bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Digital 3D Brain MRI Arterial Territories Atlas

Chin-Fu Liu[1], Johnny Hsu[2], Xin Xu[2], Nicholas G Kim[3], Shannon M. Sheppard[4,5], Erin L. Meier[4,6], Michael Miller[1,7], Argye E. Hillis[4,8], Andreia V. Faria[2*].

[1]Center for Imaging Science, Institute for Computational Medicine, Johns Hopkins University; Department of Biomedical Engineering, Johns Hopkins University.

[2]Department of Radiology, School of Medicine, Johns Hopkins University.

[3]Department of Neurosciences, Johns Hopkins University.

[4]Department of Neurology, School of Medicine, Johns Hopkins University.

[5]Department of Communication Sciences & Disorders, Chapman University.

[6]Department of Communication Sciences & Disorders, Northeastern University.

[7]Kavli Neuroscience Discovery Institute, Johns Hopkins University.

[8]Department of Physical Medicine & Rehabilitation, and Department of Cognitive Science, Johns Hopkins University.

***Corresponding Author**: Andreia V. Faria, MD, PhD

Associate Professor of Radiology, Johns Hopkins University.

217B Traylor Bldg, 720 Rutland Ave, Baltimore, MD 21205, USA

Email: afaria1@jhmi.edu

Phone: +1 410-955-4215

Figures: 5 **Tables**: 1

Keywords: atlas, arterial, vascular territories, MRI, stroke

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ABSTRACT

The locus and extent of brain damage in the event of vascular insult can be quantitatively established quickly and easily with vascular atlases. Although highly anticipated by clinicians and clinical researchers, no digital MRI arterial atlas is readily available for automated data analyses. We created a digital arterial territory atlas based on lesion distributions in 1,298 patients with acute stroke. The lesions were manually traced in the diffusion-weighted MRIs, binary stroke masks were mapped to a common space, probability maps of lesions were generated and the boundaries for each arterial territory was defined based on the ratio between probabilistic maps. The atlas contains the definition of four major supra- and infra-tentorial arterial territories: Anterior, Middle, Posterior Cerebral Arteries and Vertebro-Basilar, and sub-territories (thalamoperforating, lenticulostriate, basilar and cerebellar arterial territories), in two hierarchical levels. This study provides the first publicly-available, digital, 3D deformable atlas of arterial brain territories, which may serve as a valuable resource for large-scale, reproducible processing and analysis of brain MRIs of patients with stroke and other conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

BACKGROUND AND SUMMARY

Vascular atlases provide means to quickly establish the locus and extension of brain damage in the event of vascular insult [15], can assist in stroke treatment planning and determining recovery prognosis [2,3], and can be used to inform future clinical trials of relevant scores (e.g., ASPECTS [4]). Research-wise, atlas-based analysis increases the statistical power and greatly improves the feasibility and reproducibility of lesion-based studies that usually require large scale data. Prior to the advent and use of modern neuroimaging techniques, vascular atlases were derived mainly from post-mortem injection studies [6-8] and depicted schematic representations of the vascular territories. Such studies typically included single cases or very small samples, and thus, these atlases did not account for the high inter-individual variability in vascular distributions. Two-dimensional, manually-drawn atlases [9-13] also suffer from a lack of flexibility; if the template slices provided in the atlas do not match the slice thickness or orientation of acquired clinical neuroimages, the atlas utility is diminished [7,10,14-16]. Even with well-aligned slices, printed templates lack precision and do not allow for quantitative assessment of vascular territories in a single patient.

Contemporary, digital atlases often derived from *in vivo* perfusion studies in healthy individuals [16-18] or analysis of stroke lesions [14,15,19-21] overcome the aforementioned limitations and can provide a fast and objective measure of vascular damage in individual patients. In particular, digital atlases that incorporate data from a large number of stroke patients are likely most representative of the vascular anatomy of older adults who are most likely to experience infarct. Recently, Kim et al. [20] used a large dataset of stroke patients from academic centers in Korea to map supratentorial cerebral arterial territories. Despite its strengths, infratentorial and deep perforating arterial territories were not defined, even though approximately 20% of all strokes occur in both the infratentorial [26] and deep perforating arterial zones [27]. Furthermore, a publicly accessible electronic version of the atlas is not available, thus reducing applicability. Wang et al.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[28] later used MRIs of 458 patients to create a "stroke atlas" using voxel-wise data-driven hierarchical density clustering. Although this atlas represents major and end-arterial supply territories, the difficulties in accessing the stability of small clusters and the lack of unitary-defined boundaries reduces applicability for individual brain mapping.

To remedy the limitations of existing vascular atlases, we created a novel whole-brain arterial territory atlas based on lesion distributions in 1,298 patients with acute stroke. Our atlas covers supra- and infra-tentorial regions, includes straightforward and robustly-defined watershed border zones, and contains hierarchical segmentation levels created by a fusion of vascular and classical anatomical criteria. The atlas was developed using diffusion weighted MRIs (DWIs) of patients with acute and early subacute ischemic strokes, expert annotation (lesion delineation), and metadata (radiological evaluation) of a large and heterogeneous sample, both in biological terms (e.g.., lesion pattern, population phenotypes) and technical terms (e.g., differences in MRI protocols and scanners). The biological and technical heterogeneity of the sample is advantageous in that it increases the generalization and robustness of the atlas, and consequently its utility. This first publicly-available, digital, 3D deformable atlas can be readily used by the clinical and research communities, enabling to explore large-scale data automatically and with high levels of reproducibility.

METHODS

Cohort

This study included Magnetic Resonance Images (MRIs) of patients admitted to the Johns Hopkins Stroke Center with the clinical diagnosis of acute stroke, between 2009 and 2019. Baseline MRIs adequate for clinical analysis and evidence of ischemic stroke in DWI were included. From the 1,878 patients identified under these criteria, 580 individuals were excluded because their stroke lesion was 1) characteristic of embolic etiology, 2) bilateral, 3) multifocal / involving more than one major arterial territory, 4) exclusively within "watershed" areas, and/or 5)

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

hemorrhagic. Ultimately, 1,298 MRIs were included with acute or early subacute ischemic strokes affecting exclusively one major arterial territory (anterior, middle, and posterior cerebral, or vertebro-basilar arteries).

The characteristics of the cohort are summarized in Table 1. The distribution of strokes according to arterial territories (MCA > PCA > VB > ACA) and the demographic and clinical characteristics reflect the general population of stroke patients. We note that our population includes a higher percentage of Black, and lower percentage of Hispanic/Latinx and Asian patients than many urban stroke centers, therefore a regional bias may exist. The vast majority of scans were performed more than 2 hours after the onset of symptoms, reducing the odds of significant further change in the lesion volume [25].

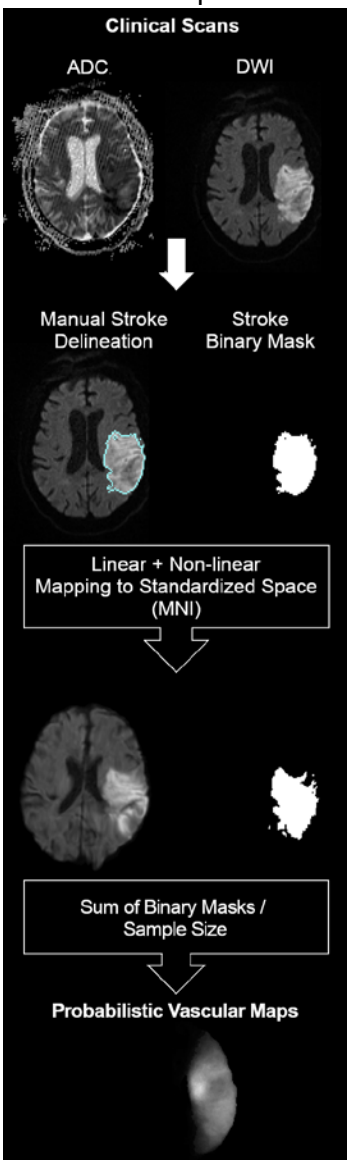


Image analysis

MRIs were obtained on seven scanners from four different vendors, in different magnetic fields, with dozens of protocols. The DWIs had high in plane (axial) resolution (1x1mm, or less), and typical clinical large slice thickness (ranging from 3 to 6 mm). The delineation of the stroke core was defined in the DWI by 2 experienced evaluators and revised by a neuroradiologist until reaching a final decision by consensus. Details are described in Technical Validation. The binary stroke masks (ischemic lesion = 1; brain and background = 0) were then used to create probabilistic maps, as shown in Figure 1.

Creating probabilistic maps and atlases requires transforming images to a common space for consistency among image coordinates and biological structures. Clinical stroke images offer three additional challenges to image mapping: 1) the typical high slice

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

thickness, often associated with large rotations out-of-plane, 2) the presence of drastic changes in image morphology and intensity, caused by the stroke as well as associated chronic conditions (e.g., white matter microvascular chronic lesions), and 3) the variable and often remarkable degree of brain atrophy in elderly participants who primarily constitute the stroke population. As linear deformations are not enough to accurately map internal brain shapes, and high elastic transformations (e.g., diffeomorphic) are very sensitive to focal abnormalities in morphology and contrast, we applied sequential steps of linear and non-linear deformation, using Broccoli [22]. To minimize the effects of the acute stroke, we used the "B0" images (less affected by the acute lesion intensity) for mapping. As a template, we used the JHU_SS_MNI_B0 [23], a subject in standard MNI coordinates. The optimization of the parameters and other details of the image mapping are described in the Technical Validation.

Probabilistic maps

The cohort was divided into four groups based on the major arterial territory that was exclusively affected in each patient, according to radiological classification: 1) Anterior Cerebral Artery (ACA), 2) Middle Cerebral Artery (MCA), 3) Posterior Cerebral Artery (PCA), or 4) Vertebro-Basilar artery (VB). The MCA infarcts affected the lateral lenticulostriate territory, or the "lobar" (cortex and adjacent subcortical white matter) MCA territory, or both; the PCA infarcts affected the posterior and anterior thalamoperforating territory, or the "lobar" PCA territory, or a combination of both; the ACA infarcts affected the medial lenticulostriate territory, or the "lobar" ACA territory, or both; the VB infarcts affected the inferior or superior cerebellar areas, or the basilar territory, or a combination of both.

Probabilistic maps were generated by summing the stroke binary mask in the common space and dividing by the sample size for each of these territories. Each subject has an annotated pair (s_i, X_i), where s_i is the group label, $s_i \in$ {ACA, MCA, PCA, VB}, and X_i is the stroke mask. Then, we can define the probabilistic map for MCA region as following,

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

$$pMCA = \frac{\sum_{Vs_i=MCA} X_i}{N_{MCA}}$$

, where N_{MCA} is the sample size in the MCA group. Similarly, we compute *p*ACA, *p*PCA, and *p*VB. These probabilistic maps can be considered as an estimate for the likelihood of voxels belonging to a certain territory. The more samples in a group, the more accurate the estimate is.

For the "border zone" between the major cerebral arteries (ACA, MCA, PCA), we considered each voxel that had a probability greater than zero to belong to two arterial territories. By dividing the probabilities, we estimated the likelihood that voxels in the border zone belong to one or another major arterial territory. For example, a ratio of *p*MCA/*p*PCA > 1 indicates that the given "border zone" voxel is more often affected in MCA strokes than in PCA strokes while *p*MCA/*p*PCA <1 indicates the opposite. We acknowledge the sample imbalance (the ACA group is smaller than the others) as a possible bias for the conditional probability ratio. Yet, because this imbalance reflects the regular distribution of strokes, and given the lack of adequate means to measure the possible bias, we adopted the simplest estimation.

Definition of arterial territories

Our ultimate goal was to create a 3D deformable atlas of the major arterial territories, which requires sharp boundaries between neighbor regions. After adding the major probabilistic maps (ACA, MCA, PCA, VB), the vast majority of voxels belonged to an exclusive vascular territory and were attributed an integer intensity that identified that territory. When a voxel belonged to more than one vascular territory (which occurred in the border zone) it was attributed to the territory with the highest probability. The resulting atlas had a relatively coarse contour and was smoothed by a kernel filter followed by manual editing by an expert neuroradiologist. The manual adjusts were "cosmetic" (no more than 3 voxels, in plane, at 1x1x1 mm resolution).

The same procedure was applied to define the sub-territories of lateral lenticulostriate (within MCA), thalamoperforating (within PCA), and anterior and posterior cerebellar and basilar

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(within VB) arteries. The "lobar" sub-areas of MCA (frontal, parietal, temporal, insular, occipital) and PCA (occipital, temporal), were defined based on classic anatomical landmarks (gyri, as defined in our previous atlas [24]). In less stroke-prone territories, such as medial lenticulostriate and anterior thalamopeforating (that have less than 10 cases of exclusive strokes each), the probabilistic maps would be less reliable as they derive from small samples. For those territories, we chose to rely on anatomical knowledge derived from prior studies; future studies can refine the boundaries using probabilistic maps once we accumulate additional cases. The medial lenticulostriate territory, within ACA, corresponds to the topography of corpus callosum. The anterior thalamoperforating territory, within PCA, corresponds to the topography of the mesial temporal area and hippocampus. As a result, we created a hierarchical (2-level) brain atlas of arterial territories (Figure 2).

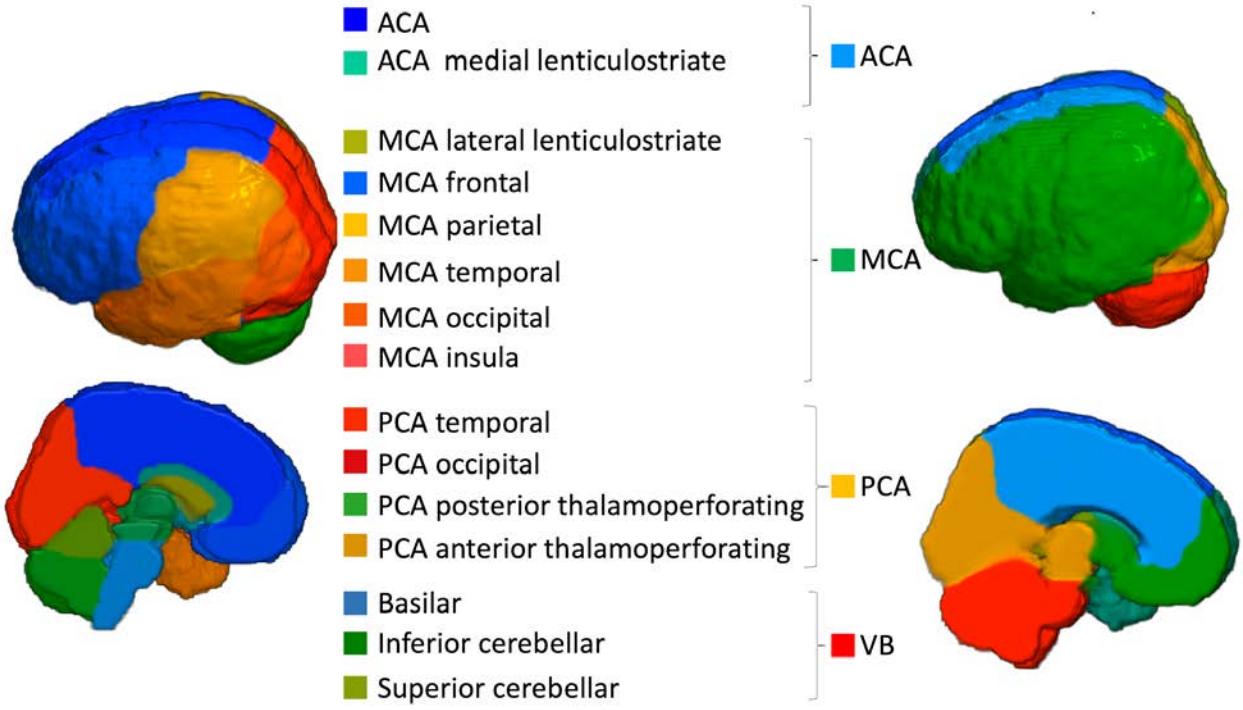
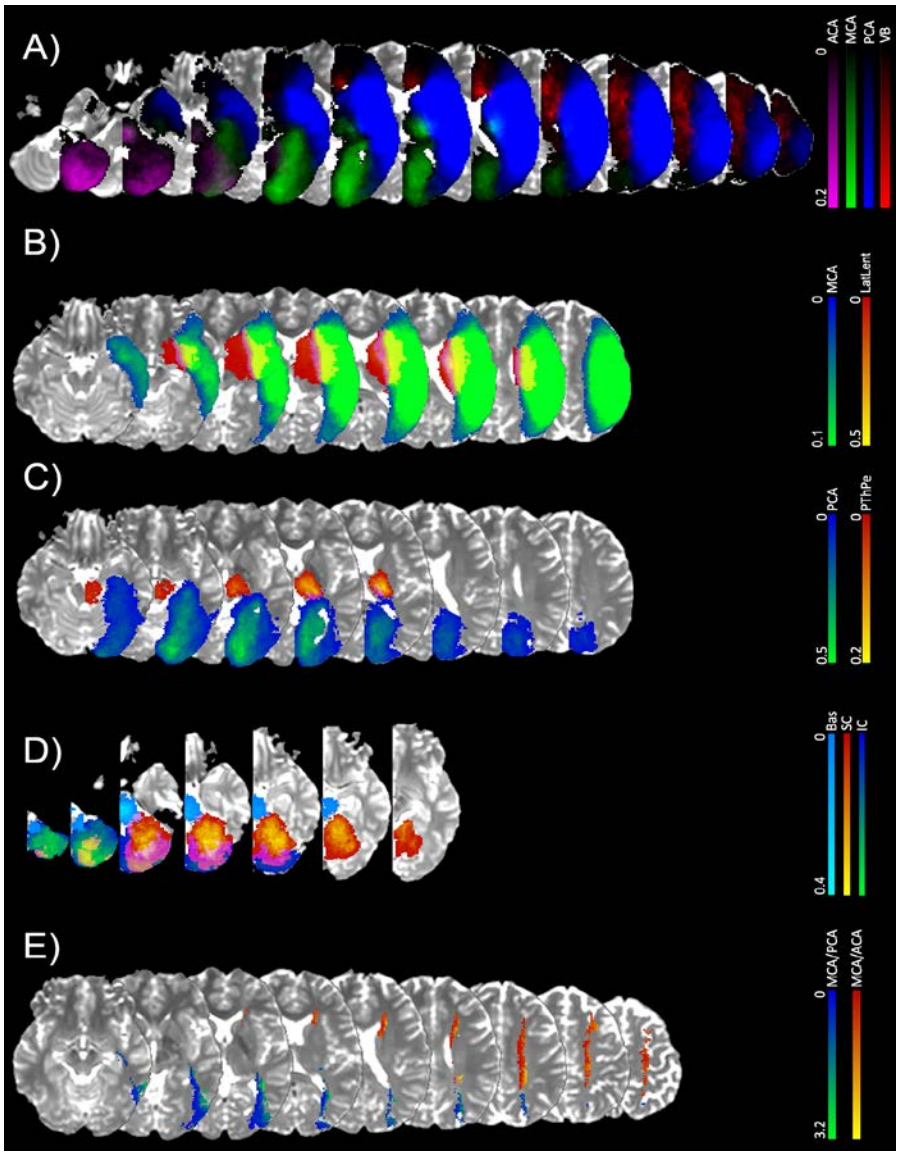


Figure 2: 3D reconstruction of the arterial atlas. The right columns represent the major vascular territories; the left columns show subdivisions of biological importance, based on probabilistic maps and / or anatomical landmarks.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Arterial Atlas

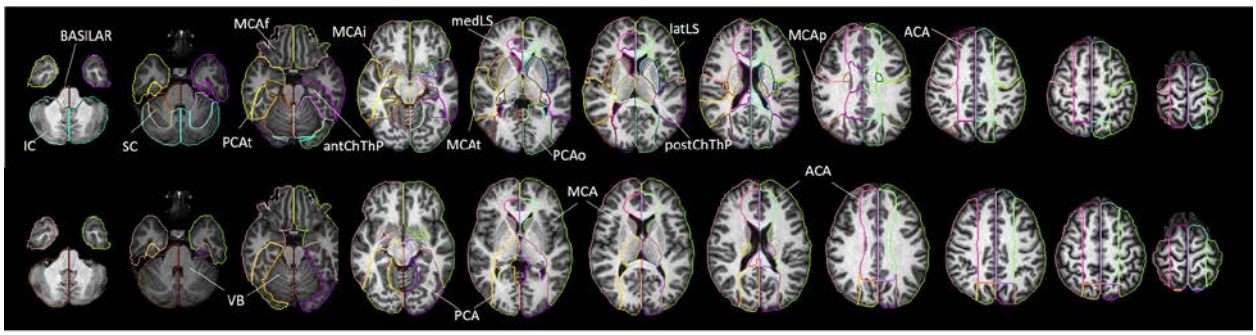
Figure 3A shows the probabilistic maps of the four major arterial territories overlaid in the template. Figures 3B-D show the probabilistic maps of strokes in "sublevels" of MCA (lateral lenticulostriates, Fig 3B), PCA (posterior thalamoperforating, Fig 3C), and VB (anterior and posterior cerebellar, and basilar; Fig 3D). The sample size employed for each probabilistic map is summarized in Venn Diagrams in Table 1. Figure 3E shows the border zone between the MCA-PCA and the MCA-ACA territories, represented by the ratio between the probabilistic maps. Values higher than 1 indicate that the given "border zone" voxel is more often affected by MCA strokes than by PCA or ACA strokes.



bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 3: A) Probabilistic maps of the four major arterial territories: anterior cerebral artery (ACA), medium cerebral artery (MCA), posterior cerebral artery (PCA), and vertebra-basilar (VB). C-D) Probabilistic maps of strokes in "sublevels" of the major territories, with lesion affecting: C) the territories of lateral lenticulostriate arteries exclusively, or the "lobar" (cortex and adjacent white matter) MCA; D) the territories of posterior thalamoperforating arteries exclusively, or the "lobar" (cortex and adjacent white matter) PCA; D) the vertebro-basilar territory, composed of the anterior cerebellar, posterior cerebellar and basilar arteries. Color scales for each territory indicate the frequency of lesions. E) Border zone between MCA-PCA (blue-green) and MCA-ACA (red-yellow). The color intensity represents the ratio between the respective probabilistic maps. Values higher than 1 indicate that the given "border zone" voxel is more often affected by MCA strokes than by PCA or ACA strokes.

Figure 4 is the 2D version of Figure 2 and shows our arterial territory atlas overlaid in T1-weighted axial slices of a brain in standardized MNI space. The bottom row shows the major arterial territories (ACA, MCA, PCA, VB) and is the result of the combination of the respective probabilistic maps (Figure 3A) followed by categorical definition of the border zones plus smoothing and manual cosmetic editing. The top row represents the sublevels defined in the major territories by probabilistic maps (lateral lenticulostriate, posterior thalamoperforating, anterior and posterior cerebellar, and basilar strokes; as in Figures 3B-D) or by prior knowledge of structural anatomy (medial lenticulostriate, anterior thalamoperforating, and "lobar" segments).



bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 4: Digital arterial atlas overlaid on axial T1WIs of the template in standardized space (MNI). At the bottom, the major arterial territories: anterior (ACA), middle (MCA), and posterior (PCA) cerebral arteries, and vertebro-basilar (VB). At the top, sub segments of ACA (frontal (ACAf) and medial lenticulostriate (medLS)), MCA (frontal (MCAf), parietal (MCAp), temporal (MCAt), occipital (MCAo), insula (MCAi) and lateral lenticulostriate (latLS)), PCA (temporal (PCAt), occipital (PCAo), Posterior Choroidal and Thalamoperforating (postChThp), Anterior Choroidal and Thalamoperforating (antChThp)), and VB (inferior cerebellar (IC), superior cerebellar (SC), and basilar). See 3D reconstruction in Figure 2

DATA RECORDS

The atlas of the arterial territories is available in NITRC (https://www.nitrc.org/projects/arterialatlas). Any questions, concerns, or comments related to the atlas are welcome and can be publicly reported, so users are aware and we can resolve any issues in a timely manner. The images are in Montreal Neurological Institute (MNI) coordinates and the data format is “Nifti”, as recommended by the Brain Imaging Data Structure (BIDS). The description of the files is as follows:

- ArterialAtlas.nii: Image defining 30 arterial territories and ventricles. The intensities of the parcels correspond to their labels IDs, listed in the lookup table “ArterialAtlasLabels.txt”, that accompanies the documentation. If visualized in 3D, the ArterialAtlas.nii corresponds to the left of Figure 2 of this paper; if visualized in 2D over an anatomical slide, it corresponds the top row of Figure 4.

- ArterialAtlas_level2.nii: The combination of ArterialAtlas.nii parcels in 4 major territories (ACA, MCA, PCA, VB). Again, the intensities of the parcels correspond to their labels, listed in the ArterialAtlasLabels.txt. If visualized in 3D, the ArterialAtlas_level2.nii corresponds to right

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

column of Figure 2 of this paper; if visualized in 2D over an anatomical slide, it corresponds the bottom row of Figure 4.

- ArterialAtlasLabels.txt: is the "lookup table". It contains the labels (descriptive and acronyms) for the regions defined in ArterialAtlas.nii and ArterialAtlas_level2.nii. The intensity of the region, in the images, corresponds to the label ID, in the lookup table.

TECHNICAL VALIDATION

Lesion Delineation

Although there is no perfect method for defining the stroke core, we chose to use DWI and Apparent Diffusion Coefficient maps (ADC), based on the fact that DWI is the most informative and most common sequence performed for acute stroke detection. Likewise, prior acute stroke studies and trials have defined lesions in DWI, so our data will be broadly comparable to those investigations. The delineation of the stroke core was defined in the DWI by 2 experienced evaluators and revised by a neuroradiologist until reaching a final decision by consensus. The evaluators looked for hyper intensities in DWI and / or hypo intensities (<30% average brain intensity) in ADC. Additional contrasts were used to rule out chronic lesions or microvascular white matter disease. A "seed growing" tool in ROIEditor (MRIStudio.org) was often used to achieve a broad segmentation, followed by manual adjustments. In a subset of 130 training cases (random selected 10% of cases) traced twice in a two-week interval, the index of agreement, DICE, was 0.76±0.14, inter-evaluator, and 0.8±0.13, intra-evaluator. Although this is a satisfactory level of agreement (Dice values range from 0 to 1; where 1 indicates total agreement), it demonstrates that human segmentation has suboptimal reproducibility, even when performed by experienced, trained evaluators, reinforcing the importance of the consensus agreement and the multiple revisions performed.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Image Mapping

The parameters for the non-linear deformation were chosen to maximize the accuracy on mapping structures such as the ventricles and the brain contour. For the linear registration we used the default setting of the affine transformation in BROCCOLI RegisterTwoVolumes function (https://github.com/wanderine/BROCCOLI/blob/master/code/Bash_Wrapper/RegisterTwoVolumes.cpp). As for the non-linear transformation, there are two important parameters, sigma and iterationsnonlinear (N_iter), to determine the smoothness of the displacement field and the number of iteration steps. To find the optimal sigma and N_iter, we "hyperparameter-searched" all the combinations of sigma in [1,3,5] and N_iter in [10, 15, 20] to register a set of 470 subjects with "not visible" strokes to the template. We evaluated the deformation results for each combination of parameters across all subjects, as follows:

a) Three regions of 5 voxels bandwidth were defined in the template: the outside strip of the brain mask (OSBM), the inside strip of the brain mask (ISBM), and the outside strip of the lateral ventricles (OSLV), as shown in Figure 5, top.

b) The ratio of the mis-deformed voxels in OSBM, ISBM and OSLV for each subject, defined as rOSBM, rISBM, and rOSLV, was calculated as follows:

$$r_{OSBM} = \frac{the\ number\ of\ the\ deformed\ B0\ voxels\ whose\ intensity\ is\ larger\ than\ 0\ in\ OSBM}{the\ number\ of\ voxels\ in\ OSBM}$$

$$r_{ISBM} = \frac{the\ number\ of\ the\ deformed\ B0\ voxels\ whose\ intensity\ is\ \ 0\ in\ ISBM}{the\ number\ of\ voxels\ in\ ISBM}$$

$$r_{OSLV} = \frac{the\ number\ of\ the\ deformed\ B0\ voxels\ whose\ intensity\ is\ larger\ than\ \lambda\ in\ OSLV}{the\ number\ of\ voxels\ in\ OSLV}$$

$$\lambda = \mu_{deformed\ B0\ in\ LV} - 0.5 * \sigma_{deformed\ B0\ in\ LV}$$

rOSBM indicates the ratio of the deformed B0 voxels aligned outside the template brain mask and rISBM indicates the ratio of the background voxels aligned inside the template brain mask.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

High rOSBM or rISBM mean the global deformation on B0 boundary is not good enough. rOSLV indicates the ratio of voxels from a subject's deformed ventricles that exist outside the template lateral ventricle boundaries. High rOSLV is common in this population since aged subjects' lateral ventricles are usually larger than the template's lateral ventricles and indicated the need for local deformation. Figure 5, bottom, shows the statistics of rOSBM, rISBM, and rOSLV for all parameter's combinations tested.

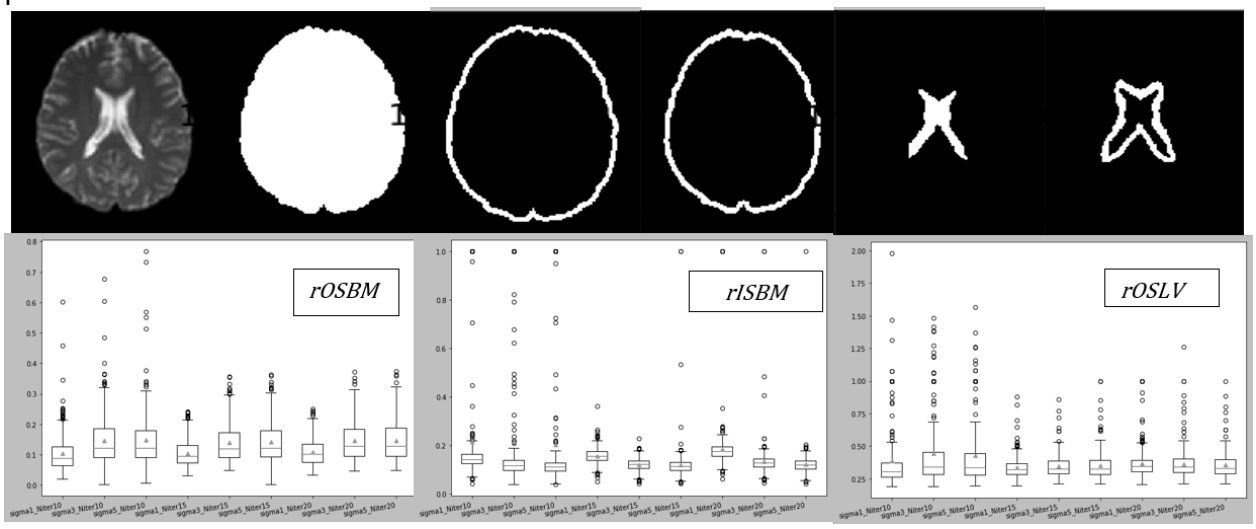


Figure 5: Optimization of parameters for the non-linear brain deformation. At the top, areas used for optimization, from left to right: the brain template, the brain mask, the outside strip of the brain mask (OSBM), the inside strip of the brain mask (ISBM), the lateral ventricles mask, and the outside strip of the lateral ventricles (OSLV). B) Boxplots of ratios of OSBM, ISBM, and OSLV for all parameter's combinations tested.

c) The optimal parameters were sigma=3 and N_iter=15. The registration with these parameters achieved better global deformation around the brain mask and comparable (to sigma=1 and N_iter=15) local deformation around the lateral ventricles.

Visual quality control followed parameter optimization. The deformation matrix of each image was then applied to the respective stroke mask using near neighbor interpolation to preserve the binary nature. Images of patients with strokes in the right hemisphere were flipped along the x-axis so that all the stroke masks were considered in the left hemisphere in the standard space.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

REFERENCES

1. Hennerici M, Daffertshofer M, Jakobs L. Failure to identify cerebral infarct mechanisms from topography of vascular territory lesions. *AJNR American journal of neuroradiology.* 1998;19(6):1067-1074.
2. Ay H, Benner T, Arsava EM, et al. A computerized algorithm for etiologic classification of ischemic stroke: the Causative Classification of Stroke System. *Stroke; a journal of cerebral circulation.* 2007;38(11):2979-2984.
3. Hillis AE, Gold L, Kannan V, et al. Site of the ischemic penumbra as a predictor of potential for recovery of functions. *Neurology.* 2008;71(3):184-189.
4. Albers GW, Wald MJ, Mlynash M, et al. Automated Calculation of Alberta Stroke Program Early CT Score: Validation in Patients With Large Hemispheric Infarct. *Stroke; a journal of cerebral circulation.* 2019;50(11):3277-3279.
5. Meier EL, Breining BL, Sheppard SM, et al. White Matter Hyperintensities Contribute to Language Deficits in Primary Progressive Aphasia. *Cognitive and behavioral neurology : official journal of the Society for Behavioral and Cognitive Neurology.* 2020;33(3):179-191.
6. Tatu L, Moulin T, Bogousslavsky J, Duvernoy H. Arterial territories of human brain: brainstem and cerebellum. *Neurology.* 1996;47(5):1125-1135.
7. Tatu L, Moulin T, Bogousslavsky J, Duvernoy H. Arterial territories of the human brain: cerebral hemispheres. *Neurology.* 1998;50(6):1699-1708.
8. van der Zwan A, Hillen B, Tulleken CA, Dujovny M, Dragovic L. Variability of the territories of the major cerebral arteries. *Journal of neurosurgery.* 1992;77(6):927-940.
9. Berman SA, Hayman LA, Hinck VC. Correlation of CT cerebral vascular territories with function: I. Anterior cerebral artery. *AJR American journal of roentgenology.* 1980;135(2):253-257.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

10. Berman SA, Hayman LA, Hinck VC. Correlation of CT cerebral vascular territories with function: 3. Middle cerebral artery. *AJR American journal of roentgenology.* 1984;142(5):1035-1040.

11. Bogousslavsky J, Regli F. Unilateral watershed cerebral infarcts. *Neurology.* 1986;36(3):373-377.

12. Damasio H. A computed tomographic guide to the identification of cerebral vascular territories. *Archives of neurology.* 1983;40(3):138-142.

13. Hayman LA, Berman SA, Hinck VC. Correlation of CT cerebral vascular territories with function: II. Posterior cerebral artery. *AJR American journal of roentgenology.* 1981;137(1):13-19.

14. Nowinski WL, Qian G, Kirgaval Nagaraja BP, et al. Analysis of ischemic stroke MR images by means of brain atlases of anatomy and blood supply territories. *Academic radiology.* 2006;13(8):1025-1034.

15. Phan TG, Fong AC, Donnan G, Reutens DC. Digital map of posterior cerebral artery infarcts associated with posterior cerebral artery trunk and branch occlusion. *Stroke; a journal of cerebral circulation.* 2007;38(6):1805-1811.

16. van Laar PJ, Hendrikse J, Golay X, Lu H, van Osch MJ, van der Grond J. In vivo flow territory mapping of major brain feeding arteries. *NeuroImage.* 2006;29(1):136-144.

17. Kansagra AP, Wong EC. Mapping of vertebral artery perfusion territories using arterial spin labeling MRI. *Journal of magnetic resonance imaging : JMRI.* 2008;28(3):762-766.

18. Lee JS, Lee DS, Kim YK, et al. Probabilistic map of blood flow distribution in the brain from the internal carotid artery. *NeuroImage.* 2004;23(4):1422-1431.

19. Kim DE, Jang J, Schellingerhout D, et al. Supratentorial Cerebral Arterial Territories for Computed Tomograms: A Mapping Study in 1160 Large Artery Infarcts. *Scientific reports.* 2019;9(1):11708.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442478; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

20. Kim DE, Park JH, Schellingerhout D, et al. Mapping the Supratentorial Cerebral Arterial Territories Using 1160 Large Artery Infarcts. *JAMA neurology.* 2019;76(1):72-80.

21. Phan TG, Fong AC, Donnan GA, Srikanth V, Reutens DC. Digital probabilistic atlas of the border region between the middle and posterior cerebral arteries. *Cerebrovascular diseases (Basel, Switzerland).* 2009;27(6):529-536.

22. Eklund A, Dufort P, Villani M, Laconte S. BROCCOLI: Software for fast fMRI analysis on many-core CPUs and GPUs. *Frontiers in neuroinformatics.* 2014;8:24.

23. Mori S, Oishi K, Jiang H, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage.* 2008;40(2):570-582.

24. Oishi K, Zilles K, Amunts K, et al. Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. *NeuroImage.* 2008;43(3):447-457.

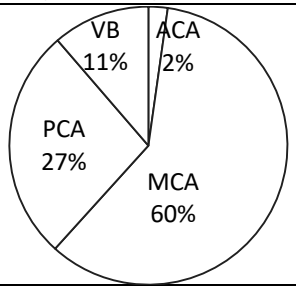
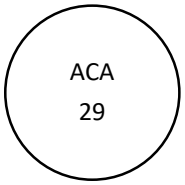
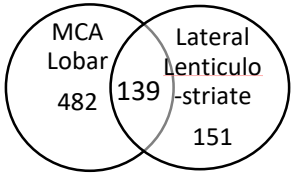
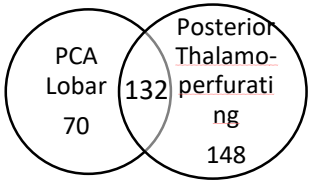
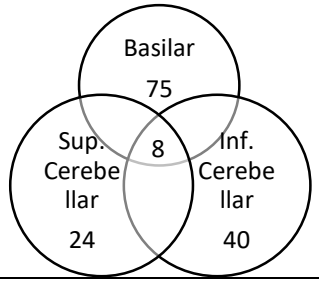
25. Wheeler HM, Mlynash M, Inoue M, et al. The growth rate of early DWI lesions is highly variable and associated with penumbral salvage and clinical outcomes following endovascular reperfusion. *International journal of stroke : official journal of the International Stroke Society.* 2015;10(5):723-729.

26. Lekic T, Krafft PR, Coats JS, Obenaus A, Tang J, Zhang JH. Infratentorial strokes for posterior circulation folks: clinical correlations for current translational therapeutics. *Translational stroke research.* 2011;2(2):144-151.

27. Sudlow CL, Warlow CP. Comparable studies of the incidence of stroke and its pathological types: results from an international collaboration. International Stroke Incidence Collaboration. *Stroke; a journal of cerebral circulation.* 1997;28(3):491-499.

28. Wang Y, Juliano JM, Liew SL, McKinney AM, Payabvash S. Stroke atlas of the brain: Voxel-wise density-based clustering of infarct lesions topographic distribution. *NeuroImage Clinical.* 2019;24:101981.

Table 1. Demographic and clinical characteristics of the cohort, and imaging and lesion characteristics.

	Total (n = 1298)	**ACA (n = 29)**	**MCA (n = 772)**	**PCA (n = 350)**	**VB (n = 147)**
					
Age in years	62.3 ± 14.6 [99 , 18]	65.8 ± 15.7 [85 , 22]	63.0 ± 15.0 [99 , 18]	61.3 ± 13.6 [96 , 20]	60.5 ± 14.4 [93 , 21]
Sex					
Male	686 (52.8%)	13 (44.8%)	400 (51.8%)	191 (54.6%)	82 (55.8%)
Female	612 (47.2%)	16 (55.2%)	372 (48.2%)	159 (45.4%)	65 (44.2%)
Race					
African American	600 (63.6%)	16 (69.5%)	320 (58.6%)	186 (70%)	78 (63.4%)
Caucasian	334 (34.8%)	7 (30.5%)	212 (38.8%)	74 (28%)	41 (33.3%)
Asian	24 (0.2%)	0	14 (2.5%)	6 (0.2%)	4 (0.3%)
Missing Data	340	6	226	84	24
NIHSS	5.4 ± 5.9 [31 , 0]; 551	3.4 ± 5.4 [19 , 0]; 10	* 7.1 ± 6.7 [31 , 0]; 329	3 ± 2.7 [14 , 0]; 146	3 ± 3.8 [22 , 0]; 66
Days of Hospital Stay	5.2 ± 5.5 [39 , 0]; 416	4.8 ± 3.3 [13 , 1]; 8	5.8 ± 5.7 [39 , 0]; 261	* 3.6 ± 4.3 [38 , 0]; 106	5.5 ± 6.5 [34 , 0]; 41
Onset to MRI in hours					
<2	79 (9.5%)	4 (19.1%)	52 (10.9%)	18 (7.9%)	5 (4.8%)
2-6	170 (20.5%)	5 (23.8%)	94 (19.7%)	42 (18.5%)	29 (27.8%)
6-12	166 (20.0%)	6 (28.6%)	88 (18.5%)	55 (24.2%)	17 (16.3%)
12-24	346 (41.7%)	4 (19.1%)	204 (42.7%)	95 (41.9%)	43 (41.3%)

>24	68 (8.2%);	2 (9.5%);	39 (8.2%);	17 (7.5%)	10 (9.6%);
Missing Data	469	8	295	123	43
MRI Magnetic Field (1.5T / 3T)	869 (66.9%) / 429 (33.1%)	16 (55.2%) / 13 (44.8%)	524 (67.9%) / 248 (32.1%)	231 (66%) / 119 (34%)	98 (66.7%) / 49 (33.3%)
Hemisphere					
Left	681 (52.5%)	* 23 (79.3%)	390 (50.5%)	183 (52.3%)	85 (57.8%)
Right	617 (47.5%)	6 (20.7%)	382 (49.5%)	167 (47.7%)	62 (42.2%)
Stroke volume (mm^3)	21636.2 ± 47141.7 [403460 , 28]	7629.9 ± 10945.9 [44355 , 150]	* 31841.4 ± 57883.4 [403460 , 60]	6504.2 ± 14112.1 [131615 , 28]	6833.0 ± 13810.5 [70940 , 84]
Relative Stroke Volume (% of arterial territory / % of brain volume)	8.18% / 3.2%	6.80% / 1.25%	10.11% / 4.71%	4.21% / 0.93%	7.73% / 0.97%

* Statistical difference at *p*-value < 0.05. Continued variables showed as mean ± standard deviation [maximum , minimum]; missing data