UCLA
UCLA Previously Published Works

Title

Toward automated classification of pathological transcranial Doppler waveform morphology via spectral clustering.

Permalink

Journal

Authors

Thorpe, Samuel
Thibeault, Corey
Canac, Nicolas
et al.

Publication Date

DOI

Peer reviewed

Toward automated classification of pathological transcranial Doppler waveform morphology via spectral clustering

Samuel G. Thorpe[1]*, **Corey M. Thibeault[1]**, **Nicolas Canac[1]**, **Kian Jalaleddini[1]**, **Amber Dorn[1]**, **Seth J. Wilk[1]**, **Thomas Devlin[2]**, **Fabien Scalzo[3]**, **Robert B. Hamilton[1]**

1 Department of Research, Neural Analytics, Inc., Los Angeles, California, United States of America,
2 Department of Neurology, Erlanger Medical Center, Chattanooga, Tennessee, United States of America,
3 Department of Neurology, University of California Los Angeles, Los Angeles, California, United States of America

* samuel.thorpe@neuralanalytics.com

Abstract

Cerebral Blood Flow Velocity waveforms acquired via Transcranial Doppler (TCD) can provide evidence for cerebrovascular occlusion and stenosis. Thrombolysis in Brain Ischemia (TIBI) flow grades are widely used for this purpose, but require subjective assessment by expert evaluators to be reliable. In this work we seek to determine whether TCD morphology can be objectively assessed using an unsupervised machine learning approach to waveform categorization. TCD beat waveforms were recorded at multiple depths from the Middle Cerebral Arteries of 106 subjects; 33 with Large Vessel Occlusion (LVO). From each waveform, three morphological features were extracted, quantifying onset of maximal velocity, systolic canopy length, and the number/prominence of peaks/troughs. Spectral clustering identified groups implicit in the resultant three-dimensional feature space, with gap statistic criteria establishing the optimal cluster number. We found that gap statistic disparity was maximized at four clusters, referred to as flow types I, II, III, and IV. Types I and II were primarily composed of control subject waveforms, whereas types III and IV derived mainly from LVO patients. Cluster morphologies for types I and IV aligned clearly with Normal and Blunted TIBI flows, respectively. Types II and III represented commonly observed flow-types not delineated by TIBI, which nonetheless deviate from normal and blunted flows. We conclude that important morphological variability exists beyond that currently quantified by TIBI in populations experiencing or at-risk for acute ischemic stroke, and posit that the observed flow-types provide the foundation for objective methods of real-time automated flow type classification.

2. Introduction

Transcranial Doppler ultrasound (TCD) is a noninvasive methodology for measuring Cerebral Blood Flow Velocity (CBFV) through the large arteries of the brain [1–4]. The pulsatile CBFV waveform can provide information concerning numerous cerebrovascular pathologies [5–8], including stroke [9–13], intracranial hypertension [14–17], sickle cell disease [18–20], and mild Traumatic Brain Injury [21–23]. In the context of acute ischemic stroke, the leading

cause of long-term disability in the United States [24], TCD is commonly used to detect occluded and stenosed cerebral arteries [25–30]. When blood flow in these arteries is occluded, impaired oxygen supply can cause rapid brain tissue death, and permanent neurological dysfunction [31–33]. In particular, Large Vessel Occlusions (LVO) involving partial or total blockage of the Middle Cerebral and/or Internal Carotid Arteries (MCA/ICA) have disproportionately high morbidity and mortality due to the large volume of brain tissue which these vessels supply [34–36]. Because TCD provides specific information about the status of flow through the cerebral vasculature, TCD examinations are routinely conducted as standard of care at many comprehensive stroke centers [37,38].

Numerous methodologies for evaluating stroke with TCD have been published. Occlusion can often be determined by straightforward comparison of mean blood flow velocities across multiple vessels in both cerebral hemispheres [12,39–41]. However, as discussed further in subsequent paragraphs, it can also be determined by evaluating specific aspects of waveform morphology [25–30]. The morphology of the TCD signal is that of a pulsatile waveform with dominant period dictated by heart rate. Typically, each cycle consists of an initial systolic upstroke, followed by three consecutive peaks (referred to as P1, P2, and P3, respectively) which correspond to specific events in the cardiac cycle, and subsequent return to diastolic minimum [42–44]. Absence or distortion of these features may indicate flow irregularities associated with various cerebral hemodynamic pathologies such as LVO. Indeed, a growing number of studies have demonstrated changes in the timing and amplitude of waveform peaks associated with medical conditions ranging from intracranial hypertension [43,45,46] to various age-related changes and pharmacological interventions [47,48].

The most widely-cited method for assessing stroke with TCD is the evaluation of waveform morphology using Thrombolysis in Brain Ischemia (TIBI) flow grades [25,27]. In this framework, waveform features such as the presence of identifiable pulses/peaks and the onset of maximum velocity are used to determine grade assignment [26,28–30]. The TIBI categories range from grades 0–5, with 5 indicating normal flow, and grades 0 and 1 designating absent (0) or minimal (1) flow associated with complete or partial vascular occlusion. The grades between are associated with characteristic morphologies which are used to determine blunting (grade 2), dampening (grade 3), or stenotic (grade 4) flows. Several foundational studies have shown TIBI assessment to be a valuable tool for LVO diagnosis, possessing sensitivity and specificity often exceeding 90% [5,25,30] relative to gold standard Computed Tomography Angiography (CTA). However, reliable determination of TIBI grades requires subjective assessment by expert evaluators, severely limiting their utility for prehospital stroke assessment by less specialized personnel.

The complexity and subjectivity inherent in the assignment of TIBI categories, along with their demonstrable utility for stroke assessment, make them a natural candidate problem for automation via machine learning. That is, the clinical relevance of the morphological information captured by subjective TIBI grading heuristics has now been clearly established, but an objective and computationally tractable framework for extracting this information does not currently exist. Though the traditional TCD Pulsatility Index constitutes an objective metric which describes waveform morphology to some degree, the information it contains is too coarse to effectively detect the presence of occluded or stenosed vessels [28,49]. Moreover, machine learning approaches to extracting information from TCD waveforms have already proven fruitful for multiple clinical applications [50–56], including the diagnosis of cerebrovascular stenosis [57]. In particular, recent work by our group has shown that a TCD-derived morphological biomarker termed Velocity Curvature Index (VCI) may provide a robust, objectively computable metric for detecting Large Vessel Occlusion (LVO) [54,58]. Though VCI readily identifies waveforms with pathologically deviant curvature, it does not

differentiate between pathological morphologies such as those delineated by the TIBI scale. An objective means of waveform categorization could thus provide additional information concerning stroke etiology to better inform stroke triage and transfer decisions.

Before approaching waveform categorization explicitly as a classification problem, a more foundational issue must first be addressed, in that it is currently unclear to what degree the TIBI categories actually capture the natural variance inherent in empirical data. That is, we know these categories are useful, but we do not know that they are comprehensive. Additional waveforms may be present which are also informative, and perhaps a different subset may better explain variability across subjects. In this work we take a data-driven approach to waveform categorization, retrospectively applying an unsupervised learning algorithm to a dataset comprised of multiple subject groups, including patients experiencing acute LVO, as well as control subjects collected both in and out of hospital. For these purposes we employ spectral clustering, which does not make strong assumptions about inherent cluster density, and thus performs well when clusters are connected but potentially non-convex [59–61]. Combined with gap statistic criterion for choosing the optimal cluster number [62], this approach yields a natural partitioning of waveforms into groups with distinct morphological characteristics.

For clustering, we sought to include information of clinical relevance to the presence of LVO, thus we chose morphological features which quantify important aspects of TIBI evaluation criteria (see section 3.3 for details). The *onset* feature encodes information concerning where in the cardiac cycle the maximal velocity is attained, which can be important in distinguishing blunted waveforms (TIBI grade 2) from healthy (TIBI grade 5) [25]. This feature is a time-normalized version of the "latency to P1" variable used in previous studies of TCD morphology [45,46]. Likewise, the *canopy* feature provides information about the size of the systolic complex, which is helpful for distinguishing both blunted and stenotic waveforms from healthy. This feature is similar to the "fractional time in systole" used to evaluate waveform morphology in [43]. Finally, the number and prominence of observed peaks/troughs (quantified by our *peaks* feature) are also critical in distinguishing blunted and stenotic waveforms from healthy. We hypothesize that LVO subject waveforms will fall into clusters which are mostly distinct from non-LVO controls, for which we can subsequently compare alignment with established TIBI categories.

3. Materials and methods

3.1. Subjects

We compared TCD waveform morphology across three subject groups: one with CTA-confirmed LVO, a second non-LVO control group collected in-hospital, and a third group of control subjects collected out of hospital. LVO and in-hospital controls (IHC) were enrolled at Erlanger Southeast Regional Stroke Center in Chattanooga, TN, between October 2016 and October 2017. As detailed in [54,58], subjects who arrived at the hospital presenting with stroke symptoms received TCD examinations along with standard care, pharmaceuticals and CT imaging. CTA was performed using a GE Lightspeed VCT 64-section multidetector scanner (GE Healthcare, Milwaukee, WI) with a slice thickness of 0.625 mm, and bolus injection of 70–150 mL of Omnipaque 350 (GE Healthcare, Milwaukee, WI) contrast material (4.0 mL/s). CTA images were reformatted in the coronal and sagittal plane, and 10-mm maximum intensity projection reconstructions were rendered for review. Occlusion location was determined by the radiologist on call, and reviewed/confirmed independently by the authors. TCD examinations were performed during available time between patient testing/treatment, and in no way impacted patient care. Subjects for whom an acceptable exam was obtained within 4 hours of imaging, and to whom no Table 1 (left column) exclusion criteria applied, were

eligible for enrollment in either the LVO group (if CTA confirmed occlusion of the proximal extracranial or terminal intracranial ICA segments, or M1/M2 branches of the MCA), or the IHC group (if no LVO were detected). Out of hospital controls (OHC) were scanned at multiple locations in Los Angeles, CA, and did not undergo CTA examination. Volunteers were eligible for enrollment if between 40–85 years of age, and all Table 1 (right column) exclusion criteria were absent. Experiment protocols for LVO and IHC subjects were approved by the University of Tennessee College of Medicine Institutional Review Board (ID: 16–097), and for OHC subjects by Western Institutional Review Board (ID: 20151682). The study was carried out in accordance with the recommendations of the Declaration of Helsinki, with written informed consent from all subjects.

3.2. TCD waveform recording & processing

CBFV signals were acquired by trained technicians using 2 MHz handheld probes in conjunction with either DWL Doppler Box-X (DWL Inc., USA), or Lucid M1 TCD System (Neural Analytics Inc., USA), to transtemporally insonate the left/right MCA. The technician was instructed to obtain recordings for as many depths as possible between 45–60 mm in both the left/right cerebral hemispheres, with the depth range enforced to help ensure errant recordings of other vessels (e.g. ACA/PCA or M2) were not included in the analysis. Once signal was identified and optimized at a specific depth, waveform recordings were made in 30-second intervals. The choice of 30 seconds was intended to guarantee the acquisition of a sufficiently large ensemble of at least 15 individual beat waveforms from which to compute an average beat waveform for analysis. The choice of 15 individual beats is a free parameter chosen to optimize the trade off between the size of the acquired ensemble and the technician's available time and ability to maintain continuous steady contact between the probe and the subject's head.

To construct the average waveform for each recorded depth, individual beat waveforms from each recording were extracted offline (post-recording) using an automated beat identification algorithm [63] to identify the samples marking the onset of each beat (and thus the end

Table 1. Subject exclusion criteria.

Exclusion Criteria	
LVO, IHC subjects	**OHC subjects**
1. Head CT findings consistent with acute primary intracranial hemorrhage (SAH, ICH, etc.).	**1.** Individuals taking any psychoactive medication.
2. Hemodynamically unstable patients requiring pharmacological support for hypotension.	**2.** Systolic arterial blood pressure greater than 140 mmHg (Hypertensive).
3. Subjects who underwent partial or full craniotomy.	**3.** Individuals presenting with an external wound in head.
4. Additional intracranial pathologies present (tumor, hydrocephalus, etc.).	**4.** Pregnant women.
5. Anticipated insufficient time to acquire a complete set of scan as described by the protocol.	**5.** Individuals who have a known history of: severe TBI, moderate TBI (within previous 3 years), vascular disease, stroke, sickle cell anemia, brain tumor, epidural or subdural hematoma, abnormal MRI or CT scan of the brain, cardiovascular disease.
6. Significant hemodynamic pharmacological agent (cocaine, amphetamine, etc.).	
7. Subjects who are under arrest for a felony.	

Subject exclusion criteria. Note the following abbreviations: SAH–Subarachnoid Hemorrhage, ICH–Intracranial Hemorrhage, TBI–Traumatic Brain Injury, MRI–Magnetic Resonance Imaging, CT–Computerized Tomography.

of the previous beat). Outlier beats with excessive artifact and/or poor signal quality were auto-matically identified and rejected via Iterated Interquartile Range exclusion with cross-correla-tion and beat length as primary comparators [64]. To be included in this analysis, exams were required to contain at least one bilateral pair of left/right MCA scans at depths between 45–60 mm, each containing at least 15 accepted beats. To obtain a rectangular ensemble array for averaging, the accepted beats were aligned at the onset of systolic upstroke and individually padded with each beat's final value to the length of the longest beat. The resulting ensemble array was averaged across beats, resulting in a single representative average beat waveform for each recorded 30 second interval. It is very important to note that waveforms were only aver-aged over beats obtained from individual depth recordings (i.e. consecutive beats recorded in a single contiguous 30 second interval), and were not averaged across different recordings or across the left/right hemispheres. OHC waveforms, digitally sampled at 400 Hz, were resam-pled to 125 Hz to match the native sampling rate of LVO and IHC waveforms. To reduce the potential effects of artifactual noise in the individual beats of higher frequency than the rele-vant hemodynamics (roughly 10 Hz), all waveforms were smoothed via convolution with a 90 ms Hanning window. Since this analysis sought to evaluate morphological commonality regardless of underlying heart rate or velocity scale, each waveform was normalized with respect to both time and velocity. These normalizations help ensure that waveforms with gen-erally similar shapes, but potentially different heart rates or velocity ranges will still likely be clustered together. The velocity normalization was accomplished for each waveform by first subtracting the minimum, then subsequently dividing by the resultant maximum, thus rescal-ing to the interval [0, 1] along the velocity axis. The temporal normalization was accomplished by resampling each waveform to 100 total samples (via cubic spline), effectively enforcing a common heart rate across waveforms.

3.3. Cluster feature extraction

For clustering, we sought to construct a feature space of minimal dimension which would nonetheless capture enough stroke related morphological variance to produce a meaningful clustering partition for comparison to TIBI. Clearly, there are a vast number of morphological variables which might be used for this purpose. For example, the Morphological Clustering Analysis of Intracranial Pressure Pulses (MOCAIP) algorithm [65] defines 128 morphological features which can be extracted from TCD and pressure wave pulses [45]. However, most of these assume that the standard peaks/troughs associated with healthy waveforms are present and identifiable. For pathological waveforms such as those associated with LVO, this assump-tion is rarely met. As noted in Section 1, our strategy was to start with the features we deemed most relevant to flow grade assignment using the TIBI criterion. These features must also be readily computable for all pulsatile TCD waveforms, regardless of the presence of stroke pathology. The three features decided upon for this analysis meet these criteria, but admittedly represent a first approach to the problem which could be expanded or refined in future work.

From each waveform, denoted $x(t)$ in Eqs 1–3, we extracted each of the three morphological features (depicted in Fig 1). The first, termed *onset* (Eq 1), marked the temporal onset of maxi-mal velocity. The second feature, termed *canopy* (Eq 2), was defined as the number of samples comprising the systolic complex, i.e. the "beat canopy", given by the cardinality of the set of samples with velocity greater than 25% of the diastolic-systolic range (see [58] for details). The final feature, termed *peaks* (Eq 3), quantified the number and "weight" of waveform peaks/troughs. First we identified the set of true-peaks/troughs (TP, approximate zeros of the first derivative) as points in the canopy corresponding to a sign-change in the difference between successive samples. True peaks were each assigned a weight of one. Next we identified the set
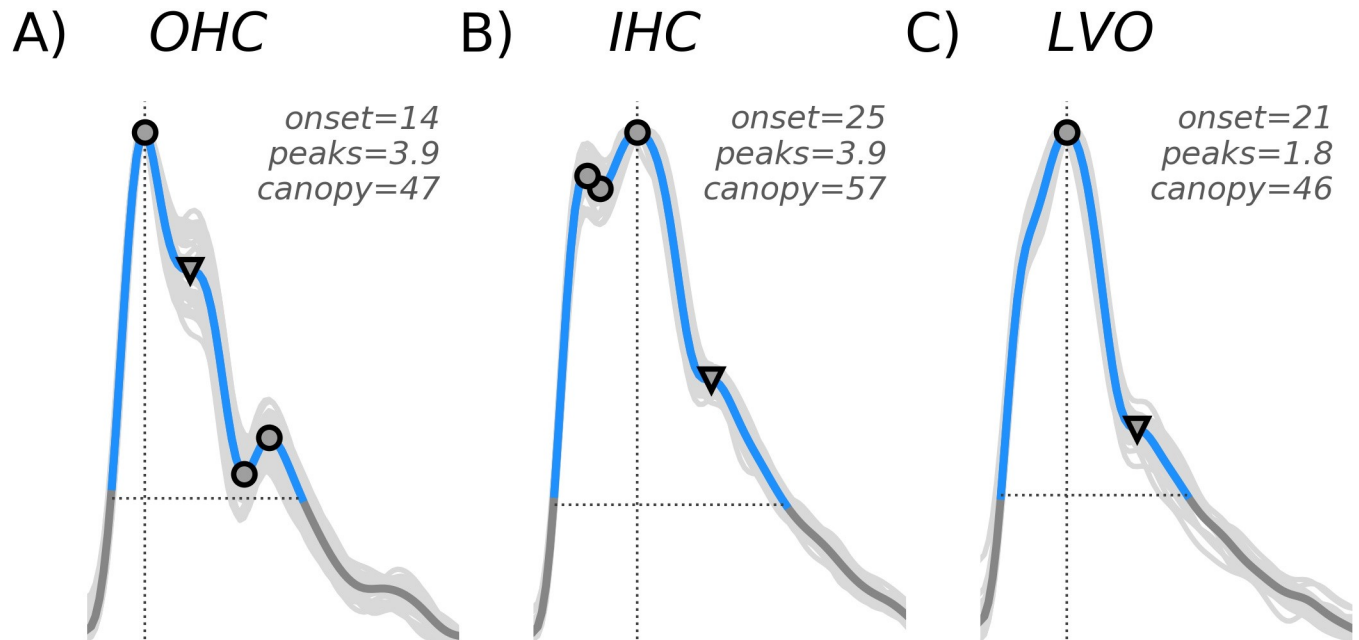
Fig 1. Cluster features. Cluster features are depicted for three example waveforms taken from each of the subject groups; OHC (A), IHC (B), and LVO (**C**). All waveforms were normalized in both time and velocity, so as to span 100 total samples ranging from zero to one on the y-axis. The onset variable (vertical line) marks the time sample where maximum velocity is attained. The canopy variable (horizontal line) marks the length (in samples) of the systolic canopy. The *peaks* variable is a weighted sum of waveform peaks, both true (indicated by circles) and pseudo (triangles).

of "pseudo-peaks/troughs" (PP, points where the derivative is small but non-vanishing) where the difference magnitude between successive samples dropped below a chosen critical threshold of 0.01 (choosing the point with smallest difference magnitude in any group of adjacent sub-threshold samples). For pseudo-peaks we assigned weights corresponding to one minus the ratio of the associated difference magnitude to the threshold (0.01), such that those with the smallest derivative were weighted most heavily. The *peaks* feature was then computed as the sum over all corresponding weights. The *onset*, *canopy*, and *peaks* features for each waveform are defined explicitly by Eqs 1–3, respectively. Note that the "card" expression in Eqs 2 and 3 refers to the cardinality of a set, and the colons in Eqs 1 and 2 denote the set-theoretic notion of "having the property such that". In Eq 2, $t_{systolic}$ and $t_{diastolic}$ refer to the time samples at which the waveform attains its systolic maximum and diastolic minimum, respectively.

$$onset = n : x(t_n) = \max_{i \in \{1,2,\dots 100\}} \{x(t_i)\} \tag{1}$$

$$canopy = card(\{i : x(t_i) > x(t_0) + 0.25(x(t_{systolic}) - x(t_{diastolic}))\}) \tag{2}$$

$$peaks = card(TP) + \sum_{k \in PP} 1 - \frac{|x(t_k) - x(t_{k-1})|}{0.01} \tag{3}$$

With regard to specification of the threshold 0.01 in Eq 3, the idea is to allow for adequate consideration of "pseudo peaks/troughs" for which the derivative gets very small, but doesn't quite vanish, and weight them towards the total *peaks* count. Our experience with the data has shown that there are indeed numerous cases for which this threshold adds meaningful variance to the peaks feature distribution. For example, each waveform depicted in Fig 1 possesses

a pseudo-peak, occurring at roughly the location expected for either the P2 (Fig 1A), or P3 (Fig 1B and 1C); we refer the reader to the introduction for background on the peak structures (P1, P2, and P3) typically observed sequentially in standard TCD beat morphology. If these were idealized waveforms with true peaks and corresponding true troughs at these locations, the combined structures would count two units toward the *peaks* value. Without consideration of pseudo-peaks, these structures would contribute zero units towards the total, which we believe to be less descriptive and ultimately less meaningful, whereas the specification we have made allows them an intermediate contribution of 0.9, 0.9, and 0.8 for 1A, B, C, respectively. The smaller this threshold is chosen, the closer the peaks feature comes to approximating the discrete count of the true peaks, *TP*, for which the derivative does vanish.

3.4. Spectral clustering and gap analysis

In order to standardize the ranges of the resultant feature distributions for clustering, z-score normalization (not to be confused with the z-transform used in signal processing) was applied across observations [66,67]. That is, the feature distributions were rescaled by subtracting from each feature value the mean taken across all observations, and subsequently dividing by the standard deviation across observations. The resultant three-dimensional feature space was then partitioned via spectral clustering implemented in scikit-learn, an open-source machine learning library for Python [68], specifically the *SpectralClustering* module with default parameters and radial basis kernel. To determine the optimal number of clusters, we employed the gap statistic methodology of Tibshirani et al. [62], which is widely used for cluster number selection [69,70]. Conceptually, the gap statistic was designed to formalize the statistical notion of finding the "elbow" in the graph of cluster dissimilarity as a function of cluster number within a widely applicable nonparametric framework; the critical underlying idea being that consideration of additional clusters beyond the optimal number should yield diminishing returns in terms of minimizing cluster dispersion (i.e. the sum across clusters of pair-wise distances between members of each cluster). For our analysis, we compared gap statistics (G_k) for total clusters (k) ranging from two to seven. In this procedure, gap statistics are computed as the difference between observed log intra-cluster dispersion pooled across k clusters (denoted W_k), and the analogous expected dispersion bootstrapped from a null distribution incorporating the covariance structure of the observed data. As detailed in method b (section 4) of [62], each of the 1000 bootstrap iterations was generated by sampling uniformly over the range of the columns of the observed data transformed by its right-singular vectors, and back-transforming the resultant sample to feature space via the right-singular transpose. The optimal number of clusters were selected as the smallest k such that $G_k > G_{k+1} - S_{k+1}$, where S_k is the standard deviation of the k-cluster bootstrap distribution corrected to account for simulation error.

3.5. Cluster archetypes

To visualize the characteristic morphology of the resultant clusters, we first computed for each cluster the matrix of squared Euclidean distances between all cluster member waveforms (i.e. the entities denoted as d_{ii} in [62]), and ranked each by average distance to other members. For each cluster, the individual waveform with the smallest mean intra-cluster distance was designated as the most representative exemplar. To further visualize commonalities in cluster morphology independent of an individual exemplar, the five waveforms with smallest mean intra-cluster distance were averaged together to obtain the waveform archetype for each cluster.

4. Results

4.1. Subject demographics

The current analyses included 33 LVO subjects (16 female), 33 IHC subjects (13 female), and 40 OHC subjects (25 female), with average ages of 66.9 (SD = 15.7), 56.4 (SD = 16.3), and 58.4 (SD = 10.9) years, respectively. A total of 996 average beat waveforms were included in this analysis, with 354, 445, and 196 contributed by LVO, IHC, and OHC subjects, respectively, resulting in a clustering feature space of dimension 996 x 3.

4.2. Gap analysis & cluster archetypes

Fig 2A shows expected and observed log intra-cluster dispersion as a function of total clusters, with the gap statistic determined by their difference. The elbow in observed dispersion at four clusters corresponds to both the maximum gap statistic (Fig 2B) and optimal number of clusters determined by the Tibshirani et al. criteria [19]. The associated feature space (Fig 3A) is shown partitioned into the four resultant clusters, with membership indicated by color, and associated cluster archetypes shown in Fig 3B–3E. The largest cluster (type I), containing 400 waveforms, was characterized by early max velocity onset with wide canopy and strong peaks (Fig 3B). Of waveforms in this cluster, 18% came from LVO subjects, with 82% from controls (52% IHC, 30% OHC). The second largest cluster (type II), containing 257 waveforms, was characterized by later max velocity onset, with wide canopy and strong peaks (Fig 3C). Of waveforms in this cluster, 20% came from LVO subjects versus 80% from controls (68% IHC, 12% OHC). The smallest cluster (type III), containing 83 waveforms, was characterized by early max velocity onset, with narrow canopy and weak peaks (Fig 3D). Of waveforms in this
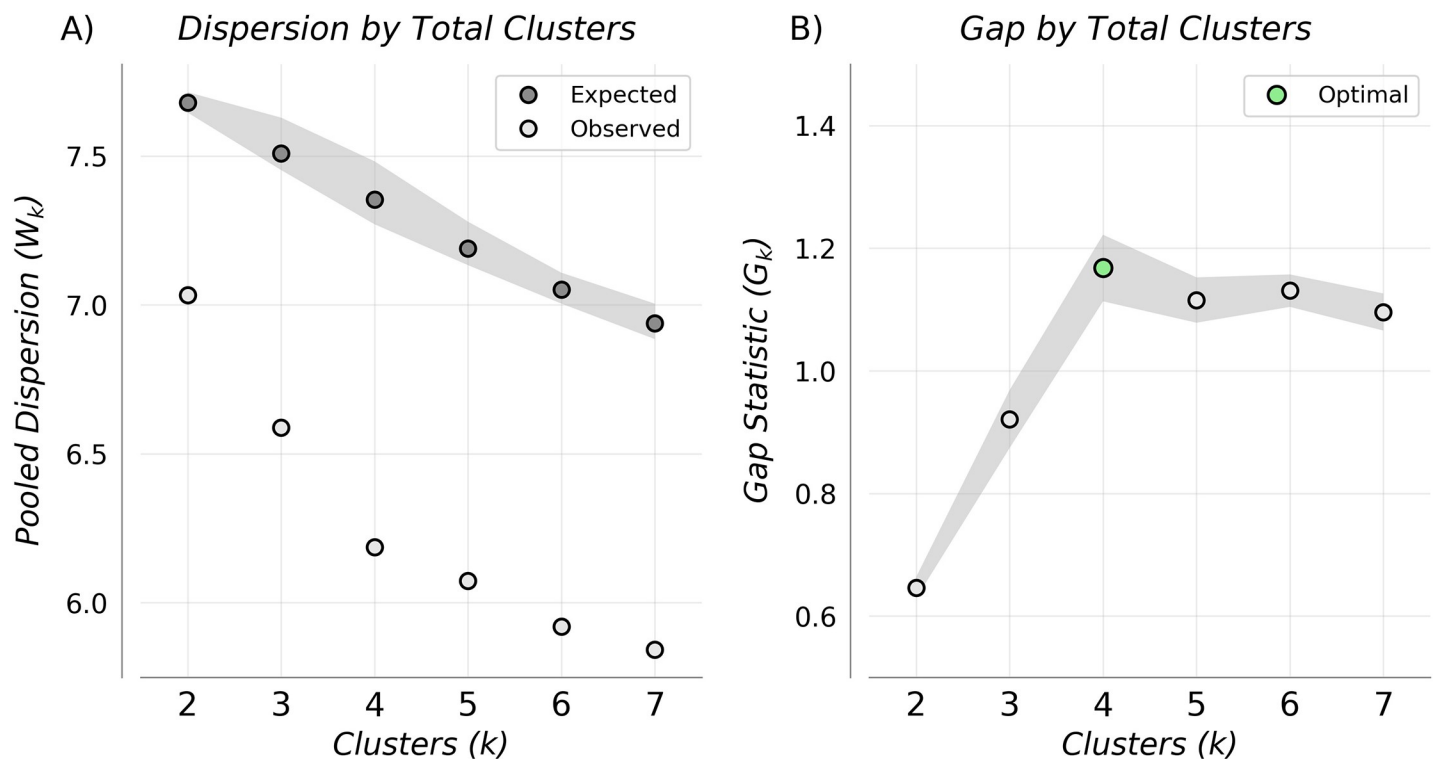

Fig 2. Cluster dispersion and gap statistic. Pooled Intra-cluster dispersion (A), and associated Gap Statistics (B) as a function of cluster number. Gap-statistic disparity was maximized at four clusters which also corresponded to the optimal number given by the selection criteria in Tibshirani et al. [19].
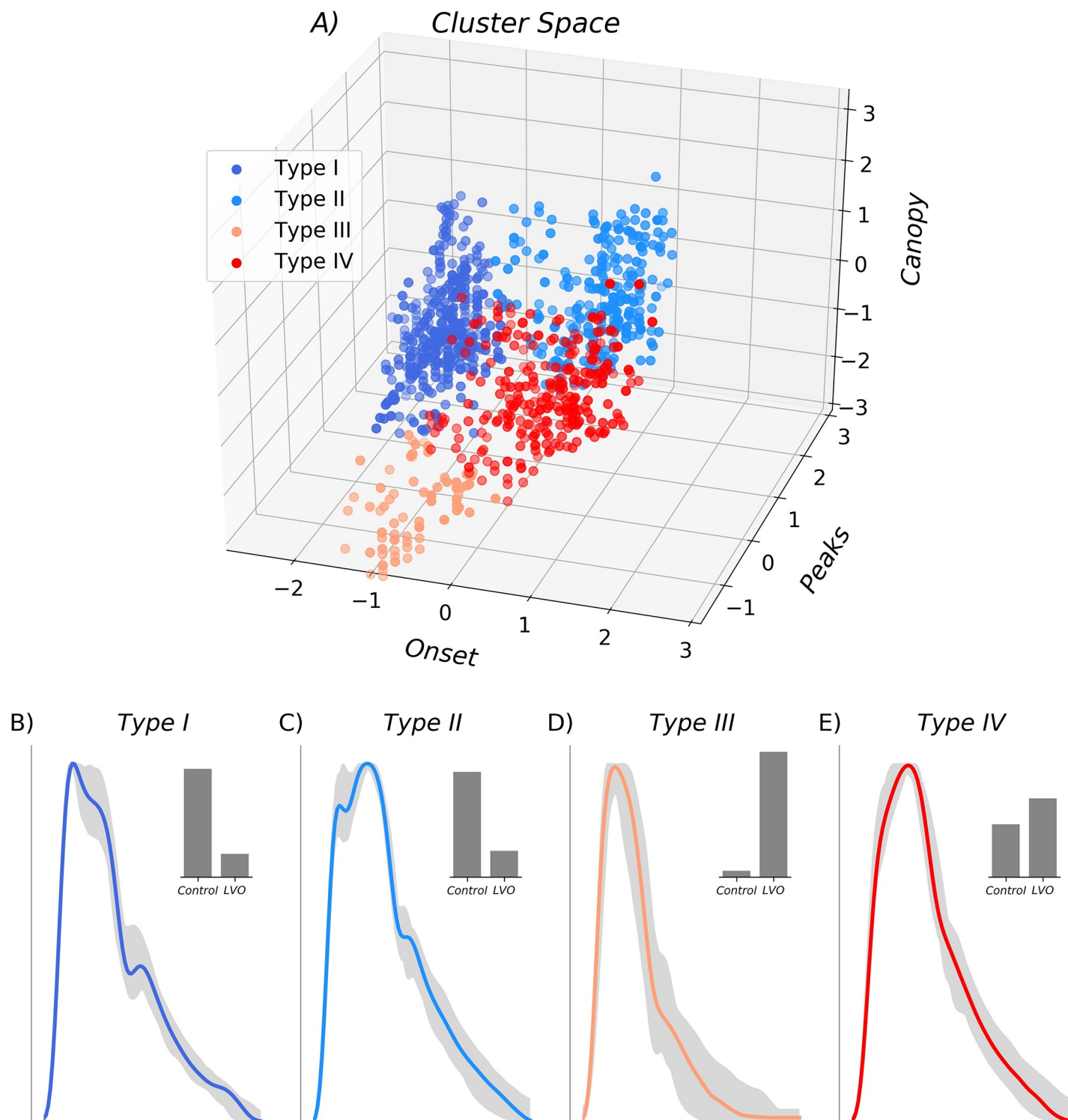
Fig 3. Cluster feature space. Three-dimensional feature space is shown in Z-scored coordinates with the optimal four clusters indicated by color. Associated cluster morphologies are shown for each beat type (B-E), with each cluster archetype shown in color, and the range of the next 50 most representative exemplars depicted in gray. Associated histograms in the upper right (B-E) demonstrate that Type I and II clusters were primarily composed of waveforms from control patient populations, whereas Type III and IV clusters were primarily composed of LVO patient waveforms.

https://doi.org/10.1371/journal.pone.0228642.g003

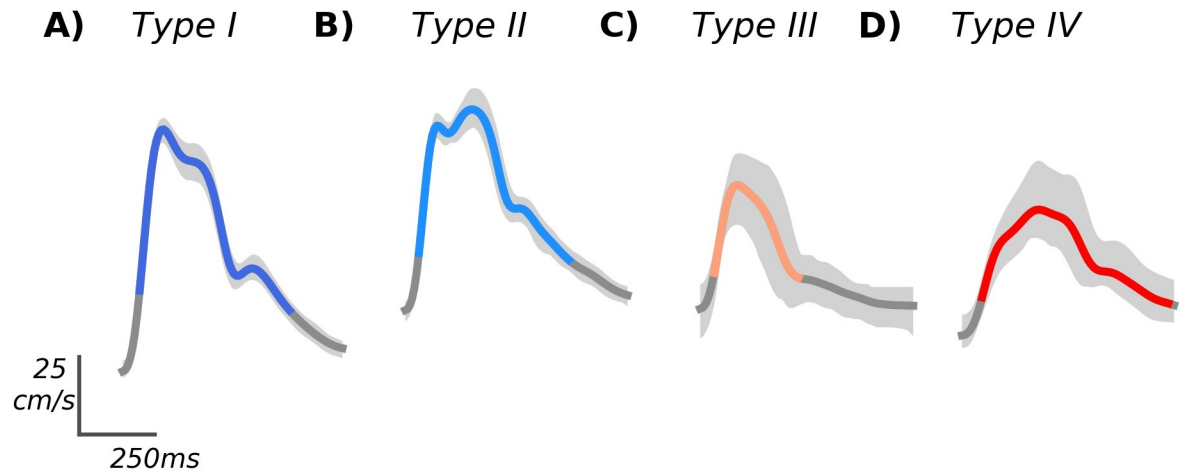**Fig 4. Example waveforms.** The individual example waveforms most representative of each cluster are shown in standard units (unnormalized), with variability across individual beats depicted in light gray. The Types I and II examples (A and B) originated from the OHC and IHC groups, respectively, whereas the Types III and IV examples (C and D) originated from LVO, and IHC subjects, respectively.

https://doi.org/10.1371/journal.pone.0228642.g004

cluster, 95% came from LVO subjects versus 5% from controls (1% IHC, 4% OHC). The final cluster (type IV), containing 256 waveforms, was characterized by late max velocity onset, with wide canopy but weak peaks (Fig 3E). Of waveforms in this cluster, 60% came from LVO subjects versus 40% from controls (28% IHC, 12% OHC).

5. Discussion

5.1. Comparison to TIBI flow grades

Of the clusters we observed, two have definitive analogues on the TIBI scale. Specifically, our type I cluster showed the early systolic maximum and recognizable peak structure associated with TIBI grade 5 normal flow. Similarly, our type IV cluster exhibited the delayed flow acceleration with no discernible early peak, and maximum velocity in mid-to-late systole, characteristic of TIBI grade 2 blunted flow; features which can also be clearly observed in the most representative cluster waveforms, shown unnormalized in Fig 4. Accordingly, we observed type I flows more often in control subjects, whereas type IV flows were more commonly associated with LVO. The remaining clusters did not have unambiguous TIBI analogues, though their contrasting subject group compositions suggest type II is more commonly observed in controls, and type III nearly always associated with LVO. The type II cluster, characterized by late onset maximal velocity but otherwise normal peaks, may reflect differences in peripheral vascular resistance relative to type I, which could conceivably impact either or both the initial systolic upstroke and/or the timing of reflected waves affecting the amplitude of the mid-systolic peak.

Interpretation of the pathological type III morphology represents the most speculative aspect of this work. We infer from its strong association with LVO subjects that it likely results from occlusion or stenosis of the cerebral vessels, though in a manner distinct from typically blunted waveforms, leaving the initial systolic acceleration unaffected while suppressing all subsequent morphological structure. These features can be seen clearly in both the associated cluster archetype (Fig 3D), as well as the most representative cluster exemplar (Fig 4C). This pattern is intriguingly similar to previously observed waveforms associated with vascular stenosis (see, for example, Fig 4 in [71]), suggesting this flow type may rightly be thought of as the

analog of the TIBI stenotic grade 4. However, multiple reasons compel us to be cautious with this claim. First, this flow type was by far the smallest cluster, which was manifest in only a small number of subjects. CTA case reports for most of these subjects did indicate significant, but non flow-limiting stenoses. However, such moderate stenoses are not uncommon in the LVO subject population in general. Moreover, the TIBI grade 4 categorical definition given by Demchuk et al. [27] refers both to elevated velocity relative to the adjacent (unaffected) hemisphere, as well as evidence for flow turbulence, which were not quantified as part of this analysis and thus require further work to assess. Finally, previous studies have shown that qualitatively similar TCD morphologies can be associated with intracranial hypertension and/ or elevated cerebral perfusion pressure [43], which were not controlled for in this study. For the time being we reserve judgment as to whether future studies incorporating richer spectral data might confirm this flow type to have stenotic etiology.

5.2. Limitations and future work

Considering the remaining TIBI flow grades, the lowest are not associated with sufficiently pulsatile CBFV waveforms, and thus could not be represented in our data set. Specifically, grade zero is defined as the absence of flow, whereas grade 1 (minimal flow) is so weakly pulsatile as to give rise to nearly flat waveforms when averaged over successive beats. In this study we required subjects to exhibit bilateral pulsatile data to be included in analysis. This was done to ensure that absence of temporal acoustic windows would not be mistaken for absence of MCA flow. However, future experimental protocols could be modified to require TCD examination of all cerebral vessels in the anterior circulation. Evidence of flow in any of these vessels in a given hemisphere without corresponding evidence of adjacent MCA flow would allow for confident assessment of absent or minimal flow grades, likely bringing our current results further into alignment with the TIBI scale. The remaining TIBI grade 3 flow (Dampened), is not solely morphologically defined, requiring comparison of velocity magnitude relative to a control waveform for assignment, and thus cannot clearly align with our clusters. Future work could explore whether our clustering framework might be extended for application to sets of waveforms, including relative velocities as features, which might help reconstruct these other TIBI categories.

From the 3-dimensional cluster space (Fig 3A) it is clear that adjacent cluster types II and IV share a fuzzy boundary primarily determined by the peaks variable, suggesting the two can be difficult to differentiate when systolic peaks are not clear. Indeed, the type IV cluster had the least homogeneous group composition, with 40% of waveforms originating from control subjects; a fact which would negatively impact specificity were we to use these clusters alone to classify LVO. Underlining this point is the fact that the most representative exemplar for both the type II and type IV waveforms originated from the IHC group. Clearly, further work is needed to determine whether additional or refined clustering features, perhaps derived from waveform spectrograms and associated M-mode, might help disambiguate these groups.

5.3. Conclusions

To our knowledge, this study presents the first unsupervised learning analysis of LVO pathology evident in the TCD signal. We have shown that spectral clustering can readily recover meaningful TCD flow types bearing clear relation to known morphological categories. Moreover, the resultant cluster archetypes provide definitive morphological templates, enabling automated categorization of novel waveforms via numerous potential comparative methods, the refinement of which should provide fruitful avenues for future work. Ultimately, we will explore whether such automated labels can be combined with other established metrics, such as VCI and velocity asymmetry, to improve LVO classification efficacy.

Acknowledgments

Author Contributions

Conceptualization: Corey M. Thibeault, Seth J. Wilk, Thomas Devlin, Robert B. Hamilton.

Data curation: Samuel G. Thorpe, Corey M. Thibeault, Nicolas Canac, Thomas Devlin.

Formal analysis: Samuel G. Thorpe, Kian Jalaleddini.

Investigation: Corey M. Thibeault, Seth J. Wilk, Thomas Devlin.

Methodology: Samuel G. Thorpe, Corey M. Thibeault, Nicolas Canac, Kian Jalaleddini, Seth J. Wilk, Thomas Devlin, Fabien Scalzo, Robert B. Hamilton.

Project administration: Corey M. Thibeault, Amber Dorn, Seth J. Wilk, Thomas Devlin, Robert B. Hamilton.

Resources: Amber Dorn, Seth J. Wilk, Thomas Devlin, Fabien Scalzo, Robert B. Hamilton.

Software: Samuel G. Thorpe, Corey M. Thibeault, Nicolas Canac, Kian Jalaleddini.

Supervision: Corey M. Thibeault, Seth J. Wilk, Thomas Devlin, Robert B. Hamilton.

Validation: Samuel G. Thorpe, Nicolas Canac, Kian Jalaleddini, Amber Dorn, Fabien Scalzo, Robert B. Hamilton.

Visualization: Samuel G. Thorpe, Nicolas Canac.

Writing – original draft: Samuel G. Thorpe.

Writing – review & editing: Samuel G. Thorpe, Corey M. Thibeault, Nicolas Canac, Kian Jalaleddini, Amber Dorn, Seth J. Wilk, Thomas Devlin, Fabien Scalzo, Robert B. Hamilton.

References

1. Spencer MP, Reid JM. Cerebrovascular Evaluation with Doppler Ultrasound. Springer, Dordrecht; 1981.

2. Aaslid R, Markwalder TM, Nornes H. Noninvasive transcranial Doppler ultrasound recording of flow velocity in basal cerebral arteries. J Neurosurg. 1982; 57: 769–774. https://doi.org/10.3171/jns.1982.57.6.0769 PMID: 7143059

3. Aaslid R, Huber P, Nornes H. Evaluation of cerebrovascular spasm with transcranial Doppler ultrasound. J Neurosurg. 1984; 60: 37–41. https://doi.org/10.3171/jns.1984.60.1.0037 PMID: 6689726

4. Aaslid R. Transcranial Doppler Examination Techniques. In: Aaslid R, editor. Transcranial Doppler Sonography. Vienna: Springer Vienna; 1986. pp. 39–59.

5. Sloan MA, Alexandrov AV, Tegeler CH, Spencer MP, Caplan LR, Feldmann E, et al. Assessment: transcranial Doppler ultrasonography: report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Neurology. 2004; 62: 1468–1481. https://doi.org/10.1212/wnl.62.9.1468 PMID: 15136667

6. Alexandrov AV, Sloan MA, Tegeler CH, Newell DN, Lumsden A, Garami Z, et al. Views and Reviews Practice Standards for Transcranial Doppler (TCD) Ultrasound. Part II. Clinical Indications and Expected Outcomes. 2010; 215–224.

7. Alexandrov AV, Sloan MA, Tegeler CH, Newell DN, Lumsden A, Garami Z, et al. Practice standards for transcranial Doppler (TCD) ultrasound. Part II. Clinical indications and expected outcomes. J Neuroimaging. 2012; 22: 215–224. https://doi.org/10.1111/j.1552-6569.2010.00523.x PMID: 20977531

8. D'Andrea A, Conte M, Cavallaro M, Scarafile R, Riegler L, Cocchia R, et al. Transcranial Doppler ultrasonography: From methodology to major clinical applications. World J Cardiol. 2016; 8: 383–400. https://doi.org/10.4330/wjc.v8.i7.383 PMID: 27468332

9. Alexandrov AV. Cerebrovascular Ultrasound in Stroke Prevention and Treatment. John Wiley & Sons; 2011.

10. Alexandrov AV. Neurovascular Examination: The Rapid Evaluation of Stroke Patients Using Ultrasound Waveform Interpretation. John Wiley & Sons; 2013.

11. Wijnhoud AD, Koudstaal PJ, Dippel DWJ. The prognostic value of pulsatility index, flow velocity, and their ratio, measured with TCD ultrasound, in patients with a recent TIA or ischemic stroke. Acta Neurol Scand. 2011; 124: 238–244. https://doi.org/10.1111/j.1600-0404.2010.01462.x PMID: 21198447

12. Aghoram R, Narayan SK. Patterns of Transcranial Doppler Flow Velocities in Recent Ischemic Stroke Patients. Ann Indian Acad Neurol. 2018; 21: 193–196. https://doi.org/10.4103/aian.AIAN_417_17 PMID: 30258261

13. Minhas JS, Panerai RB, Ghaly G, Divall P, Robinson TG. Cerebral autoregulation in hemorrhagic stroke: A systematic review and meta-analysis of transcranial Doppler ultrasonography studies. J Clin Ultrasound. 2019; 47: 14–21. https://doi.org/10.1002/jcu.22645 PMID: 30270441

14. Klingelhöfer J, Conrad B, Benecke R, Sander D. Intracranial flow patterns at increasing intracranial pressure. Klin Wochenschr. 1987; 65: 542–545.s https://doi.org/10.1007/bf01727619 PMID: 2957547

15. Hassler W, Steinmetz H, Gawlowski J. Transcranial Doppler ultrasonography in raised intracranial pressure and in intracranial circulatory arrest. J Neurosurg. 1988; 68: 745–751. PMID: 3282040

16. Schmidt B, Klingelhöfer J, Schwarze JJ, Sander D, Wittich I. Noninvasive prediction of intracranial pressure curves using transcranial Doppler ultrasonography and blood pressure curves. Stroke. 1997; 28: 2465–2472. https://doi.org/10.1161/01.str.28.12.2465 PMID: 9412634

17. Cardim D, Robba C, Bohdanowicz M, Donnelly J, Cabella B, Liu X, et al. Non-invasive Monitoring of Intracranial Pressure Using Transcranial Doppler Ultrasonography: Is It Possible? Neurocrit Care. 2016; 25: 473–491. https://doi.org/10.1007/s12028-016-0258-6 PMID: 26940914

18. Adams RJ, Nichols FT, Figueroa R, McKie V, Lott T. Transcranial Doppler correlation with cerebral angiography in sickle cell disease. Stroke. 1992; 23: 1073–1077. https://doi.org/10.1161/01.str.23.8.1073 PMID: 1636180

19. Adams RJ, McKie VC, Carl EM, Nichols FT, Perry R, Brock K, et al. Long-term stroke risk in children with sickle cell disease screened with transcranial Doppler. Ann Neurol. 1997; 42: 699–704. https://doi.org/10.1002/ana.410420505 PMID: 9392568

20. Kral MC, Brown RT, Nietert PJ, Abboud MR, Jackson SM, Hynd GW. Transcranial Doppler ultrasonography and neurocognitive functioning in children with sickle cell disease. Pediatrics. 2003; 112: 324–331. https://doi.org/10.1542/peds.112.2.324 PMID: 12897282

21. Len TK, Neary JP, Asmundson GJG, Goodman DG, Bjornson B, Bhambhani YN. Cerebrovascular reactivity impairment after sport-induced concussion. Med Sci Sports Exerc. 2011; 43: 2241–2248. https://doi.org/10.1249/MSS.0b013e3182249539 PMID: 21606867

22. Bailey DM, Jones DW, Sinnott A, Brugniaux JV, New KJ, Hodson D, et al. Impaired cerebral haemodynamic function associated with chronic traumatic brain injury in professional boxers. Clin Sci. 2013; 124: 177–189. https://doi.org/10.1042/CS20120259 PMID: 22913765

23. Thibeault CM, Thorpe S, O'Brien MJ, Canac N, Ranjbaran M, Patanam I, et al. A Cross-Sectional Study on Cerebral Hemodynamics After Mild Traumatic Brain Injury in a Pediatric Population. Front Neurol. 2018; 9: 200.

24. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, et al. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. Circulation. 2017; 135: e146–e603. https://doi.org/10.1161/CIR.0000000000000485 PMID: 28122885

25. Demchuk AM, Christou I, Wein TH, Felberg R a., Malkoff M, Grotta JC, et al. Specific transcranial Doppler flow findings related to the presence and site of arterial occlusion. Stroke. 2000; 31: 140–146. https://doi.org/10.1161/01.str.31.1.140 PMID: 10625729

26. Demchuk AM, Christou I, Wein TH, Felberg RA, Malkoff M, Grotta JC, et al. Accuracy and Criteria for Localizing Arterial Occlusion With Transcranial Doppler. J Neuroimaging. 2000; 10: 1–12. https://doi.org/10.1111/jon20001011 PMID: 10666975

27. Demchuk AM, Scott Burgin W, Christou I, Felberg RA, Barber PA, Hill MD, et al. Thrombolysis in Brain Ischemia (TIBI) transcranial Doppler flow grades predict clinical severity, early recovery, and mortality in patients treated with intravenous tissue plasminogen activator. Stroke. 2001; 32: 89–93. https://doi.org/10.1161/01.str.32.1.89 PMID: 11136920

28. Saqqur M, Uchino K, Demchuk AM, Molina CA, Garami Z, Calleja S, et al. Site of arterial occlusion identified by transcranial Doppler predicts the response to intravenous thrombolysis for stroke. Stroke. 2007; 38: 948–954. https://doi.org/10.1161/01.STR.0000257304.21967.ba PMID: 17290031

29. Tsivgoulis G, Sharma VK, Lao AY, Malkoff MD, Alexandrov AV. Validation of transcranial Doppler with computed tomography angiography in acute cerebral ischemia. Stroke. 2007; 38: 1245–1249. https://doi.org/10.1161/01.STR.0000259712.64772.85 PMID: 17332465

30. Brunser AM, Lavados PM, Hoppe A, Lopez J, Valenzuela M, Rivas R. Accuracy of transcranial doppler compared with CT angiography in diagnosing arterial obstructions in acute ischemic strokes. Stroke. 2009; 40: 2037–2041. https://doi.org/10.1161/STROKEAHA.108.542704 PMID: 19359640

31. Sheth SA, Jahan R, Gralla J, Pereira VM, Nogueira RG, Levy EI, et al. Time to endovascular reperfusion and degree of disability in acute stroke. Ann Neurol. 2015; 78: 584–593. https://doi.org/10.1002/ana.24474 PMID: 26153450

32. Saver JL, Goyal M, van der Lugt A, Menon BK, Majoie CBLM, Dippel DW, et al. Time to Treatment With Endovascular Thrombectomy and Outcomes From Ischemic Stroke: A Meta-analysis. JAMA. 2016; 316: 1279. https://doi.org/10.1001/jama.2016.13647 PMID: 27673305

33. Zerna C, Hegedus J, Hill MD. Evolving Treatments for Acute Ischemic Stroke. Circ Res. 2016; 118: 1425–1442. https://doi.org/10.1161/CIRCRESAHA.116.307005 PMID: 27126651

34. Smith WS, Tsao JW, Billings ME, Johnston SC, Hemphill JC 3rd, Bonovich DC, et al. Prognostic significance of angiographically confirmed large vessel intracranial occlusion in patients presenting with acute brain ischemia. Neurocrit Care. 2006; 4: 14–17. https://doi.org/10.1385/ncc:4:1:014 PMID: 16498189

35. Malhotra K, Liebeskind DS. Imaging in Endovascular Stroke Trials. J Neuroimaging. 2015; 25: 517–527. https://doi.org/10.1111/jon.12272 PMID: 26179500

36. Malhotra K, Gornbein J, Saver JL. Ischemic Strokes Due to Large-Vessel Occlusions Contribute Disproportionately to Stroke-Related Dependence and Death: A Review. Front Neurol. 2017; 8: 651. https://doi.org/10.3389/fneur.2017.00651 PMID: 29250029

37. Alberts MJ, Latchaw RE, Selman WR, Shephard T, Hadley MN, Brass LM, et al. Recommendations for comprehensive stroke centers: a consensus statement from the Brain Attack Coalition. Stroke. 2005; 36: 1597–1616. https://doi.org/10.1161/01.STR.0000170622.07210.b4 PMID: 15961715

38. Latchaw RE, Alberts MJ, Lev MH, Connors JJ, Harbaugh RE, Higashida RT, et al. Recommendations for imaging of acute ischemic stroke: A scientific statement from the american heart association. Stroke. 2009; 40: 3646–3678. https://doi.org/10.1161/STROKEAHA.108.192616 PMID: 19797189

39. Kushner MJ, Zanette EM, Bastianello S, Mancini G, Sacchetti ML, Carolei A, et al. Transcranial Doppler in acute hemispheric brain infarction. Neurology. 1991; 41: 109–113. https://doi.org/10.1212/wnl.41.1.109 PMID: 1985274

40. Zanette EM, Fieschi C, Bozzao L, Roberti C, Toni D, Argentino C, et al. Comparison of cerebral angiography and transcranial Doppler sonography in acute stroke. Stroke. 1989; 20: 899–903. https://doi.org/10.1161/01.str.20.7.899 PMID: 2665206

41. Xiang LI, Han X, Hua Y, Yang J. Correlation between cerebral blood flow and cerebral infarction in patients with chronic middle cerebral artery occlusion detected by ultrasound. Chinese Journal of. 2017. Available: http://wprim.whocc.org.cn/admin/article/articleDetail?WPRIMID=613963&articleId=613963

42. Cardoso ER, Rowan JO, Galbraith S. Analysis of the cerebrospinal fluid pulse wave in intracranial pressure. J Neurosurg. 1983; 59: 817–821. https://doi.org/10.3171/jns.1983.59.5.0817 PMID: 6619934

43. Aggarwal S, Brooks DM, Kang Y, Linden PK, Patzer JF. Noninvasive monitoring of cerebral perfusion pressure in patients with acute liver failure using transcranial doppler ultrasonography. Liver Transpl. 2008; 14: 1048–1057. https://doi.org/10.1002/lt.21499 PMID: 18581484

44. Wagshul ME, Eide PK, Madsen JR. The pulsating brain: A review of experimental and clinical studies of intracranial pulsatility. Fluids Barriers CNS. 2011; 8: 5. https://doi.org/10.1186/2045-8118-8-5 PMID: 21349153

45. Kim S, Hu X, McArthur D, Hamilton R, Bergsneider M, Glenn T, et al. Inter-Subject Correlation Exists Between Morphological Metrics of Cerebral Blood Flow Velocity and Intracranial Pressure Pulses. Neurocrit Care. 2011; 14: 229–237. https://doi.org/10.1007/s12028-010-9471-x PMID: 21136207

46. Kim Sunghan, Hamilton R, Pineles S, Bergsneider M, Xiao Hu. Noninvasive Intracranial Hypertension Detection Utilizing Semisupervised Learning. IEEE Transactions on Biomedical Engineering. 2013; 60: 1126–1133. https://doi.org/10.1109/TBME.2012.2227477 PMID: 23193303

47. Kurji A, Debert CT, Whitelaw WA, Rawling JM, Frayne R, Poulin MJ. Differences between middle cerebral artery blood velocity waveforms of young and postmenopausal women. Menopause. 2006; 13: 303–313. https://doi.org/10.1097/01.gme.0000177905.94515.24 PMID: 16645544

48. Lockhart CJ, Gamble AJ, Rea D, Hughes S, McGivern RC, Wolsley C, et al. Nitric oxide modulation of ophthalmic artery blood flow velocity waveform morphology in healthy volunteers. Clin Sci. 2006; 111: 47–52. https://doi.org/10.1042/CS20050365 PMID: 16499475

49. Schaafsma A. Improved parameterization of the transcranial Doppler signal. Ultrasound Med Biol. 2012; 38: 1451–1459. https://doi.org/10.1016/j.ultrasmedbio.2012.03.016 PMID: 22579541

50. Uğuz H. A hybrid system based on information gain and principal component analysis for the classification of transcranial Doppler signals. Comput Methods Programs Biomed. 2012; 107: 598–609. https://doi.org/10.1016/j.cmpb.2011.03.013 PMID: 21524813

51. Purvis D, Aldaghlas T, Trickey AW, Rizzo A, Sikdar S. A Novel Decision Tree Approach Based on Transcranial Doppler Sonography to Screen for Blunt Cervical Vascular Injuries. J Ultrasound Med. 2013; 32: 1023–1031. https://doi.org/10.7863/ultra.32.6.1023 PMID: 23716524

52. Kim S, Scalzo F, Hu X. Noninvasive intracranial hypertension diagnosis using ensemble sparse classifiers. Proceedings of the 2011 workshop on Data mining for medicine and healthcare—DMMH '11. New York, New York, USA: ACM Press; 2011. p. 68.

53. Scalzo F, Hamilton R, Asgari S, Kim S, Hu X. Intracranial hypertension prediction using extremely randomized decision trees. Med Eng Phys. 2012; 34: 1058–1065. https://doi.org/10.1016/j.medengphy.2011.11.010 PMID: 22401795

54. Thorpe SG, Thibeault CM, Canac N, Wilk SJ, Devlin T, Hamilton RB. Decision Criteria for Large Vessel Occlusion Using Transcranial Doppler Waveform Morphology. Front Neurol. 2018; 9: 847. https://doi.org/10.3389/fneur.2018.00847 PMID: 30386287

55. Cardim D, Schmidt B, Robba C, Donnelly J, Puppo C, Czosnyka M, et al. Transcranial Doppler Monitoring of Intracranial Pressure Plateau Waves. Neurocrit Care. 2017; 26: 330–338. https://doi.org/10.1007/s12028-016-0356-5 PMID: 28000131

56. Elzaafarany K, Aly MH, Kumar G, Nakhmani A. Cerebral Artery Vasospasm Detection Using Transcranial Doppler Signal Analysis. J Ultrasound Med. 2019; 38: 2191–2202. https://doi.org/10.1002/jum.14916 PMID: 30593699

57. Seera M, Lim CP, Tan KS, Liew WS. Classification of transcranial Doppler signals using individual and ensemble recurrent neural networks. Neurocomputing. 2017; 249: 337–344.

58. Thorpe SG, Thibeault CM, Wilk SJ, O'Brien M, Canac N, Ranjbaran M, et al. Velocity Curvature Index: a Novel Diagnostic Biomarker for Large Vessel Occlusion. Transl Stroke Res. 2018. https://doi.org/10.1007/s12975-018-0667-2 PMID: 30293170

59. Ng AY, Jordan MI, Weiss Y. On Spectral Clustering: Analysis and an algorithm. In: Dietterich TG, Becker S, Ghahramani Z, editors. Advances in Neural Information Processing Systems 14. MIT Press; 2002. pp. 849–856.

60. Zelnik-Manor L, Perona P. Self-tuning spectral clustering In: Advances in Neural Information Processing Systems (NIPS). MIT Press, Cambridge; 2004.

61. von Luxburg U. A tutorial on spectral clustering. Stat Comput. 2007; 17: 395–416.

62. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J R Stat Soc Series B Stat Methodol. 2001; 63: 411–423.

63. Canac N, Ranjbaran M, O'Brien MJ, Asgari S, Scalzo F, Thorpe SG, et al. Algorithm for Reliable Detection of Pulse Onsets in Cerebral Blood Flow Velocity Signals. Front Neurol. 2019; 10: 1072. https://doi.org/10.3389/fneur.2019.01072 PMID: 31681147

64. Jalaleddini K, Canac N, Thorpe S, O'Brien M, Ranjbaran M, Delay B, et al. Objective Assessment of Beat Quality in Transcranial Doppler Measurement of Blood Flow Velocity in Cerebral Arteries. IEEE Trans Biomed Eng. 2019. https://doi.org/10.1109/TBME.2019.2923146 PMID: 31217091

65. Hu X, Xu P, Scalzo F, Vespa P, Bergsneider M. Morphological clustering and analysis of continuous intracranial pressure. IEEE Trans Biomed Eng. 2009; 56: 696–705. https://doi.org/10.1109/TBME.2008.2008636 PMID: 19272879

66. Larose DT. An introduction to data mining. Discovering Knowledge in Data, John Wiley & Sons Publication, Hoboken, New Jersey, USA. 2005; 1–25.

67. Mohamad IB, Usman D. Standardization and its effects on K-means clustering algorithm. Research Journal of Applied Sciences, Engineering and Technology. 2013; 6: 3299–3303.

68. Pedregosa F, Weiss R, Brucher M. Scikit-learn: Machine Learning in Python. 2011; 12: 2825–2830.

69. Kumar S, Toshniwal D. A data mining approach to characterize road accident locations. Journal of Modern Transportation. 2016; 24: 62–72.

70. Rathod RR. Design of electricity tariff plans using gap statistic for K-means clustering based on consumers monthly electricity consumption data. Int J Energy Sect Manage. 2017; 11: 295–310.

71. Bathala L, Mehndiratta MM, Sharma VK. Transcranial doppler: Technique and common findings (Part 1). Ann Indian Acad Neurol. 2013; 16: 174–179. https://doi.org/10.4103/0972-2327.112460 PMID: 23956559