MarkVCID2 MRI Arteriosclerosis (ARTS)
Biomarker Kit Protocol

1. Brief description of the biomarker kit

The ARTS biomarker kit is a software package that takes as input MRI DICOM data and outputs a score representing the likelihood a human subject suffers from arteriolosclerosis. It requires the collection of MPRAGE, FLAIR and DTI data. It also requires the collection of simple demographic information, i.e. age and sex. Once this information has been entered into the software of the biomarker, the software outputs a score.

2. Summary of kit instrumental validation results

A. Inter-rater reliability: The goal of this study was to assess the consistency of ARTS scores when different sites (raters) run ARTS on the same raw data. All 7 MarkVCID sites ran ARTS on the same raw data from 20 MarkVCID participants and shared the scores with us. No data were excluded. Figure 1 shows Bland-Altman plots and scatter plots of ARTS scores for all pairs of sites. **In the scatter plots, all ARTS scores are essentially on the identity line (y=x) indicating excellent agreement in scores from different sites**. The intraclass correlation was computed between sites using a two-way random effects model with absolute agreement and single measurement:

$$ICC = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}, \quad \text{(Equation 1)}$$

where MS_R is the mean square for participants, MS_E is the mean square for error, MS_C is the mean square for sites, k is the number of sites and n is the number of measures. **The ICC over all sites was ICC=0.9999, $p<10^{-200}$ (CI: [0.99987, 0.99997]) demonstrating that ARTS has excellent inter-rater reliability.** The ICCs for all pairs of sites are shown in Table 1. **In summary, this study shows that different sites running ARTS on the same data but using different computers, operating systems and personnel will generate essentially identical scores, and in a fully automated fashion.**

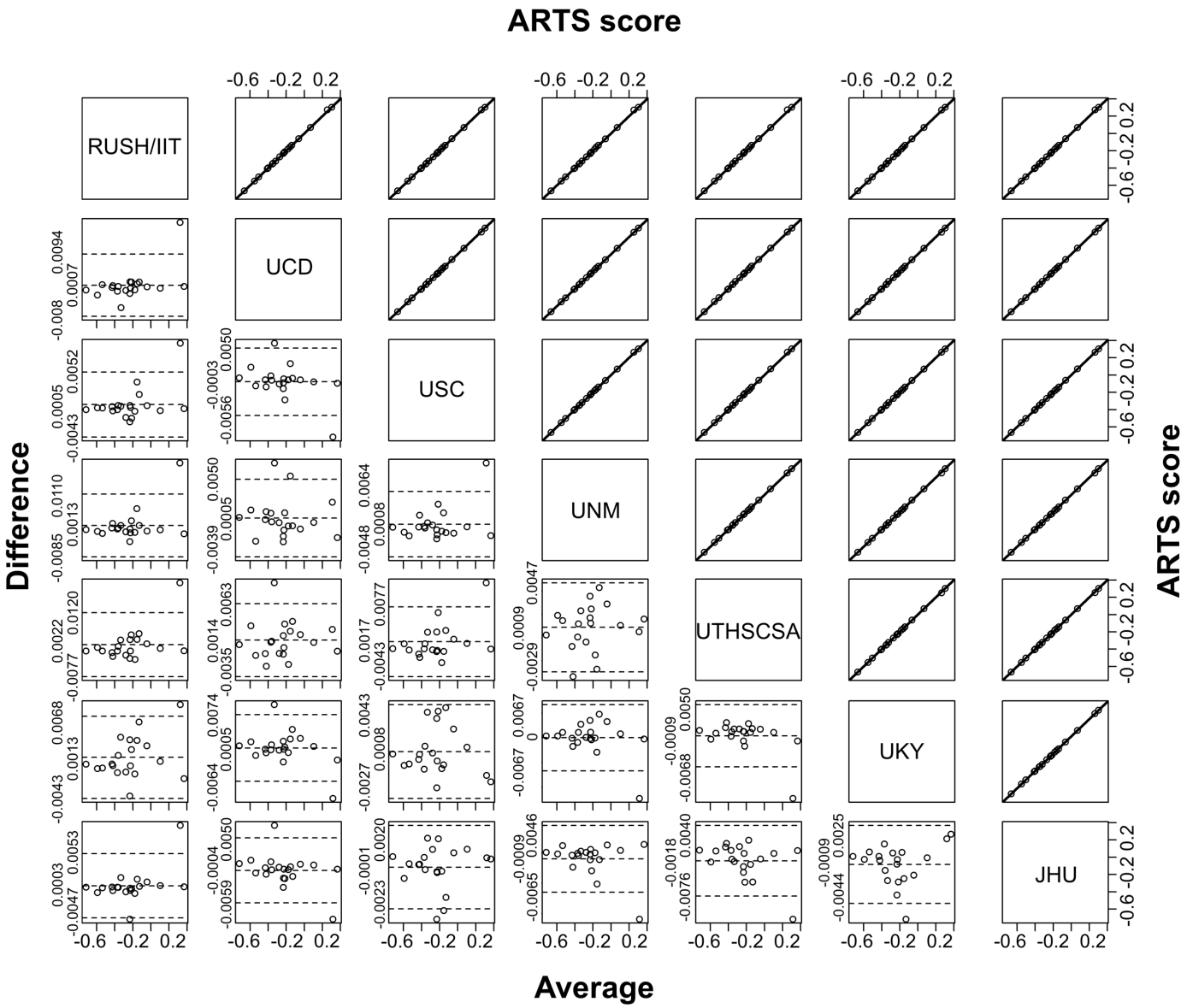

Figure 1. Bland-Altman plots (lower left triangle) and scatter plots (top right triangle) of ARTS scores on 20 MarkVCID participants of the inter-rater reliability study for all pairs of MarkVCID sites. The identity (black) and linear regression (gray) lines are also shown in the scatter plots but cannot be distinguished because they overlap. The dashed lines in the Bland-Altman plots represent the mean and 95% confidence interval of the difference in ARTS scores across sites.

	UCD	USC	UNM	UTHSCSA	UKY	JHU
RUSH/IIT	0.9999, $p<10^{-36}$ [0.9997, 0.9999]	1, $p<10^{-41}$ [0.9999, 1]	0.9998, $p<10^{-35}$ [0.9996, 0.9999]	0.9998, $p<10^{-28}$ [0.9995, 0.9999]	0.9999, $p<10^{-31}$ [0.9998, 1]	1, $p<10^{-40}$ [0.9999, 1]
UCD		1, $p<10^{-39}$ [0.9999, 1]	1, $p<10^{-42}$ [0.9999, 1]	0.9999, $p<10^{-23}$ [0.9998, 1]	0.9999, $p<10^{-38}$ [0.9998, 1]	0.9999, $p<10^{-40}$ [0.9999, 1]
USC			0.9999, $p<10^{-39}$ [0.9999, 1]	0.9999, $p<10^{-22}$ [0.9997, 1]	1, $p<10^{-33}$ [0.9999, 1]	1, $p<10^{-47}$ [1, 1]
UNM				1, $p<10^{-33}$ [0.9999, 1]	0.9999, $p<10^{-37}$ [0.9998, 1]	0.9999, $p<10^{-38}$ [0.9999, 1]
UTHSCSA					0.9999, $p<10^{-39}$ [0.9998, 1]	0.9999, $p<10^{-19}$ [0.9997, 1]
UKY						1, $p<10^{-27}$ [0.9999, 1]

Table 1. ICCs, p-values and confidence intervals of ARTS scores generated on the 20 MarkVCID participants of the inter-rater reliability study for all pairs of sites.

B. Test-retest repeatability: The goal of this study was to assess the repeatability of ARTS scores generated from two MRI datasets collected within a period of one week on the same older adults. All MarkVCID sites collected two MRI datasets within one week on a total of 41 older adults (6 persons were imaged at RUSH/IIT, UCD, USC, UNM, UTHSCSA, JHU, and 5 persons at UKY). One participant imaged at JHU was excluded from this study because the scanner software was updated between the two MRI sessions. Each site ran ARTS on the test-retest data collected locally and shared the scores with us. Figure 2 shows the Bland-Altman plot and scatter plot of the test and retest ARTS scores. **In the scatter plot, all ARTS scores closely follow the identity line (y=x) indicating excellent test-retest agreement**. The intraclass correlation was computed between test and retest ARTS scores on the remaining 40 participants using a two-way random effects model with absolute agreement and single measurement[1] as described in Equation 1. **The ICC was ICC=0.996, $p<10^{-42}$ (CI: [0.993, 0.998]) demonstrating that ARTS has excellent test-retest repeatability. In summary, this study suggests that use of ARTS in longitudinal studies will provide increased sensitivity to small changes.**

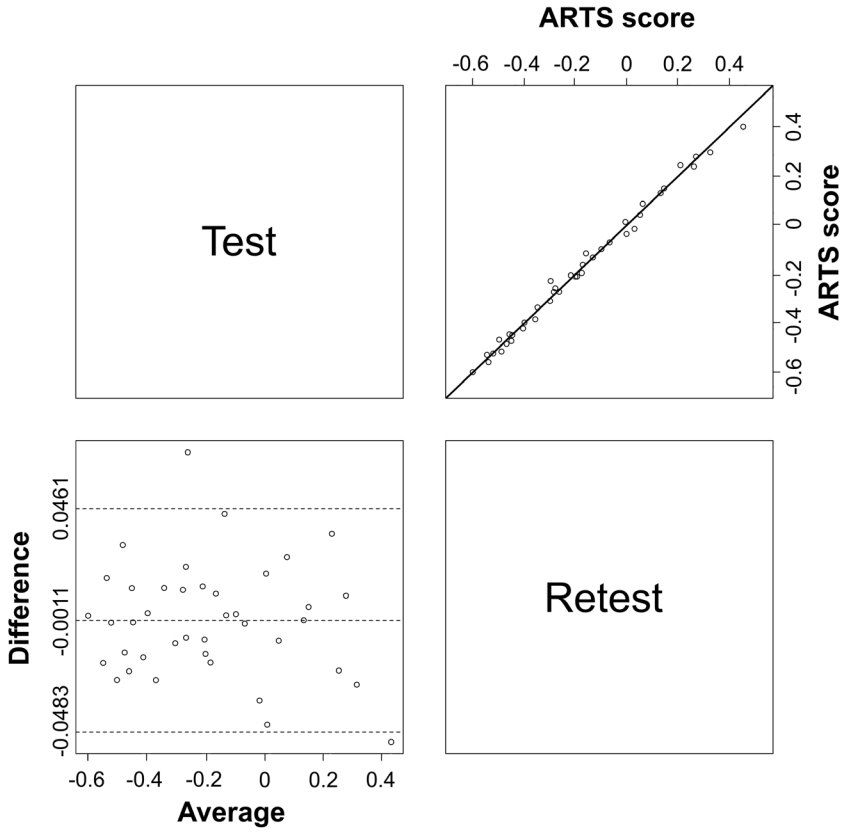

Figure 2. Bland-Altman plot (lower left) and scatter plot (top right) of test-retest ARTS scores on 40 MarkVCID participants of the test-retest repeatability study. The identity (black) and linear regression (gray) lines are also displayed in the scatter plot but cannot be distinguished because they overlap. The dashed lines in the Bland-Altman plot represent the mean and 95% confidence interval of the difference between test and retest ARTS scores.

C. Inter-scanner reproducibility: The goal of this study was to assess the reproducibility of ARTS scores across scanners from different manufacturers/models. Twenty older adults were imaged on four different scanners: Philips Achieva, Siemens Trio, Siemens Prisma, General Electric (GE) 750w (Note: The image reconstruction on the GE 750w included an unnecessary interpolation which had a measurable impact on the values of DTI metrics compared to those from other scanners; this shortfall is unrelated to the performance of ARTS and therefore the GE750w should not be considered in assessing the inter-scanner reproducibility of ARTS). Four participants were excluded from this study, three because of missing DTI data and one because of poor image quality in the diffusion-weighted data. We downloaded the data on the remaining 16 participants and generated ARTS scores for all 4 scanners. Figure 3 shows Bland-Altman plots and scatter plots of the ARTS scores for all pairs of scanners. **The scatter plots show that the ARTS scores for all pairs of scanners (including the GE 750w) closely follow lines with a slope≈1, which indicates that a simple correction by a constant is sufficient to correct ARTS scores from different scanners. This greatly simplifies pooling of data in multi-site studies using scanners from all major manufacturers (GE, Siemens, Philips)**. The scatter plots also show a larger offset between the scores from the GE 750w and those of other scanners that is attributed to the

unnecessary interpolation during image reconstruction mentioned above. The intraclass correlation of ARTS scores from different scanners was computed using a two-way random effects model with a consistency form and single measurement:

$$ICC = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E}, \quad (Equation\ 2)$$

where MS_R is the mean square for participants, MS_E is the mean square for error, k is the number of scanners. **The ICC over all scanners (excluding the GE 750w) was ICC=0.955, p=0.00018 (CI: [0.622,0.989]) indicating that ARTS has excellent inter-scanner reproducibility.** The ICCs for all pairs of scanners are shown in Table 2. **After a simple correction by a constant, the ICC over all scanners (also including the GE 750w) was ICC=0.990, p<10^{-40} (CI: [0.978,0.996]). In summary, this study suggests that ARTS is highly appropriate for multi-site studies using scanners from all major manufacturers.**

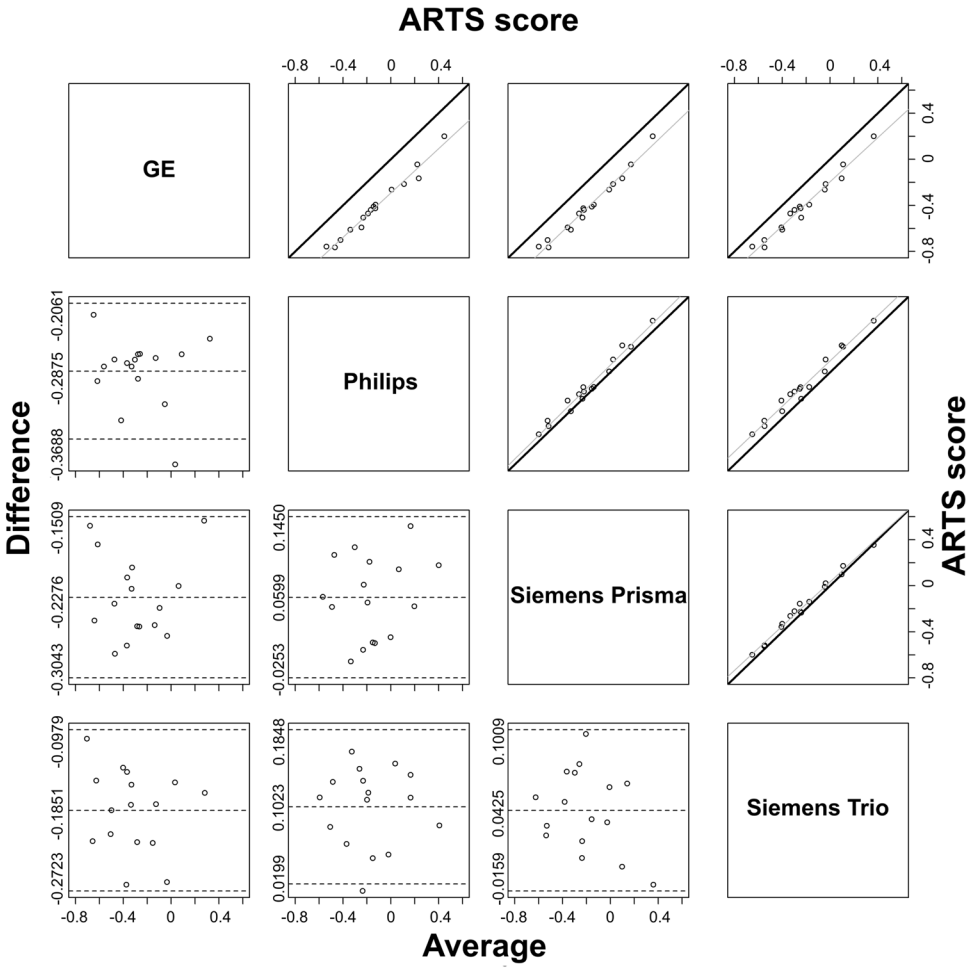

Figure 3. Bland-Altman plots (lower left triangle) and scatter plots (top right triangle) of ARTS scores on 16 MarkVCID participants of the inter-scanner reproducibility study for all pairs of scanners. The identity (black) and linear regression (gray) lines are also shown in the scatter plots. The dashed lines in the Bland-Altman plots represent the mean and 95% confidence interval of the difference in ARTS scores across scanners.

GE 750w	0.627, p=0.054 [-0.008, 0.911]	0.720, p=0.049 [-0.011, 0.940]	0.797, p=0.049 [-0.019, 0.959]
Philips		0.963, p=0.004 [0.424, 0.992]	0.923, p=0.026 [-0.005, 0.985]
Siemens Prisma			0.982, p=0.002 [0.593, 0.996]

Table 2. ICCs, p-values and confidence intervals of ARTS scores generated on the 16 MarkVCID participants of the inter-scanner reproducibility study for all pairs of scanners. Note that the GE750w should not be considered in evaluating the inter-scanner reproducibility of ARTS because the GE 750w data collection had an error that affected the DTI data collected on that scanner, which is unrelated to the performance/reproducibility of ARTS.

3. Protocol for image acquisition

3D MPRAGE, 3D FLAIR and DTI sequences per MarkVCID MRI protocol.

4. Additional data collection required for analysis

The following demographic information needs to be collected for each participant: age at the time of the scan (years), and sex (binary; female=0, male=1). This information is used as input to the software of the biomarker.

Cognitive assessments at two timepoints (at least) are required for testing the performance of the biomarker, one at the time of the MRI scan and at least a second two-three years later. The cognitive assessments should follow the MarkVCID protocol. The cognitive assessments are not used as input to the biomarker, but only for the evaluation of the performance of the biomarker.

5. Protocol for image processing

For each participant, the following information should be available:

- Demographics
  - ✓ Age at time of MRI (years)
  - ✓ Sex (binary; female=0, male=1)
- Imaging
  - ✓ MPRAGE DICOM folder
  - ✓ FLAIR DICOM folder
  - ✓ DTI DICOM folder
  - ✓ DTI reverse polarity DICOM folder

In order to run ARTS, the Singularity software must first be installed. With Singularity software, imaging data are processed inside a virtual environment that has all of the necessary programs and files. This means that the user does not need to install any additional neuroimaging software. Instructions for installing Singularity can be found at:

- Linux: https://sylabs.io/guides/3.5/user-guide/quick_start.html#quick-installation-steps
- MacOS: Appendix A.1 of this document

Prepare the data that will be input to the biomarker

Information about the demographic and imaging data must be entered into a comma separated value (.csv) or excel spreadsheet (.xls or .xlsx) file which will be input to the biomarker script; this file is referred to as the input file. In this input file, each row corresponds to a single participant, and multiple participants can be included in the same csv file (for batch processing). The information in each row must be as follows:

Column 1: Participant ID (can have letters, numbers, underscores, but no symbols; must be unique)
Column 2: Age at the time of MRI (in years)
Column 3: Sex (binary; 1 for male, 0 for female)
Column 4: Absolute path to MPRAGE DICOM folder
Column 5: Absolute path to FLAIR DICOM folder
Column 6: Absolute path to DTI DICOM folder
Column 7: Absolute path to DTI reverse polarity DICOM folder*

* Although not advisable, ARTS can also run without DTI reverse polarity data. If such data is not available for a participant, leave this column blank for that participant, and DTI processing for that participant will be performed without EPI distortion correction.

Run the biomarker

The script that runs ARTS can be found in the downloaded ARTS folder and can be executed as follows:

```
run_ARTS.sh <Path to input file> <Path to output directory>
```

Note that all arguments with a path are absolute paths. Running the script without any input arguments also displays information about the expected input arguments.

Once ARTS has finished running, the output directory will contain A) one subdirectory per participant (containing intermediate pre-processing files, final results, and files necessary for quality checks) and B) a "score_batch" csv file. A) The subdirectory of each participant includes an "analysis" folder which contains a file named "score.csv" that includes the biomarker score of that participant. B) The "score_batch" csv file in the output folder contains the Participant IDs and biomarker scores of all those participants that successfully completed ARTS. The "score_batch" csv file in the output folder has a suffix that indicates the date and time when the ARTS script was executed: [yyyymmdd]T[hhmmss]. For example, the filename "score_batch_20200216T144658.csv" indicates that ARTS was started on February 16th 2020, at 2:46:58 pm.

Important information:

- ARTS can be used on High Performance Clusters that have Singularity installed, to quickly process large amounts of data. To install Singularity on a Linux cluster, see this guide: https://sylabs.io/guides/3.5/admin-guide/installation.html
- Inside the output directory of ARTS, there is a folder for each participant. These folders contain the intermediate files for ARTS. They also contain a QC folder that includes useful images for assessing quality of the preprocessing steps, e.g. the quality of FA maps, quality of registration of FA to a template, quality of brain segmentation, quality of WMH segmentation etc. It is advisable to review these files for each participant.
- Participants that have an error during ARTS processing will not have a biomarker score. In this case, the score_batch csv file will contain fewer participants than the input file.
- The runtime of ARTS depends on the number of participants in the input file, the resolution of the MRI data, and the number of processing cores used for processing. By default, ARTS will use all available cores on the local workstation. You can limit the number of cores used by modifying the N_cpu variable inside the run_ARTS.sh script. N_cpu must be an integer.
- Ensure that a minimum of 6 GB of RAM is available during ARTS processing to avoid out of memory errors.

FAQ:

Q: ARTS was interrupted in the middle of processing an input file. I know which participant was the last one to be completed. How should I restart ARTS?

A:

1. Create a new input file containing only the participants that did not complete processing.
2. Run ARTS with the new input file. For example,

   run_ARTS.sh <Path to new input file> <Path to output directory>
3. There will be two score_batch files at the end of this process. One containing the results generated before ARTS was interrupted, and one containing the results of the second run. The time-stamps in the suffix of the two score_batch filenames reveal which file corresponds to each run.

Q: The terminal printed that a participant was completed, but with errors. How do I find out what the error is?

A: The participant's log file contains the console output of ARTS. This log file is located in the participant's folder inside the output folder and is named "processing.log". If the text near the bottom of the log file has notifications that a process was killed, it is likely that this is an out of memory error. Run ARTS again with at least 6 GB of RAM available.

For other errors or if the error persists, please send the log file to: Konstantinos_Arfanakis@rush.edu

Q: For one of the participants, the WMH mask generated by ARTS is inaccurate. I have my own WMH mask

that is more accurate. How can I use my WMH mask instead?

A: Firstly, small errors in the WMH mask are not a concern since they have insignificant effects on the biomarker score. In case the WMH mask generated automatically by ARTS is largely inaccurate, you can use your own (more accurate) WMH mask. Just check that your WMH mask has the same orientation information as the FLAIR image inside the participant's QC folder (located at <Output directory>/<Participant ID>/QC/WMH_processing/flair.nii.gz). This check is important, as different orientation information may cause your WMH mask to not overlap with voxels affected by WMH in the FLAIR image. For reference, orientation information in the FLAIR image is generated from dcm2niix.

To use your WMH mask with ARTS:

1. Copy your WMH mask named "WMH_mask_custom.nii.gz" into the participant's WMH folder located at "<Output directory>/<Participant ID>/user_input". The "WMH_mask_custom.nii.gz" must be a binary image and must be in the participant's FLAIR space.
2. Run the run_ARTS.sh command with the usual input arguments, but this time add a third input argument, "custom". For example,
   run_ARTS.sh <Path to input file> <Path to output directory> custom
3. All participants in the input file will be quickly reprocessed, since all the processing is already completed. Those with a custom WMH mask in their WMH folder will have an updated biomarker score in the new score_batch file.

APPENDIX

A.1) Installing Singularity on MacOS

These instructions demonstrate how to install Singularity on a Mac via a virtual machine (VM) that is running Ubuntu. Although installing "Singularity Desktop" is an easy option for macOS (through a conventional dmg package), "Singularity Desktop" cannot run ARTS due to its limited functionality. While most of the steps provided here are in the official Singularity guide (https://sylabs.io/guides/3.5/admin-guide/installation.html#mac), additional steps are provided for setting up the VM for ARTS.

1. Install Homebrew using the following command:

```
/usr/bin/ruby -e "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/master/install)"
```

2. With Homebrew, install Virtualbox and Vagrant:

```
brew cask install virtualbox

brew cask install vagrant

brew cask install vagrant-manager
```

3. Set the current directory to a directory where the VM settings can be saved. Once in this directory, execute the following commands:

```
mkdir singularity-vm

cd singularity-vm

vagrant init sylabs/singularity-3.5-ubuntu-bionic64 --box-version 20191206.0.0

vagrant up
```

4. With a text editor, change the properties of the virtual machine by modifying the configuration file called "Vagrantfile" in the same directory (singularity-vm).

   First navigate to line 58 of the Vagrantfile and modify that section as follows:

```
config.vm.provider "virtualbox" do |vb|
   vb.customize ["modifyvm", :id, "--ioapic", "on"]
   vb.memory = "10000"
   vb.cpus = "2"
  end

# At least 6 GB of memory is required to run ARTS. Also set number of CPUs to your preference.
```

Then navigate to line 48 and modify that section of the Vagrantfile as follows:

```
config.vm.synced_folder "/Users/Downloads/ARTS", "/ARTS"
config.vm.synced_folder "/MarkVCID/data", "/input_data"

# This assumes that the downloaded and unzipped ARTS folder is located in /Users/Downloads, and
# the imaging data you will be working on are located in /MarkVCID/data.
```

5. Restart the current VM to apply the changes to the Vagrantfile.

```
vagrant reload
```

A.2) Running ARTS on MacOS

The following step assumes that the folder /MarkVCID/data is synced to the /input_data folder inside the VM based on settings in the Vagrantfile. Similarly, the folder /Users/Downloads/ARTS is synced to the /ARTS folder inside the VM.

1. The paths to the MRI DICOM data defined inside the input csv file should refer to the absolute paths inside the VM. For example, the absolute path /MarkVCID/data/participant1/T1W should be changed to /input_data/participant1/T1W.

2. Save the modified input csv file to a folder that is synced to the VM, like /MarkVCID/data.

3. Change the current directory to the vagrant folder (singularity-vm), and enter into the VM.

```
vagrant ssh
```

4. Run the ARTS command with the appropriate input csv file and output directory relative to the VM. The output directory should be in a folder that is synced.

```
/ARTS/run_ARTS.sh <Path to input csv file in VM> <Path to output directory in the VM>
```