An echo from the past: open access repository of over 10,000 annotated Doppler audio recordings of venous gas emboli

S Lesley Blogg[1,2], Arian Azarang[3], Rachel Lance[4,5], Frauke Tillmans[4,6], Richard Moon[5], Virginie Papadopoulou[3], Peter Lindholm[1,7]

[1] *Department of Emergency Medicine, University of California San Diego, USA*
[2] *SLB Consulting, c/o Home Park Barn, Kirkby Stephen, Cumbria, UK*
[3] *Department of Radiology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA*
[4] *Divers Alert Network, Durham, North Carolina, USA*
[5] *Departments of Anesthesiology and Medicine, Center for Hyperbaric Medicine and Environmental Physiology, Duke University, Durham, North Carolina, USA*
[6] *Lampe Joint Department of Biomedical Engineering, The University of North Carolina at Chapel Hill and North Carolina State University, Chapel Hill, North Carolina, USA*
[7] *Department of Physiology and Pharmacology, Karolinska Institute, Sweden*

Corresponding author: *Professor Peter Lindholm, 200 W Arbor Drive, MC8676, San Diego, CA 92103, USA*
peter.lindholm@ki.se

Keywords
Artificial intelligence; Bubbles; Decompression sickness; Measurement; Venous gas embolism

Abstract
(Blogg SL, Azarang A, Lance R, Tillmans F, Moon R, Papadopolou V, Lindholm P. An echo from the past: open access repository of over 10,000 annotated Doppler audio recordings of venous gas emboli. Diving and Hyperbaric Medicine. 2026 31 March;56(1):8−12. doi: 10.28920/dhm56.1.8-12. PMID: 41875437.)
Introduction: Doppler ultrasound measurements have been recorded since the 1970s across the world and provide a valuable data resource for learning, analysis, and potential training of deep learning algorithms to recognise and grade venous gas emboli (VGE) allowing assessment of decompression sickness (DCS) risk.
Methods: We collected a 'big database' of Doppler recordings and associated metadata. Audio tapes with recorded Doppler data were converted to digital files, then cut into individual recordings and matched with their metadata, including subject and pressure profile information. The audio signals and their Doppler grades were then processed further for suitability to train an algorithm to identify VGE.
Results: A total of 10,099 Doppler ultrasound recordings were compiled. Divers (n = ≤ 311; 170 identified, ≤ 141 unidentified) were male, with a median age of 31.5 years among the 170 identified divers. The maximum depth of the dives included ranged from 24 m (80 feet) to 91.4 m (300 feet). The timing of the Doppler measurements ranged from two minutes post-dive to 594 min post-dive, with a median time of 52 min. Breathing gases included air, nitrox, and heliox. DCS was noted in only 12 individuals. The dataset centred around lower VGE loads (Spencer Grades 0, I, and II).
Conclusions: This database represents a landmark in DCS investigation as the audio dataset and metadata collected have been released under a public domain license for further use. The large number of data points has also allowed the development of a deep learning algorithm that can grade bubble loads without a human operator.

Introduction

Decompression sickness (DCS) remains a major concern for commercial, military, and recreational diving. In diving research, circulating bubbles detected with ultrasound, which are also termed venous gas emboli (VGE), are used as markers to estimate the level of decompression stress a diver has been subject to during a dive and then further gain a measure of the likelihood of DCS occurring.[1] Once the diver returns to atmospheric pressure, VGE are monitored and recorded using precordial (PC) or subclavian (SC) Doppler ultrasound (audio), or with transthoracic 2D cardiac imaging to determine the load circulating in the body. The more VGE in the circulation, then generally the higher the risk of DCS.[1,2]

Doppler ultrasound measurements have been made and recorded since the 1970s by varying dive research bodies across the world. The technique uses a transducer to transmit sound waves into the body, and the resulting echo allows operators to detect moving objects in the blood vessels and heart.[1] Bubbles strongly reflect sound waves and so can be detected by this method.[1] Most of these measurements made historically were recorded onto audio cassettes and digital audio tapes (DAT) for storage. In 2020, the US Office of Naval Research supported an effort from various US academic institutions to gather these recordings from around the world, then digitise and store the data online for safeguarding and future research efforts. The University of California at San Diego (UCSD) and University of North Carolina at Chapel Hill along with other partners including the Divers Alert Network (DAN), Duke University (all US-based), the Karolinska Institute, Sweden, and QinetiQ, UK, have been involved in the efforts to collate this large database of previously acquired Doppler audio recordings and metadata. It represents a vast undertaking in terms of locating data, obtaining permissions for inclusion, digitising, cutting, and pairing recordings to their metadata with the end goal to make these historic and valuable Doppler data available to the diving community. These data include PC (over the heart) and / or SC (over the subclavian veins) measurements during rest and limb movement (flex) for each diver.

One obstacle to making use of such data is that a high degree of training and practice is necessary for an operator to assess and grade audio bubble signals correctly. To the untrained ear, little can be made of the audio signals thus rendering the data worthless. However, efforts have been made to develop computer-based programs to do the job of grading bubble loads. The development of machine learning algorithms requires well-curated data sets for training, validation, and testing. Large amounts of data are key to producing reliable, high-performance algorithms and this was another reason that the collection of these data was recognised to be of benefit and actioned. This paper aims to document these collection and collation efforts, to briefly describe the data and how they have been used to date, and to show the potential for further studies to come.

Methods

ETHICAL APPROVAL

This study was determined exempt from Institutional Review Board requirements under US Federal Policy for the Protection of Human Subjects ("*Common Rule*") 45CFR 46.104(d) (4) by the University of California at San Diego (UCSD) Institutional Review Board (IRB) (determination #805229).

COLLECTION AND ORGANISATION OF RAW DATA FOR DATABASE REPOSITORY

Both analogue audio tapes and digital audio tapes (DAT) tapes were available with recorded Doppler data. To convert audio tape signals to digital files, an audio cassette player with an audio line out socket, and left and right male Radio Corporation of America (RCA) coaxial stereo phono leads was used (Yamaha MT400 multitrack cassette mixing/recording unit). The RCA leads were then connected to a USB audio interface with a digital output device (Behringer UCA202, Behringer Switzerland) to digitise the analogue audio signal and transfer it to a personal computer. DAT tapes were treated similarly, using a Sony DAT player (Sony Group Corp, Tokyo, Japan) attached to the audio interface device.

The open-source program Audacity (https://www.audacityteam.org/) was used to record the signal, which could then be cut and manipulated for further use and storage. Settings in Audacity for this program were as follows: sampling rate at 44100 Hz, Microsoft Multimedia Environment input set to microphone (USB Audio CODEC), 2 (stereo) recording channel. Files were saved in the free lossless audio codec (FLAC) format, whose files are roughly 50% size of waveform audio file format (WAV).

DATA SEGMENTATION

Following digitisation, each Doppler audio recording was segmented into individual clips corresponding to discrete measurements for each diver. Typically, one clip (~20 s) represented a resting Doppler recording at a specific time after surfacing, followed by a second clip recorded during movement at the same session. Each segment was then verified against the available metadata and catalogued accordingly.

Each Doppler recording originally contained examiner commentary superimposed on the audio channel, which was routinely provided during acquisition to document the measurement site, subject identifier, date, and physiological condition (e.g., movement or rest). In order to isolate the Doppler ultrasound signal alone (necessary for developing or testing computer-automated Doppler ultrasound analysis software for example), these spoken annotations were systematically trimmed, leaving only the Doppler ultrasound component.[3] This ensured that all recordings are deidentified and represent physiological signals only without overlapping speech, thereby facilitating signal processing, machine learning analysis, and reproducibility. The trimming was performed without altering the acoustic characteristics of the Doppler trace itself, preserving both signal fidelity and grading validity.

METADATA COMPILATION

For storage in the repository, the dive series' original filenames were used where possible with sequential numbering, for example, THe4008_00001, where THe4008 is the name of the test dive recorded on the audio tape. These files would be labelled THe4008_00001 and THe4008_00002 and so on, and represented the sequential recordings made for each diver over time that was captured on the audio tape. Metadata for each clip, which included the original Kisman Masurel (KM) grades as given by the operators monitoring the dives and making the measurements at the time, were then recorded into an Excel file and saved for use.[4]

KM grades are non-linear measures of the relative bubble load detected in the venous circulation, with the KM code ranging from Grade 0 (no detectable bubbles) through grades I, II, and III, up to IV sequentially reflecting higher loads of bubbles present.[4] For quality control purposes the original KM grades assigned were randomly checked by an experienced Doppler operator (SLB) as they were saved

Table 1
Correspondence between the Spencer and Kisman-Masurel venous gas emboli assessment scales[1]

Spencer grade	Kisman-Masurel grades
0	[000]
1	[111] [112] [113] [211] [212] [213]
2	[121] [122] [123] [221] [222] [223]
3	[232] [233] [242] [243] [332] [333] [342] [343]
4	[444]

into the Excel sheet, to make sure that the files matched the metadata grades originally assigned to them.

The metadata collected included (although not in every case): filename, anatomical site (PC or SC), SC left or right, rest or flex, KM code, KM grade, KM decimal grade, dive maximum depth (metres or feet), gas, oxygen partial pressure, helium partial pressure, atmospheres absolute (atm abs) or Bar, time to maximum depth, total dive time, age, sex, data origin, DCS occurrence, type of DCS, DCS symptom(s), Doppler post-dive time, wet or dry dive, in water, study identifier, and subject identifier. All Doppler measurements had been made using a Techno Scientific Doppler monitor (Techno Scientific Inc., Ontario, Canada). These structured metadata enabled secondary analyses linking Doppler findings to decompression stress, physiological conditions, and instrumentation characteristics.

For user navigation and data retrieval, audio recordings were curated into two primary anatomical site folders (PC and SC). Within each, subfolders correspond to bubble grade classifications derived from the KM code and expressed as equivalent Spencer grades, which can be achieved by conversion from the KM codes, as shown in Table 1.[3] This hierarchical organisation enables efficient browsing and targeted download of recordings by both site and VGE grade.

The KM code provides a very detailed multi-parameter description of VGE load, but this level of complexity makes it harder to use directly for a deep learning model development as it expands the number of categories and reduces the number of samples per class. In contrast, the Spencer grading system is a simpler ordinal scale that summarises the VGE load into five distinctive classes. To facilitate reuse for algorithm benchmarking or VGE classification, the recordings were also organised according to both anatomical site and converted Spencer grades derived from the KM code (see Table 1). This secondary categorization was introduced solely to streamline machine learning development based on Spencer grades and does not affect the structure of the raw repository.

REPOSITORY STRUCTURE AND ACCESSIBILITY

The completed database was thoroughly validated for internal consistency by cross-checking, for each recording, the metadata across subject, dive, and bubble grade information using automated column-wise verification routines. All initial discrepancies were subsequently reviewed and manually double-checked (authors SLB, AA, VP) before finalising the dataset. In addition, each entry was verified to ensure that its associated audio file adhered to the correct naming convention and was stored in the appropriate folder corresponding to its anatomical location and grade classification.

To ensure reproducibility and transparency, the complete curated dataset has been made publicly available via Zenodo (https://doi.org/10.5281/zenodo.16877955) under a public domain license. Files are organised by anatomical site (PC or SC) and labeled consistently with their original recording series. Metadata enable linkage of Doppler findings with experimental context and decompression parameters, supporting a wide range of secondary analyses.

Results

PUBLIC AVAILABILITY

The dataset and metadata collected in this study have been released under a public domain license to encourage widespread use within the scientific community and beyond. Researchers may freely download, analyse, and incorporate the recordings for applications such as algorithm benchmarking, training of machine learning models for automated VGE detection, decompression risk modelling, and educational purposes.

DATASET COMPOSITION BY RECORDING SITE

The data collection process, which included identification of suitable recordings, discovery and reconciliation of matching metadata, digitisation of files, cutting of files, insertion into

Table 2
Distribution of recordings by site and Spencer grade

Spencer grade	Precordial (*n*)	Subclavian (*n*)	Total (*n*)
0	2,586	4,407	6,993
I	406	1,022	1,428
II	359	441	800
III	504	280	784
IV	63	31	94
Total	3,918	6,181	10,099

the database, and finally quality control of the historical grades assigned to the files, took over four years of work. A total of 10,099 Doppler ultrasound recordings were compiled into the final database. Of these, 3,918 recordings (38.7%) were obtained from the precordial site, and 6,181 recordings (61.2%) were acquired from the subclavian sites (right and left). This distribution provides broad coverage across both clinically relevant monitoring locations, which differ in their sensitivity to VGE and in their practical applicability for decompression monitoring.

DESCRIPTIVE DATA

All the identified divers (n = 170) whose data were included were male, with a median age of 31.5 years. For these individuals, the originally assigned de-identified subject IDs were retained to maintain consistency with source records. In some cases, subjects participated in both air and heliox dives and were therefore assigned multiple identifiers in the original records, typically one ID for their air dives, another for their heliox dives, and an additional unique identifier that cross-referenced both sets of dives. In such cases, only the primary subject ID was retained in the 'Subject ID' column of the database, while the additional original identifiers were recorded in a separate metadata column for reference.

There were an additional 141 sets of data that belonged to unidentified subjects (i.e., those with missing or inaudible IDs in the audio recordings). A small subset of these datasets may have derived from the same subject (repeated dives) and so would reduce this number from 141 but as there was no way to determine this, these data were treated as standalone IDs. To ensure traceability within the database, unidentified subjects were assigned dummy identifiers consisting of the capital letter 'U' followed by a unique three-digit number. These identifiers were allocated systematically based on the recording structure and reconciliation with original experimental documentation.

The total number of distinct subjects represented in the database is therefore estimated at 311. Based on available historical records, we are confident that all unidentified subjects were also male and had demographic characteristics comparable to those of the identified group. All subjects were drawn from military personnel and scientific staff working with the military, or commercial tunnelling employees.

The maximum depth of the dives included ranged from 24 m (80 ft) to 91.4 m (300 ft). Total dive time ranged from six minutes (91.4 m dives) to 4 h 11 min (tunnelling exposures). For 15 dive profiles, data were available for time to maximum depth; this parameter ranged between nine and 180 min.

The timing of the Doppler measurements ranged from two minutes post-dive to 594 minutes post dive, with a median time of 52 minutes. Dives were made in water, in a dry hyperbaric chamber, in a dry tunnelling environment, or in a wet hyperbaric chamber.

Breathing gases included air, oxygen in nitrogen (all with a partial pressure of oxygen [PPO_2] of 0.7 atmospheres absolute [atm abs]), and heliox (also PPO_2 0.7 atm abs)

Decompression sickness was noted in only 12 individuals within this database, with 10 subjects reporting Type 1 symptoms (pain), one reporting both Type I and Type II (pain and motor deficits), and one reporting Type II symptoms only (sensory and motor deficits).

DISTRIBUTION ACROSS SPENCER GRADES

Bubble grades occurred from 0 to KM grade IV, that is, across the whole grading range. Subsequently, each recording was annotated according to the Spencer grading system, resulting in five levels of VGE load.[5] The distribution of recordings across grades is shown in Table 2. The dataset was predominantly composed of lower VGE loads (Grades 0, I, and II), which accounted for 91.4% of the audio files, while higher VGE loads (Grades III and IV) represented only 8.6% of the total.

Discussion

This database represents a landmark in DCS investigation and open resources, as it is the first publicly available collection to include the original Doppler audio recordings rather than only the bubble grades assigned by investigators. By sharing the underlying audio clips, this resource enables independent verification of VGE assessments and supports the development of automated analysis methods. The dataset has already facilitated the creation of a deep learning algorithm capable of grading bubble loads without a human operator.[6] Although other datasets may contain larger numbers of participants or DCS cases, none have previously made the raw Doppler signals accessible for unrestricted scientific use.

Up until this point, deep learning has not been utilised for the detection and classification of VGE in post-dive Doppler audio data. However, with the realisation of this large data base, a bubble classification model developed previously using lab-generated synthetic bubble data was then fine-tuned on a subset of these real-world data. This resulted in a model that achieved an average ordinal accuracy of 83.8% for precordial and 90.4% for subclavian Doppler across all five Spencer grades. Importantly, an average of 77.8% and 93.8% binary classification accuracy was demonstrated between low and high VGE grades for PC and SC, respectively, and 84.6% and 80.9% between 'no VGE' and 'VGE present' for PC and SC, respectively.[6]

With regards to the data themselves, perhaps the most notable point is that of the ~311 subjects included in the data set, only 12 experienced DCS following a dive. Most of the dives included in the database were to test various tables or help develop computer algorithms, so may not have been very conservative in their dive profile. The one contrast to this was the 46 data points included that represented monitoring in three subjects working on the Belfast tunnelling project, where KM grades ranged from 0 to IV, but as would be hoped in the workplace, there was no occurrence of DCS. Although the overall number of DCS cases is low, reflecting the rarity of the condition, the dataset remains highly valuable for studies focused on VGE characterisation, Doppler signal analysis, and algorithm benchmarking. However, analyses aimed at identifying or modelling DCS risk factors should account for the limited number of clinical DCS events and the resulting statistical constraints.

The database may be limited in the fact that no female divers' data are included, but unfortunately this is the nature of using historical data drawn mostly from military sources, an environment where only now are women becoming more common in the work force. Positively, there was a relatively good range in diver age, from 17 to 45 years, so giving a good representation of active divers across age groups.

Overall, the metadata are rich enough that many other investigations and studies could be made from this database. We have already published a study comparing PC and SC bubble grades in an attempt to clarify which bubble monitoring site is preferential.[7] We welcome other workers using the data for their own research and the audio files can also be used as a training resource for operators wishing to hone their skills.

References

1 Nishi RY, Brubakk AO, Eftedal O. Bubble detection. In: Brubakk AO, Neuman TS, editors. Bennett and Elliott's physiology and medicine of diving. Edinburgh: Saunders; 2003. p. 501–29.

2 Currens JB, Doolette DJ, Murphy FG. Venous gas emboli (VGE) in 2-D echocardiographic images following movement: grading and association with cumulative incidence of decompression sickness. Diving Hyperb Med. 2025;55:44–50. doi: 10.28920/dhm55.1.44-50. PMID: 40090025. PMCID: PMC12263276.

3 Azarang A, Blogg SL, Currens J, Lance RM, Moon RE, Lindholm P, et al. Development of a graphical user interface for automatic separation of human voice from Doppler ultrasound audio in diving experiments. PLoS One. 2023;18(8):e0283953. doi: 10.1371/journal.pone.0283953. PMID: 37561745. PMCID: PMC10414643.

4 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.

5 Spencer MP, Johanson DC, editors. Investigation of new principles for human decompression schedules using doppler ultrasonic blood bubble detection. 1974.

6 Azarang A, Le DQ, Hoang AH, Blogg SL, Dayton PA, Lance RM, et al. Deep learning-based venous gas emboli grade classification in doppler ultrasound audio recordings. IEEE Trans Biomed Eng. 2023;70:1436–46. doi: 10.1109/TBME.2022.3217711. PMID: 36301781.

7 Blogg SL, Azarang A, Papadopoulou V, Lindholm P. Agreement of precordial and subclavian Doppler ultrasound venous gas emboli grades in a large diving data set. Diving Hyperb Med. 2025;55:2–10. doi: 10.28920/dhm55.1.2-10. PMID: 40090020.

Acknowledgements

Thanks to Geoff Loveman and QinetiQ, UK, for enabling the inclusion of their datasets in the repository. Thanks to Helena Housley for assistance in cutting the recordings.

Conflicts of interest and funding

This work was funded by: The Department of the Navy, Office of Naval Research (ONR award numbers N000142012590 and N000142312548). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The Karolinska Institute, Fraenkel fund for aviation research Dnr: 2019-01046. Dr Blogg is the Deputy Editor of *Diving and Hyperbaric Medicine* Journal, and a member of the journal's editorial board. She played no role in managing the peer review process for this manuscript, nor in the decision to publish.

Submitted: 27 October 2025
Accepted after revision: 7 December 2025