Consensus Development Conference

Consensus guidelines for the use of ultrasound for diving research

Andreas Møllerløkken, S Lesley Blogg, David J Doolette, Ronald Y Nishi and Neal W Pollock

Abstract

(Møllerløkken A, Blogg SL, Doolette DJ, Nishi RY, Pollock NW. Consensus guidelines for the use of ultrasound for diving research. *Diving and Hyperbaric Medicine*. 2016 March;46(1):26-32.)
The International Meeting on Ultrasound for Diving Research produced expert consensus recommendations for ultrasound detection of vascular gas bubbles and the analysis, interpretation and reporting of such data. Recommendations for standardization of techniques to allow comparison between studies included bubble monitoring site selection, frequency and duration of monitoring, and use of the Spencer, Kisman-Masurel or Eftedal-Brubakk scales. Recommendations for reporting of results included description of subject posture and provocation manoeuvres during monitoring, reporting of untransformed data and the appropriate use of statistics. These guidelines are available from <www.dhmjournal.com/>.

Key words
Doppler; echocardiography; bubbles; cardiovascular; right-to-left shunt; decompression illness; diving research; meetings; review article

Introduction

The International Meeting on Ultrasound for Diving Research was held on 25–26 August 2015 in Karlskrona, Sweden. It brought together an international group of 27 physicians and scientists from 12 countries with the goal of developing consensus guidelines to aid investigators in designing research protocols and reviewers who may evaluate submitted reports. Topics addressed both Doppler ultrasound and newer two-dimensional imaging modalities. Discussion areas included the strengths and limitations of different techniques, technician training, monitoring and grading protocols, data handling and reporting. The following consensus guidelines were agreed upon through discussions during the meeting and during a post-meeting period when draft documents were circulated to the delegates. The guidelines produced by the panel are not exhaustive, but may aid in standardizing and, in some cases, improving experimental techniques. Future efforts can refine these guidelines and incorporate new and emerging technologies and procedures.

Bubbles and decompression stress

Some of the bubbles which form as a consequence of decompression can be detected by ultrasonic methods. Although technology is evolving, the most common technique is the detection of intravascular bubbles using either a Doppler flow transducer or two-dimensional echocardiography. The detection of bubbles in any individual is not diagnostic for decompression sickness (DCS). However, the bubble load detected in large systemic veins and, in particular, in the mixed venous blood is considered to be correlated with the probability of DCS. In large compilations of data, the number of venous bubbles is correlated with the observed incidence of DCS.[1,2] Therefore, ultrasonically-detected bubbles can be a useful outcome measure for some research questions.

The ability of bubble measurements to answer specific research questions should be considered carefully. If bubble studies are appropriate, they must be designed and conducted such as to produce useful results and should be reported in a manner that can be compared meaningfully to the rest of the scientific literature. A wide variety of monitoring protocols and data analyses can be found in the literature and in manuscripts submitted for publication. Whilst some variants are well founded, others reflect weaknesses in methodology that would best not be perpetuated. Ideally, well established protocols should be employed for ultrasonic monitoring. Variations should be clearly justified, should be based on scientific merit and with consideration of the value of comparison with other studies. Investigators who are new to ultrasonic detection of bubbles are encouraged to seek assistance from experienced peers to develop effective protocols.

The purpose of these guidelines is to present recommendations for best practice and standardization of protocols for ultrasonic detection of bubbles for diving research. The goal is not to stifle scientific creativity or thoughtful differences; protocols are expected to continue to be refined, or new ones developed, to improve utility or take advantage of new technological capabilities and developments. These are designed to help investigators develop and implement useful protocols. Journal editors and reviewers may also find this information useful to consider when evaluating manuscripts submitted with bubble data.

Table 1

The Spencer Scale[4] is an ordinal scale developed to facilitate semi-quantitative grading of intravascular bubble signals identified with aural Doppler ultrasound technology; Roman numerals are used to remind users that these are non-parametric data

Grade	
0	No bubble signals;
I	Occasional bubble signals; great majority of cardiac cycles signal free;
II	Many, but less than half, of the cardiac cycles contain bubble signals;
III	Most cardiac cycles contain bubble signals, but not obscuring signals of cardiac motion;
IV	Bubble signals sounding continuously throughout systole and diastole, obscuring normal cardiac signals.

Table 2

The Kisman-Masurel scale[7] was developed to allow bubble signals identified with aural Doppler ultrasound technology to be evaluated on multiple parameters; the individual parameter codes (scored with Arabic numerals) and then combined and converted to yield a single semi-quantitative ordinal, non-parametric grade (Roman numerals)

Code	Frequency (f), bubbles/cardiac period
0	0
1	1–2
2	several, 3–8
3	rolling drumbeat, 9–40
4	continuous sound

Code	Rest % (p)	Movement duration (d)
0	0	0
1	1–10	1–2
2	10–50	3–5
3	50–99	6–10
4	100	> 10

Code	Amplitude (A)
0	No bubbles discernable
1	Barely perceptible, $A_b << A_c$
2	Moderate amplitude, $A_b < A_c$
3	Loud, $A_b \approx A_c$
4	Maximal, $A_b > A_c$

Technician training

There is no credentialing standard for certifying the competency of ultrasonography technicians involved with decompression research. Obtaining interpretable ultrasound bubble signals requires practice, and grading of these signals is subjective. The reliability of research data can be enhanced by documentation of technical skill and assurance of inter-rater reliability between laboratories. Researchers who are new to ultrasonic detection of bubbles should seek training with an established laboratory or undertake an independent, blinded review of their data.[3] It is expected that 10% of the total recordings from a study, or at least 30 recordings, would constitute a minimum review effort. An independent data reviewer should be able to request and evaluate any recording reported in a study; an inability to provide the requested recordings would be cause for concern and could prompt the call for a more comprehensive review.

RECOMMENDATION 1

Ultrasound technician training and/or level of experience should be described in research reports. It is to be encouraged that research teams without established records with these techniques include the results of independent, blind reviews of their data by established investigators. These should identify the reviewer, the absolute number of records reviewed, the percentage of total measures reviewed, and the agreement between researcher and reviewer scores.

Signal grading – Doppler

While many Doppler grading scales have been described in the literature, the two most widely accepted are the Spencer and Kisman-Masurel (KM) ordinal grading scales.[4–7] Both have been used sufficiently over several decades to warrant recognition as standards of practice. The KM scale does offer the advantage that KM grades can be converted to Spencer grades. Spencer grades cannot be converted to KM values. The Spencer scale consists of five grades (0–IV) representing increasing numbers of bubbles in the Doppler signal (Table 1). The KM scale has 12 grades (0, I-, I, I+, II-, II, II+, III-, III, III+, IV-, IV), and grading is a two-step procedure. First, the Doppler signal is assigned a three-digit code, *fpA* for at rest and *fdA* for movement conditions (Table 2), where *f* (frequency) is the number of bubbles per cardiac period; *p* is the percentage of cardiac periods with specified bubble frequency at rest or *d* is the number of cardiac cycles with elevated bubble sounds after movement and *A* is the amplitude of bubble sounds (A_b) in comparison to normal blood flow/cardiac sounds (A_c).[7,8] Next, the three-digit code is converted to its corresponding KM grade (Table 3).

Signal grading – two-dimensional echocardiography

Two-dimensional imaging is gaining popularity over aural Doppler scanning. The grading scales are still evolving, as is appropriate for advances in the technology. Again, while a number of scales have been published, the original and expanded forms of the Eftedal-Brubakk (EB) scale are most widely used (Tables 4 and 5).[9–11] There are published data showing the association of Spencer and KM grades with the incidence of DCS[1,2] and demonstrating the correspondence between the EB scale and the Spencer and KM scales.[2,12]

Modifications that subdivide existing grades within well-established grading scales are potentially useful to take advantage of future, improved detection methodologies. Such expanded scales can be collapsed back to the original grades for comparison with previous studies and validation data.

Table 3
Conversion of KM codes (fpA/fdA) to KM Bubble Grades

fpA fdA	Bubble grade	fpA fdA	Bubble grade	fpA fdA	Bubble grade	fpA fdA	Bubble grade
111	I-	211	I-	311	I	411	II-
112	I	212	I	312	II-	412	II
113	I	213	I+	313	II	413	II+
114	I	214	II-	314	II	414	III-
121	I+	221	II-	321	II	421	III-
122	II	222	II	322	II+	422	III
123	II	223	II+	323	III-	423	III
124	II	224	II+	324	III	424	III+
131	II	231	II	331	III-	431	III
132	II	232	III-	332	III	432	III+
133	III-	233	III	333	III	433	IV-
134	III-	234	III	334	III+	434	IV
141	II	241	III-	341	III	441	III+
142	III-	242	III	342	III+	442	IV
143	III	243	III	343	III+	443	IV
144	III	244	III+	344	IV-	444	IV

Table 4
Eftedal-Brubakk scale[10]

Grade
0 – no observable bubbles
I – occasional bubbles
II – at least one bubble every four cardiac cycles
III – at least one bubble every cardiac cycle
IV – at least one bubble·cm^{-2} in every image
V – whiteout; single bubbles cannot be discriminated

Table 5
Expanded Eftedal-Brubakk scale (fairly widely published [11,12])

Grade
0 – no observable bubbles
I – occasional bubbles
II – at least one new bubble every four cardiac cycles
III – at least one new bubble every cardiac cycle
IV a – at least one bubble·cm^{-2} in every image
b – at least three bubbles·cm^{-2} in every image
c – near whiteout; individual bubbles still discerned
V – whiteout; individual bubbles cannot be discerned

RECOMMENDATION 2

Doppler signal grading should employ either the Spencer or KM scales. When the KM scale is used, ideally the KM grades converted to Spencer grades should also be reported. Two-dimensional imaging should use an original or expanded EB scale. Modifications of these scales or alternative scales should be clearly explained and validated to justify use.

Subject selection

There is a high degree of inter-subject variability in intravascular bubble development; some individuals bubble readily while others are relatively resistant to bubbling.[13,14] This reality is best handled by study designs in which individuals serve as their own controls. With this approach, the relative risk of different exposures can be more effectively assessed. Bubble data are far less appropriate to establish absolute risk.

RECOMMENDATION 3

Employ repeated measures designs, with subjects serving as their own controls to improve the assessment of relative risk.

Monitoring site selection

The standard site for Doppler monitoring of venous gas bubbles in decompression studies is the precordium, as this captures the entire systemic venous return. Subclavian monitoring is sometimes used for additional information. The standard for two-dimensional echocardiographic imaging of the heart is the apical long-axis view, which allows assessment of bubbles in the entire systemic venous return and any subsequent systemic arterialization of bubbles. Subcostal monitoring may be appropriate for smaller individuals. Parasternal views do not provide comparable fields to the apical or subcostal views for bubble grading. Optimal windows for ultrasonic measures can vary on an individual basis, requiring technicians to adjust their approach on a case-by-case basis.

RECOMMENDATION 4

The precordial site should be used as the standard for Doppler monitoring. Subclavian monitoring may be useful in providing additional information. The apical window should be used as the standard for two-dimensional imaging.

Body position

Numerous scanning positions have been reported: standing, seated, supine, and left lateral decubitus. Variation does make cross-study comparison more difficult.

RECOMMENDATION 5

Body position during monitoring should be standardized where practical and fully described in reports.

Provocation

Bubble measurements can be made at the end of a period during which subjects remain at rest or following active provocations that can promote showers of detectable bubbles. These provocations include intentional coughing, deep knee bends, and single, paired, or sequential limb movements. Separate measurements may be made after different provocations, particularly separate upper and lower limb movements, which can produce distinctly different results. Resting bubble measurements and provocation bubble measurements have different associations with the probability of DCS; ideally, measurements should be made following both rest and provocation.[1,2]

RECOMMENDATION 6

Resting measurements should always be made. The minimum period of rest prior to the measurement should be standardized and reported. When measurements following provocation are collected, the provocation should be standardized and clearly described. Irrespective of whether the analysis focuses on rest or provocation measurements, both should be reported.

Monitoring duration

The period following decompression during which bubble measurements are made should be designed to ensure capture of maximum bubble grade and other metrics of interest. These other metrics may include times of onset and disappearance of detectable bubbles (the latter often demonstrated by two consecutive grade zero scans). The duration of monitoring can vary appreciably as a function of the exposure variables, including: the dive profile; physical exercise; thermal stress and breathing gases. The time course for bubble onset, maximum grade and waning is not always predictable.[15]

RECOMMENDATION 7

As a standard rule, measurements should be conducted for 120 minutes from the completion of the decompression period. Shorter monitoring periods should be clearly justified. Consideration should be given to extending monitoring periods if bubbles persist at the end of the planned period. Pilot trials may be warranted to establish appropriate monitoring endpoints for exposure profiles known or expected to produce bubbles beyond 120 minutes.

Frequency of measurements

The frequency of measurements during the monitoring period is important to establish confidence that a meaningful assessment has been made.[15] The substantial variability of frequency of measurements between published reports has been problematic. Infrequent measurements are operationally easier but increase the likelihood of missing periods of active bubbling and maximum grade. Frequent measurements are more operationally demanding but much more likely to capture maximum grade and temporal patterns of detectable bubbles.

RECOMMENDATION 8

The first measurements should be made within 15 minutes following decompression. During the first 120 minutes following decompression, measurement intervals should be no greater than 20 minutes. Sampling frequency may be reduced after 120 minutes following decompression. Shorter or longer sampling intervals may be warranted for some exposures and depending on the objective of the study.

Data pooling

Grade pooling may be appropriate for analyzing and reporting bubble data. A wide range of data handling practices have been employed and they are often idiosyncratic. The pooling of bubble grades should reflect meaningful clusters.[1] Grade 'zero' has a high negative predictive value for DCS and should not be pooled with other grades.[2]

RECOMMENDATION 9

Given evidence of an increased association between DCS and the highest Spencer/KM grades, pooling grades I–II and III–IV may be appropriate. Zero grades should be reported but not pooled with other grades. Wherever possible, unpooled data should be included to allow reanalysis.

Data reporting

A variety of parameters can be reported from ultrasonic imaging. Reporting multiple parameters and raw data facilitates reanalysis and potentially comparison between studies.

RECOMMENDATION 10

Standard parameters to report include time to onset of non-zero grades, time to maximum grade reached, and maximum grade for individual subjects. In addition, median grade, grade range and mode can be reported; all measured zero grades should be included in calculated summary statistics. Wherever feasible, raw data should be reported. If deemed appropriate, data transformation may be used to allow time integration of non-zero grades to be computed. Otherwise, data transformations should be used judiciously with clear justification and, in all cases, the untransformed data should also be reported.

Statistics

Bubble grades represent nonlinear ordinal data for which nonparametric analysis is appropriate. Roman numerals are frequently employed with grading scales as a reminder that computation of means and associated measures of variability are not valid with ordinal data. Transformations purported to linearize bubble data do not make the data suitable for parametric hypothesis testing. Such transformations may be useful to compute time integrals,[13,16] or for some forms of linear modelling. There is substantial inter- and intra-individual variability in maximum bubble grade produced after identical exposures, so comparative studies should be designed with enough subjects to ensure appropriate power to detect a difference of interest. One analysis of two-dimensional echocardiographic data indicated a paired sample size of 50 subjects was required for 80% power to detect a one-grade difference in VGE (two-sided $\alpha = 0.05$).[17]

RECOMMENDATION 11

Bubble grade data are most appropriately analyzed non-parametrically. Attempts to linearize bubble data should be employed cautiously. Consideration should also be given to ensure that studies are powered appropriately.

Fair interpretation

Interpretation of bubble data should be appropriately constrained, for a number of reasons:

- bubbles do not equal DCS;
- the intravascular focus of current technology provides an incomplete picture of conditions in the body;
- the standard techniques of aural Doppler and two-dimensional cardiac imaging do not allow bubble sizing;
- Doppler technology captures only a limited three-dimensional space and two-dimensional images only a slice of the three-dimension field.

Most measures are made intermittently, capturing a small percentage of total time.

While recognition of limitations is the responsibility of authors, peer reviewers should critically evaluate manuscripts for shortcomings.

RECOMMENDATION 12

The limitations of bubble data should be considered as part of any interpretation of study results. Peer reviewers must ensure that a reasonable standard has been met to justify publication.

Data preservation

Research standards typically require preservation of raw data.

RECOMMENDATION 13

Ideally, measurements conducted for research publication should be recorded and preserved for future review. This includes audio and visual files, as appropriate for the technology employed.

Evolving technology

Evolving technology is increasing instrument sensitivity, particularly with two-dimensional imaging.[18] Caution is required in pooling data between studies or in single studies employing different instruments or when comparing data taken with earlier-generation instruments.

RECOMMENDATION 14

The validity of comparing or pooling data collected by different machines must be considered cautiously. Both equipment and protocols used should be clearly described.

Ultrasound safety/influence

Clinical ultrasound is generally well tolerated by subjects/patients but the potential impact should be considered when directing ultrasound energy into any person.[19]

RECOMMENDATION 15

The intensity of sound energy introduced during ultrasonic monitoring should be kept as low as reasonably achievable (ALARA) during ultrasonic scanning. Both the mechanical and thermal indices should be considered. Scan duration should be as short as necessary.

Conclusions

The International Meeting on Ultrasound for Diving Research brought together representatives from around the world to discuss procedures used to study the effects of diving decompression. Integration of the recommendations is expected to help researchers improve the robustness of their data, improving standardization and utility. Those reviewing relevant research that uses ultrasound procedures may also benefit, recognizing issues identified as being of concern to the meeting participants. In the future, the guidelines may be refined and perhaps new methodologies developed for new and emerging technologies.

References

1 Sawatzky KD. *The relationship between intravascular Doppler-detected gas bubbles and decompression sickness after bounce diving in humans* [MSc Thesis]. Toronto: York University; 1991.

2 Nishi RY, Brubakk AO, Eftedal OS. Bubble detection. In: Brubakk AO, Neuman,TS, editors. *Bennett and Elliott's physiology and medicine of diving*, 5th ed. Edinburgh: Saunders; 2003. p. 501-29.

3 Cooper PD, Van den Broek C, Smart DR, Nishi RY, Eastman D. Hyperbaric chamber attendant safety I: Doppler analysis of decompression stress in multiplace chamber attendants. *Diving Hyperb Med.* 2009;39:63-70.

4 Spencer MP, Johanson DC. *Investigation of new principles for human decompression schedules using the Doppler blood bubble detector*. Office of Naval Research Tech Report. ONR Contract N00014-73-C-0094; 1974. Available from: http://archive.rubicon-foundation.org/3788 [cited 2015 December 02].

5 Kisman KE, Masurel G, Guillerm R. Bubble evaluation code for Doppler ultrasonic decompression data [abstract]. *Undersea Biomedical Research.* 1978a;5:28.

6 Kisman K, Masurel G. *Method for evaluating circulating bubble grades detected by means of the Doppler Ultrasonic Methods using the "K.M. Code"*. (English translation of "Masurel G. Methode d'Evaluation des Degrés de Bulles Circulantes Révélées par la Détection Ultrasonore à Effet Doppler "Code K.M". Repère 283 CERTSM 1983"). Toulon, France: Centre d'Études et de Reserches Techniques Sous-Marines; 1985.

7 Eatock BC, Nishi RY. *Procedures for Doppler ultrasonic monitoring of divers for intravascular bubbles.* DCIEM Report No. 86-C-25. Downsview (ON): Department of National Defence: Defence and Civil Institute of Environmental Medicine; 1986. Available from: http://archive.rubicon-foundation.org/7899 [cited 2015 December 02].

8 Nishi RY. Doppler Evaluation of Decompression Tables. In: Lin YC, Shida KK, editors. *Man in the sea, Volume I.* San Pedro, CA: Best Publishing. 1990. p. 297-316.

9 Eftedal O, Brubakk AO. Agreement between trained and untrained observers in grading intravascular bubble signals in ultrasonic images. *Undersea Hyperb Med.* 1997;24:293-9.

10 Thom SR, Milovanova TN, Bogush M, Yang M, Bhopale VM, Pollock NW, et al. Bubbles, microparticles and neutrophil activation: changes with exercise level and breathing gas during open-water scuba diving. *J Appl Physiol.* 2013;114:1396-405.

11 Zanchi J, Ljubkovic M, Denoble PJ, Dujic Z, Ranapurwala SI, Pollock NW. Influence of repeated daily diving on decompression stress. *Int J Sports Med.* 2014;35:465-8.

12 Brubakk AO, Eftedal O. Comparison of three different ultrasonic methods for quantification of intravascular gas bubbles. *Undersea Hyperb Med.* 2001;28:131-6.

13 Nishi RY, Kisman KE, Eatock BC, Buckingham IP, Masurel G. Assessment of decompression profiles and divers by Doppler ultrasonic monitoring. In: Bachrach AJ, Matzen MM, editors. *Underwater physiology VII.* Proceedings of the 7th Symposium on Underwater Physiology. Bethesda (MD): Undersea Medical Society; 1981. p. 717-27.

14 Cialoni,D, Pieri M, Balestra C, Marroni, A. Flying after diving: should recommendations be reviewed? In-flight echocardiographic study in bubble-prone and bubble-resistant divers. *Diving Hyperb Med.* 2015;45:10-15.

15 Blogg SL, Gennser M. The need for optimisation of post-dive ultrasound monitoring to properly evaluate the evolution of venous gas emboli. *Diving Hyperb Med.* 2011;41:139-46.

16 Kisman KE, Masurel G, LaGrue D. Evaluation de la qualité d'une décompression basée sur la détection ultrasonore des bulles. *Méd Aéro Spat Méd Sub Hyp.* 1978b;17:293-7. French.

17 Doolette DJ, Gault KA, Gutvik CR. Sample size requirement for comparison of decompression outcomes using ultrasonically detected venous gas emboli (VGE): power calculations using Monte Carlo resampling from real data. *Diving Hyperb Med.* 2014;44:14-9.

18 Germonpré P, Papadopoulou V, Hemelryck W, Obeid G, Lafère P, Eckersley RJ, et al. The use of portable 2D echocardiography and 'frame based' bubble counting as a tool to evaluate diving decompression stress. *Diving Hyperb Med.* 2014;44:5-13.

19 Powell MR, Spencer MP, von Ramm O. Ultrasonic surveillance of decompression. In: Bennett PB, Elliott DH, editors. *Physiology and medicine of diving*, 3rd ed. San Pedro (CA): Best Publishing Co.; 1982. p. 404-34.

Acknowledgments

The Swedish Armed Forces Diving and Naval Medicine Centre kindly provided host facilities for the meeting. Support was also provided by the Norwegian Labour Inspectorate and the Norwegian Research Council through the Petromaks Program project number 220546/E30. The Scott Haldane Foundation provided CME accreditation for the meeting. Milton Delves of Semacare provided sponsorship for a session. GE Healthcare provided several different ultrasound systems for demonstration purposes.

In addition to the listed authors, the following individuals participated in the August 2015 workshop and contributed to these consensus findings: Pieter Bothma, Antoine Boutros, Jan J Brandt Corstius, Alf Brubakk, Mattijn Buwalda, Jorge C Villarroel, Danilo Cialoni, Johan Douglas, Christian Fabricius, Ian Gawthrope, Mikael Gennser, Peter Germonpré, Peer-Ulrik Haagerup, Marianne Bjordal Havnes, Angelica Lodin-Sundström, Bo-Valentin Nielsen,

Nicklas Oscarsson, Christine Penny, Massimo Pieri, Roswitha Prohaska, Celine Quinsac, Anders Rosen, Olle Sandelin, Göran Sandström, Fiona Seddon, Milan Sevaljevic, David Smart, Julian Thacker, Corry Van den Broek, Claire Walsh, Wilhelm Welslau, Jurg Wendling, Peter Wilmshurst and Tomi Wuorimaa.

Conflicts of interest: The authors declare no competing interest.

Submitted: 23 December 2015; revised 17 January 2016
Accepted: 18 January 2016

Andreas Møllerløkken[1], S Lesley Blogg[2], David J Doolette[3], Ronald Y Nishi[4], Neal W Pollock[5,6]

[1] *Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway*
[2] *SLB Consulting, Winton, Cumbria, UK*
[3] *US Navy Experimental Diving Unit, Panama City, Florida, USA*
[4] *Defence Research and Development Canada – Toronto (retired), Toronto, Ontario, Canada*
[5] *Center for Hyperbaric Medicine and Environmental Physiology, Duke University Medical Centre Durham, NC, USA*
[6] *Divers Alert Network, Durham, North Carolina, USA*

Address for correspondence:
Andreas Møllerløkken
NTNU, Medical Faculty
Department of Circulation and Medical Imaging
P O Box 8905
7491 Trondheim, Norway
E-mail: *<andreas.mollerlokken@ntnu.no>*
