Studying Cerebral Vasculature Using Structure Proximity and Graph Kernels

Roland Kwitt[1], Danielle Pace[1], Marc Niethammer[2], and Stephen Aylward[1]

[1] Kitware Inc., Carrboro, NC, USA
[2] University of North Carolina (UNC), Chapel Hill, NC, USA

Abstract. An approach to study population differences in cerebral vasculature is proposed. This is done by 1) extending the concept of encoding cerebral blood vessel networks as spatial graphs and 2) quantifying graph similarity in a kernel-based discriminant classifier setup. We argue that augmenting graph vertices with information about their proximity to selected brain structures adds discriminative information and consequently leads to a more expressive encoding. Using graph-kernels then allows us to quantify graph similarity in a principled way. To demonstrate our approach, we assess the hypothesis that gender differences manifest as variations in the architecture of cerebral blood vessels, an observation that previously had only been tested and confirmed for the Circle of Willis. Our results strongly support this hypothesis, i.e, we can demonstrate non-trivial, statistically significant deviations from random gender classification in a cross-validation setup on 40 healthy patients.

1 Motivation

The human circulatory system, with blood vessels transporting nutrients and waste from one location to another, appears to lend itself naturally to a graph-based representation. A question that immediately arises is whether we can leverage that graph structure to study population differences that manifest as topological changes. In this work, we focus on the cerebrovascular network. Several studies indicate an association between selected mental diseases and irregular cerebral blood vessel topology. Preliminary to our study of such associations, herein we assess the hypothesis that gender-associated differences exist in the architecture of cerebral vessel networks. A recent study [**?**] demonstrates gender-related differences in *local* geometry, e.g., radius; however, evidence in support of *general* architectural differences across genders has only been reported for the Circle of Willis [**?**]. Beyond serving as a preliminary study, identifying gender-related differences could also help to explain why certain vascular pathologies, such as aneurysms, have a higher incidence rate in women than in men [**?**].

Although it is straight-forward to compare individuals or populations in terms of local vessel network properties, e.g., tortuosity measures or radius estimates [**?**], a general framework for comparing and characterizing vascular topology at a global scale does not exist. Furthermore, vascular graph comparisons are typically performed using summary statistics in order to circumvent

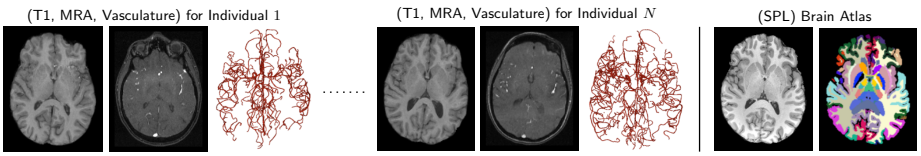


Fig. 1: Input data: example (mid-transversal) slices of skull-stripped T1-weighted MRI and MRA as well as extracted vasculature from different individuals. The right-hand side shows the atlas of brain structures [**?**] that we use as a reference space.

computational complexity issues inherent to many graph similarity metrics. Additionally, while progress has been made in representing cerebral vasculature as graphs [**?**] and even as graphs that span intracranial space [**?**], no prior work, to our knowledge, has augmented vascular graph encodings with *brain structure information*, e.g., hippocampus, thalamic nuclei and so forth.

Our contribution is two-fold: First, we augment the spatial graph representation of [**?**] with brain structure information at each vertex to enhance expressiveness. Second, we draw upon recent advances in machine learning with graph-structured data [**?**] to quantify gender-related differences in cerebrovascular architecture using a graph-kernel based discriminant classifier.

2 Functional Graphs

In this section, we introduce the concept of *functional graphs*, an extension of spatial graphs [**?**] with vertex labels that encode information about proximal brain structures. Our approach has the following requirements: For a population of N individuals, we require MRA and T1-weighted MRI images, as well as extracted cerebral vasculature in a centerline + radius representation. Fig. **??** shows some example images from the dataset we use throughout this work (see Sect. **??** for details). To study architectural differences of blood vessel networks, effects arising from translation, rotation, etc. need to be removed. This is typically done by aligning all images with respect to a common brain atlas and applying the corresponding transform(s) on the blood vessels to map them into the atlas space. We further require that the atlas has an associated voxel-wise labeling of brain structures (e.g., see [**?**]). For clarity of presentation, we briefly recapitulate spatial graph construction [**?**] and then discuss our contribution.

From Centerlines to Spatial Graphs. The representation of blood vessels as centerlines is difficult to use directly for comparing individuals, primarily due to extensive, naturally occurring spatial and topological variations among individuals. To study the global blood vessel network, we favor a more compact and less noisy representation.

To compute a *spatial graph*, the brain is parcellated into regions that are equally likely to contain a vessel; the vessel paths across those regions thereby form a compact graph representation of the individual, and those graphs can be summarized and compared within and across individuals and populations.

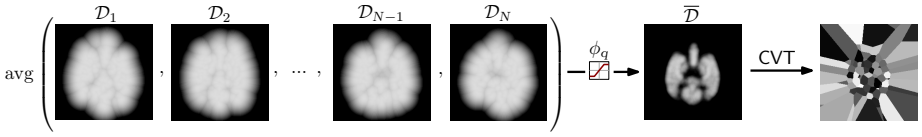


Fig. 2: Construction of the vascular density atlas, i.e., the spatial density image $\overline{\mathcal{D}}$ of blood-vessels, from N (inverted) distance maps $\mathcal{D}_1, \ldots, \mathcal{D}_N$. The centroidal Voronoi tessellation is then driven by the density information of the $[0, 1]$-normalized image $\overline{\mathcal{D}}$.

Regarding parcellation, first, an atlas of vessel density is constructed for a population of individuals. The atlas is established by computing an Euclidean distance map for each individual's vascular network[3]. Averaging the (inverted) distance maps $\{\mathcal{D}_i\}_{i=1}^N$, followed by a normalization of distances above the q-th percentile (e.g., $q = 80$) to $[0, 1]$ (denoted by ϕ_q in Fig. **??**), allows the resulting image $\overline{\mathcal{D}}$ to be interpreted as an atlas of spatial blood vessel density.

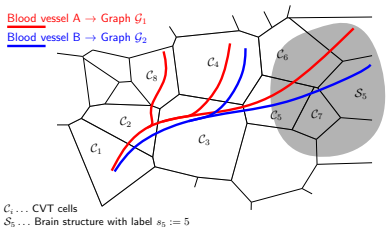
Second, [**?**] proposes a centroidal Voronoi tessellation (CVT) of the density atlas to identify regions of approximately equal probability. Specifically, CVT is applied to $\overline{\mathcal{D}}$ to generate C cells $\mathcal{C}_1, \ldots, \mathcal{C}_C$. The cells are computed by running Lloyd's algorithm on a subset of all voxels, selected with probability proportional to the vessel density in $\overline{\mathcal{D}}$. The algorithm is initialized by C randomly chosen cell centers which are then iteratively refined. This computation scheme results in a higher number of cells in regions of high spatial density and essentially approximates an equiprobable (in terms of expected probability of observing a blood vessel in a cell) tessellation of $\overline{\mathcal{D}}$, illustrated in Fig. **??**.

The graph representation of an individual's blood vessel network is then encoded in an adjacency matrix $\boldsymbol{A} = [a_{ij}]_{C \times C}$, where $a_{ij} = 1$ signifies that a blood vessel traverses from cell $\mathcal{C}_i$ to cell $\mathcal{C}_j$. This encoding leads to an (unlabeled) undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with vertex set $\mathcal{V} = \{v_1, \ldots, v_C\}$ and edge set $\mathcal{E} = \{(v_i, v_j) : a_{ij} = 1\}$. While $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ captures blood vessel topology, no information about brain anatomy is included. We hypothesize that adding information regarding adjacent brain structures will produce graphs with more discriminative power.

Graph Augmentation by Vertex Labels. A label function $\gamma : \mathcal{V} \to E_0 := \{l_1, \ldots, l_L\}$ assigns one of L discrete labels from an alphabet E_0 to a vertex $v \in \mathcal{V}$. Several variations of E_0 and γ are possible. In fact, the identifier $i \in \mathbb{N}$ of a CVT cell $\mathcal{C}_i$ is already a valid label, since it uniquely identifies the spatial position of the graph vertices in atlas space. Another strategy is to use the vertex degree, i.e., the number of incident edges, as a vertex label. We refer to those two strategies as the *default* labeling strategies (a) and (b) with $E_0^a := \{1, \ldots, C\}$, $E_0^b := \{d_i : d_i := \deg(v_i)\}$ and corresponding label functions $\gamma^a(v_i) = i, \gamma^b(v_i) = \deg(v_i) = d_i$.

To augment the graph with more informative labels, we introduce the concept of *label transfer*. Given a set of S segmented brain structures $\mathcal{S}_0, \ldots, \mathcal{S}_{S-1}$,

[3] i.e., an approximate voxel-wise distance to the closest blood vessel, e.g., computed using the N-d extension of Daniellson's algorithm, cf. [**?**]



	Labeling		
Cell / Vertex	γ^a	γ^c	γ^d
$\mathcal{C}_1$ / v_1	1	0	1
$\mathcal{C}_2$ / v_2	2	0	2
$\mathcal{C}_3$ / v_3	3	0	3
$\mathcal{C}_4$ / v_4	4	0	4
$\mathcal{C}_5$ / v_5	5	5	5
$\mathcal{C}_6$ / v_6	6	5	5
$\mathcal{C}_7$ / v_7	7	5	5
$\mathcal{C}_8$ / v_8	8	0	6

Fig. 3: Exemplary CVT tessellation with two overlayed blood vessels. *Without* label transfer (γ^a), vertices are labeled sequentially. *With* label transfer using non-unique background labels (γ^c), vertices $v_{5,6,7}$ are labeled as $s_5 := 5$, vertices $v_{1,2,3,4,8}$ as background $s_0 := 0$; unique background labels (γ^d) allow to distinguish vertices $v_{1,2,3,4,8}$.

represented by a voxel-wise label map, we propose to transfer those labels to the graph vertices in a way that abstractly encodes the function of a vessel in the brain, i.e., the structures it may influence. We refer to this extension of spatial graphs, i.e., augmentation with labels encoding information about proximal brain structures, as *functional graphs*.

Formally, structure $\mathcal{S}_u = \{i : \xi(\boldsymbol{x}_i) = s_u\}$ is given as the index set of voxels $\boldsymbol{x}_i$ with segmentation label s_u from the brain atlas, where ξ denotes the mapping from voxels to segmentation labels. One possible strategy for label transfer, i.e., the assignment of a discrete label $l_i \in \{s_0, \dots, s_{S-1}\}$ to vertex v_i, is to compute the intersection of $\mathcal{C}_i$ with each segmentation region $\mathcal{S}_j$ and tally the occurrences of the corresponding segmentation labels. Given that $h_{in} := |\mathcal{C}_i \cap \mathcal{S}_n|$ denotes the cardinality of the index set for the intersection of $\mathcal{C}_i$ with $\mathcal{S}_n$, we can define a label histogram for vertex v_i as $\boldsymbol{h}_i := [h_{i0}, \dots, h_{iS-1}]^\top$. This label histogram can then be used extract one discrete label $\gamma(v_i) = s_u$, e.g., by means of max-pooling $u = \arg\max_s h_{is}$. The strategy has one drawback, though. Since the label map is *sparse* in the sense that voxels which are not part of segmented structures are labeled as *background* s_0, max-pooling can lead to a significant number of s_0-labeled vertices and, in further consequence, to undesirable ambiguities. For instance, edges in two graphs with start/end vertex labeled as s_0 are topologically equivalent, presumably leading to a loss in expressiveness. To avoid this effect, we propose to track vertices labeled as s_0 and assign a new, *unique* label $b_i \in \mathcal{B} := \{b_j\}_{j=1}^B, \mathcal{B} \cap \{s_0, \dots, s_{S-1}\} = \varnothing$ each time. We refer to label transfer with / without unique background labels as labeling strategies (c) and (d), resp., with $E_0^c := \{s_0, \dots, s_{S-1}\}$, $E_0^d := E_0^c \setminus \{s_0\} \cup \mathcal{B}$ and label functions $\gamma^c(v_i) = s_u$ (u defined as above), and $\gamma^d(v_i) = \{s_u$, if $u \neq 0$, $b_j, j \to j+1$, otherwise$\}$.

3 Measuring Graph Similarity via Graph Kernels

Quantifying similarity between graphs is an inherently difficult problem due to the fact that no polynomial-time algorithm is known to compare graphs in a strict graph-isomorphism sense. Other strategies, such as comparing graphs by

means of topological descriptors, however, are often too restrictive or ignore important aspects. A promising alternative is to rely on *graph-kernels*, a concept that has recently gained considerable interest in the machine learning community, not least because kernels are the key ingredient of many learning algorithms.

The Weisfeiler-Lehman (WL) Subtree Kernel. In this work, we exploit the Weisfeiler-Lehman subtree kernel (WLS) [**?**], a recently introduced graph-kernel with an efficient computation scheme that relies on comparing *subtree patterns*. A *subtree* is a subgraph of a graph with a designated root and no cycles. The maximum distance between the root and any other vertex in the subtree is referred to as the *subtree height* h. A *subtree pattern* extends the notion of a subtree by allowing equal vertices. The key concept of the WLS kernel is to count all pairs of matching subtree patterns, rooted at the vertices of both graphs. For graphs with $|\mathcal{V}| = n$, this can be done in $\mathcal{O}(hn)$.

Assuming graphs with discrete vertex labels, the algorithm proceeds as follows: For each vertex v in both graphs, a signature is created by 1) collecting the vertex labels $\{\gamma(u)|u \in \mathcal{N}(v)\}$ of all neighbors of v, 2) sorting the resulting multiset and 3) adding the sorted multiset to the existing label (e.g., as a string). A new, compressed label is then assigned to all vertices, ensuring that vertices with the *same* signature get the same compressed label. This procedure is iterated until we reach the desired subtree height h. After h iterations, we tally the common multisets in both graphs. Fig. **??** illustrates this concept on the toy example of the two vessels in Fig. **??**. By letting E_i denote the set of all vertex labels in iteration i with $E_i \cap E_{i+1} = \varnothing$ and defining $c_i(\mathcal{G}, l_{ij}), l_{ij} \in E_i$ to be a function that counts the number of occurrences of the j-th vertex label l_{ij} in graph $\mathcal{G}$, we can define the WLS kernel between $\mathcal{G}_u$ and $\mathcal{G}_v$ as:

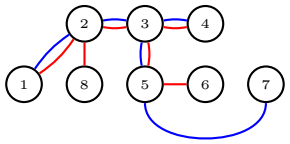
$$k(\mathcal{G}_u, \mathcal{G}_v) = \langle [c_0(\mathcal{G}_u, l_{01}), \ldots, c_h(\mathcal{G}_u, l_{h|E_h|})], [c_0(\mathcal{G}_v, l_{01}), \ldots, c_h(\mathcal{G}_v, l_{h|E_h|})] \rangle. \quad (1)$$

The inner product in Eq. (**??**) is taken over vectors that contain the concatenated counts of the original and compressed labels after h iterations. In our toy example of Fig. **??** for $h = 1$, $k(\mathcal{G}_1, \mathcal{G}_2) = 11$ in case of γ^a and $k(\mathcal{G}_1, \mathcal{G}_2) = 19$ in case of γ^d (see suppl. material). This demonstrates the benefit of label transfer (with unique background labels), since similarity increases due to the fact that both blood vessels traverse the same cells and feed into the same structure.

4 Experiments

The goal of our experimental study is to assess, using graph-kernel based classification, whether there is evidence supporting the hypothesis of gender-related differences in cerebrovascular architecture. This is done by studying classifier predictions in terms of non-trivial deviations from random choice. Along those lines, we show that functional graphs demonstrate preferred behavior compared to spatial graphs with default vertex labelings.

Dataset. We use a dataset of 40 healthy patients , 18 male and 22 female. For all individuals, MRA and T1-weighted MRI scans were acquired by a Siemens Allegra head-only 3T MR system. Voxel spacing is $0.5 \times 0.8 \times 0.8$mm for the MRA



		using γ^a		using γ^d	
v_i	γ^a/γ^d	$\mathcal{G}_1, h=1$	$\mathcal{G}_2, h=1$	$\mathcal{G}_1, h=1$	$\mathcal{G}_2, h=1$
1	1/1	1, 2 (9)	1, 2 (9)	1, 2 (7)	1, 2 (7)
2	2/2	2, 138 (10)	2, 13 (16)	2, 136 (8)	2, 13 (14)
3	3/3	3, 245 (11)	3, 245 (11)	3, 245 (9)	3, 245 (9)
4	4/4	4, 3 (12)	4, 3 (12)	4, 3 (10)	4, 3 (10)
5	5/5	5, 36 (14)	5, 37 (17)	5, 35 (11)	5, 35 (11)
6	6/5	6, 5 (15)	– (–)	5, 5 (12)	– (–)
7	7/5	– (–)	7, 5 (18)	– (–)	5, 5 (12)
8	8/6	8, 2 (13)	– (–)	6, 2 (13)	– (–)

Fig. 4: Graph structure for both blood vessels of Fig. **??**, with edge colors corresponding to blood vessel colors, i.e. $\mathcal{G}_1$, $\mathcal{G}_2$. The table illustrates multiset construction and label compression (numbers in parentheses) for $h = 1$ and γ^a, γ^d (i.e., sequential labels and label transfer with unique background labels).

($448 \times 448 \times 128$ voxel), and 1mm isotropic for the T1 images ($176 \times 256 \times 176$ voxel). Cerebral blood vessels were extracted by an expert (cf. [**?**]). All blood vessels are transformed into the space of the SPL 2012 brain atlas [**?**] (1mm isotropic, $256 \times 256 \times 256$ voxel), using a combination of two transforms, $t_1 \circ t_0$, where t_0 denotes a *rigid* transform to align the MRA and T1 images and t_1 denotes an *affine* transform to align the (skull-stripped) T1 image to the (skull-stripped) brain atlas[4].

Evaluation Setup. In all experiments, we use a C-SVM classifier [**?**] with the cost parameter optimized by cross validation (CV) on the training data. All reported results are averaged over five CV runs. As a common measure of problem difficulty, we also report the average fraction of selected support vectors (listed next to the classification accuracies in parentheses). The ratio of training to testing samples is set to 0.7/0.3, resulting in 12 test examples in each CV fold. To avoid biasing the results, we run the full atlas formation pipeline + CVT in each CV run. Consequently, the graph representations change slightly due to differences in vessel density atlas formation and partitioning. We additionally perform a set of permutation tests (cf. [**?**]) to assess whether a classification result could have been obtained by chance. In other words, the classifier could have picked up spurious patterns in the data that correlate with the labels. Our null-hypothesis, $\mathcal{H}_0$, is that the data is independent from its labels. The setup is as follows: for each CV split, we choose $P = 10^3$ random permutations of the training labels, train the classifier and compute an estimate of the classification error using the testing portion of the data. This gives an empirical estimate e_i of the error under $\mathcal{H}_0$ and allows the computation of a p-value estimate as $\hat{p} = (\#\{e_i < e^*\} + 1)/(P + 1)$, where e^* is the prediction error *without* label permutation. We *cannot* reject $\mathcal{H}_0$ if $\hat{p} > \alpha$, indicated by a † next to the classification result (for $\alpha = 0.05$). To study the impact of the vascular atlas tessellation granularity as well as the WLS kernel height h, we vary the number of CVT cells from 256 to 2048 and h from $h = 2$ to $h = 6$. Increasing h affects the kernel in that longer vessel paths are taken into account.

[4] Skull-stripping is done using `bet2` (FMRIB); Registration is done using 3D Slicer.

#Cells	$h=2$	$h=4$	$h=6$
256 $_{\lvert E_0^a\rvert=256}$	53 (1.00)	55 (1.00)	53 (1.00)
512 $_{\lvert E_0^a\rvert=512}$	63 (1.00)	58 (1.00)	58 (1.00)
1024 $_{\lvert E_0^a\rvert=1024}$	58 (1.00)	57 (1.00)	57 (1.00)
2048 $_{\lvert E_0^a\rvert=2048}$	63 (1.00)	67 (1.00)	65 (1.00)

(a) $\gamma^a : \mathcal{V} \to E_0^a$

#Cells	$h=2$	$h=4$	$h=6$
256 $_{\lvert E_0^b\rvert=12}$	$57^{\dagger}$ (0.65)	67 (0.79)	67 (0.80)
512 $_{\lvert E_0^b\rvert=11}$	70 (0.56)	70 (0.69)	70 (0.73)
1024 $_{\lvert E_0^b\rvert=11}$	$62^{\dagger}$ (0.51)	67 (0.65)	70 (0.69)
2048 $_{\lvert E_0^b\rvert=10}$	62 (0.50)	68 (0.59)	70 (0.63)

(b) $\gamma^b : \mathcal{V} \to E_0^b$

Table 1: Gender classification accuracies (in %) *without* label transfer for our two *default* vertex labeling strategies: **??** sequential labels and **??** vertex degrees.

Evaluation of Default Labeling Strategies. To establish a baseline, Tables **??** and **??** list the classification results for spatial-graphs with nodes labeled using γ^a and γ^b. In case of γ^a, the label alphabet $|E_0^a|$ grows linearly with the number of CVT cells, whereas in case of γ^b, $|E_0^b|$ is limited by the maximum vertex degree. The results of Table **??** indicate that sequential vertex labeling leads to accuracies often close to random choice. Essentially, all samples are chosen as support vectors, indicating a lack of structure in the data, most likely due to uninformative subtree patterns. Labeling by vertex degree, cf. Table **??**, produces slightly better result in general, yet the limited variety of the label alphabet also seems to restrict discriminative power.

Evaluation of Label Transfer. Tables **??** and **??** list the classification accuracies for functional graphs, *with* and *without* unique background labels. Overall, we notice an increase in classification performance, with top rates reaching up to 77% and 72%, respectively. Table **??** indicates that, without unique background labels, increasing the number of CVT cells does not necessarily improve performance. In fact, we notice a slight overall decrease in accuracy as we move from 256 to 1024 cells, presumably due to the increase in non-informative subtree patterns caused by an increasing number of non-unique background labels. A comparison to Table **??** supports this conclusion, since background labels are unique and increasing the number of cells from 256 to 1024 leads to consistent improvements. As we further increase the number of cells from 1024 to 2048, the high number of unique, but nonetheless arbitrary labels (665 on avg., cf. Table **??**) seems to confound the positive effect of label transfer.

Discussion. While our results indicate gender-related architectural differences in cerebral vasculature, consistent with [**?**], several points are worth noting. First, the resolution of MRA is limited to capture blood vessels of size (in diameter) equal to the voxel dimension. This essentially prevents us from quantifying capillaries and determining exact vessel function. Second, our approach is designed to quantify similarity of blood vessel graphs at a *global* level, to determine "*if*" there are differences. Answering the question of "*where*" differences occur requires a different approach and is left for future research. Third, while the classification accuracies are lower than originally reported in [**?**], we remark that our results are obtained via *cross-validation*. In fact, some splits do lead to perfect

#Cells	$h = 2$	$h = 4$	$h = 6$
256 $_{\lvert E_0^d\rvert = S+132}$	60† (1.00)	62 (1.00)	62 (1.00)
512 $_{\lvert E_0^d\rvert = S+226}$	70 (0.93)	70 (0.94)	70 (0.93)
1024 $_{\lvert E_0^d\rvert = S+386}$	75 (0.91)	77 (0.92)	77 (0.92)
2048 $_{\lvert E_0^d\rvert = S+665}$	73 (0.87)	73 (0.88)	73 (0.89)

(a) $\gamma^d : \mathcal{V} \to E_0^d$

#Cells	$h = 2$	$h = 4$	$h = 6$
256 $_{\lvert E_0^c\rvert = S}$	68 (0.85)	72 (0.87)	72 (0.87)
512 $_{\lvert E_0^c\rvert = S}$	72 (0.74)	70 (0.79)	72 (0.79)
1024 $_{\lvert E_0^c\rvert = S}$	63 (0.79)	68 (0.81)	70 (0.84)
2048 $_{\lvert E_0^c\rvert = S}$	68 (0.71)	68 (0.74)	72 (0.74)

(b) $\gamma^c : \mathcal{V} \to E_0^c$

Table 2: Gender classification accuracies (in %) using label transfer **??** with and **??** without unique background labels. Numbers in red show the cardinality of the set of added background labels, i.e., $|\mathcal{B}|$ and $S = 155$ (#segmented structures, see [**?**]).

classification. Finally, most permutation test outcomes actually support the rejection of $\mathcal{H}_0$ at the 5% significance level. At 1%, the overall picture remains the same, however, some results of the default labeling strategies (e.g., in Table **??**) are not statistically significant. In summary, the presented approach facilitates the study of vascular differences among populations of healthy individuals, and may facilitate the study of cerebrovascular differences between individuals and populations involving selected mental disorders.

Acknowledgments. This work was funded in part by the following grants: NIH/NCI 1R01CA138419, 1R43CA165621; NIH/NIMH 2P41EB002025-26A1, 1R01MH091645-01A1. The software is available as part of `TubeTK` (http://www.tubetk.org); the full dataset is available from http://midas3.kitware.com/midas/folder/8051.