Intra-Individual Test-Retest Variation Regarding Venous Gas Bubble Formation During High Altitude Exposures

Rickard Ånell; Mikael Grönkvist; Ola Eiken; Antonis Elia; Mikael Gennser

INTRODUCTION: Hypobaric decompression sickness remains a problem during high-altitude aviation. The prevalence of venous gas emboli (VGE) serves as a marker of decompression stress and has been used as a method in evaluating the safety/risk associated with aviation profiles and/or gas mixtures. However, information is lacking concerning the variability of VGE formation when exposed to the same hypobaric profile on different occasions. In this paper, intra-individual test-retest variation regarding bubble formation during repeated hypobaric exposures is presented. The data can be used to determine the sample size needed for statistical power.

METHOD: A total of 19 male, nonsmoking subjects volunteered for altitude exposures to 24,000 ft (7315 m). VGE was measured using ultrasound scanning and scored according to the Eftedal-Brubakk (EB) scale. Intraindividual test-retest variation in bubble formation (maximum VGE) was evaluated in subjects exposed more than once to hypobaric pressure. The statistical reliability was examined between paired exposures using the Intraclass Correlation test. G*Power version 3.1.9.6 was used for power calculations.

RESULTS: During repeated 20–30 and 70-min exposures to 24,000 ft, 42% ($N = 19$, CI 23–67%) and 29% ($N = 7$, CI 5–70%) of the subjects varied between maximum EB scores < 3 and ≥ 3. The sample size needed to properly reject statistical significance of 1 EB step nominal difference between two paired exposures varied between 29–51 subjects.

CONCLUSION: The large intraindividual test-retest variations in bubble grades during repeated hypobaric exposures highlight the need for relatively large numbers of subjects to reach statistical power when there are no or small differences in decompression stress between the exposures.

KEYWORDS: Altitude decompression sickness, decompression sickness risk, gas bubble formation, hypobaric repeated altitude decompression, intra-individual variation, test-retest, venous gas emboli.

Ånell R, Grönkvist M, Eiken O, Elia A, Gennser M. *Intra-individual test-retest variation regarding venous gas bubble formation during high altitude exposures.* Aerosp Med Hum Perform. 2022; 93(1):46–49.

Decompression sickness (DCS) is a risk associated with high-altitude aviation and diving.[5] During these activities, decompression leads to supersaturation of inert gas dissolved in body tissues and, subsequently, release of free gas and formation of bubbles.[15] Ultrasonic methods are used to evaluate circulating bubbles or venous gas emboli (VGE), which are considered a measure of decompression stress.[4] The amount of VGE detected during and after decompression is linked to the risk of developing DCS where low VGE scores correspond to a low risk of DCS.[12] The advantage with VGE compared to just using symptoms of DCS is that VGE represent an objective measurable variable and decrease the number of subjects needed for each experiment.[12] The disadvantage using VGE is that it is very time consuming to count the bubbles and, therefore, different ordinal scales are used instead. There is an interindividual as well as an intraindividual variation when it comes

From the Swedish Aerospace Physiology Centre, Division of Environmental Physiology, Royal Institute of Technology, KTH, Stockholm, Sweden.

This manuscript was received for review in May 2021. It was accepted for publication in November 2021.

Address correspondence to: Rickard K. J. Ånell, M.D., Department of Environmental Physiology, Swedish Aerospace Physiology Centre, School of Chemistry, Bioengineering and Health, Royal Institute of Technology, KTH, Berzelius väg 13, Solna, Stockholm 17176, Sweden; rickard.anell@saabgroup.com.

DOI: https://doi.org/10.3357/AMHP.5938.2022

to test-retest measurements during both hypo- and hyperbaric exposures. It is well documented that there is a large interindividual variation in bubble formation.[11,12,13] When it comes to variation of occurrence of bubbles post-decompression after repeated hyperbaric exposures, a few studies have been published.[8,14] Contrastingly, to date, there is a dearth of literature concerning the intraindividual variation in decompression effects during repeated hypobaric exposures. Instead, studies with repeated similar hypobaric exposures for the same subjects have mostly aimed to examine the presence of a possible training effect for DCS.[16] Because of this, we considered that our repeated altitude exposures with the same breathing gas mixtures[1,2,3] could be used to evaluate the test-retest variability. The data could also be used to determine the sample size needed for statistical power from paired exposures. Thus, we aimed to investigate the intraindividual test-retest variation regarding bubble formation during repeated hypobaric exposures, based on earlier results collected, to evaluate short periods of recompressions during stable hypobaric pressure to affect tendencies to form VGE.[1,2,3]

METHODS

Subjects

In total, 19 male nonsmoking subjects [mean age 46 (range 23–58 yr); body mass index (BMI) 27.1 (23.1–34.4) $kg \cdot m^{-2}$] with mixed experience of hypobaric exposures took part in the experiments. The subjects were divers (N =10), pilots (N = 4), military (N = 2), and medical students (N = 3). Five of the subjects had experienced altitude DCS with total remission of symptoms upon descent to ground during the exposures. No exposures with DCS were used for the present comparisons. All subjects were well informed about DCS symptoms and told to immediately report any symptoms to the inside experimenter. The pilots and divers had performed an annual medical exam and the remaining subjects were examined by a flight surgeon before admittance. The subjects gave their written consent and were informed according to the Helsinki declaration regarding their right to terminate the exposure at any time. The experiments were approved by regional or national ethics review boards.

Procedures

The subjects were exposed to hypobaric pressure equivalent to 24,000 ft (7315 m) altitude.[1,2,3] During these exposures, subjects were either breathing 54% O_2 in nitrogen (N_2) (N = 13) or 90% O_2 in nitrogen (N = 13) and VGE scores were assessed using transthoracic cardiac ultrasound at 5-min intervals and scored according to the Eftedal-Brubakk scoring scale (EB). All subjects were exposed at least two times using the same breathing mixture. The two first exposures with each breathing mixture were used to assess the test-retest reliability. Although the total exposure time at altitude was more than 60 min and differed in pressure profiles and breathing mixtures,[1,2,3] the first 30 min at 24,000 ft were equivalent for all exposures with the same breathing gas mixture. The shortest and longest intervals between two consecutive exposures were 3–31 d (mean 4.7/median 4 d). In short, comparisons of the initial period at 24,000 ft were carried out on 13 subjects breathing 54% oxygen (30-min periods in 10 subjects, 20-min periods in 3 subjects), and on 13 subjects breathing 90% oxygen (7 of those subjects did also take part in the 54% oxygen/46% nitrogen trials). Of the 13 subjects with exposures using 54% oxygen, 7 had repeated exposures also for 70 min.

Statistical Analyses

All data were statistically analyzed using the IBM SPSS statistics software version 21. Comparison of the effect of different breathing gases (54% oxygen or 90% oxygen) on the distribution of ranking steps was made using the Kolmogorov-Smirnov two-sample test. Reliability measure of bubble scores during repeated exposures was determined using the intraclass correlation test [ICC(1,1)]. Power calculations were done using G*Power,[9] version 3.1.9.6. Unless otherwise stated, data are reported as means ± SD. Significance was accepted at $P < 0.05$.

RESULTS

Seven subjects each carried out two 70-min exposures to 24,000 ft (7315 m) breathing 54% O_2 in N_2. The median difference in maximum EB score between two corresponding runs was 1 (range 0–2). The ICC(1,1) was 0.32 (CI −0.51–0.86). In two of the seven subjects, the maximum EB score was below 3 in one run and 3 or above in the other run [i.e., the fraction of subjects who had runs with EB scores < 3 and EB scores ≥ 3 was 28.6% (N = 7, CI 5.1–69.8%)].

Bubble scores were also measured during the initial 30-min exposures to 24,000 ft for 13 subjects breathing 54% oxygen (3 of whom were only exposed for 20-min periods) and for 13 subjects breathing 90% oxygen. The median paired difference in maximum EB score for subjects breathing 54% O_2 was EB 1 (range 0–3). The ICC(1,1) was 0.34 (CI −0.24–0.75). Also for the exposures with 90% O_2, the median difference in maximum EB score was 1 (range 0–4); thus, the range of differences almost covered the complete range of the EB grading system. The ICC(1,1) for the subjects breathing 90% oxygen was −0.05 (CI −0.58–0.51). The low ICC scores indicate a low test-retest reproducibility of maximum VGE scores during these altitude exposures. The correlation between repeated measurements and the spread in maximum VGE during the various exposures were used to calculate the power to detect 1 EB scale step between two different exposures. The minimum number of paired observations required varied between 29–51 subjects, depending on exposure time and fraction of O_2. The longest exposures needed the least number of subjects and the short exposure with 90% oxygen needed the largest number.

A Kolmogorov-Smirnov test did not show any significant difference in distribution of maximum difference in scores

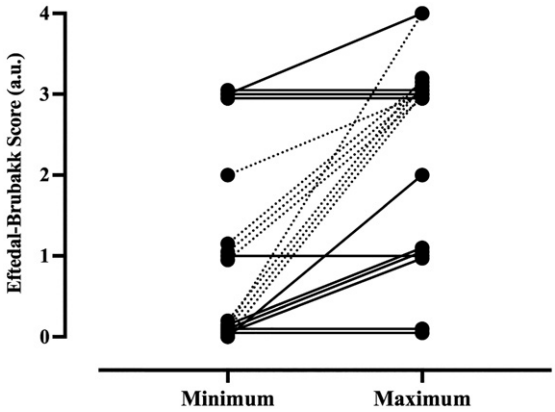

Fig. 1. Maximum EB and minimum EB score in the same individuals. Solid lines indicate subjects whose scores do not cross between high (EB 3, 4) and low (EB 0, 1, 2) scores. Dotted lines indicate scores that varied between high and low scores. *N* = 19.

between individual runs with the different O_2 levels. Adding the results for the two different O_2 levels gave the following distribution: 32% had a maximum EB-score difference (*Δ*EBmax) of 0; 26% had a *Δ*EBmax of 1; 21% a *Δ*EBmax of 2; 16% a *Δ*EBmax of 3; and 5% a *Δ*EBmax of 4. The median maximum variation in EB scores for each individual doing several exposures with either 54% and/or 90% O_2 was EB 1 (range 0–4). If one considers EB scores 0, 1, and 2 low bubble scores with low risk of DCS, and EB 3 and 4 high scores with high risk of DCS,[11] then 8 out of the 19 subjects tested (42.1%, CI 23.2–66.18%) showed scores on both sides of the boundary between low and high risk of DCS in the two runs (**Fig. 1**). Barometric pressure was controlled for the 24-h period before each run, but there was no significant relation between barometric pressure and levels of VGE for any of the different exposures.

DISCUSSION

The aim of the study was to evaluate the test-retest variation within subjects. This can be used to estimate sample size needed to enable detection of small differences in VGE scores during hypobaric exposures. The studies from which the data points were taken were not initially designed for that specific reason.[1,2,3] However, although we are dealing with a relatively small experimental sample, it was decided to evaluate the intraindividual variation and compare the results with previous reports (VGE caused by both hypobaric exposures and decompression from hyperbaric pressures). The evaluation showed that it takes a sample size of between 29 to 51 subjects to gain sufficient power (> 85%) to rule out that no difference exists in a one-grade median difference comparing maximum VGE during paired exposures. We are aware that this conclusion is based on a relatively small sample of repeated exposures, but our findings are in accordance with the findings of Doolette et al.[7] In the latter study, different decompression schedules from hyperbaric pressure were used and were investigated using resampling and a Monte Carlo simulation technique instead of measuring directly repeated similar exposures for the same individuals. Their conclusion was that a study can be considered well powered if the sample size is around 50 paired measurements even if only one grade difference of median VGE grade is of interest.[7] A similar pattern of intraindividual variation was also shown by Dixon et al.[6] during three 6-h exposures to 7.8 psia [equivalent to 16,506 ft (5031 m) altitude] on consecutive days. The purpose of their study was to determine the minimum space suit pressure needed to avoid DCS. VGE were recorded using transthoracic ultrasound Doppler and scored according to a scale similar to the Spencer scale.[6,12] We used their published data on maximum VGE grades in Table I to make an intraclass correlation analysis for all subjects [*N* = 28, ICC(1.1) = 0.51 (CI 0.29–0.71)] and then calculate the required sample size with a double sided alpha of 0.05 for a one-grade difference. The sample size obtained was 40 subjects. It should, however, be noted that in the study by Dixon and coworkers, eight of the subjects did not show any bubbles during any of the exposures. In studies where different exposures or conditions are compared, it is customary to remove subjects who do not show any VGE at all, since the provocations appear to be subliminal. Removing the "consistent zeroes" from the calculation increased the calculated required sample size to 58 subjects. Nevertheless, the results for the three different studies show that the sample size needed for statistical power during VGE measurement is approximately 30–60 subjects. The number and frequency of measurements could possibly affect the variation and sensitivity of the method, but even with our frequent ultrasound measurements every fifth minute, relatively large groups of subjects would be needed to reach statistical power.

As shown above, a relatively large proportion of our subjects who carried out more than two repeated 24,000 ft (7315 m) exposures were observed to occasionally convert from "low-bubblers" to "high bubblers" or from high bubblers to low bubblers. Activities in daily life might play a role in intraindividual variation. In one case where the subject changed from low bubbler to high bubbler, he admitted that he suffered from muscle soreness after hard physical training the previous day and in another case the subject reported flu-like symptoms on the day following the exposure. Muscle soreness not mentioned before the exposures or infections could possibly increase the risk of high VGE.[14] Unfortunately no structured inquiry regarding these issues was carried out.

The intraindividual variation should be considered when designing repeated altitude experiments with the purpose of comparing different strategies, tables, or protocols to ensure safety. Even when median and/or mean values of VGE for a specific exposure seems totally safe for the whole group, certain individuals could at times be at higher risk due to their actual tendency to form bubbles during that particular exposure.

It should be noted that hypobaric exposures are to be considered decompressions from a saturation condition, where the subjects are saturated with inert gas (N_2) at sea level before ascent. Therefore, it is very unlikely that the test-retest differences were due to a different amount of gas and supersaturation. Instead, it seems reasonable to assume that the observed

test-retest variations were due to the ease with which decompression bubbles were formed on any given day.

In conclusion, there is a large intraindividual test-retest variation with respect to formation of venous gas emboli during altitude exposures. The test-retest variation within the same subject was of such a magnitude that to exclude the existence of small differences (e.g., 1 level on the EB scale) during our experiments, it would take approximately 29–51 subjects to obtain statistical power depending on the length of observation and the breathing gas. The estimated sample size is in concordance with previous observations on both hypobaric exposures[6] and post-hyperbaric decompressions.[7]

This study was carried out using the Eftedal-Brubakk scale to assess the amount of VGE. Given reported data that a framed-base bubble counting method has a better external consistency than bubble scoring using the EB scale, the number of observations needed to gain an acceptable power ought to be the same or lower for the counting method.[10] The intraindividual variation of VGE between boundaries of high and low risks for DCS during similar exposures makes the predictions of individual susceptibility to DCS unreliable. There is need for further research to explore what biological mechanisms (for instance, whether related to gas loading, number of bubble precursors, or ease of bubble formation) give rise to the different levels of VGE during or after similar exposures in the same individual.

ACKNOWLEDGMENTS

Financial Disclosure Statement: The project was supported by a grant from the Swedish Armed Forces (No: 922: 0905). The authors have no competing interests to declare.

Authors and Affiliation: Rickard Ånell, M.D., Dr. Med. Sci., Mikael Grönkvist, Ph.D., Ola Eiken, M.D., Ph.D., Antonis Elia, Ph.D., and Mikael Gennser, M.D., Ph.D., Swedish Aerospace Physiology Centre, Division of Environmental Physiology, Royal Institute of Technology, KTH, Stockholm, Sweden.

REFERENCES

1. Ånell R, Grönkvist M, Eiken O, Gennser M. Nitrogen washout and venous gas emboli during sustained vs. discontinuous high-altitude exposures. Aerosp Med Hum Perform. 2019; 90(6):524–530.
2. Ånell R, Grönkvist M, Gennser M, Eiken O. Evolution and preservation of venous gas emboli at alternating high and moderate altitude exposures. Aerosp Med Hum Perform. 2020; 91(1):11–17.
3. Ånell R, Grönkvist M, Gennser M, Eiken O. Hyperoxic effects on decompression strain during alternating high and moderate altitude exposures. Aerosp Med Hum Perform. 2021; 92(4):223–230.
4. 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(1):35–44.
5. Conkin JR, Andrew FJ, Wessel III JH, Klein JS, Dervay JP, Gernhardt ML. Risk of decompression sickness (DCS). Houston (TX): NASA JSC; 2016.
6. Dixon GA, Adams JD, Harvey WT. Decompression sickness and intravenous bubble formation using a 7.8 psia simulated pressure-suit environment. Aviat Space Environ Med. 1986; 57(3):223–228.
7. 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(1):14–19.
8. Eckenhoff RG, Hughes JS. Acclimatization to decompression stress. In: Bachrach AJ, Matzen MM, editors. Underwater physiology III, proceedings of the eight symposium on underwater physiology. Bethseda (MD): Undersea Medical Society Inc.; 1984.
9. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009; 41(4):1149–1160.
10. Germonpré P, Papadopoulou V, Hemelryck W, Obeid G, Eckersly 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(1):5–13.
11. Hugon J, Metelkina A, Barbaud A, Nishi R, Bouak F, et al. Reability of venous gas embolism detection in the subclavian area for decompression stress assessment following scuba diving. Diving Hyperb Med. 2018; 48(3):132–140.
12. Nishi RY, Brubakk AO, Eftedal OS. Bubble detection. In: Brubakk AO, Neuman TS, editors. Bennett and Elliott's physiology and medicine of diving. London: Saunders, Elsevier Science Ltd; 2003:501–529.
13. Nishi R, Kisman K, Eatock B, Buckingham I, Masurel G. Assessment of decompression profiles and divers by ultrasonic monitoring. In: Underwater physiology VII. Proceedings of the seventh symposium on underwater physiology. Bethesda (MD): Undersea Medical Society; 1981:717–727.
14. Papadopoulou V, Germonpré P, Cosgrove D, Eckersley RJ, Dayton PA, et al. Variability in circulating gas emboli after a same scuba diving exposure. Eur J Appl Physiol. 2018; 118(6):1255–1264.
15. Papadopoulou V, Tang MX, Balestra C, Eckersley RJ, Karapantsios TD. Circulatory bubble dynamics: from physical to biological aspects. Adv Colloid Interface Sci. 2014; 206:239–249.
16. Smedal HA. Effects of repeated decompression of 25 subjects to a simulated altitude of 20,000 feet during approximately one month. Pensacola (FL): Naval School of Aviation Medicine; 1948. Contract No.: Project X-762(Av-390-K).