decompression illness' (DCI) will be employed, as suggested by Francis and Smith,[2] to embrace the clinical presentations of both the above pathological processes.

Introduction

Scoring systems for severity of disease are applied in many medical disciplines, usually with one or more of the following goals:

- prediction of outcome
- support of clinical decision making
- stratification of patients into comparable severity subgroups to enable valid comparisons of facilities or therapeutic strategies.

Examples of such systems that have achieved wide acceptance in their respective disciplines include the Acute Physiology and Chronic Health Evaluation (APACHE) score in either of its iterations,[3] which is applied to predict mortality in intensive care patients, and the Glasgow Coma Score,[4] which is applied in the setting of brain injury for all of the above purposes.

Genesis of a scoring system for severity of disease usually involves a process with some or all of the following components.[5]

1 Patient selection. The patient population to which the system will be applied is defined.
2 Outcome selection. Most scoring systems are designed to allow prediction of a particular outcome, for example, the presence of symptoms at discharge from hospital, and this must be chosen and defined.
3 Predictor variable selection and data collection. Clinical factors considered likely to influence the chosen outcome(s) are identified, and relevant data pertaining to these variables in a patient population of known outcome are recorded.
4 Assembly of the model. Combinations of mathematical techniques (such as multivariate logistic regression) applied to the data collected at 3, and clinical judgement are used to derive weightings for the chosen variables in the scoring paradigm.
5 Transformation of the score into a probability for the outcome(s) of interest. Various approaches can be used to establish the prognostic implications of a particular score or score interval.
6 Validation of the model. Ideally, the prognostic power of the system is validated (predicted versus actual outcomes) in a separate group of patients who were not involved in the derivation of the system.

Despite the proposal of several systems for scoring severity of DCI,[6–10] some of which have been developed through a similar process to that outlined above, no method has achieved universal or even widespread acceptance. There are many difficulties in designing such a system, not least of which is a profound lack of prospectively gathered data that relate manifestations to outcome.[8] As a result, it becomes very difficult to objectively rank the protean manifestations of DCI in order of prognostic importance, and to derive appropriate weightings for their respective contributions to a severity score. Moreover, since the clinical spectrum of severity in DCI is so broad (it can range from feeling 'off colour' to paralysis, circulatory collapse and death), the derivation of a scoring system that can both embrace the entire spectrum of disease and be considered linearly related is a substantial challenge. Finally, the matter is further complicated by the existence of two pathophysiological processes (DCS and AGE) that are sometimes indistinguishable clinically but which have different natural histories. Not surprisingly, most of the proposed systems limit their scope to subsets of DCI patients (such as those with neurological disease), and some specifically exclude patients suspected of suffering AGE.

This paper will review the scoring systems that have been proposed to this point with discussion of their strengths and weaknesses. Possible future directions will be considered.

Extant scoring systems

DICK AND MASSEY 1985[6]

This appears to be the first attempt to ascribe a numerical score to indicate severity in DCS. The score is derived by adding two 'grades' chosen from descriptive scales of sensory and motor symptoms respectively. The weighting of the grades for the various manifestations appears to have been based on clinical experience rather than an objective process designed to determine their relative importance. The maximum grade in each scale is 5, giving a maximum possible severity grade of 10 (Table 1). This simple structure makes it quick and easy to apply. However, the system can be applied only to patients with neurological DCS, not those suffering from AGE, and takes no account of non-neurological manifestations. Moreover, the score descriptors (Table 1) are open to interpretation. "Weakness" and "paralysis" can overlap as can "paraesthesia" and "numbness", and thus there is a risk of inconsistent application.[7]

Table 1
The neurologic decompression sickness severity scale proposed by Dick and Massey.[6] The scores for each scale are added to give a maximum possible score of 10

Sensory symptoms

Grade	Symptom	Extent
1	Paraesthesia	Single limb or area
2	Paraesthesias	Multiple regions
3	Numbness	Single region or limb
4	Numbness	Two regions or limbs
5	Numbness	Three or more limbs

Motor symptoms

Grade	Symptom	Extent
1	Weakness	Single limb or muscle group
2	Weakness	Multiple limbs or muscle groups
3	Paralysis	Single limb or muscle group
4	Paralysis	Two limbs
5	Paralysis	Three or more limbs

Table 2
Severity as graded by the Dick and Massey system[6] versus outcome as reported in two studies

Source	Severity	Complete relief/ total cases	%
Dick & Massey[6]	Mild (1–3)	24/24 (treated)	100
		10/11 (untreated)	91
	Moderate (4–6)	10/14	71
	Severe (7–10)	1/6	16
Ball[11]	Mild (1–3)	13/14	93
	Moderate (4–6)	4/11	36
	Severe (7–10)	2/24	8

Not surprisingly, when this system was applied to a population of DCS patients, it was demonstrated that greater proportions of those with low grades exhibited either spontaneous recovery or complete relief with recompression than those with higher grades.[6] This was also observed by Ball (1993)[11] who utilised the Dick and Massey system in an investigation of the effect of severity and time to recompression on outcome in DCS.

The severity-versus-outcome data reported in these two studies are summarised in Table 2. Ball also used percentage changes in severity grade (referenced to the admission score) to track recovery during treatment and at discharge.[11]

BOND ET AL 1990[7]

This group compared treatment outcomes in divers treated with different initial recompression tables. They developed a descriptive system for categorising DCI severity in order to control for this parameter in their assessment of recompression efficacy. The severity categories (Table 3) were arbitrarily defined on the basis of the authors' experience. There was no attempt to validate the categories and, surprisingly, data presented in the original paper suggested that the categories were not very predictive of outcome defined as resolution or persistence of symptoms 24 hours after the last recompression treatment.

KELLEHER ET AL 1996[8]

This group accumulated 214 case reports of neurological DCI in standardised format from multiple hyperbaric units. No attempt was made to separate these cases into sub-diagnoses of DCS and AGE, and patients without neurological manifestations were excluded. Each case was classified according to: the modalities affected

Table 3
Severity 'categories' for DCI as proposed by Bond et al[7]

Code	Severity	Descriptors
0	No symptoms	Nil
1	DCS I (mild)	Pain, rash, itching
2	DCS II (mild/moderate)	Pain, numbness/tingling, restlessness, headache, skin – sensation, muscle twitch
3	DCS II (moderate)	Ringing ears, pain, fatigue, reflex change
4	DCS II (moderate/severe)	Weakness, numbness/tingling, breathing, nausea, vomiting, hearing loss, skin – sensation, personality change, walk – standing
5	DCS II (severe)	Visual disturbance, speech disturbance, weakness, paralysis, bladder, bowel
6	AGE	Semiconscious, unconscious, paralysis, convulsions

(consciousness, sensory, motor); the number of sites involved in any sensory or motor deficit; and the anatomical locations of the manifestations. Post-treatment outcomes were recorded. In particular, the focus in development of the subsequent predictive model was on outcomes after the first recompression treatment. A relationship between incomplete resolution following the first treatment and the pattern of pre-treatment sensory and/or motor manifestations was apparent, and this relationship was further described in terms of a linear logistic model. This model allowed calculation of predicted outcome (incomplete resolution after the first treatment) according to the pattern of sensory and motor manifestations as shown in Table 4. It should be noted that this is not a severity scoring system *per se* (no score is actually derived); rather, it is a manifestation-based model for predicting outcome after the first treatment.

BOUSSUGES ET AL 1996[9]

This group reviewed 96 cases of DCS treated at a single hospital over an eight-year period. Patients suffering from non-neurological DCS or suspected AGE were excluded. Univariate analyses examined the relationship between various parameters (characteristics of the dive, the evacuation and the clinical presentation) and adverse outcome defined as incomplete resolution at one month after treatment. Those parameters found to influence outcome were: progression of symptoms prior to recompression; the presence and degree of motor impairment; the presence of urinary function disturbance; and the presence of objective sensory disturbance. Interestingly, delay to treatment was not found to influence outcome (though this analysis was not adjusted for disease severity). Coefficients of importance were assigned to those parameters that appeared to have a significant influence on outcome. Repetitive diving was also included in the group of influential factors even though its influence was non-significant in the univariate analysis. The value of the assigned coefficients appears to have been influenced by the significance of the parameter in the univariate analysis, and by the 'experience' of the authors. The severity score is calculated by adding all applicable coefficients (Table 5) at presentation.

The system was validated in a population of divers treated at a separate institution. Sixty-six divers were scored at presentation, and their outcomes (complete or incomplete recovery at one month) were noted. The positive and negative predictive values (for the presence of any sequelae) for a score greater than 7 were 86% and 89% respectively (Table 6). Subsequently, Pitkin et al retrospectively applied this system to a further 217 cases of neurological DCI recorded on the Institute of Naval Medicine database for the years 1995–96.[12] In a similar analysis to that performed by Boussuges et al, they recorded outcomes in patients whose score was less than or equal to 7, and greater than 7. However, instead of classifying outcomes in terms of presence or absence of any sequelae, they used categories designated as "no or mild sequelae" and "severe sequelae". The validity of this approach in assessing the prognostic

Table 4
Observed and 'expected' (from application of the model) numbers of cases with incomplete resolution of any manifestations after first treatment in relation to initial sensory/motor manifestations[8]
(X = symptoms present)

Sensory		Motor		Total	Cases with incomplete resolution	
Arm	Leg	Arm	Leg		Actual n (%)	Expected n
X				56	10 (18)	10.3
	X			43	18 (42)	16.4
X	X			20	7 (35)	8.5
		X		12	1 (8)	1.6
			X	2	0 (0)	0.5
		X	X	2	0 (0)	0.5
X		X		10	3 (30)	1.6
	X		X	32	18 (56)	17.1
X			X	0		
	X	X		2	0 (0)	0.7
X	X	X		0		
X	X		X	2	1 (50)	1.2
X		X	X	2	1 (50)	0.5
	X	X	X	2	1 (50)	1.0
X	X	X	X	4	2 (50)	2.2

Table 5
Attribution of 'coefficients' of importance for use in calculating the Boussuges gravity score[9]
(severity score is calculated by summing all relevant coefficients for the patient)

Parameter	**Descriptor**	**Coefficient or score**						
		0	**1**	**2**	**3**	**4**	**5**	**6**
Repetitive dive	No	X						
	Yes			X				
Clinical course prior to HBO	Improvement	X						
	Stable				X			
	Deterioration						X	
Objective sensory disorder	No	X						
	Yes					X		
Motor impairment	No	X						
	Monoparesis, paraparesis, or tetraparesis					X		
	Paraplegia							X
	Hemiplegia				X			
Urinary disturbance	No	X						
	Yes						X	

power of the system is open to debate since the majority of patients with a score of 7 or less would actually have had to deteriorate during treatment to find themselves in the "severe sequelae" group. Not surprisingly, the negative predictive value of a score 7 or less for severe sequelae is 99%. In contrast, the positive predictive value of a score greater than 7 for severe sequelae is only 16% (Table 6). In comparing their results with those published by Boussuges, the authors appear to have overlooked these discrepancies in outcome group definitions.

Table 6
Prognostic value of the Boussuges scoring system as assessed by the original authors[9] and by Pitkin et al[12]

As assessed by the original authors[9]

Score	**Sequelae**	**Recovery**	**Total**
>7	18	3	21
≤7	5	40	45
Totals	23	43	66

sensitivity = 78%; specificity = 93%; positive predictive value = 86%; negative predictive value = 89%

As assessed by Pitkin et al[12]

Score	**Severe sequelae**	**No/mild sequelae**	**Total**
>7	16	71	87
≤7	1	129	130
Totals	17	200	217

sensitivity = 94%; specificity = 65%; positive predictive value = 18%; negative predictive value = 99%

MITCHELL ET AL 1998[10]

In the mid 1990s the Royal New Zealand Navy (RNZN) group planned a randomised trial of lignocaine as an adjuvant to recompression in DCI patients. For a number of compelling reasons,[13] an alternative trial was subsequently conducted in cardiac surgery patients.[14] However, the initial intention to involve divers raised issues of severity stratification, trial group homogeneity, and comparative outcome measurement. The last of these was a particularly vexing problem. The intention was to involve divers with serious neurological DCI, and since many such patients are left with residual symptoms despite maximal therapy, it was felt that complete versus incomplete recovery was too insensitive a measure of outcome for the purposes of the trial. For example, lignocaine might have provided greater degrees of recovery without necessarily altering the complete–incomplete recovery ratio between the respective trial groups. There was no system in existence which allowed scoring of DCI on a scale that implied linear separation of the various presentations according to their 'severity'. A new system was developed with the key goal being an ability to assess relative severity and recovery of the trial patients, rather than to facilitate prognosis.[10]

In application of the RNZN system, the first step is to score each of 22 possible manifestations of DCI on a scale of 0 (absent) to 3 (maximal). Structured guidelines on how to derive this score for each manifestation are specified. In the second step, the unweighted score for each manifestation is multiplied by conversion factors intended to reflect its relative importance, and the progression of the manifestation prior to treatment. The conversion for relative importance reflects the fact that some manifestations are less likely to resolve and more likely to result in long-term problems than others. Once each manifestation has been scored, and the score 'converted' to reflect its relative importance and progression, the individual manifestation scores are summed to give the patient's overall severity score. Thus, all manifestations are scored (not just neurological), and all contribute to the overall severity score, even though the relative contribution of minor, naturally evanescent symptoms would be quite small.

Since the overall severity score is assumed to lie on a linear continuum of relative severity, then re-scoring during recovery and subtracting the new score ("progress score") from the initial score ("admission score") provides a numerical index of recovery (or deterioration) that, in theory, can be used to compare relative recovery between patients. If a patient's progress score is 0 (that is, they have fully recovered), then their recovery score will be equal to their admission score.

The derivation of the importance of conversion factors was a complex process that is best appreciated by reference to the original paper.[10] There was (and still is) no database describing the outcome and quality-of-life implications of individual manifestations of DCI. It follows that three experienced diving physicians independently rated 22 manifestations of DCI on scales describing their respective specificity for DCI, their natural history if untreated, and their potential to cause incapacity in daily living if unresolved. This allowed generation of a numerical index of relative importance of the individual manifestations. Unfortunately, an objective validation of these essentially subjective rankings would be an extraordinarily difficult task.

Holley (2000)[15] conducted a validation study of the RNZN system. One hundred consecutive case files for divers treated for DCI at the RNZN hospital were retrospectively reviewed. Sufficient data were available for application of the system in 79 patients. Although it was not intended as a prognostic scoring system, Holley used threshold scores of ≤25 and >25 to conduct a similar analysis of positive and negative predictive values to those reported by Boussuges et al and Pitkin et al (Table 7).[9,12] The positive and negative predictive values were 77% and 89% respectively. In support of the assumption that scores were linearly related to severity (and in the absence of any other more appropriate marker of severity), Holley correlated the number of treatments required to reach complete resolution or plateau with the admission score ($r = 0.80$, $r^2 = 0.64$, Figure 1).

Table 7
Prognostic value of the RNZN scoring system as assessed by Holley[15]

Score	Sequelae	No sequelae	Total
>25	10	3	13
≤25	7	59	66
Totals	17	62	79

sensitivity = 58%; specificity = 95%; positive predictive value = 77%; negative predictive value = 89%

Discussion

Scoring systems for DCI severity are desirable for two purposes:

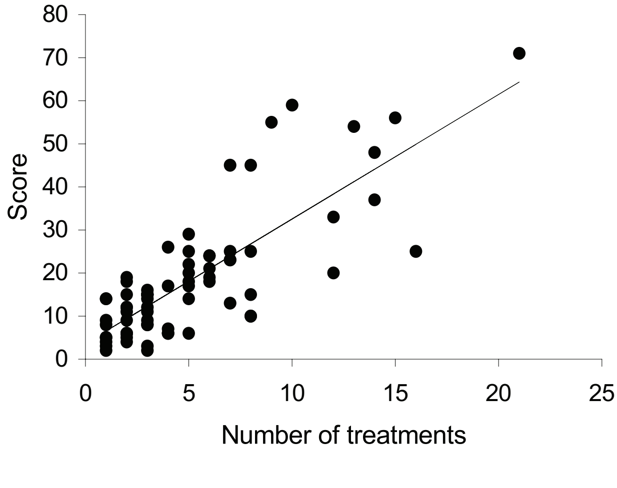
1 to prognosticate and guide clinical decision making
2 in the context of research, to measure severity and recovery in divers participating in trials of therapy.

A third goal might be a scoring system to facilitate DCI diagnosis, but this is unrelated to severity and is not considered further here.

PROGNOSTICATION AND CLINICAL DECISION MAKING

It is self-evident that a capacity for prognostication is useful for advising patients of their likely outcome. The expectations of badly injured divers are frequently in need of temperance, and outcome prediction based on objective data is always better than what might appear to the patient as a vague clinical impression. Prognostication might also prove highly valuable in clinical decision making. For example, a system that predicted outcomes for mildly injured divers in the absence of recompression would be invaluable in the management of remote DCI cases. Unfortunately, given the current fashion for recompressing all DCI patients no matter how trivial their disease, it is hard to see how such a system could be developed and validated objectively for contemporary populations of sport divers. A more realistic example of prognostic scoring in clinical decision making might include earlier withdrawal of repetitive recompression therapy and avoidance of delays in accessing

Figure 1
Correlation of RNZN scoring system admission score versus the number of recompressions required to achieve resolution or plateau in recovery[15]



rehabilitation in cases whose score indicates a poor prognosis even with extended courses of multiple recompressions. It is also conceivable that should superior efficacy for more aggressive recompression protocols ever be demonstrated, a severity scoring system might be useful in determining which patients should be allocated to such treatments.

To be useful in a busy clinical environment, a scoring system for prognostication and clinical decision making should be relatively simple and quick to apply. Several of the extant systems fit this description.[6,7,9] In particular, the system proposed by Boussuges et al is simple and quick, and it appears to perform well in prediction of outcome. However, patients with non-neurological disease cannot be scored using this system (nor one of the others that could be considered as 'quick and simple').[6] Non-neurological DCS may still result in sequelae and it follows that if prediction of outcome is the goal, then a separate or all embracing system will need to be developed. The Bond system does allow categorisation of all patients,[7] but we have no idea of the validity of the categories. The RNZN system can be used to score patients with most manifestations of DCS, and performs almost as well prognostically as the Boussuges system, yet over the entire range of disease (neurological and non-neurological).[10] However, it is not so simple and readily applied, and is probably better suited to the research situation (see below).

MEASUREMENT OF SEVERITY AND RECOVERY IN RESEARCH

A method of ensuring equivalence between study groups is vital in any study involving a disease as variable as DCI.[8,9] Even more important is a sensible measure of outcome. The use of incomplete versus complete recovery is hopelessly insensitive and potentially very misleading; for example, how does one compare a dense paraplegic who eventually walks out of hospital with a limp (incomplete recovery) with a patient with upper-limb sensory change who recovers completely. Both have neurological disease. The paraplegic has made a functionally more important but incomplete recovery, whilst the other patient has recovered completely. The use of percentage recovery as reported by Ball is similarly flawed since the recovery percentages take no account of the initial 'absolute' severity.

The RNZN system attempts to avoid these pitfalls by generating overall severity scores that are weighted to reflect the natural history and potential for functional handicap of the component manifestations. If it can be assumed that the scores are truly linearly related, then the changes in score during treatment provide a valid index of the size and significance of recovery that has occurred. However, despite the authors' attempts to objectify the allocation of importance weightings for the various manifestations of DCI, the process used must still be considered somewhat arbitrary. In addition, the system is cumbersome to apply, although an automated spreadsheet recently developed by James Francis has simplified matters considerably.

Future directions

Large databases of DCI patients whose manifestations and outcomes are carefully recorded are the key to developing scoring systems that both embrace non-neurological manifestations and more accurately predict outcome. However, even with such databases the task may not be easy. Ideally, discrete clinical DCI syndromes could be identified and the prognostic implications of these syndromes could be defined. However, although most clinicians probably perceive several 'typical' DCI presentations, it is noteworthy that principal component analysis of the DAN database (incorporating 2,822 case records) failed to identify any clear syndromes.[16]

At the present time, the Boussuges system probably represents the 'standard' in terms of prognostication, but is limited by being applicable only to those patients with neurological disease. The target should be a relatively simple algorithm similar to the Boussuges system, but which includes score components for important predictors of outcome in non-neurological disease also. Preferably, the validation process for such a system would examine its application (for prognostication) not only on initial assessment of the diver, but also at progressive stages of treatment.

For assessing severity and recovery, and for comparing these parameters between groups, a system designed to reflect true differences in significance of symptoms and to accurately reflect their progress is more appropriate, despite the almost inevitable increase in its complexity. At this time, the RNZN system probably represents the 'standard', but there is almost certainly room for improvement. In particular, a more objective determination of the relative importance of the various symptoms is almost certainly possible.

References

1 Francis TJR, Mitchell SJ. Manifestations of decompression disorders. In: Brubakk AO, Neumann TS, editors. *Bennett and Elliott's physiology and medicine of diving*. 5th ed. London: Harcourts Publishers; 2003. p. 578-99.
2 Francis TJR, Smith DH, editors. *Describing decompression illness*: The 42nd Undersea and Hyperbaric Medical Society Workshop. Bethesda, MD: Undersea and Hyperbaric Medical Society; 1991.
3 Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991; 100: 1619-36.
4 Teasdale G, Jennett B. Assessment of coma and impaired

consciousness. A practical scale. *Lancet*. 1974; 2: 81-4.

5 Moreno R, Matos R. The “new” scores: what problems have been fixed, and what remain? *Curr Opin Crit Care*. 2000; 6: 158-65.

6 Dick AP, Massey EW. Neurologic presentation of decompression sickness and air embolism in sport divers. *Neurology.* 1985; 35: 667-71.

7 Bond JG, Moon RE, Morris DL. Initial table treatment of decompression sickness and arterial gas embolism. *Aviat Space Environ Med*. 1990; 61: 738-43.

8 Kelleher PC, Pethybridge RJ, Francis TJ. Outcome of neurological decompression illness: development of a manifestation-based model. *Aviat Space Environ Med*. 1996; 67: 654-8.

9 Boussuges A, Thirion X, Blanc P, Molenat F, Sainty JM. Neurologic decompression illness: a gravity score. *Undersea Hyperb Med*. 1996; 23: 151-5.

10 Mitchell SJ, Holley T, Gorman DF. A new system for scoring severity and measuring recovery in decompression illness. *SPUMS J*. 1998; 28: 84-94.

11 Ball R. Effect of severity, time to recompression with oxygen, and re-treatment on outcome in forty-nine cases of spinal cord decompression sickness. *Undersea Hyperb Med*. 1993; 20: 133-45.

12 Pitkin AD, Benton PJ, Broome JR. Outcome after treatment of neurological decompression illness is predicted by a published clinical scoring system. *Aviat Space Environ Med*. 1999; 70: 517-21.

13 Mitchell SJ, Pellett O, Gorman DF. Open chamber cardiac surgery: a clinical injury model for arterial gas embolism. *SPUMS J*. 1997; 27: 230-5.

14 Mitchell SJ, Pellett O, Gorman DF. Cerebral protection by lidocaine during cardiac operations. *Ann Thorac Surg*. 1999; 67: 1117-24.

15 Holley T. Validation of the RNZN system for scoring severity and measuring recovery in decompression illness. *SPUMS J*. 2000; 30: 75-80.

16 Dear G deL, Vann RD, Pieper BP, et al. DCI symptoms do not appear to form syndromes. *Undersea Hyperb Med*. 1997; 24: 33.

Simon Mitchell, PhD, is Honorary Senior Lecturer, Occupational Medicine Unit, Department of Medicine, Faculty of Medicine & Health Sciences, The University of Auckland.

Address for correspondence:
“Aromahana”
45 Opanuku Road, Henderson Valley,
Auckland, New Zealand
Phone: *+64-(0)9-814-9999*
Fax: *+64-(0)9-814-9982*
E-mail: *<dr.m@xtra.co.nz>*

This paper was originally presented at the Workshop on Consistent Diagnosis of Decompression Illness, Duke University, NC, USA, 17–18 April 2003.