

1 **A large, open source dataset of stroke anatomical brain images and manual**
2 **lesion segmentations**

3 **Authors**

4 Sook-Lei Liew^{1†*}, Julia M. Anglin^{1*}, Nick W. Banks¹, Matt Sondag¹, Kaori L. Ito¹,
5 Hosung Kim¹, Jennifer Chan¹, Joyce Ito¹, Connie Jung¹, Nima Khoshab², Stephanie
6 Lefebvre¹, William Nakamura¹, David Saldana¹, Allie Schmiesing¹, Cathy Tran¹,
7 Danny Vo¹, Tyler Ard¹, Panthea Heydari¹, Bokkyu Kim¹, Lisa Aziz-Zadeh¹, Steven C.
8 Cramer², Jingchun Liu³, Surjo Soekadar⁴, Jan-Egil Nordvik⁵, Lars T. Westlye^{6,7},
9 Junping Wang³, Carolee Winstein¹, Chunshui Yu³, Lei Ai⁸, Bonhwang Koo⁸, R.
10 Cameron Craddock^{8,9}, Michael Milham^{8,9}, Matthew Lakich¹⁰, Amy Pienta¹¹, Alison
11 Stroud¹¹

12 **Affiliations**

- 1 1. University of Southern California, Los Angeles, California, USA
- 2 2. University of California, Irvine, Irvine, California, USA
- 3 3. Tianjin Medical University General Hospital, Tianjin, China
- 4 4. University of Tübingen, Tübingen, Germany
- 5 5. Sunnaas Rehabilitation Hospital HT, Nesodden, Norway
- 6 6. NORMENT and KG Jebsen Centre for Psychosis Research, Division of Mental
7 Health and Addiction, Oslo University Hospital, Oslo, Norway
- 7 7. Department of Psychology, University of Oslo, Oslo, Norway
- 8 8. Child Mind Institute, New York, New York, USA
- 9 9. Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York, USA
- 10 10. University of Texas Medical Branch, Galveston, Texas, USA
- 11 11. University of Michigan, Ann Arbor, Michigan

27 [†] Corresponding author: Sook-Lei Liew (sliew@usc.edu)

28 * Denotes equal contributions

1 **Abstract**
2

3 Stroke is the leading cause of adult disability worldwide, with up to two-thirds
4 of individuals experiencing long-term disabilities. Large-scale neuroimaging
5 studies have shown promise in identifying robust biomarkers (e.g., measures
6 of brain structure) of long-term stroke recovery following rehabilitation.
7 However, analyzing large rehabilitation-related datasets is problematic due to
8 barriers in accurate stroke lesion segmentation. Manually-traced lesions are
9 currently the gold standard for lesion segmentation on T1-weighted MRIs, but
10 are labor intensive and require anatomical expertise. While algorithms have
11 been developed to automate this process, the results often lack accuracy.
12 Newer algorithms that employ machine-learning techniques are promising, yet
13 these require large training datasets to optimize performance. Here we
14 present ATLAS (Anatomical Tracings of Lesions After Stroke), an open-source
15 dataset of 304 T1-weighted MRIs with manually segmented lesions and
16 metadata. This large, diverse dataset can be used to train and test lesion
17 segmentation algorithms and provides a standardized dataset for comparing
18 the performance of different segmentation methods. We hope ATLAS release
19 1.1 will be a useful resource to assess and improve the accuracy of current
20 lesion segmentation methods.

21
22

1 **Background & Summary**
2

3 Approximately 795,000 people in the United States suffer from a stroke
4 every year, resulting in nearly 133,000 deaths¹. In addition, up to 2/3 of stroke
5 survivors experience long-term disabilities that impair their participation in
6 daily activities^{2,3}. Careful clinical decision making is thus critical both at the
7 acute stage, where interventions can spare neural tissue or be used to
8 promote early functional recovery⁴, and at the subacute/chronic stages, where
9 effective rehabilitation can promote long-term functional recovery. Enormous
10 efforts have been made to predict outcomes and response to treatments at
11 both acute and subacute/chronic stages using brain imaging.

12 At the acute stage, within the first 24 hours or so after stroke onset,
13 clinicians face important, time-sensitive decisions such as whether to
14 intervene to save damaged tissue (e.g., administer thrombolytic drugs,
15 perform surgery). Clinical brain images such as magnetic resonance imaging
16 (MRI) and computerized tomography (CT) scans are routinely acquired to help
17 diagnose and make these urgent clinical decisions. Images obtained often
18 include lower-resolution CT scans or structural MRIs (e.g., T2-weighted,
19 FLAIR, diffusion weighted, or perfusion weighted MRIs), and impressive
20 efforts have been made to use these images to automatically detect the lesion
21 volume, predict responses to acute interventions, and predict general
22 prognosis. As clinical scans are typically a mandatory part of acute stroke
23 care, there has been excellent progress in using large-scale datasets of the
24 acquired images to relate to outcomes and build automated lesion detection
25 algorithms and predictive models over the past few decades⁵. In addition,
26 using imaging to assess the extent of neural injury within the first few days
27 after stroke can be helpful for informing entry criteria and stratification
28 variables for enrollment in clinical trials of early recovery therapies, which
29 have specific time windows shortly after stroke onset⁴.

30 On the other hand, there have been fewer advances in large-scale
31 neuroimaging-based stroke predictions at the subacute and chronic stages.
32 Here, clinicians must triage patients and assign scarce rehabilitation
33 resources to those who are most likely to benefit and recover. Brain imaging,
34 such as MRI, is primarily acquired as part of research studies to understand
35 brain-related changes in response to different therapeutic interventions or to
36 provide valuable additional information, beyond what can be gleaned from
37 bedside exams, that can be used to predict rehabilitation outcomes⁶. As stroke
38 is a leading cause of adult disability worldwide, there is a large emphasis
39 placed on predicting and understanding how to best promote long-term
40 rehabilitation in these individuals. Although there are fewer MRIs acquired
41 during this time, the most common research scan is a high-resolution T1-
42 weighted structural MRI, which is often acquired along with functional MRI and
43 high-resolution diffusion MRI scans and can show infarcts at the post-acute
44 stage. Research using these types of images at this stage of stroke have
45 shown promising biomarkers that could potentially provide additional
46 information, beyond behavioral assessments, to predict an individual's
47 likelihood of recovery for specific functions (e.g., motor, speech) and response
48 to treatments⁷⁻⁹. Thus far, measures that include the size, location, and
49 overlap of the lesion with existing brain regions or structures, such as the
50 corticospinal tract, have been successfully used as predictors of long-term
51 stroke recovery and rehabilitation⁹⁻¹⁵. However, to date, this has only been

1 done in smaller-scale studies, and results may conflict across studies or be
2 limited to each sample. Examining lesion properties with larger datasets at the
3 subacute and chronic stages could lead to the identification of more robust
4 biomarkers for rehabilitation that are widely applicable across diverse
5 populations. Recently, efforts for creating large-scale stroke neuroimaging
6 datasets across all time points since stroke onset have emerged and offer a
7 promising approach to achieve a better understanding of the long-term stroke
8 recovery process (e.g., ENIGMA Stroke Recovery;
9 <http://enigma.ini.usc.edu/ongoing/enigma-stroke-recovery/>).

10 However, a key barrier to properly analyzing these large-scale stroke
11 neuroimaging datasets to predict rehabilitation outcomes is accurate lesion
12 segmentation. While many acute neuroimaging stroke studies bypass manual
13 lesion segmentation by using a visual scoring of lesion characteristics with
14 validated scoring tools applied by expert raters, research studies that wish to
15 examine the overlap of the lesion with specific brain structures (e.g., in voxel-
16 based lesion symptom mapping, or lesion load methods) require an accurate
17 and detailed lesion map. In T1-weighted MRIs, which are often used in
18 research, the gold standard for delineating these lesions is manual
19 segmentation, a process that requires skilled tracers and can be prohibitively
20 time consuming and subjective¹⁶. A single large or complex lesion can take up
21 to several hours for even a skilled tracer. As a result of this demand on time
22 and effort, this method, which has been used in previous smaller
23 neuroimaging studies, is not suitable for larger sample sizes. Based on the
24 literature, most studies with manually segmented brain lesions on T1-weighted
25 MRIs use smaller sample sizes between 10 to just over 100 brains¹⁶⁻²⁰.
26 Accurately segmenting hundreds or thousands of stroke lesions from T1-
27 weighted MRIs may thus present a barrier for larger-scale stroke
28 neuroimaging studies.

29 Many stroke neuroimaging studies have utilized semi- or fully-
30 automated lesion segmentation tools for their analyses. Semi-automated
31 segmentation tools employ a combination of automated algorithms, which
32 detect abnormalities in the MR image, and manual corrections or inputs by an
33 expert. Fully-automated algorithms rely completely on the algorithm for the
34 lesion segmentation. While these require little human input or expertise, they
35 still may require significant computational resources and processing time.
36 Many of these fully-automated algorithms employ machine learning
37 techniques that require training and testing on large datasets²¹, and the
38 performance of the algorithm is highly dependent on the size and diversity of
39 the training dataset. While there have been several exciting initiatives
40 regarding lesion segmentation in acute clinical imaging, discussed below,
41 there are few publically available large training/test datasets of manually
42 segmented stroke lesion masks on research-grade T1-weighted images that
43 could be used for improving such algorithms. Thus, while both semi- and fully-
44 automated lesion segmentation tools have the potential to greatly reduce the
45 time and expertise needed to analyze stroke MRI data²², it is unclear whether
46 they provide the accuracy needed for rigorous stroke lesion-based analyses.

47 In addition, it is difficult to compare the performance of automated
48 lesion segmentation tools as they are often not evaluated for performance on
49 the same dataset. Recently, some exciting initiatives have emerged to
50 develop better segmentation algorithms using standardized datasets and
51 metrics. In particular, the Ischemic Stroke Lesion Segmentation (ISLES)

1 challenge is an annual satellite challenge of the Medical Image Computing
2 and Computer Assisted Intervention (MICCAI) meeting that provides a
3 standardized multimodal clinical MRI dataset of approximately 50-100 brains
4 with manually segmented lesions²³. The ISLES competition encourages
5 research groups to use the dataset to evaluate their lesion segmentation
6 algorithms and predict acute outcomes to inform clinical decision making. This
7 approach is promising for developing better lesion segmentation algorithms
8 and predictive models for acute imaging. However, past ISLES challenge
9 datasets have traditionally focused more on using multimodal clinical MRIs to
10 predict more acute results, and these algorithms are not easily translatable to
11 the high-quality T1-weighted MRIs typically found in subacute/chronic stroke
12 rehabilitation research. Thus, here, we aimed to develop a complementary
13 large dataset using only anatomical T1-weighted MRIs, which are typically
14 acquired in research studies after the acute stage to assess rehabilitation
15 outcomes. We anticipate this dataset could be useful for enhancing lesion
16 segmentation methods for T1-weighted images often used in medical
17 rehabilitation research.

18 Here, we present ATLAS (Anatomical Tracings of Lesions After Stroke),
19 Release 1.1, an open-source dataset consisting of 304 T1-weighted MRIs with
20 manually segmented diverse lesions and metadata. The goal of ATLAS is to
21 provide the research community with a standardized training and testing
22 dataset for lesion segmentation algorithms on T1-weighted MRIs. We note
23 that this dataset is not representative of the full range of stroke, as this data
24 was acquired through research studies in which individuals with stroke
25 voluntarily participated, and all participants had to be eligible for a research
26 MRI session. However, this dataset may be useful for testing and comparing
27 the performance of different lesion segmentation techniques and identifying
28 key barriers hindering the performance of automated lesion segmentation
29 algorithms. We believe that this diverse set of manually segmented lesions will
30 serve as a valuable resource for researchers to use in assessing and
31 improving the accuracy of lesion segmentation tools.

32 **Methods**

33

34 *Data Overview*

35 304 MRI images from 11 cohorts worldwide were collected from research
36 groups in the ENIGMA Stroke Recovery Working Group consortium. Images
37 consisted of T1-weighted anatomical MRIs of individuals after stroke. These
38 images were collected primarily for research purposes and are not
39 representative of the overall general stroke population (e.g., only including
40 individuals who opt in to participate in a research study, and excluding
41 individuals with stroke who cannot undergo MRI safely).

42 For each MRI, brain lesions were identified and masks were manually
43 drawn on each individual brain in native space using MRIcron²⁴, an open-
44 source tool for brain imaging visualization and defining volumes of interest
45 (<http://people.cas.sc.edu/rorden/mricron/index.html>). At least one lesion mask
46 was identified for each individual MRI. If additional, separate (non-contiguous)
47 lesions were identified, they were traced as separate masks. An expert
48 neuroradiologist reviewed all lesions to provide additional qualitative
49 descriptions of the type of stroke, primary lesion location, vascular territory,
50 and intensity of white matter disease. Finally, a separate tracer performed

1 quality control on each lesion mask. This included assessing the accuracy of
2 the lesion segmentations, revising the lesion mask if needed, and categorizing
3 the lesions to generate additional data such as the number of lesions in left
4 and right hemispheres, and in cortical and subcortical regions. This dataset is
5 provided in native subject space and archived (n=304). A subset of this
6 dataset was also defaced, intensity normalized, and provided in standard
7 space (normalized to the MNI-152 template, n=229; for an overview of the
8 dataset and archives, see Figure 1).

9
10 All ATLAS contributions were based on studies approved by local ethics
11 committees and were conducted in accordance with the 1964 Declaration of
12 Helsinki. Informed consent was obtained from all subjects. The receiving site's
13 ethics committee at the University of Southern California approved the receipt
14 and sharing of the de-identified data. Data were fully de-identified by removing
15 all 18 HIPAA (Health Insurance Portability and Accountability)-protected
16 health information identifiers, lesions were manually segmented on each MRI,
17 and all data were visually inspected before release. In addition, the terms of
18 the data sharing agreements were approved by the University of Southern
19 California's technology transfer office.

20
21 *Data Characteristics*

22 All T1-weighted MRI data were collected on 3T MRI scanners at a resolution
23 of 1 mm³ (isotropic), with the exception of data from cohorts 1 and 2 which
24 were collected on a 1.5T scanner with a resolution of 0.9 mm x 0.9 mm x 3.0
25 mm (excluded from the normalized dataset). Scanner information (scanner
26 strength, brand) and image resolution are included in the ATLAS meta-data,
27 and sample image header information for a subject from each of the cohorts
28 can be found in *Supplementary Information*.

29
30 Characteristics of the ATLAS dataset include an average lesion volume
31 across all cohorts of $2.128 \pm 3.898 \times 10^4$ mm³, with a minimum lesion size of 10
32 mm³ and a maximum lesion size of 2.838×10^5 mm³. Information regarding
33 the distribution of lesions in the ATLAS dataset (e.g., single versus multiple
34 lesions per individual, percent of lesions that are left versus right hemisphere,
35 or subcortical versus cortical) can be found in Tables 1 and 2. Overall, slightly
36 more than half of the subjects had only one lesion (58%) while the rest had
37 multiple lesions (42.1%). Lesions were roughly equally distributed between left
38 and right hemispheres (48.4% left hemisphere, 43.8% right hemisphere, 7.7%
39 other location such as brainstem or cerebellum). In this dataset, there were
40 more subcortical lesions than cortical lesions (70.7% subcortical, 21.5%
41 cortical, 7.7% other).

42
43 **Table 1.** A total of 304 subjects within 11 cohorts were included in the full
44 ATLAS Release 1.1 native dataset. The number of brains in which only one
45 lesion was found (left/right hemispheres and other locations found within the
46 brainstem and cerebellum, etc.), and the number of brains in which multiple
47 lesions were found, are shown.

Cohort	Number of Subjects	Brains with One Lesion			Brains with
		Left	Right	Other	

					Multiple Lesions
c0001	6	3	2	0	1
c0002	25	6	12	1	6
c0003	55	14	19	0	22
c0004	34	7	3	1	23
c0005	30	9	3	6	12
c0006	12	4	1	2	5
c0007	36	9	9	0	18
c0008	32	8	11	0	13
c0009	12	8	0	0	4
c0010	47	9	10	7	21
c0011	15	1	11	0	3
Total	304	78	81	17	128
% Total Subjects	25.7%	26.7%	5.6%	42.1%	

Table 2. The number of lesions found in each location (i.e. cortical vs. subcortical; left vs. right hemispheres), and other locations (i.e., brainstem, cerebellum, etc.) are shown. Here we have included primary lesions as well as additional lesions, resulting in 521 total lesion masks across n=304.

Cohort	Cortical Lesions		Subcortical Only Lesions		Other Lesion Locations
	Left	Right	Left	Right	
c0001	2	2	2	1	0
c0002	1	7	14	11	4
c0003	3	0	38	44	0
c0004	7	6	31	32	7
c0005	9	1	20	14	7
c0006	3	2	8	3	3
c0007	8	10	28	18	4
c0008	6	13	20	13	1
c0009	5	3	12	2	0
c0010	11	8	21	24	13
c0011	0	5	3	10	1
Total	55	57	197	172	40
%Total Lesions	10.6%	10.9%	37.8%	32.9%	7.7%

Training Individuals Performing Lesion Tracing

Eleven individuals were carefully trained in identifying and segmenting lesions. Individuals had a range of backgrounds, including undergraduate students, graduate students, and postdoctoral fellows. All tracers were given detailed information regarding neuroanatomy, and underwent standardized training, which utilized a detailed protocol as well as an instructional video. All tracers were guided through the training process with extensive feedback on lesion tracing performance by an expert tracer and in consultation with an expert neuroradiologist. The detailed protocol, with pictures of example tracings, is

1 freely available and can be found on the ATLAS GitHub website
2 (<https://github.com/npl/ATLAS/>). All individuals were trained on an initial set
3 of 5 brains with varying lesion sizes and locations (size range: min: 1,871
4 mm³, max: 162,015 mm³; location: cortical [Left: 0, Right: 1], subcortical [Left:
5 4, Right: 0]). After tracing the first set of 5 lesions, tracings were reviewed by
6 an experienced tracer and differences in the lesion masks were discussed
7 with the tracer. One week later, individuals retraced the lesions on the same
8 set of 5 brains, but were blinded to their first segmentation attempt to examine
9 intra-tracer reliability. After this, each lesion segmentation was reviewed by a
10 separate tracer. In addition, the primary lesion location was identified by an
11 expert neuroradiologist, who also created the meta-data (see *Metadata*
12 below). Any questions regarding lesion masks were referred to the
13 neuroradiologist. Inter- and intra-rater reliability measures and additional
14 technical validation of the lesion tracings can be found in *Technical Validation*
15 below. Finally, we note that lesion tracing is a subjective process, even across
16 trained individuals. As mentioned in *Usage Notes* below, any problems,
17 questions, or issues with specific lesion masks can be publically reported on
18 our ATLAS GitHub under the Issues page
19 (<https://github.com/npl/ATLAS/issues>) so that the community of users can
20 make comments and be aware of any identified issues. We will work to
21 resolve any issues in a timely manner.
22

23 *Identifying and Tracing Lesions*

24 To identify lesions, each T1-weighted MRI image was displayed using the
25 multiple view option in MRIcron²⁰, which displays the brain in the coronal,
26 sagittal, and axial view (see Figure 2). To identify lesions, tracers looked for
27 darker intensities within typically healthy tissue. For lesions that were more
28 difficult to detect with the grayscale setting, colored look-up table settings
29 (e.g., “cardiac”, “NIH”, or “spectrum” settings in MRIcron) were used to provide
30 additional insight. Once the lesion or lesions were identified, the lesion mask
31 was traced using either the coronal or axial view, using either a mouse, track
32 pad, or tablet (i.e. Wacom Intuos Draw). A combination of MRIcron tools was
33 used to draw the lesion masks, which included the 3D fill tool, the pen tool and
34 the closed pen tool. Typically, and especially for larger sized lesions, tracers
35 used the 3D fill tool to begin the segmentation. Crosshairs were placed in the
36 center of the identified lesion and the tool would fill in voxels similar to the one
37 at the point of origin with the selected radius and at the sensitivity specified by
38 the difference from origin and difference at edge tools. The pen and closed
39 pen tool, typically was used to fill in (or remove) the areas that the 3D tool had
40 missed or was used to trace smaller lesions slice by slice. Once completed,
41 lesion masks were saved in the volume of interest (VOI) file format with the
42 identifier name “cXXXXsXXXXtXX_LesionRaw” (see *Data Records* and Table
43 3 below for full naming conventions). Lesions masks were then checked for
44 correctness by a separate tracer, who made additional corrections to the
45 lesion mask, if needed. After lesions were identified as being correct, masks
46 were smoothed using MRIcron’s smooth VOI tool where the full width half
47 maximum parameter was set to 2 mm and the threshold was set to 0.5. These
48 masks were saved in both VOI and NIfTI file formats with the identifier name
49 “cXXXXsXXXXtXX_LesionSmooth”.

50 Any additional lesions that were not contiguous with the primary lesion
51 mask were drawn as separate lesion masks and labeled. As described in *Data*

1 Records, any secondary lesions followed the same procedures as the primary
2 lesion mask, but were labeled as Lesion_1, Lesion_2, Lesion_3 and so on,
3 with the naming convention moving from the largest to smallest mask (e.g.,
4 Lesion_1 is the largest secondary lesion mask). In general, the primary lesion
5 mask was the largest lesion, with any secondary lesion masks subsequently
6 named and ordered by size (largest to smallest). The only exception to this
7 was, in the case of multiple lesions, if the neuroradiologist identified a primary
8 stroke location as a different lesion from the largest lesion mask. In these
9 cases, we used the lesion identified by the neuroradiologist as the primary
10 mask. This occurred in less than 5% of the subjects.

11
12 *Metadata*

13 For each lesion, we also provided metadata on the lesion properties to give
14 the user additional qualitative information, beyond the binary lesion mask. This
15 information can be used to quickly sort the dataset based on specific lesion
16 characteristics (e.g., only left hemisphere lesions, or only subcortical lesions).
17 It can also provide additional insight into the types of lesions that succeed or
18 fail for a given lesion segmentation algorithm. The lesion properties were
19 manually reported for each individual lesion mask. These include the number
20 of lesions identified and traced, and the location of each lesion (i.e. right/left,
21 subcortical, cortical, or other). In order to count each lesion only once, we
22 defined subcortical lesions as lesions that are contained completely in the
23 white matter and subcortical structures. Any lesion that extends beyond this
24 area and into the cortex is considered a cortical lesion. In this way, cortical
25 lesions may extend into the subcortical space, but subcortical lesions do not
26 extend into the cortical space. “Other” includes the brainstem and cerebellum.
27 An experienced neuroradiologist also identified the following information for
28 each individual brain: the type of stroke (e.g., embolic, hemorrhagic), primary
29 stroke location, vascular territory, and intensity of white matter disease
30 (periventricular hyperintensities, or PVH, and deep white matter
31 hyperintensities, or DWMH). White matter hyperintensities were graded using
32 the Fazekas scale²⁵. For periventricular hyperintensities, the following grades
33 were applied: 0 = absence, 1 = “caps” or pencil-thin lining, 2 = smooth “halo”,
34 3 = irregular PVH extending into the deep white matter. For deep white matter
35 hyperintensities, the following grades were applied: a = absence, 1 = punctate
36 foci, 2 = beginning confluence of foci, 3 = large confluent areas. The white
37 matter hyperintensity ratings are included because areas of white matter
38 hyperintensity often pose challenges for lesion segmentation algorithms.
39 Finally, scanner strength, brand/model, and image resolution are included in
40 the metadata as well.

41
42 *Normalization to a Standard Template, Intensity Normalization, and Defacing*
43 To expand access to the dataset, we have also provided a subset of the data
44 that is defaced, intensity-normalized, and normalized to standard (MNI-152)
45 space. Lesion segmentation algorithms vary in whether the input should be in
46 native (subject) space or a standardized space. Therefore, to provide this
47 option for users, we also generated a version of the ATLAS dataset in
48 standard space. To convert the images to standard space, MRI images first
49 underwent automated correction for intensity non-uniformity and intensity
50 standardization using custom scripts derived from the MINC-toolkit²⁶
51 (<https://github.com/BIC-MNI/minc-toolkit>). These corrected images were

1 linearly registered to the MNI-152 template using a version that was
2 nonlinearly constructed and symmetric (version 2009;
3 <http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009>) to normalize
4 their intracranial volume in a standardized stereotaxic space²⁷. Using the
5 resulting transformation matrix, the labels drawn on the MRI images were also
6 registered to the MNI template. The MRI images were resampled using the
7 linear interpolation whereas their labels used a nearest neighborhood
8 interpolation to keep their binary nature. Finally . Freesurfer's mri_deface tools
9 were used to perform the defacing (e.g., to remove any facial structures)
10 (https://surfer.nmr.mgh.harvard.edu/fswiki/mri_deface) on all T1-weighted
11 images.

12 Due to technical difficulties and differences in scanner image quality, a
13 subset of brains is not included in the standard space conversion, resulting in
14 a total of n=229 ATLAS brains converted into standard MNI space. Scans
15 from the two cohorts with 0.9 x 0.9 x 3.0 mm resolution images, collected on
16 1.5T scanners, were excluded from this standardized dataset due to their
17 lower resolution. In addition, any images that failed registration were excluded.
18 Primary reasons for failed registration include large lesion volumes or poor
19 image quality (e.g., image artifacts, motion artifacts). We are currently working
20 on manually editing the registrations for these images, which will be released
21 in the future. This dataset can be widely accessed from the FCP-INDI archive
22 (see Figure 1 and Table 3 for archive details). All images were named in
23 accordance with the INDI data policy, following the Brain Imaging Data
24 Structure (BIDS), and a meta-data sheet using the INDI naming convention is
25 included with this dataset.

26

27 *Probabilistic Spatial Mapping of ATLAS Lesion Labels*

28 We also created a probabilistic spatial mapping of the lesion labels solely to
29 visualize the distribution of lesion masks across the normalized ATLAS
30 dataset. We note that this does not provide a representation of a true stroke
31 distribution, but rather shows the distribution of lesions included in this
32 dataset. To do this, we performed a population-based averaging of all the
33 individual primary lesion labels in MNI space, producing a voxel-wise map
34 where values can range from 0 at each voxel (always background for all
35 subjects) to 1 (100% presence of the lesion label across subjects). A
36 probabilistic spatial map of the primary lesions can be found in Figure 3 and a
37 3D visualization of the lesion map can be found in the following video link:
38 <https://www.youtube.com/watch?v=Ag5CUsRNY9Q>. In addition, this map has
39 also been provided in NIfTI format (.nii.gz) and uploaded to NeuroVault.org,
40 an open-source database for neuroimaging data where it can be freely
41 accessed (<https://neurovault.org/collections/3073/>).

42

43 **Data Records**

44 The full raw dataset (*native dataset*, n=304) is archived with the Archive
45 of Disability Data to Enable Policy research at the Inter-university Consortium
46 for Political and Social Research (ICPSR). ICPSR is the world's largest social
47 science data archive that supports several substantive-area archive
48 collections including disability and rehabilitation. ICPSR provides access to
49 the data and provides technical assistance to individuals accessing the data.
50 In addition, a standardized, defaced subset of the dataset (*standardized*
51 dataset, n=229) is archived with the International Data Sharing Initiative, which

1 hosts many widely available neuroimaging datasets such as the Functional
2 Connectome Project (FCP-INDI). See *Usage Notes* for more details regarding
3 access.

4 For the full dataset archived with ICPSR, the naming convention and
5 description of the files in ATLAS R1.1 can be found in Table 3. Within the
6 ATLAS R1.1 main folder, there is an excel file with the metadata for the entire
7 dataset. The data in this archive is in native space (i.e., original subject space;
8 n=304). Throughout the dataset, MRIs are named and sorted based on each
9 cohort (c); each cohort is in the format of cXXXX where XXXX is the number
10 that the cohort was assigned (e.g., c0001). There are 11 total cohorts. Within
11 each cohort folder are the individual subject(s) folders. Subject folders are
12 named based on the cohort that they are in (cXXXX), the subject number that
13 they were assigned (sXXXX) and the time point at which they were taken
14 (tXX) (e.g. c0001s0004t01). For instance, participants with data taken two
15 weeks apart would have two time points, where t01 is the first time point and
16 t02 is the second. Every image starts with the subject identifier of
17 cXXXXsXXXXtXX.

18 Each subject folder has several components: at minimum, each will
19 have the original T1-weighted MRI image (*.nii.gz) and three masks for the
20 main lesion: the unsmoothed lesion mask (*LesionRaw.voi), and two
21 smoothed lesion masks in .voi and .nii.gz formats (*LesionSmooth.voi;
22 *LesionSmooth.nii.gz). The LesionRaw volume is the original hand-traced
23 lesion volume, while the LesionSmooth volume used a Gaussian smoothing
24 kernel (full width half maximum parameter set to 2 mm, threshold set to 0.5, to
25 overcome small errors between slices in tracing; see *Methods* above). We
26 anticipated that most researchers would use the LesionSmooth volume as it is
27 slightly more robust to small slice-by-slice human errors, and therefore
28 created the .nii.giz version from this. Notably, the .voi files are in an MRICron
29 format so the masks can be further edited in MRICron if desired. The .nii.gz
30 files use the standard NIfTI format²⁸ (<http://nifti.nimh.nih.gov/nifti-1/>), which
31 can be opened, edited, and viewed by most standard neuroimaging software.

32 If a particular subject had multiple lesions, for each additional lesion,
33 there would be three additional lesion masks (e.g. *LesionRaw_1.voi,
34 *LesionSmooth_1.voi, *LesionSmooth_1.nii.gz). In general, lesions were
35 ranked based on size where the largest lesion was considered the main
36 lesion. As mentioned previously, if the largest lesion differed from the primary
37 lesion identified by the neuroradiologist, we deemed the primary lesion to be
38 the one identified by the neuroradiologist. This occurred in less than 5% of
39 cases.

40 Finally, in the FCP-INDI archive (standardized dataset, n=229), there is
41 a separate naming convention, following the Brain Imaging Data Structure
42 (<http://bids.neuroimaging.io/>), adopted by FCP-INDI. Images in this dataset
43 have been normalized to a standard MNI-152 template, intensity normalized,
44 and defaced. Table 3 provides a list of all naming conventions and filenames,
45 along with descriptions.

46
47 **Table 3.** Filenames and file descriptions for ATLAS R1.1 dataset. * represents
48 a wildcard.
49

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36684>

native folder (n=304)

Filename or Identifier	Description
cXXXXsXXXXtXX.nii.gz	Raw T1-weighted MRI for each subject, where c = cohort number, s = subject number, and t = time point
*LesionRaw.voi	Raw primary lesion mask, drawn as a volume of interest in MRICron
*LesionSmooth.voi	Smoothed primary lesion mask, drawn as a volume of interest in MRICron
*LesionSmooth.nii.gz	Smoothed primary lesion mask, saved as a nifti file
*LesionRaw/Smooth_1(or 2, 3, ...).voi/.nii.gz	Raw and smoothed secondary lesion masks (same as the three above, but for additional lesions)

INDI ARCHIVE

http://fcon_1000.projects.nitrc.org/indi/retro/atlas.html

standard folder (n=229)

Filename or Identifier	Description
Site, Subject ID, Session	Naming convention follows Brain Imaging Data Structure (BIDS) recommendations

1

2

Technical Validation

3 Each trained tracer created lesion masks for the same five brains twice, 4 one week apart, to assess both inter- and intra-tracer reliability. Training 5 lesions ranged in size and difficulty (see *Methods*). Each tracer's lesion masks 6 were compared, providing both inter- and intra-rater reproducibility measures. 7 We first calculated inter- and intra-rater reliability measures using the lesion 8 volumes. Based on lesion volumes, the inter-rater reliability was 0.76 ± 0.14 , 9 while the intra-rater reliability was 0.84 ± 0.09 .

10 In addition, we also calculated inter-and intra-rater reliability using the 11 Dice similarity coefficient (DC), which is a segmentation accuracy metric, and 12 Hausdorff's distance (HD), which is a metric of the maximum distance 13 between two volumes surface points. DC allows us to examine not only if the 14 volumes are similar, but also if the same voxels are being selected as part of 15 the lesion mask or not. This is particularly useful for comparing neuroimaging 16 volumes, such as lesion masks. DC is calculated by the formula:

$$DC = \frac{2|X \cap Y|}{|X| + |Y|}$$

17 where X and Y represent the voxels from each lesion segmentation, and DC 18 ranges from 0 to 1 (where 0 means there were no overlapping voxels and 1 19 means that the segmentations were completely the same). HD allows us to 20 examine the distance between the surfaces of two images and thus can be 21 used to identify outliers, providing another useful metric for comparing 22 neuroimaging volumes. HD is calculated by the formula:

$$HD(X, Y) = \max\{\max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(y, x)\}$$

23

24

25

1
2 where x and y represent the surface points from the volumes X and Y
3 respectively. HD is measured in millimeters, and a lower value denotes that
4 the maximum distance between the two images is smaller.

5 Inter-rater scores (DC, HD) were calculated for each manual
6 segmentation by comparing each individual tracer's lesion mask to the rest of
7 the tracers' lesion masks. Inter-rater DC and inter-rater HD scores were then
8 averaged to obtain one final DC score and one final HD score for the initial
9 segmentations (average inter-rater DC for first segmentation: 0.73 ± 0.20 ;
10 average inter-rater HD for first segmentation: $22.57\pm21.36\text{mm}$) and for the
11 second segmentations (average inter-rater DC for second segmentation:
12 average inter-rater HD for second segmentations: $25.29\pm23.53\text{mm}$).
13 Furthermore, intra-rater DC and HD scores were calculated for each brain
14 traced by comparing the initial segmentation to the secondary segmentation
15 for each tracer; these scores were then averaged to obtain a final intra-rater
16 DC score (0.83 ± 0.13) and a final intra-rater HD value ($21.02\pm22.66\text{mm}$).

17 Trained tracers segmented all lesion masks. In addition, each lesion
18 mask was checked by a separate tracer, and changes were made to the
19 lesion mask as needed. Any difficulty identifying the lesion was discussed with
20 the expert neuroradiologist. Lastly, after the completion of the dataset, lesion
21 masks were checked a second time to ensure correct segmentation and data
22 descriptors. It is important to note that while tracers did participate in a
23 thorough training process and segmentations were checked multiple times,
24 this is still a subjective process. Comments regarding the lesion masks can be
25 submitted as issues on the ATLAS GitHub site
26 (<https://github.com/npnl/ATLAS/issues>), and we plan to publish updated and
27 expanded versions of this dataset based on feedback and comments from
28 users (see *Usage Notes*).

30 **Usage Notes**

31 The full native-space archived dataset (n=304) can be found at ICPSR:
32 <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36684>. For more information
33 on the data archive, visit the ICPSR website
34 (<https://www.icpsr.umich.edu/icpsrweb>).

35
36 In addition, a standardized, intensity-normalized, defaced subset of the data
37 (n=229) can be found at FCP-INDI:
38 http://fcon_1000.projects.nitrc.org/indi/retro/atlas.html.

39
40 Data is accessible under a standard Data Use Agreement, under which users
41 must agree to only use the data for purposes as described in the agreement.
42 Users of the ATLAS dataset should acknowledge the contributions of the
43 original authors and research labs by properly citing this article and the data
44 repository link from which they accessed the data.

45 As described above, lesions were segmented using the NITRC open
46 source software MRICron which can be downloaded from the NITRC website
47 (<https://www.nitrc.org/projects/mricron>). Users can also quickly and easily view
48 the brains on BrainBox (<http://brainbox.pasteur.fr/>), an open-source Web
49 application to collaboratively annotate and segment neuroimaging data
50 available online²⁹. For additional quick quantification, our group has also
51 created a small package of scripts called SRQL (Semi-automated Robust

1 Quantification of Lesions), which provide three features: it uses a semi-
2 automated white matter intensity correction to further correct for human errors
3 in lesion tracing, outputs a report of descriptive statistics on lesions
4 (hemisphere and volume of lesion), and gives users the option to perform
5 analyses in native or standard space (<https://github.com/npl/SRQL>)³⁰. In
6 addition, as we plan to grow this dataset in the future, additional releases of
7 data or software will be announced on our ATLAS GitHub page
8 (<https://github.com/npl/ATLAS/>). Any issues or feedback can also be
9 submitted on the ATLAS GitHub page under “issues,” and a team of
10 researchers will address these in a timely manner. Finally, as a general note
11 regarding the usage of this dataset, we strongly encourage users to be
12 cautious of overfitting training algorithms to this particular dataset. We note
13 that this data is relatively diverse, given the data collection across 11 research
14 sites worldwide. However, we caution users against overfitting to only a
15 particular cohort or subset of this data. Future work will aim to provide
16 additional test datasets for users to properly test their algorithms on untrained
17 data.
18
19

1 **Acknowledgments**

2 We thank Dr. Mauricio Reyes for insightful conversations and would
3 like to acknowledge the following people for their assistance on this effort:
4 Anthony Benitez, Xiaoyu Chen, Cristi Magracia, Ryan Mori, Dhanashree
5 Potdar, Sandyha Prathap. The archiving of this dataset was specifically
6 supported by the NIH-funded Center for Large Data Research and Data
7 Sharing in Rehabilitation (CLDR; <https://www.utmb.edu/cldr>) under a Category
8 2 Pilot Grant (P2CHD06570) and this work was also funded by an NIH K01
9 award (K01HD091283).

10
11 **Author contributions**

12 S.-L.L. conceptualized the study, reviewed lesions, analyzed data, established
13 archives, and contributed to the writing and editing of the manuscript. J.M.A.
14 segmented and reviewed lesions, oversaw the organization to the
15 segmentation process and contributed to the writing and editing of the
16 manuscript. N.W.B. organized, segmented and reviewed lesions. M.S.
17 provided the neuroradiology expertise and information. K.L.I. and H.K.
18 performed data analysis. H.K. also performed data processing and generated
19 the standardized dataset and probabilistic lesion maps. T.A. provided data
20 visualization expertise and generated the figures/videos. J.C., D.S., A.S. J.I.,
21 C.J., W.N., D.V. and S.L. segmented and/or reviewed lesions. P.H., B.K.,
22 N.K., L.A.-Z., S.C.C., J.L., S.S., L.T.W., J.W., C.W., C.Y. collected and
23 provided the MRI data. M.L., A.P., and A.S. handled the archiving of the data.
24
25

26 **Competing interests**

27 The authors have no conflict of interest.
28
29
30
31
32
33
34
35
36
37
38

1 **Figure Legends**
2

3 Figure 1. A schematic diagram showing the steps performed on the data for
4 each archive release.

5 Figure 2. An example of lesion segmentation in MRICron.

6 Figure 3. A probabilistic lesion overlap map for the primary lesions from the
7 ATLAS R1.1 dataset. A 3D visualization of the lesion overlap map can be
8 found at <https://www.youtube.com/watch?v=Ag5CUUsRNY9Q>.

9
10
11
12

1 **References**
2

3 1. Benjamin, E.J., *et al.* Heart Disease and Stroke Statistics-2017
4 Update: A Report From the American Heart Association.
5 *Circulation* **135**, e146-e603 (2017).

6 2. Feigin, V.L., *et al.* Global and regional burden of stroke during
7 1990-2010: findings from the Global Burden of Disease Study
8 2010. *Lancet (London, England)* **383**, 245-254 (2014).

9 3. Kwakkel, G., Kollen, B.J., Van der Grond, J.V. & Prevo, A.J.H.
10 Probability of regaining dexterity in the flaccid upper limb: Impact
11 of severity of paresis and time since onset in acute stroke. *Stroke*
12 **34**, 2181-2186 (2003).

13 4. Ren, J., Kaplan, P.L., Charette, M.F., Speller, H. & Finklestein,
14 S.P. Time window of intracisternal osteogenic protein-1 in
15 enhancing functional recovery after stroke. *Neuropharmacology*
16 **39**, 860-865 (2000).

17 5. Group, I.-C. Association between brain imaging signs, early and
18 late outcomes, and response to intravenous alteplase after acute
19 ischaemic stroke in the third International Stroke Trial (IST-3):
20 secondary analysis of a randomised controlled trial. *The Lancet
21 Neurology* **14**, 485-496 (2015).

22 6. Burke Quinlan, E., *et al.* Neural function, injury, and stroke
23 subtype predict treatment gains after stroke. *Ann Neurol* **77**, 132-
24 145 (2015).

25 7. Marie-Hélène, M. & Cramer, S.C. Biomarkers of recovery after
26 stroke. *Current opinion in neurology* **21**, 654-659 (2008).

27 8. Nijland, R.H.M., van Wegen, E.E.H., Harmeling-van der Wel, B.C.
28 & Kwakkel, G. Presence of finger extension and shoulder
29 abduction within 72 Hours after stroke predicts functional
30 recovery. *Stroke* **41**, 745-750 (2010).

31 9. Riley, J.D., *et al.* Anatomy of stroke injury predicts gains from
32 therapy. *Stroke* **42**, 421-426 (2011).

33 10. Cramer, S.C., *et al.* Predicting functional gains in a stroke trial.
34 *Stroke* **38**, 2108-2114 (2007).

35 11. Jongbloed, L.Y.N. Prediction of function after stroke: a critical
36 review. *Stroke* **17**, 765-776 (1986).

37 12. Nouri, S. & Cramer, S.C. Anatomy and physiology predict
38 response to motor cortex stimulation after stroke. *Neurology* **77**,
39 1076-1083 (2011).

40 13. Prabhakaran, S., *et al.* Inter-individual variability in the capacity
41 for motor recovery after ischemic stroke. *Neurorehabilitation and
42 Neural Repair* **22**, 64-71 (2007).

43 14. Stinear, C. Prediction of recovery of motor function after stroke.
44 *The Lancet Neurology* **9**, 1228-1232 (2010).

45 15. Zhu, L.L., Lindenberg, R., Alexander, M.P. & Schlaug, G. Lesion
46 load of the corticospinal tract predicts motor impairment in chronic
47 stroke. *Stroke* **41**, 910-915 (2010).

- 1 16. Fiez, J.A., Damasio, H. & Grabowski, T.J. Lesion segmentation
- 2 and manual warping to a reference brain: Intra- and interobserver
- 3 reliability. *Human Brain Mapping* **9**, 192-211 (2000).
- 4 17. Montaner, J., et al. Plasmatic level of neuroinflammatory markers
- 5 predict the extent of diffusion-weighted image lesions in
- 6 hyperacute stroke. *Journal of Cerebral Blood Flow & Metabolism*
- 7 **23**, 1403-1407 (2003).
- 8 18. Sakamoto, Y., et al. Early ischaemic diffusion lesion reduction in
- 9 patients treated with intravenous tissue plasminogen activator:
- 10 infrequent, but significantly associated with recanalization.
- 11 *International Journal of Stroke* **8**, 321-326 (2013).
- 12 19. Thomas, R.G.R., et al. Apparent diffusion coefficient thresholds
- 13 and diffusion lesion volume in acute stroke. *Journal of Stroke and*
- 14 *Cerebrovascular Diseases* **22**, 906-909 (2013).
- 15 20. Wittsack, H.-J., et al. MR Imaging in Acute Stroke: Diffusion-
- 16 weighted and Perfusion Imaging Parameters for Predicting Infarct
- 17 Size. *Radiology* **222**, 397-403 (2002).
- 18 21. Pustina, D., et al. Automated segmentation of chronic stroke
- 19 lesions using LINDA: Lesion identification with neighborhood data
- 20 analysis. *Human Brain Mapping* **37**, 1405-1421 (2016).
- 21 22. de Haan, B., Clas, P., Juenger, H., Wilke, M. & Karnath, H.-O.
- 22 Fast semi-automated lesion demarcation in stroke. *NeuroImage.*
- 23 *Clinical* **9**, 69-74 (2015).
- 24 23. Maier, O., et al. ISLES 2015 - A public evaluation benchmark for
- 25 ischemic stroke lesion segmentation from multispectral MRI.
- 26 *Medical Image Analysis* **35**, 250-269 (2017).
- 27 24. Rorden, C. & Brett, M. Stereotaxic display of brain lesions. *Behav*
- 28 *Neurol* **12**, 191-200 (2000).
- 29 25. Fazekas, F., Chawluk, J., Alavi, A., Hurtig, H. & Zimmerman, R.
- 30 MR signal abnormalities at 1.5 T in Alzheimer's dementia and
- 31 normal aging. *American Journal of Roentgenology* **149**, 351-356
- 32 (1987).
- 33 26. Sled, J.G., Zijdenbos, A.P. & Evans, A.C. A nonparametric
- 34 method for automatic correction of intensity nonuniformity in MRI
- 35 data. *IEEE Transactions on Medical Imaging* **17**, 87-97 (1998).
- 36 27. Collins, L., D., Neelin, P., Peters, T., M. & Evans, A., C. Automatic
- 37 3D intersubject registration of MR volumetric data in standardized
- 38 Talairach space. *Journal of Computer Assisted Tomography* **18**,
- 39 192-205 (1994).
- 40 28. Cox, R.W., et al. A (sort of) new image data format standard: Nifti-
- 41 1. *Neuroimage* **22**, e1440 (2004).
- 42 29. Heuer, K., Ghosh, S., Robinson Sterling, A. & Toro, R. Open
- 43 Neuroimaging Laboratory. *Research Ideas and Outcomes* **2**,
- 44 e9113 (2016).
- 45 30. Ito, K., Anglin, J. & Liew, S.-L. Semi-automated Robust
- 46 Quantification of Lesions (SRQL) Toolbox. *Research Ideas and*
- 47 *Outcomes* **3**, e12259 (2017).

Manual lesion
segmentation performed
on native (raw) T1-
weighted MRIs

ARCHIVE #1:

Full raw dataset archived on
ICPSR (n=304)

Data processed: Defacing,
intensity normalization,
normalization to standard
space (MNI-152 template)

ARCHIVE #2:

Normalized dataset archived
on FCP/INDI (n=229)

44x4x25= 69.00000



