Agreement between ultrasonic bubble grades using a handheld self-positioning Doppler product and 2D cardiac ultrasound

Oscar Plogmark[1,2], Carl Hjelte[1,2,3], Magnus Ekström[1], Oskar Frånberg[2,4]

[1] *Lund University, Faculty of Medicine, Department of Clinical Sciences Lund, Respiratory Medicine and Allergology, Lund, Sweden*
[2] *Swedish Armed Forces Diving and Naval Medicine Center, Swedish Armed Forces, Karlskrona, Sweden*
[3] *Sahlgrenska University Hospital, Anesthesia and Intensive Care. Gothenburg, Sweden*
[4] *Blekinge Institute of Technology, Department of Mathematics and Natural Science, Karlskrona, Sweden*

Corresponding author: *Oscar Plogmark, Sten Bergmans väg 21, 121 46 Johanneshov, Sweden*
o_plogmark@hotmail.com

Keywords
Decompression; Decompression illness; Decompression sickness; Diving research; Echocardiography; Ultrasound; Venous gas emboli

Abstract
(Plogmark O, Hjelte C, Ekström M, Frånberg O. Agreement between ultrasonic bubble grades using a handheld self-positioning Doppler product and 2D cardiac ultrasound. Diving and Hyperbaric Medicine. 2022 December 20;52(4):281–285. doi: 10.28920/dhm52.4.281-285. PMID: 36525686.)
Introduction: Intravascular bubble load after decompression can be detected and scored using ultrasound techniques that measure venous gas emboli (VGE). The aim of this study was to analyse the agreement between ultrasonic bubble grades from a handheld self-positioning product, the O'Dive™, and cardiac 2D ultrasound after decompression.
Methods: VGE were graded with both bilateral subclavian vein Doppler ultrasound (modified Spencer scale) and 2D cardiac images (Eftedal Brubakk scale). Agreement was analysed using weighted kappa (K_w). Analysis with K_w was made for all paired grades, including measurements with and without zero grades, and for each method's highest grades after each dive.
Results: A total of 152 dives yielded 1,113 paired measurements. The K_w agreement between ultrasound VGE grades produced by cardiac 2D images and those from the O'Dive was 'fair'; when zero grades were excluded the agreement was 'poor'. The O'Dive was found to have a lower sensitivity to detect VGE compared to 2D cardiac image scoring.
Conclusions: Compared to 2D cardiac image ultrasound, the O'Dive yielded generally lower VGE grades, which resulted in a low level of agreement (fair to poor) with K_w.

Introduction

Bubbles in the bloodstream and tissue can form when the surrounding pressure decreases below the pressure of dissolved inert gas in the body, as during decompression from a dive. This is generally accepted to be a potential instigator for decompression sickness (DCS).[1] Intravascular bubble load after decompression can be detected and graded using ultrasound techniques that measure venous gas emboli (VGE). Two different ultrasound techniques are used to quantify VGE; Doppler audio and two dimensional (2D) echocardiography.[1] Doppler was the first of the techniques used, and VGE grades have been correlated to the risk of DCS.[2] In recent years, a device called the O'Dive™ has been developed by Azoth Systems (Ollioules, France), which is designed for use in the field to perform subclavian ultrasonic Doppler detection of bubbles.

When using this device as a lay person/diver, Azoth System's method integrates the bubble grade with depth and time information to calculate a 'severity index' for the dive. With this severity index, the manufacturer has attempted to let the diver simulate what changes in time and depth would have done for that score and what it might have done for the risk of DCS.[3] Ultrasonic 2D echocardiography grades have also been proven to be related to DCS,[4] and a good agreement with audio Doppler grades has been reported.[5] Two dimensional cardiac image grading is easier to perform for an untrained rater compared to audio Doppler grading.[6] However, it can prove significantly more challenging to get an apical four-chamber cardiac view, than collecting subclavian Doppler audio data, so both methods are challenging in order to collect reliably secure high-quality data.

The key element contributing to the O'Dive's severity index is the Doppler grade assigned by the device and therefore, we aimed to study the comparability of this score to a previously established method. We evaluated the level of agreement between bubble grades from the O'Dive and 2D ultrasonic cardiac images after wet chamber dives.

Methods

The study was approved by the Swedish Ethical Review Authority (Dnr: 2020-06865) and all subjects provided their informed, written consent to participate before the start of the study.

DESIGN AND SUBJECTS

This study investigated a cohort of divers performing experimental air dives (ValTKLHN2021) as per the EL-DCM Thalman dive table (SWEN 21B, unpublished), which was developed to yield an overall risk of DCS of approximately 1%. All the dives were performed in a hyperbaric wet chamber (water temperature 10°C ± 1 degree). Two divers wearing dry suits performed each dive, with several dives at each time/depth combination.

Inclusion criteria were healthy subjects that were eligible for diving in the Swedish Armed Forces, which meant they had passed the fitness-to-dive standard. Exclusion criteria were diving within the previous 48 hours, and any ongoing infection.

ASSESSMENTS

Two dimensional cardiac images were obtained using an UltraSound EDGE II ultrasound machine (Fujifilm SonoSite, Bothell WA, USA) using a cardiac probe (rP19x5-1MHz) with the subject lying on the left side (left lateral decubitus position) giving an apical four-chamber view; in one case, the subject was shifted to the supine position and the probe positioned in the subcostal position in order to attain a view that had otherwise been unattainable. Images were graded using the Eftedal Brubakk (EB) scale[6] by two physicians in real time. All 2D cardiac recordings were preserved for review. The grading system was as follows: 0 = no bubbles; 1 = occasional bubbles; 2 = at least one bubble every 4th cycle; 3 = at least one bubble every cycle; 4 = continuous bubbling at least one bubble /cm^2; 5 = chamber white-out.

Doppler measurements over the left and right subclavian veins were obtained using O'Dive's Doppler transducer, VISION (2 MHz), with the recommended interface and the subject in a sitting position. Doppler assessments were graded blindly with Azoth Systems' proprietary algorithm based on a modified Spencer scale:[7] 0 = a complete lack of bubbles; 1 = an occasional bubble signal discernible with the cardiac motion signal with the great majority of cardiac periods free of bubbles; 2 = many, but less than half, of the cardiac periods containing bubble signals, singly or in groups; 3 = most of the cardiac periods contain showers of single-bubble signals but not dominating or overriding the cardiac motion signals; 4 = the maximum detectable bubble signals sounding continuously throughout systole and diastole of every cardiac period and over-riding the amplitude of the normal cardiac signals. Proprietary software (Azoth Systems) automatically determined 0, 1, 2 grades and 'high grades, 3–4' from the frequency, bubbles over time (with the assumption of a heart rate of 60). The high grades were then manually graded by a technician at Azoth System and differentiated to grade 3 or 4. We chose to use Azoth Systems grades, not grading the sound files ourselves, as we wanted to see if the semiautomatic grading system agreed with 2D ultrasound grades. The highest grade from the right or left subclavian registration was used.

The dives were conducted in pairs, and the diver who removed his diving suit first was taken for 2D cardiac imaging grading, while the second diver was then sent for evaluation with the O'Dive device. The initial measurement for all divers was made within five to 15 minutes after surfacing, and every 15 minutes thereafter. The period between the two measurements was three to ten minutes. Between four and nine paired measurements were performed after each individual dive.

STATISTICAL ANALYSES

Agreement between bubble grades from the O'Dive and cardiac 2D images was analysed using weighted kappa (K_w) and reported with standard error (SE).[2,5,8] Weighted kappa is used to evaluate agreement between two grading methods when the scale is categorical and has more than two categories.[9] It ranges between 0 (worst agreement, equal to chance) and 1 (perfect agreement). In accordance with earlier studies,[5,8] we weighted deviations so complete agreement gave 1.0 credit, 0.75 credit for one category disagreement, 0.5 credit for two category disagreements and so on, down to 0 credit for four categories of disagreement (B-E category 5 were never used). Weighted deviation for the analysis with no zeros gives 1.0 credit for complete agreement, 0.67 credit for one category disagreement and so on, down to 0 credit for three categories of disagreement. O'Dive grades were paired chronologically (within ± 10 minutes) to the 2D image grades. O'Dive grades with no 2D image score within ± 10 minutes were excluded (less than 1%). Weighted kappa was calculated for the highest grades from each dive. The level of agreement (based on K_w) was evaluated by the following categories: poor = < 0.2; fair = 0.21–0.40; moderate = 0.41–0.60; good = 0.6–0.80; and very good = 0.81–1.00.[5,6]

Because Azoth Systems refers to an article[10] using a binary scoring system to characterise the amplification of the risk of precordial measurement compared to subclavian measurements, we also performed a complementary binary agreement analysis with Cohen's kappa.[9] Per the referenced article,[10] the adopted categories were high bubble score (3–4) or low bubble score (0–2).

To evaluate if any methodological disagreements between the two methods could be explained by scattered grades

Table 1
Agreement between paired bubble grades from the O'Dive and 2D cardiac imaging including zero grades; weighted Kappa 0.24 (SE 0.017)

Grades O'Dive	Paired 2D cardiac imaging grades					Total
	0	1	2	3	4	
0	**256**	141	107	132	17	653
1	36	**44**	35	103	13	231
2	11	13	**19**	69	13	125
3	1	6	6	**54**	12	79
4	0	0	0	15	**10**	25
Total	304	204	167	373	65	**1,113**

Table 2
Agreement between paired bubble grades from the O'Dive and 2D cardiac imaging excluding zero grades; weighted Kappa 0.16 (SE: 0.026)

Grades O'Dive	Paired 2D cardiac imaging grades				Total
	1	2	3	4	
1	**44**	35	103	13	195
2	13	**19**	69	13	114
3	6	6	**54**	12	78
4	0	0	15	**10**	25
Total	63	60	241	48	**412**

Table 3
Agreement between highest bubble grades from the O'Dive and 2D cardiac imaging; weighted Kappa: 0.30 (SE 0.045)

Highest grades O'Dive	Highest grades 2D cardiac imaging					Total
	0	1	2	3	4	
0	**8**	13	11	12	2	46
1	4	**9**	5	16	3	37
2	2	1	**5**	16	5	29
3	0	0	3	**16**	9	28
4	0	0	0	3	**9**	12
Total	14	23	24	63	28	152

Table 4
Binary categorical agreement analysis of low bubble grades (0–2) and high bubble grades (3–4); Cohen's Kappa 0.31 (SE: 0.064)

Highest grades O'Dive	Highest grades 2D cardiac image		Total
	Low	High	
Low	**58**	54	112
High	3	**37**	40
Total	61	91	152

or systematically biased grades, the Wilcoxon signed rank test was used, with $P < 0.05$ indicating a clearly biased disagreement with lower grades for one of the methods.

Results

A total of 162 individual dives were performed by 48 divers with eight depth/time combinations. In 152 dives, we were able to grade bubbles with both 2D cardiac ultrasound and bilateral Doppler over the subclavian veins. The mean period between each measurement was 4.6 min (SD 1.9, range 3–10). The number of dives made by each diver varied from one to 12 (11 divers with one dive, 22 divers with two dives, two divers with three dives, four divers with four dives, two divers with six dives, two divers with seven dives, one diver with 10 dives, one diver with 11 dives and one diver with 12 dives). Three different divers were diagnosed with DCS and received hyperbaric oxygen treatment once. Nine divers had minor cutaneous stress, three of whom were treated with normobaric oxygen.

Agreement between all 1,113 paired grades was K_w 0.24 (SE 0.017), which was equal to a fair level of agreement (Table 1). Perfect agreement was found in 383 (34%) measurements. The O'Dive had 642 grades (58%) that were lower than the 2D image grades, and only 88 grades (8%) that were higher (Table 1). When analysing the agreement between 850 paired grades that had no more than 5 min between the measurements, the same level of agreement was found K_w 0.22 (0.019) (without zero grades K_w 0.14 (0.030); a poor level of agreement).

Agreement between all 412 paired grades without the inclusion of zero grades was K_w 0.16 (0.026); a poor level of agreement (Table 2). Perfect agreement was found in 127 (31%) measurements. The O'Dive had 245 grades (59%) that were lower than the 2D cardiac ultrasound grades, and 40 grades (10%) that were higher (Table 2).

The agreement between all 152 paired highest grades was K_w 0.30 (0.045); a fair level of agreement (Table 3). In 47 cases (31%), the highest grades were the same. In 92 cases (61%) the O'Dive's highest Doppler grades were lower compared to 2D image highest grades, and in 13 cases (9%), it was higher (Table 3). In 30% (14 of 46) of the O'Dive's zero grades, the 2D image grade was 3 or 4. However, none (0 of 14) of the 2D ultrasound zeros resulted in a high grade (3 or 4) from the O'Dive.

The binary categorical agreement analysis of low bubble grades (0–2) and high bubble grades (3–4) Cohen's kappa was 0.31 (0.064) (Table 4). In 95 cases (63%), both methods produced the same category. In 54 cases (36%), the O'Dive's category was lower compared to the 2D image category, and in only 3 cases (2%) was it higher (Table 4).

All agreement analyses clearly indicated a bias, with generally lower grades given by the O'Dive, which was also shown by the Wilcoxon signed-rank test, which gave significant results in all cases.

Discussion

The main finding of the present study was that the O'Dive's subclavian Doppler bubble grades had a fair to poor agreement with 2D cardiac ultrasound images.

An earlier study comparing subclavian Doppler grades collected manually by a trained Doppler probe operator, with 2D cardiac grades showed an agreement of good to very good.[5] In contrast, in the present study the O'Dive generally yielded lower bubble grades, which means that it has a lower sensitivity to intravascular bubbles transported to the right heart. Another recently published study[11] involving 173 paired measurements reported a similarly weak correlation between the O'Dive and 2D image categorical assessment scales. That study also found a poor sensitivity to VGE for the O'Dive in comparison to 2D images made using a Vivid q™ device (GE Healthcare, Chicago IL, USA). As the O'Dive had only a fair to poor level of agreement with the established methodology, the implication of this study is that the grades from one method cannot be directly translated to the grades of the other; the degree of difference between the two clearly indicates bias.

A strength of this study is the large number of controlled dives ($n = 162$). In 152 dives, we were able to assess bubbles in four to nine post-dive measurements with both Doppler and 2D cardiac ultrasound, giving a total number of 1,113 paired measurements to include in the analysis. This method of measuring bubbles after dives with 2D ultrasound grades complies with published guidelines.[12] Furthermore, our dives had a relatively high frequency and range of bubbles; this diversity is important when analysing agreement and correlations related to bubble load across the range of the scales.

Limitations of this analysis include the two different anatomical locations. Intravascular bubbles coming from the lower body and/or the neck and head cannot be detected in the subclavian veins, as they drain blood from the arms and shoulders only. However, one large study[2] did not report that the subclavian Doppler signals had a decreased sensitivity to detect intravascular bubbles in general, compared to signals from the chest.

Another limitation is that the frequency of the O'Dive's ultrasonic Doppler device is 2 MHz, while that of the 2D cardiac images we obtained using the UltraSound EDGE II device was 5-1 MHz. This difference in probe frequencies and detection techniques could theoretically lead to different sensitivity to bubbles. The smallest bubble detectable using 2D cardiac images is thought to be between 10–20 µm[8,13] and for Doppler audio ultrasound no smaller than 30 µm.[8,14]

A third limitation is the time taken between the measurements (mean 4.6 min [SD 1.9, range 3–10 min]). This can influence agreement due to the dynamic character of bubble evolution, especially as there was no restriction to the participants in terms of movement between the measurements. A fourth limitation is that many divers did more than one dive. For example, three divers did 33 of the 152 dives. Therefore, these divers will influence the agreement between the two methods more than the other individuals. A fifth limitation is the method by which the O'Dive measurements are graded, combining an automatic and manual grading of bubbles by Azoth Systems (see methods). In this process, the heart rate is approximated to 60 beats per minute, which may certainly influence the result as the Spencer scale categorises bubbles by heart period. The fact that Azoth Systems choose to assess grades 3 and 4 manually is probably because of software limitations, making the results harder to interpret.

Conclusions

The O'Dive's grades yielded a low level of agreement compared to 2D ultrasound cardiac image grades. Generally, the grades were lower with O'Dive and the level of agreement was fair to poor.

References

1 Brubakk AO, Neuman TS, editors. Bennett and Elliott's physiology and medicine of diving. 5th ed. Edinburgh: Saunders; 2003. p. 501–556.
2 Sawatzky KD. The relationship between intravascular Doppler-detected gas bubbles and decompression sickness after bounce diving in humans. Toronto, Ontario: York University; 1991.
3 Germonpré P, Van der Eecken P, Van Renterghem E, Germonpré FL, Balestra C. First impressions: use of the Azoth systems O'Dive subclavian bubble monitor on a liveaboard dive vessel. Diving Hyperb Med. 2020;50:405–12. doi: 10.28920/dhm50.4.405-412. PMID: 33325023. PMCID: PMC7872790.
4 Eftedal OS, Tjelmeland H, Brubakk AO. Validation of decompression procedures based on detection of venous gas bubbles: a Bayesian approach. Aviat Space Environ Med. 2007;78:94–9. PMID: 17310879.
5 Brubakk AO, Eftedal O. Comparison of three different ultrasonic methods for quantification of intravascular gas bubbles. Undersea Hyperb Med. 2001;28:131–6. PMID: 12067148.
6 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. PMID: 9444060.
7 Spencer M. Investigation of new principles for human

decompression schedules using the doppler ultrasonic blood detector. Institute for Environmental Medicine and Physiology; 1974.
8 Blogg SL, Gennser M, Møllerløkken A, Brubakk AO. Ultrasound detection of vascular decompression bubbles: the influence of new technology and considerations on bubble load. Diving Hyperb Med. 2014;44:35–44. PMID: 24687484. [cited 2022 Oct 26]. Available from: https://www.dhmjournal.com/images/IndividArticles/44March/Blogg_dhm.44.1.35-44.pdf.
9 Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. Theriogenology. 2010;73:1167–79. doi: 10.1016/j.theriogenology.2010.01.003. PMID: 20138353.
10 Hugon J, Metelkina A, Barbaud A, Nishi R, Bouak F, Blatteau JE, et al. Reliability of venous gas embolism detection in the subclavian area for decompression stress assessment following scuba diving. Diving Hyperb Med. 2018;48:132–40. doi: 10.28920/dhm48.3.132-140. PMID: 30199887. PMCID: PMC6205931.
11 Karimpour K, Brenner RJ, Dong GZ, Cleve J, Martina S, Harris C, et al. Comparison of newer hand-held ultrasound devices for post-dive venous gas emboli quantification to standard echocardiography. Front Physiol. 2022;13:907651. doi: 10.3389/fphys.2022.907651. PMID: 35755430. PMCID: PMC9222333.
12 Møllerløkken A, Blogg SL, Doolette DJ, Nishi RY, Pollock NW. Consensus guidelines for the use of ultrasound for diving research. Diving Hyperb Med. 2016;46:26–32. PMID: 27044459. [cited 2022 Oct 26]. Available from: https://www.dhmjournal.com/images/IndividArticles/46March/Mollerlokken_dhm.46.1.26-32.pdf.
13 Papadopoulou V, Tang MX, Balestra C, Eckersley RJ, Karapantsios TD. Circulatory bubble dynamics: from physical to biological aspects. Adv Colloid Interface Sci. 2014;206:239–49. doi: 10.1016/j.cis.2014.01.017. PMID 24534474.
14 Lubbers J, Van Den Berg JW. An ultrasonic detector for microgasemboli in a bloodflow line. Ultrasound Med Biol. 1977;2:301–10. doi: 10.1016/0301-5629(77)90031-x. PMID 867565.

Acknowledgments

Thanks to Azoth Systems that graded all the Doppler soundfiles with the Spencer scale.

Further we greatly appreciate the support from the Swedish Armed Forces, Diving and Naval Medicine Centre.

Conflicts of interest and funding

Research regarding decompression tables, Swedish Defense Material Administration, contract nr. 430919-LB967593. Magnus Ekström was supported by unrestricted grants from the Swedish Society of Medical Research and the Swedish Research Council (Dnr: 2019-02081).

No conflicts of interest were declared.

Submitted: 29 December 2021
Accepted after revision: 22 October 2022