Original articles

The use of portable 2D echocardiography and 'frame-based' bubble counting as a tool to evaluate diving decompression stress

Peter Germonpré, Virginie Papadopoulou, Walter Hemelryck, Georges Obeid, Pierre Lafère, Robert J Eckersley, Ming-Xing Tang and Costantino Balestra

Abstract

(Germonpré P, Papadopoulou V, Hemelryck W, Obeid G, Lafère P, Eckersley RJ, Tang M-X, Balestra C. The use of portable 2D echocardiography and 'frame-based' bubble counting as a tool to evaluate diving decompression stress. *Diving and Hyperbaric Medicine*. 2014 March;44(1):5-13.)

Introduction: 'Decompression stress' is commonly evaluated by scoring circulating bubble numbers post dive using Doppler or cardiac echography. This information may be used to develop safer decompression algorithms, assuming that the lower the numbers of venous gas emboli (VGE) observed post dive, the lower the statistical risk of decompression sickness (DCS). Current echocardiographic evaluation of VGE, using the Eftedal and Brubakk method, has some disadvantages as it is less well suited for large-scale evaluation of recreational diving profiles. We propose and validate a new 'frame-based' VGE-counting method which offers a continuous scale of measurement.

Methods: Nine 'raters' of varying familiarity with echocardiography were asked to grade 20 echocardiograph recordings using both the Eftedal and Brubakk grading and the new 'frame-based' counting method. They were also asked to count the number of bubbles in 50 still-frame images, some of which were randomly repeated. A Wilcoxon Spearman *rho* calculation was used to assess test-retest reliability of each rater for the repeated still frames. For the video images, weighted kappa statistics, with linear and quadratic weightings, were calculated to measure agreement between raters for the Eftedal and Brubakk method. Bland-Altman plots and intra-class correlation coefficients were used to measure agreement between raters for the frame-based counting method.

Results: Frame-based counting showed a better inter-rater agreement than the Eftedal and Brubakk grading, even with relatively inexperienced assessors, and has good intra- and inter-rater reliability.

Conclusion: Frame-based bubble counting could be used to evaluate post-dive decompression stress, and offers possibilities for computer-automated algorithms to allow near-real-time counting.

Key words

Echocardiography, Doppler, bubbles, venous gas embolism, arterial gas embolism, decompression sickness, risk assessment, diving research

---

Introduction

Underwater diving on compressed air or other breathing gases exposes the diver to so-called 'decompression stress', caused by the release of nitrogen and/or other inert gases from the body tissues during and after ascent from depth, resulting in bubbles forming in tissues and (more commonly observable) in blood. In order to minimise this stress and decrease the risk of decompression sickness (DCS), decompression algorithms, summarised in dive tables or incorporated into dive computers, have been developed. These algorithms are not completely successful in the avoidance of every instance of DCS and, to this day, a major research effort is directed to identifying factors and interventions (pre dive, during the dive and post dive) that could make decompression safer.[1]

Evaluation of these algorithms and of the efficacy or inefficacy of other preventive measures has been done primarily on the basis of the presence or absence of clinical symptoms of DCS, as well as on the detection of bubbles in the vascular system using Doppler ultrasonic bubble detectors. Doppler bubble 'grades' were first defined by Spencer et al. in 1974, and classified into 5 grades (0 to 4), depending on the number of acoustic bubble signals audible in the precordial region:[2]

Grade 0 – Complete lack of bubbles;
Grade 1 – Occasional bubble signal, vast majority of cardiac cycles bubble-free;
Grade 2 – Many, but less than half, of cardiac cycles contain bubbles, singly or in groups;
Grade 3 – All cardiac cycles contain bubbles in showers, but not overriding heart signals;
Grade 4 – Bubbles sounding continuously during systole and diastole, overriding amplitude of normal heart signals.

In 1976, Kisman and Masurel defined a scale using three parameters (frequency, amplitude and duration) allowing for more precise classification but rendering acquisition and evaluation more complicated.[3,4] Both these scales require a skilled, experienced Doppler technician in order to be reproducible.[5,6] In 2004, Divers Alert Network (DAN) Europe Research proposed a simplified 'bubble score',

distinguishing only low, medium, high and very high bubble grades based on precordial Doppler, but this scale has not been widely adopted by others.[7,8] Modifications of the original Spencer scale have likewise been proposed, resulting in the 'Expanded Spencer Scale', with a larger number of categories and thus a more incremental grading.[7,8] Whilst the original Spencer scale has been by far the most frequently used in diving research, the Kisman-Masurel scale has been preferred for large, well-controlled, laboratory decompression research studies, and an association between bubble grade and risk for decompression sickness has been developed that can equally be used for the Spencer scale.[2] Generally, it is accepted that the higher the number of bubbles detected precordially, the higher the statistical risk for DCS after a dive.[4,6,9]

Using echocardiography, Eftedal and Brubakk in 1997 proposed a bubble score of six grades based on visual analysis of 2D precordial echo images:[10]

Grade 0 – No observable bubbles;
Grade 1 – Occasional bubbles;
Grade 2 – At least one bubble every four cardiac cycles;
Grade 3 – At least one bubble every cardiac cycle;
Grade 4 – At least one bubble per cm² in every image;
Grade 5 – 'White-out', single bubbles cannot be discriminated.

This allows a semi-quantitative evaluation in a reproducible manner, with minimal intra- and inter-observer variability. However, the scoring system as proposed does not discriminate well in the medium range of bubble scoring, with a large jump between grade 3 and grade 4, making this score less adapted for the evaluation of low to medium levels of decompression stress (classifying into either 'low' or 'severe'). Also, the use of echocardiography made this method less practical for deployment in real-life diving situations (e.g., on a dive boat with a humid, sometimes cold environment and possible lack of AC power). Only recently have good-quality, portable echocardiographs become available, that make on-site evaluation (at the waterfront) possible, by visualising decompression VGE. The use of 'harmonic imaging' (HI) decreases noise in the cardiac cavities, and Color Map application ('gold' setting instead of standard 'grey') provides better image contrast.[11,12] Thus, the detection of VGE in divers' heart cavities and large veins is easier and visualisation of smaller VGE than were detectable by older echography machines is possible.[13] Of note, this use of HI improves the signal-to-noise ratio and increases contrast, but does not aim to make VGE oscillate to emit their own harmonic frequencies, as much lower scanning frequencies would be needed for this to happen.[14–17] For a useful review of HI the reader is referred to references 11 and 12.

In this paper, we describe a newly developed method of evaluation of decompression-induced VGE, using transthoracic 2D echocardiography, which may offer significant advantages compared to current methods.

Methods

A standardised technique for evaluation of decompression stress by means of counting the number of VGE is described, using a portable echocardiography device, with hard-disk recording and *a posteriori* (off-line) evaluation of cardiac images. The technique was developed using a Vivid-i portable echograph (GE Healthcare, UK) and subsequently applied successfully using a Vivid 7 echograph (GE Healthcare, UK), both in a controlled environment (beside a swimming pool) and in the field (dressing room of a Belgian quarry dive site).

A GE 3S-RS sector array ultrasound probe (GE Healthcare, UK) is used; the machine is used in harmonic imaging mode (2.0/4.0 MHz). A four-chamber view is obtained by placing the probe at the level of the left fifth intercostal space. It is necessary to modify the standard four-chamber view by rotating the probe slightly ventrally (in the direction of the xyphoid process) so the right atrium and ventricle can be fully visualised. Three 'landmark points' are identified to aid proper positioning of the ultrasound probe: both transsections of the tricuspid ring and the top of the right ventricle should be visible in the image (Figure 1). A series of at least 15 cardiac cycles are recorded onto the internal hard disk of the echograph while keeping the probe immobile. With practice, each recording can be done in less than 3 minutes (positioning of the diver, attachment of three ECG electrodes, obtaining a good view, recording, detaching the electrodes), allowing for serial measurements on up to 10 divers within a 30-minute interval between measurements of the same diver. At the completion of the measuring period, all videos are saved onto external hard disk or USB thumb drive in the 'wmv' format (Windows Media Video, at 30 frames per second), for which GE Healthcare provided a proprietary video player (MPEGVue Player).

At a later stage, the recordings stored on portable hard disk are reviewed using the MPEGVue software (GE Healthcare, UK), which allows for easy patient and examination selection, frame-by-frame advancing of the video frames using the keyboard arrow keys and freezing of the video frames while maintaining good still-image quality. First, the pre-dive echography loops are reviewed in order to identify intra-cardiac structures that may mimic VGE (e.g., papillary muscles, valve leaflets, Chiari network, Valsalva sinus). Then, the post-dive echography is reviewed and played in a loop at real-time speed in order to rapidly assess the presence or not of circulating bubbles. In cases where bubbles are seen, a formal bubble counting procedure is performed. Using the pause button, the loop is frozen at the start, and then with the forwards and backwards buttons, an image frame is selected in end-diastolic/proto-systolic position (where the tricuspid valve leaflets are fully opened and almost invisible) (Figure 2) and bubbles are counted in both the right atrium and ventricle (Figure 3). In case the chosen view does not contain any bubbles, but bubbles are clearly present in the heart cycle, the forwards and

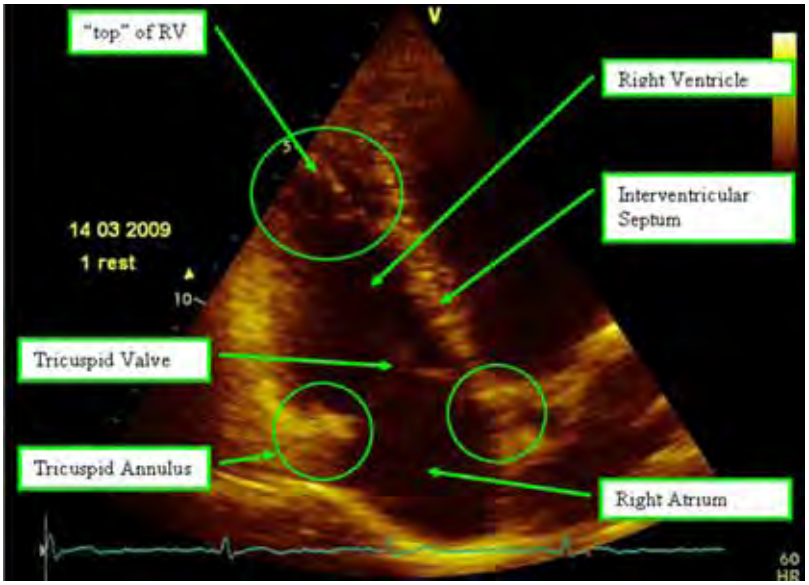


Figure 1
Landmark structures in the right heart echography image: the upper circle identifies the 'top' of the right ventricle (RV) while the lower two circles identify the section through the tricuspid annulus on either side of the right atrium and constitute the 'upper' border of the RA. (N.B., echocardiograph images are inverted)

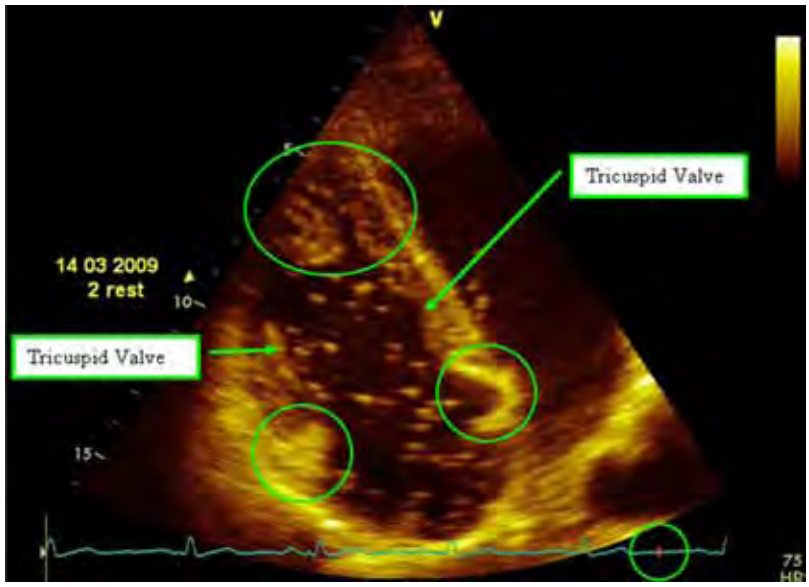


Figure 2
Choice of frame to analyse: the three landmark circles are drawn as in Figure 1. The frame chosen for analysis is indicated by the red marker on the electrocardiography trace (marked by the small green circle, bottom right). Both leaflets of the tricuspid valve are fully open and visible against the ventricular wall (points of green arrows); the right atrium and ventricle form a single cavity

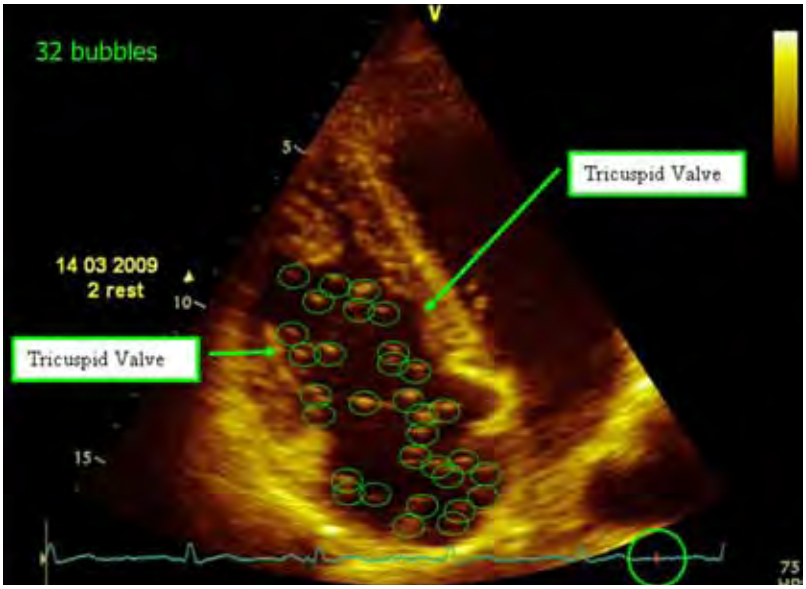


Figure 3
Bubble counting: bubble signals are identified as bright spots and counted individually; tricuspid valve leaflets and other fixed structures (e.g., papillary muscles in the top of the right ventricle) are not counted

backwards buttons are used to select another frame, within two to three frames of the frame originally chosen. Ten consecutive frames are analysed and the bubble count is averaged over these 10 frames.

The technique was developed for use during a series of standardised test dives organised by DAN Europe Research (Roseto, Italy and Brussels, Belgium), in an indoor swimming pool of 34 metres' fresh water (mfw) depth (Nemo33, Brussels, Belgium). The dives were designed to evaluate the effect of several pre-dive interventions on the number of VGE post dive. For this purpose, each diver performed one (identical) dive per week, to 33 mfw for 20 minutes. This 'standard' dive was performed at least three times under 'normal' conditions, and several times under 'experimental' conditions, when the effects of several methods of preconditioning were measured. The order of the experimental dives was randomised. Each diver was evaluated with, among other tests, precordial echocardiography at three time points: before the dive, at 30 minutes and at 90 minutes after surfacing. The study was approved by the Academic Bioethical Committee of the Free University of Brussels (CE/2008/66); all divers were unpaid volunteers who provided written informed consent.

In order to verify the internal (intra-rater) and external (inter-rater) consistency of this frame-based counting method, nine observers were asked to perform analysis of the same set of images. Three were trained cardiologists, at various times involved in diving research performed by DAN Europe. All had performed one or more image acquisition sessions during the experimental pool dives. Three were medical doctors from the Centre of Hyperbaric Oxygen Therapy of the Military Hospital Brussels, who had no formal cardiology training but were present during some or all of the diving experiments, and had some experience in viewing echocardiographic images. The third group consisted of DAN Europe researchers or certified hyperbaric technicians (CHT) from the Centre of Hyperbaric Oxygen Therapy, who had various degrees of paramedical training, allowing them to identify the major intra-cardiac structures after some instruction. All received written instructions detailing the evaluation procedure (and containing the same pictures as in this report) and a short period of hands-on training in the use of the MPEGVue software, which is simple and intuitive to use.

First, a test was administered to verify the reliability and repeatability of the VGE counting by itself. A set of 50 still-frame images was presented for static bubble counting. These images were extracted by the authors from the available video loops, and chosen so as to represent a mix of better- and worse-quality images containing between 0 and 40 VGE signals. Images were presented as a Microsoft PowerPoint presentation. No identifying elements (such as name, birthdate, acquisition date) were displayed on the images, only the slide number. No time limit was given for viewing the slides. Unknown to the test persons, several of the slides were in fact identical but spread out randomly over the presentation. Then, a selection of 20 post-dive video sequences were presented, together with their baseline pre-dive echocardiographic loop (no bubbles present) and the observers were asked to evaluate these video loops, using first the Eftedal and Brubakk score, then using frame-based counting as described above.

As there is no way to determine the exact number of VGE in the images, obviously a true 'gold standard' cannot be determined. The need to set a standard by which to compare the data from this study prompted us to define a 'reference score' as the number of visible bubbles in each image and video loop, agreed on by a priori consensus by the main authors of the study.

STATISTICAL METHODS

Internal consistency was verified on the static images; external consistency was verified on the static and video images with both scoring systems, using the following statistical methods.

Eftedal and Brubakk score

The weighted kappa statistic was chosen to evaluate the inter-rater agreement, in accordance with the discussion on the appropriateness of statistical methods to this effect by Sawatzky.[5] Cohen's kappa (κ) statistic is used to calculate the coefficient of agreement between raters for nominal grades where the outcome of agreement is binary: either agreement or disagreement.[18–20] For ordinal scales, the degree of agreement should be taken into account and this is done using the weighted kappa statistic instead. Both the kappa and weighted kappa are completely corrected for chance agreement.[18] The weights chosen to weight disagreements were defined in the same manner as the original Eftedal and Brubakk method to allow direct comparison. Since the data are ordinal (but not continuous) for the Brubakk and Eftedal method, a disagreement is 'stronger' if one rater assigns a score of 4 and another a score of 1, compared to 1 and 2 respectively. This is taken into account by using weights for characterising the degree of disagreement. In the usual contingency tables for two raters, the weights were specified as:

$$\omega_{ij} = 1 - \frac{|i-j|}{k-1} \quad (1)$$

where i and j index the rows and columns and k is the maximum number of possible ratings. The weighted kappa is then calculated from the proportional observed and expected agreements:[18,21]

$$Po(\omega) = \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{k}\omega_{ij}f_{ij} \quad (2)$$

and

$$Pe(\omega) = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\omega_{ij}r_i c_j \quad (3)$$

where f_{ij} is the number of recordings graded i by one rater

and j by the other, r_i is the row total for grade i and c_j is the column total for grade j, such that:

$$weighted\ kappa = \frac{Po(\omega) - Pe(\omega)}{1 - Pe(\omega)} \quad (4)$$

The kappa-statistic measure is a value between -1 and 1, with 0 corresponding to the value expected by chance and 1 perfect agreement. The interpretation of the values as suggested by Landis and Koch are given as:[22,23]

below 0.00 – Poor
0.00–0.20 – Slight
0.21–0.40 – Fair
0.41–0.60 – Moderate
0.61–0.80 – Substantial
0.81–1.00 – Almost perfect.

Frame-based counting method

For the frame-based counting method, both on still images and on the average over 10 video frames, the data are also ordinal but this time continuous (video) or discrete (units of bubbles). The same weighting applies and the added possibilities are factored in through the use of k so the kappa scores are comparable. The weighted kappa statistic cannot be used for continuous variables.[24] Therefore, another statistical test has to be chosen. For continuous data the intra-class correlation coefficient should be used as a measure of reliability, or Bland-Altman plots for limits of agreement and bias.[24,25] The intra-class correlation coefficient or ICC gives a measure of the proportion of total variance due to the difference between raters by penalising systematic error. For ordinal data, the intra-class correlation coefficient is comparable to the weighted kappa statistic if quadratic weights are used, which is why both weighted kappas (linear as in Sawatzky, and quadratic for comparing with the ICC) are quoted in this paper.[5,26] Note that it is exactly equivalent only for uniform marginal distributions.[25,27] The ICC scale goes from 0 to 1, with 1 representing perfect agreement and 0 no agreement. The Bland-Altman plot displays for two assessors (or groups of assessors) the difference for each assessment against the mean of each assessment.[21,28] The confidence interval is also displayed, calculated as the 95% percentiles such that the upper and lower bounds are given by:

Means of differences ± *1.96 (std of differences)* (5)

As such, the Bland-Altman plot shows any bias and the limits of agreement between two raters.

Intra-rater reliability (internal consistency)

The intra-rater reliability was assessed on the still-images test for the repeated images by theWilcoxon signed-rank test, calculating the Spearman *rho* (rank correlation coefficient ρ) for every rater on the repeated images counts (taking the maximum discrepancy for the one image repeated three times). The value of ρ lies between -1 and 1, a higher number indicating a better reliability. The calculation of the weighted kappa statistic and ICC was performed offline using the standard statistical package Stata (StataCorp. 2011. *Stata Statistical Software: Release 12.* College Station, TX: StataCorp LP). All other data processing and plotting was done by calculating the appropriate values offline as defined above directly in the commercial software package MatLab (MATLAB 6.1, The MathWorks Inc., Natick, MA, 2000).

Table 1
Static images bubble counting – identical image pairs scores Spearman ρ between raters and a reference score (see text); all comparisons non-significant (Wilcoxon test-retest $P > 0.05$) C – cardiologist, MD – physician, O – other (paramedic or hyperbaric chamber attendant)

Rater	Category	Spearman rho
1	C	0.9733
2	C	0.9487
3	C	0.2052
4	MD	0.9211
5	MD	0.7632
6	MD	0.9211
7	O	0.7632
8	O	0.9747
9	O	0.8922

Results

After some practice runs with the frame-based method, all observers reported bubble counting to be relatively easy and rapid, although the process of scrolling through video files was found to be somewhat tedious and slow (approximately 5 minutes for a video file evaluation). The static images were less confidently scored because, as the raters reported, no video images were available to help discriminate between intracardiac structures and VGE. However, the number of bubbles counted was not significantly different between observers (absolute number of bubbles 0 to 40 bubbles). As expected, a larger standard deviation was observed for larger bubble numbers. The ICC between the reference score and all raters was 0.96 (95% confidence interval (CI) from 0.92 to 0.99).

Calculated differences in scoring for identical image pairs (intra-rater or internal consistency) were non-significant (Wilcoxon test-retest, $P > 0.05$) with excellent Spearman ρ (0.76 to 0.97) except for one cardiologist, rater C3 ($\rho = 0.21$, Table 1). Further analysis showed that this observer consistently scored approximately 5 bubbles higher than the average, suggesting that a systematic error was present (see Bland-Altman plot, Figure 4). However, even in the case of this cardiologist with lower Spearman ρ, the Wilcoxon test-retest P-value showed that the differences in the test-retest counts were non-significant.

For the video sequences, the Eftedal and Brubakk scoring gave a weighted kappa of $\kappa = 0.5815$ with linear weights and $\kappa = 0.7634$ with quadratic weights, which shows a moderately good external consistency. It was found to be slightly lower than reported in the original publication ($\kappa = 0.6796$ using linear weights);[10] this may be a reflection of our study design testing and how easy the grading

Figure 4

Bland-Altman plot showing systematic over-estimating by cardiologist 3 as compared to the mean number of VGE counted by all others; X-axis: number of VGE in the image, Y-axis: difference of count vs. mean; horizontal lines – 95% confidence intervals as 1.96 (std of differences); LoA – limits of agreement

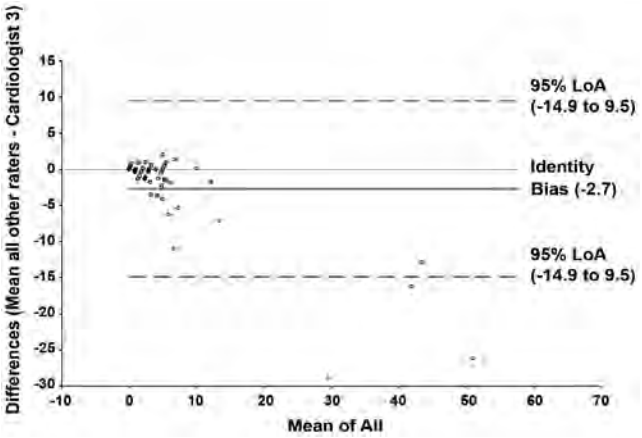
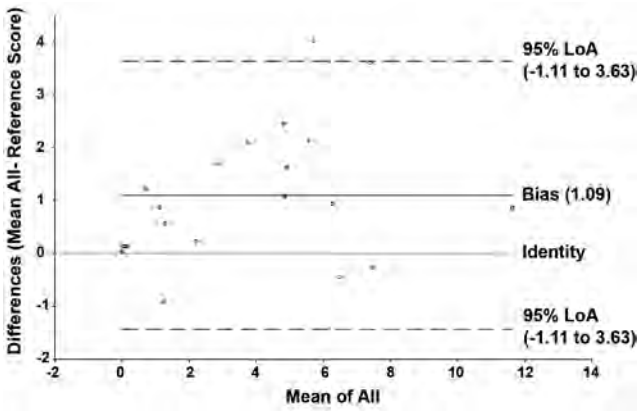


Figure 5

Bland-Altman plot showing the good consistency between the reference score (see text for explanation) and all observers for frame-based counting in the video sequences; X-axis: number of VGE counted in the video sequences (average of 10 frames); Y-axis: difference of count vs. mean; horizontal lines – 95% confidence intervals as 1.96 (std of differences); LoA – limits of agreement



methods are to learn (use of non-expert raters with only written instructions). As indicated in the methods section, all raters received only minimal instructions in the various methods: a three-page document and a short hands-on training session on the use of the video player software. Therefore, the lower external consistency may well reflect the lesser experience in grading according to this score, as none of the nine raters had ever performed an Eftedal and Brubakk scoring before. The ICC for the Eftedal and Brubakk scoring gives 0.79 (95% CI 0.54 to 1.05); as this method is similar to the weighted κ with quadratic weights, it shows a very good inter-rater agreement.

Frame-based counting gave a higher external consistency, with an ICC of 0.84 (95% CI 0.77 to 0.92). There was no significant difference between all observers and the reference score (see Bland-Altman plot, Figure 5); however, here again, the same cardiologist scored consistently approximately 5 bubbles higher on every occasion.

Discussion

(Semi-)quantitative determination of VGE is an important, if not still the only tool available for evaluation of diving decompression stress. Currently used methods suffer from either the necessity of highly skilled observers, a complicated evaluation method (Spencer and Kisman-Masurel scales) or a semi-quantitative visual evaluation that fails to discriminate well in the mid-range of VGE (Eftedal and Brubakk score), exactly the range that most interventions to improve decompression safety for recreational divers would act upon. Also, bubble counting takes place only at certain points in time after the dive, and the accuracy of estimating the total bubble load is dependent on the number of measurements and their timing. One method of estimating the bubble load out of a number of discrete bubble evaluations is the Kisman integrated severity score (KISS), which integrates bubble grades from a number of observations over a given time period into a single value; it can be considered an estimate of the 'area under the bubble grade curve', and is a relative value that can be used for comparative purposes.[29–31]

Using frame-based counting, a continuous-scale (more quantitative) evaluation of VGE presence can be done in a relatively quick, easy way, with good reproducibility. Using the bubble counts for 10 consecutive frames allows for small beat-to-beat variations in bubble numbers to be averaged out. A current drawback is that bubble counting must be done manually at a later stage, which requires additional steps (exporting the video loops in MPEGVue format) and takes some time for counting. Thus, it is not real-time analysis. However, taking into account the echogenicity of the different surrounding structures and using intelligent learning algorithms, computerised automatic counting may become possible. This would permit real-time and continuous counting of VGE, and thus make VGE evaluation independent of the timing of observations after the dive. These algorithms are currently under development.[32–34]

As 2D echocardiography permits viewing the cardiac cavities in a single plane only, the choice of plane may be of some importance. The standard four-chamber view, as used in echocardiography, shows only the basal part of the right ventricle, with the top of the right ventricular cavity out of view. This is not a problem in cardiac evaluation, as most emphasis lies on the morphology and function of the left atrium and ventricle, but may obscure significant parts of the right heart cavities, where VGE are primarily visible after the dive. To overcome this, the method described requires slight tilting of the echo probe to point more in the direction of the xyphoid region, permitting identification of the three landmarks: the top of the right ventricle, the tricuspid ring

and the left and right tricuspid valve leaflet bases, in order to maximally expose the right heart cavities (Figure 1).

The selection of the freeze frame where counting will be done is somewhat arbitrary, but based on the following considerations:

- The end-diastolic/proto-systolic time point is when atrial contraction has finished and ventricular contraction has yet to begin. This is the moment in the cardiac cycle when there is the least flow of blood. Although small areas of turbulence cannot be ruled out, there is at least no rapid movement driven by cardiac contraction.
- It is also the moment when the tricuspid valve leaflets are fully open and almost invisible, making the right atrium and ventricle into a single blood-filled cavity; this decreases the chance of erroneously interpreting valve leaflets as bubble signals.
- This moment is identified easily using the electrocardiographic trace, when recorded with the images.

Although it may be possible theoretically to analyse other frames in the cardiac cycle, these considerations make it unlikely that a better estimation of the number of bubbles might be obtained. In any case, it is important to count the same frame consistently.

Dynamic evaluation such as the Eftedal and Brubakk method seems to slightly over-estimate VGE numbers as compared to actual counting on freeze frames. This can be explained by the fact that vortices of blood exist both in the atrium and ventricle, by which VGE may be swept several times through the plane of vision.[35,36] These blood-flow patterns account for the fact that in some instances, the 'correct' freeze frame chosen for frame-based counting does not show any VGE at all, whereas the previous or next frames do show a significant number (up to 9 or 10) VGE. The procedure therefore allows choosing a frame slightly 'off' if there are obviously VGE in the heart cycle but none can be seen in the initially chosen frame. With automated computerised counting, it will be possible, using three to five frames around the optimal frame, to eventually average out these turbulence effects. Currently, the manual method is too slow to reasonably permit counting of more than 10 to 20 frames in a video loop, as a certain degree of 'observer fatigue' eventually sets in.

The counting method described here makes use of a proprietary video file player on the PC (MPEGVue) which is offered as a package by the echograph's manufacturer (GE). This offers the possibility of viewing echocardiography video files off-line on any Windows PC while offering an easy patient selection menu and the possibility to smoothly step forwards and backwards through the video file, making frame-accurate selection of the images possible. Although a large range of video-playing software that can play back 'wmv' video files on a PC is available, none of them offer this frame-accurate playback. The major drawback here is that the MPEGVue videoplayer can only play back files if the file structure is organised in a certain way – in practical terms, it limits the application to using GE echographs for acquisition and storage of the videos. All of those echographs offer MPEGVue export of the digital (DICOM) files, and once in the MPEGVue format, video files can be shared using either USB disk or sent by e-mail, with the player installation files added to the export package. Automated software will not suffer from this limitation, as it will be able to digest the individual frames of a video stream or file using proprietary software, e.g., MatLab software (MathWorks, Natick, MA, USA).

The inter-rater agreement for frame-based counting is high (ICC of 0.84), indicating there is no major difference between the individual observers and the reference score. This would permit pooling of data from different observers within the same experimental data set. As the VGE counts are an ordinal and continuous variable, mean and average VGE numbers can be calculated, which represents an obvious advantage over the use of discrete variables such as bubble grade scores for evaluating decompression stress. However, the almost perfect (ICC 0.96) intra-rater consistency for this method means that having the same assessor count VGEs for a set of experimental data would give extremely reliable results with regards to the evolution of VGE numbers post dive. Of course, it will be necessary to verify the (intra-rater) consistency of the computer automated counting software which is being developed. If confirmed, this software could be used either for off-line analysis of large numbers of files or, perhaps, directly on an ultrasound scanner (real-time evaluation). At present, the time-consuming process of counting individual bubbles and moving back and forth between frames to discriminate bubbles from their paths and movement prohibits large-scale use of the method.

It has been correctly pointed out that newer echocardiographic techniques are able to detect much smaller bubbles and that, as a result, it is impossible to compare published research using counted bubbles on echography unless exactly the same settings are used. Specifically, Eftedal and Brubakk scores will be impossible to compare among different studies, and it will be impossible to compare the effect of (pre-) diving interventions on VGE production with previous data from similar dives because of this. Recent case reports have indeed described divers with Eftedal and Brubakk grade 5 cardiac echograms (initially thought to be almost impossible without resulting severe DCS), without any symptoms of DCS.[37] This is undoubtedly a result of the better spatial resolution of modern echography, and the use of second harmonics imaging.[13]

The same applies for frame-based bubble counting; it is important to obtain baseline, control dive and post-intervention images on the same group of divers. However, the continuous-scale nature of this method will permit a quantitative evaluation of the effect of the intervention on VGE production. This way, even if the echographic method per se changes and becomes more sensitive, the relative effect observed in different studies may be compared.

Finally, using echocardiography, it may also be possible to evaluate (de)hydration state (by the degree of respiratory collapse of the inferior vena cava, IVC) and, in some subjects, decompression bubbles may even be detected in the IVC and the portal veins.[38–40] Incorporation of this information may provide additional insights into the influence of factors unrelated to the dive profile itself on the production of VGE after the dive. Using solely the degree of VGE after a dive as a measure of dive profile safety without at least trying to standardize these individual (diver-related) factors that may make a diver, either constitutionally or temporarily, less or more prone to the production and liberation of VGE after a dive, disregards a mass of scientific information already available on this subject.[41–45] The presence of VGE in the left cardiac cavities after a dive, be it by passage through a patent foramen ovale or through pulmonary arteriovenous shunts, may indicate a higher risk for cerebral or high-spinal DCS in the individual diver.[46,47] This may guide a decision as to whether a particular diver should be excluded from further participation in diving studies, especially if high risk.

Conclusions

As opposed to existing methods of evaluation, a frame-based counting method permits the investigator to define bubbles as a continuous variable, allowing more flexible and powerful statistical evaluation of the presence of VGE as an indicator of decompression stress. The method presented here shows excellent inter- and intra-rater consistencies, which can be achieved with minimal training by non-experts. Because of the linear, continuous-scale nature of the evaluation, a better discrimination of VGE levels can be achieved in the important intermediate range of bubble load. Therefore, the method seems well suited for use in interventional human diving experiments, where it is ethically impossible to subject volunteer divers to dive profiles generating extreme bubble grades. Moreover, the method is suitable for the development of automated counting software.

References

1 Pollock NW, Vann RD, Denoble PJ, Freiberger JJ, Dovenbarger JA, Nord DA, et al. *Divers Alert Network Annual Diving Report 2007 (based on 2005 data).* Durham, NC: Divers Alert Network; 2007.

2 Spencer MP, Johanson A. *Investigation of new principles for human decompression schedules using Doppler ultrasound blood bubble detection. Technical Report to ONR on Contract N00014-73-C-0094*. Seattle: Institute for Environmental Medicine and Physiology, 1974.

3 Kisman K, Masurel G. *Method for evaluating circulating bubbles detected by means of the Doppler ultrasonic method using the "K.M. code" (English translation of 283 CERTSM 1983).* Toulon, France: Centre d'Etudes et de Recherches Techniques Sous-Marines; 1983

4 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. Philadelphia, PA: WB Saunders; 2003. p. 501-29.

5 Sawatzky KD, Nishi RY. Assessment of inter-rater agreement on the grading of intravascular bubble signals. *Undersea Biomedical Research.* 1991;18:373-96.

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

7 Marroni A, Bennett PB, Cronje FJ, Cali-Corleo R, Germonpré P, Pieri M, et al. A deep stop during decompression from 82 fsw (25 m) significantly reduces bubbles and fast tissue gas tensions. *Undersea Hyperb Med.* 2004;31:233-43.

8 Bennett PB, Marroni A, Cronje FJ, Cali-Corleo R, Germonpré P, Pieri M, et al. Effect of varying deep stop times and shallow stop times on precordial bubbles after dives to 25 msw (82 fsw). *Undersea Hyperb Med.* 2007;34:399-406.

9 Eftedal OS, Lydersen S, Brubakk AO. The relationship between venous gas bubbles and adverse effects of decompression after air dives. *Undersea Hyperb Med.* 2007;34:99-105.

10 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.

11 Choudhry S, Gorman B, Charboneau JW, Tradup DJ, Beck RJ, Kofler JM, et al. Comparison of tissue harmonic imaging with conventional US in abdominal disease. *Radiographics.* 2000;20:1127-35.

12 Uppal T. Tissue harmonic imaging. *AJUM.* 2010;13:29-31.

13 Daniels C, Weytjens C, Cosyns B, Schoors D, De Sutter J, Paelinck B, et al. Second harmonic transthoracic echocardiography: the new reference screening method for the detection of patent foramen ovale. *Eur J Echocardiogr.* 2004;5:449-52.

14 Eftedal O. *PhD Thesis: Ultrasonic detection of decompression induced vascular microbubbles*. Trondheim, Norway: Norwegian University of Science and Technology; 2007.

15 Dejong N, Hoff L, Skotland T, Bom N. Absorption and scatter of encapsulated gas filled microspheres - theoretical considerations and some measurements. *Ultrasonics*. 1992;30:95-103.

16 Tang MX, Mulvana H, Gauthier T, Lim AK, Cosgrove DO, Eckersley RJ, et al. Quantitative contrast-enhanced ultrasound imaging: a review of sources of variability. *Interface Focus.* 2011;1:520-39.

17 Eatock BC, Nishi RY, Johnston GW. Numerical studies of the spectrum of low-intensity ultrasound scattered by bubbles. *J Acoust Soc Am.* 1985;77:1692-701.

18 Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70:213-20.

19 Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:207-16.

20 Kraemer HC. Extension of the kappa coefficient. *Biometrics.* 1980;36:207-16.

21 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307-10.

22 Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977;33:363-74.

23 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-74.

24 Rousson V, Gasser T, Seifert B. Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Stat Med.* 2002;21:3431-46.

25 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420-8.

26 Fleiss JL, Cohen J. Equivalence of weighted kappa and intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas.* 1973;33:613-9.

27 Fleiss JL, Shrout PE. Approximate interval estimation for a certain Intraclass Correlation-Coefficient. *Psychometrika.* 1978;43:259-62.
28 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Int J Nurs Stud.* 2010;47:931-6.
29 Kisman KE, Masurel G, Guillerm R. Bubble evaluation code for Doppler ultrasonic decompression data. *Undersea Biomedical Research.* 1978;5:28.
30 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, eds. *Underwater physiology VII. Proceedings of the 7th Symposium on Underwater Physiology.* Bethesda, MD: Undersea Medical Society; 1981. p. 717-27.
31 Kisman K, Masurel G, LaGrue D, Le Pêchon J. [Evaluation of the quality of decompression using ultrasound bubble detection]. *Méd Aéro Spat Méd Sub Hyp* 1978;67:293-7. French
32 Parlak IB, Egi SM, Ademoglu A, Balestra C, Germonpr P, Marroni A. Intelligent bubble recognition on cardiac videos using Gabor wavelet. *International Journal of Digital Information and Wireless Communications.* 2011;1:195-203.
33 Parlak IB, Egi SM, Ademoglu A, Balestra C, Germonpre P, Marroni A, et al. A neuro-fuzzy approach of bubble recognition in cardiac video processing digital information and communication technology and its applications. In: Cherifi H, Zain JM, El-Qawasmeh E, editors. *Digital information and communication technology and its applications; communications in computer and information science.* Berlin Heidelberg: Springer; 2011. p. 277-86.
34 Papadopoulou V, Hui J, Balestra C, Hemelryck W, Germonpré P, Eckersley R, et al. Evaluating counting of venous gas emboli on post-SCUBA dive echocardiographs, ID 464. *2013 IEEE Joint UFFC, EFTF and PFM Symposium.* Prague; 2013.
35 Fredriksson AG, Zajac J, Eriksson J, Dyverfeldt P, Bolger AF, Ebbers T, et al. 4-D blood flow in the human right ventricle. *Am J Physiol Heart Circ Physiol.* 2011;301:H2344-50.
36 Wigstrom L, Ebbers T, Fyrenius A, Karlsson M, Engvall J, Wranne B, et al. Particle trace visualization of intracardiac flow using time-resolved 3D phase contrast MRI. *Magnet Reson Med.* 1999;41:793-9.
37 Bakovic D, Glavas D, Palada I, Breskovic T, Fabijanic D, Obad A, et al. High-grade bubbles in left and right heart in an asymptomatic diver at rest after surfacing. *Aviat Space Environ Med.* 2008;79:626-8.
38 Romero-Bermejo FJ, Ruiz-Bailen M, Guerrero-De-Mier M, Lopez-Alvaro J. Echocardiographic hemodynamic monitoring in the critically ill patient. *Curr Cardiol Rev.* 2011;7:146-56.
39 Butler BD, Fife C, Sutton T, Pogodsky M, Chen P. Hepatic portal venous gas with hyperbaric decompression: ultrasonographic identification. *J Ultrasound Med.* 1995;14:967-70.
40 Bird N. CT finding of VGE in the portal veins and IVC in a diver with abdominal pain: a case report. *Undersea Hyperb Med.* 2007;34:393-7.
41 Papadopoulou V, Eckersley RJ, Balestra C, Karapantsios TD, Tang MX. A critical review of physiological bubble formation in hyperbaric decompression. *Adv Colloid Interface Sci.* 2013;191-192:22-30.
42 Gempp E, Blatteau JE, Pontier JM, Balestra C, Louge P. Preventive effect of pre-dive hydration on bubble formation in divers. *Br J Sports Med.* 2009;43:224-8.
43 Germonpré P, Pontier JM, Gempp E, Blatteau JE, Deneweth S, Lafère P, et al. Pre-dive vibration effect on bubble formation after a 30-m dive requiring a decompression stop. *Aviat Space Environ Med.* 2009;80:1044-8.
44 Blatteau JE, Gempp E, Balestra C, Mets T, Germonpré P. Predive sauna and venous gas bubbles upon decompression from 400 kPa. *Aviat Space Environ Med.* 2008;79:1100-5.
45 Jankowski LW, Nishi RY, Eaton DJ, Griffin AP. Exercise during decompression reduces the amount of venous gas emboli. *Undersea Hyperb Med.* 1997;24:59-65.
46 Germonpré P, Dendale P, Unger P, Balestra C. Patent foramen ovale and decompression sickness in sports divers. *J Appl Physiol.* 1998;84:1622-6.
47 Germonpré P, Hastir F, Dendale P, Marroni A, Nguyen AF, Balestra C. Evidence for increasing patency of the foramen ovale in divers. *Am J Cardiol.* 2005;95:912-5.

Conflicts of interest: None

Acknowledgements

This study is part of the PHYPODE project, financed by the European Union under a Marie Curie Initial Training Network programme. The authors would also like to acknowledge:

- GE Belgium, for the free and repeated use of a portable echo machine in harsh environments, and provision of hands-on technical support;
- the Nemo33 swimming pool management, for permitting the extended and repeated use of their diving pool outside of opening hours;
- the cardiologists and other (para)medical personnel, for the enthusiastic contribution of their time and expertise – showing that research can be exciting and enjoyable;
- volunteer divers from all over Belgium and the Netherlands, recruited by a simple call for volunteers by DAN Europe Research – indicating that sports divers do care about advancing diving medicine research and are happy to donate time and effort to it.

Submitted: 16 July 2013
Accepted: 10 December 2013

Peter Germonpré[1,2], Virginie Papadopoulou[3,4], Walter Hemelryck[1], Georges Obeid[5], Pierre Lafère[2,6], Robert J Eckersley[7], Meng-Xing Tang[4], Costantino Balestra[2,3]

[1] Centre for Hyperbaric Oxygen Therapy, Military Hospital, Brussels, Belgium
[2] Divers Alert Network Europe, Roseto, Italy and Brussels, Belgium
[3] Biophysiology and Environmental Physiology Laboratory, Haute Ecole Paul Henri Spaak, University of Brussels, Belgium
[4] Department of Bioengineering, Imperial College London, UK
[5] Department of Cardiology, Military Hospital, Brussels, Belgium
[6] Department of Anesthesiology and Hyperbaric Medicine, Hôpital de la Cavale Blanche, Brest, France
[7] Biomedical Engineering Department, Division of Imaging Sciences, King's College London, UK

Address for correspondence:
P Germonpré MD
Centre for Hyperbaric Oxygen Therapy
Military Hospital Brussels
Belgium
Phone: *+32-(0)2-264-4868*
Fax: *+32-(0)2-264-4861*
E-mail: *<peter.germonpre@mil.be>*