

---

# DIFFUSION-DRIVEN GENERATION OF MINIMALLY PREPROCESSED BRAIN MRI

---

Samuel W. Remedios<sup>1</sup>, Aaron Carass<sup>1</sup>, Jerry L. Prince<sup>1</sup>, Blake E. Dewey<sup>2</sup>,  
for the Alzheimer’s Disease Neuroimaging Initiative\*, for the Australian Imaging Biomarkers and Lifestyle Flagship  
Study of Aging<sup>†</sup>, for the Health and Aging Brain Study (HABS-HD) Study Team<sup>‡</sup>, and for the Mayo Clinic Study of  
Aging<sup>§</sup>

<sup>1</sup>Image Analysis and Communications Laboratory, Johns Hopkins University

<sup>2</sup>Department of Neurology, Johns Hopkins Hospital

## ABSTRACT

**Purpose:** To present and compare three denoising diffusion probabilistic models (DDPMs) that generate 3D  $T_1$ -weighted MRI human brain images.

**Materials and methods:** In this study, three DDPMs were trained using 80,675 image volumes from 42,406 subjects spanning 38 publicly available brain MRI datasets. These images had approximately 1 mm isotropic resolution and were manually inspected by three human experts to exclude those with poor quality, field-of-view issues, and excessive pathology. The images were minimally preprocessed to preserve the visual variability of the data. Furthermore, to enable the DDPMs to produce images with natural orientation variations and inhomogeneity, the images were neither registered to a common coordinate system nor bias field corrected. Evaluations included segmentation, Frechet Inception Distance (FID), and qualitative inspection.

**Results:** All three DDPMs generated coherent MR brain volumes. The velocity and flow prediction models achieved lower FIDs than the sample prediction model. However, all three models had higher FIDs compared to real images across multiple cohorts. In a permutation experiment, the generated brain regional volume distributions differed statistically from real data. However, the velocity and flow prediction models had fewer statistically different volume distributions in the thalamus and putamen.

**Conclusion:** This work presents and releases the first 3D non-latent diffusion model for brain data without skullstripping or registration. Despite the negative results in statistical testing, the presented DDPMs are capable of generating high-resolution 3D  $T_1$ -weighted brain images. All model weights and corresponding inference code are publicly available at <https://github.com/piksl-research/medforj>.

**Keywords** generative AI · open-source data · diffusion models · neuroimaging

## 1 Introduction

Many contemporary medical image analysis and processing techniques rely on large, pre-trained models [1]. Tasks such as segmentation and classification can rely on fine-tuning such models, allowing for exceptional results even with small amounts of training data. Medical image restoration tasks such as denoising, lesion filling, and super-resolution also rely on deep generative priors to regularize candidate solutions. Since image restoration is often the first step of any image analysis pipeline, strong priors are key to performance and reliability of the entire pipeline.

Training a deep generative model, however, is not straightforward. The current state of the art in generative modeling leverages denoising diffusion probabilistic models (DDPMs) [2], a technique that progressively transforms noise into a

---

See A.4\*, A.5<sup>†</sup>, A.7<sup>‡</sup>, and A.11<sup>§</sup>.

sample from the training distribution. Magnetic resonance (MR) images are frequently acquired in 3D or as stacks of 2D slices, resulting in 3D volumes. The corresponding generative model should accordingly produce image volumes that are coherent in 3D. This introduces several practical constraints, such as the vRAM of training GPUs, time required for the model to converge, and the number of image volumes available for training.

There are existing generative models trained to generate MR brain images. However, given the practical considerations, many methods either are trained to generate 2D slices [3], resample the data to a lower resolution (such as 2 mm [4]), use skullstripping, registration, and/or bias field correction as preprocessing steps to simplify the distribution of the learned data [5], or use latent diffusion models (LDMs) [6] to lessen hardware constraints. Many of these trained model weights are also not available online.

To address this, we trained 3D DDPMs to generate minimally preprocessed, coherent 3D human brain image volumes on a large aggregation of publicly available data. We also have made these weights available online at <https://github.com/piksl-research/medforj>.

## 2 Materials and methods

All data used in this study were retrospectively accessed under appropriate license and data use conditions. See Appendix A for data source information.

### 2.1 Datasets and pre-processing

This study used data aggregated from 38 online datasets, named and shown in Table 1. Data were omitted from the aggregated dataset if any of the following criteria were not met: 1)  $T_1$ -weighted; 2) 3D-acquisition; and 3)  $\leq 1.2$ mm resolution in all directions. After aggregation, images were oriented to the axial orientation, interpolated with bicubic interpolation to 1mm isotropic resolution, clipped to the original intensity range, padded or cropped to a size of  $192 \times 224 \times 192$  voxels, then finally linearly min-max normalized and quantized to 16-bit integer values. The pad and crop operation was performed with respect to the brain center of mass, found from the mask computed by HD-BET [7]. We consider these operations to be “minimal preprocessing” because they do not reduce the visual variability of the data.

Some datasets contained multiple scans of the same subject, indicated in Table 1. To avoid data leakage, our training and testing splits were done at the subject level. Approximately 10% of subjects were withheld from each dataset as testing data. Additionally, two datasets were entirely withheld (AIBL and SLEEP). This yielded three distinct datasets: training data (used to train the DDPMs), internal test data (subjects withheld from training but from the same datasets as the training data), and external test data (datasets withheld from training data).

### 2.2 Curation protocol

After pre-processing, data underwent human quality assurance to exclude images with excessive lesions, noise, motion artifacts, limited field-of-view (FOV), or other issues absent from the imaging header. Three expert raters (S.W.R., A.C., B.E.D.) viewed every image and excluded those that met these criteria. This process was heuristic and leveraged the raters’ intuition and experience. Representative pass and fail images are shown in Figure 1.

### 2.3 Model

Generating synthetic images from a DDPM involves progressively denoising an image over a series of steps. A time-conditioned U-net, whose parameters are learned during training, is a common choice for the denoiser. In the literature, there are different choices that can be made on the exact denoising process [32], and in this work we investigate three of them. We trained three DDPMs using a common framework [33]. These models used sample, velocity, and flow prediction targets for the denoising U-net output, respectively. Sample prediction [34, 35], trains the U-net to estimate the clean image at each timestep. Velocity and flow predictions train the U-net to estimate a weighted and unweighted sum, respectively, of the noise and clean image. Although each case has a different target for the U-net, the ultimate effect is a weighting factor on a denoising term in the loss function [32].

Each DDPM used the same U-Net backbone: five levels with 16, 32, 64, 128, and 256 channels respectively. We trained using the conventional  $T = 1000$  diffusion steps but generated images with 64 DDIM [36] steps during inference. We used the Adam optimizer with a learning rate of  $10^{-4}$  and batch size of 24, implemented as a batch of three on each of eight NVIDIA H200 GPUs. We trained for 100 epochs with random 3D rotations (with linear interpolation) of up to  $\pm 10^\circ$  and 3D translations of up to  $\pm 5$  mm.

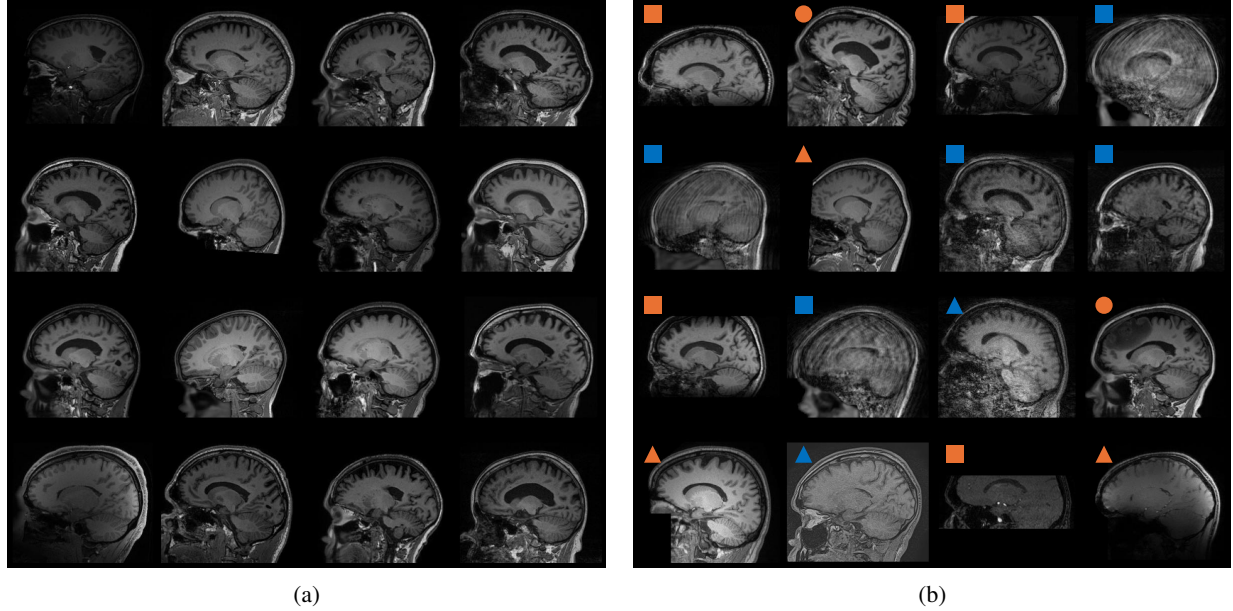


Figure 1: Parasagittal slices from representative pass and fail images. (a) Images that passed manual QA. (b) Images that failed manual QA. Failure reasons include **Pathology** (●), **Limited FOV** (■), **Aggressive defacing** (▲), **Motion artifacts** (■), and **Noise** (▲).

## 2.4 Evaluation

For each of the three trained DDPMs, we generated 1000 synthetic volumes using the exponential moving average (with 10% momentum) of end-of-epoch model weights up to epoch 100. However, the sample prediction model diverged after epoch 70, leading to weights that produced failure image volumes. Because of this, for sample prediction alone, the exponential moving average was computed up to epoch 70. We used qualitative assessment alongside the Frechet Inception Distance (FID) [37] and SynthSeg [38] segmentations to evaluate the realism of our generated images. For FID computations, axial, sagittal, and coronal 2D slices spaced by 4mm were extracted from every image volume from all real and synthetic data. Then, the library `pytorch-fid` [37] was used to compute the FID. For SynthSeg evaluations, we compared distributions of six segmented regions of interest. The Kolmogorov-Smirnov (KS) test was used to check statistical significance between the synthetic volumes and real volumes. Since there were 8 times more test volumes (8016, see Table 1) than synthetic volumes (1000), permutation testing was used to determine statistical significance. We subsampled 1000 real test volumes and repeated the KS test 1000 times, then analyzed the distribution of resulting p-values. Finally, we analyzed the time, space, and financial costs of our proposed models.

## 3 Results

### 3.1 Qualitative results

Uncurated synthetic images from each DDPM are shown as triplanar views in Fig. 2. Each model yielded coherent 3D brain volumes. In Fig. 3, we show triplanar views of the same seed alongside their two nearest neighbors in the training dataset, demonstrating that the generative models can generate coherent 3D images but did not memorize the dataset.

### 3.2 Quantitative results

The FID results are shown in Table 3. The first three columns show the FID between the pairs of real datasets as a benchmark. The last three columns show the FID between real and synthetic datasets. In all cases, there is a gap in FID score between synthetic and real FID scores, with the flow prediction model having a slightly better score.

The segmented volume distribution of six regions of interest are shown in Fig. 4. While the distributions appear close, there is apparent mean and variance shift between the real data and the synthetic data. The distribution of p-values during permutation testing is shown in Table 2. While the sample prediction shows significant differences in all six

Table 1: Distribution of subjects (S) and image volumes (V) by dataset. Note that this is not the total number of data available per dataset, but rather the amount of data after the exclusion described in Sec. 2.1 and Sec. 2.2.

| Dataset Name     | Subjects (S) | Volumes (V) | Train S | Train V | Test S | Test V |
|------------------|--------------|-------------|---------|---------|--------|--------|
| A4 A.1           | 1742         | 6903        | 1594    | 6320    | 148    | 583    |
| ABIDE A.2        | 781          | 781         | 715     | 715     | 66     | 66     |
| ABIDEII A.3      | 959          | 1221        | 878     | 1112    | 81     | 109    |
| ADNI A.4         | 2587         | 17965       | 2368    | 16505   | 219    | 1460   |
| AIBL [8]         | 673          | 1203        | —       | —       | 673    | 1203   |
| AOMIC-ID1000 [9] | 928          | 2768        | 850     | 2539    | 78     | 229    |
| AOMIC-PIOP1 [9]  | 216          | 216         | 198     | 198     | 18     | 18     |
| AOMIC-PIOP2 [9]  | 226          | 226         | 207     | 207     | 19     | 19     |
| BHRC [10]        | 609          | 902         | 558     | 825     | 51     | 77     |
| CCNP [11]        | 195          | 195         | 179     | 179     | 16     | 16     |
| COBRE [12]       | 193          | 1275        | 177     | 1165    | 16     | 110    |
| COGTRAIN [13]    | 166          | 293         | 152     | 271     | 14     | 22     |
| CORR [14]        | 1178         | 2332        | 1078    | 2113    | 100    | 219    |
| DLBS [15]        | 314          | 314         | 288     | 288     | 26     | 26     |
| FCON1000 [16]    | 563          | 563         | 516     | 516     | 47     | 47     |
| GSP [17]         | 1570         | 1639        | 1437    | 1502    | 133    | 137    |
| HABS A.7         | 4233         | 6437        | 3874    | 5880    | 359    | 557    |
| HBN [18]         | 2398         | 2398        | 2195    | 2195    | 203    | 203    |
| HCP1200 A.8      | 1112         | 1112        | 1018    | 1018    | 94     | 94     |
| ICBM A.9         | 444          | 444         | 407     | 407     | 37     | 37     |
| IXI A.10         | 581          | 581         | 532     | 532     | 49     | 49     |
| MCSA A.11        | 1799         | 3087        | 1647    | 2841    | 152    | 246    |
| MPILEMON [19]    | 226          | 226         | 207     | 207     | 19     | 19     |
| NACC A.12        | 3680         | 5472        | 3368    | 4994    | 312    | 478    |
| NADR [20]        | 301          | 301         | 276     | 276     | 25     | 25     |
| NARRATIVES [21]  | 315          | 348         | 289     | 322     | 26     | 26     |
| NEUROPHENOM [22] | 265          | 265         | 243     | 243     | 22     | 22     |
| NIMHHRVD [23]    | 249          | 499         | 228     | 456     | 21     | 43     |
| NKI [24]         | 1314         | 2254        | 1203    | 2058    | 111    | 196    |
| OASIS3 [25]      | 1340         | 3493        | 1227    | 3204    | 113    | 289    |
| OASIS4 [26]      | 655          | 723         | 600     | 663     | 55     | 60     |
| PNC [27]         | 1600         | 1600        | 1464    | 1464    | 136    | 136    |
| PPMI A.15        | 1975         | 3748        | 1808    | 3433    | 167    | 315    |
| QTIM [28]        | 1199         | 1339        | 1098    | 1222    | 101    | 117    |
| SALD [29]        | 494          | 494         | 453     | 453     | 41     | 41     |
| SCAN A.16        | 4702         | 5886        | 4303    | 5384    | 399    | 502    |
| SLEEP [30]       | 136          | 136         | —       | —       | 136    | 136    |
| SLIM [31]        | 588          | 1036        | 539     | 952     | 49     | 84     |
| Total            | 42506        | 80675       | 38174   | 72659   | 4332   | 8016   |

structures to the real data, velocity and flow predictions are not significantly different from the real data for the thalamus and putamen.

### 3.3 Memory and time

Eight NVIDIA H200 GPUs were used in parallel to achieve the training batch size of 24, each requiring 132 GiB of vRAM with a batch size of three. Training took 82 minutes per epoch with 32-bit floating point precision. We trained for 100 epochs, so the total training time was 8200 minutes, or 136.7 hours per model.

For inference, the models required 9 GiB of vRAM to generate a single sample. To analyze the real-world time cost of generating samples, we performed inference with several NVIDIA GPUs with differing numbers of DDIM steps. This is reported in Table 4.

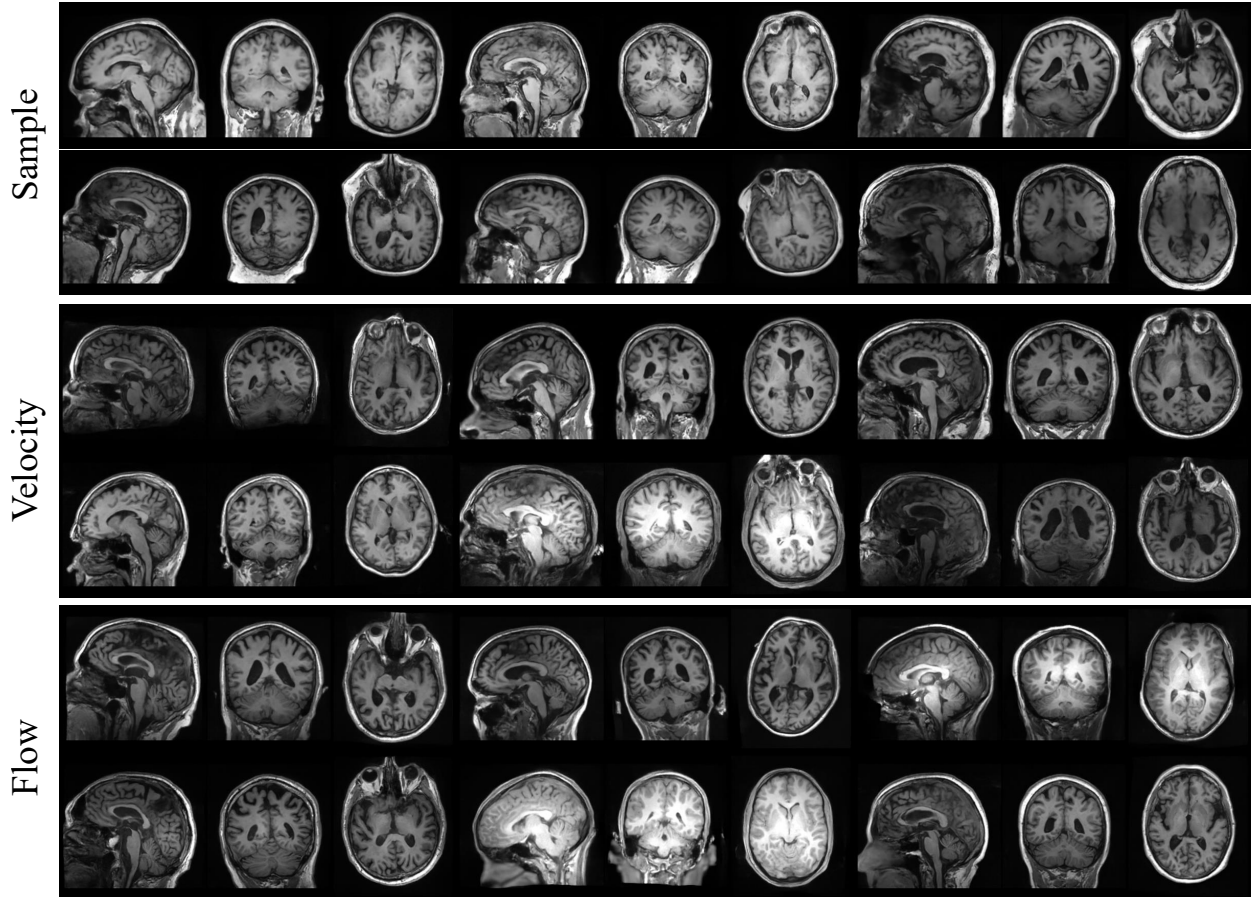


Figure 2: Uncurated synthetic volumes from each prediction type (tripplanar display). All volumes were generated with 64 DDIM steps.

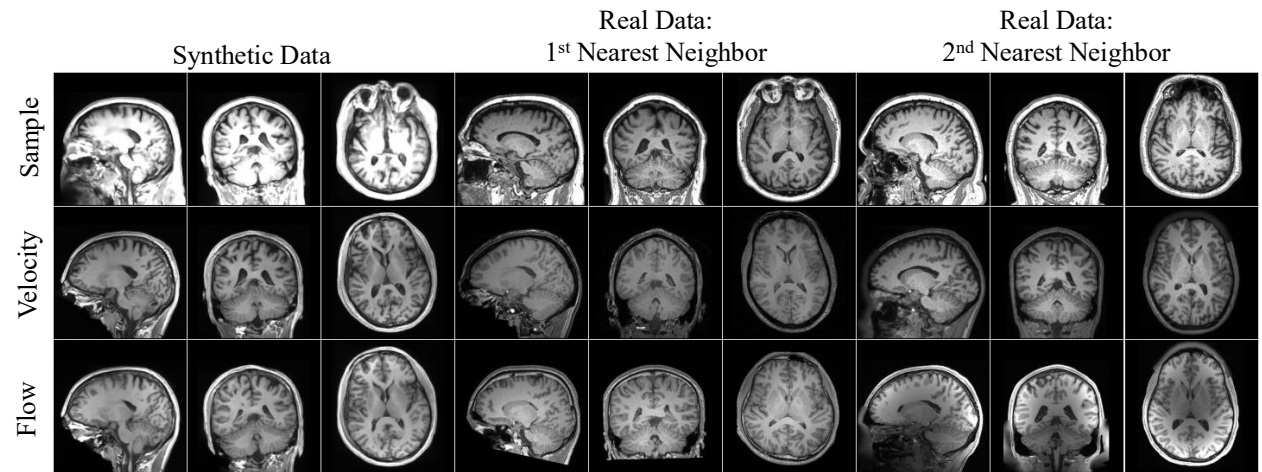


Figure 3: Triplanar views of synthetic images and their nearest neighbors. The first group of three columns shows the synthetic image, and the remaining groups of columns show the first and second nearest neighbors in the real data training set, respectively. Nearest neighbor is computed as mean squared error in voxel space.

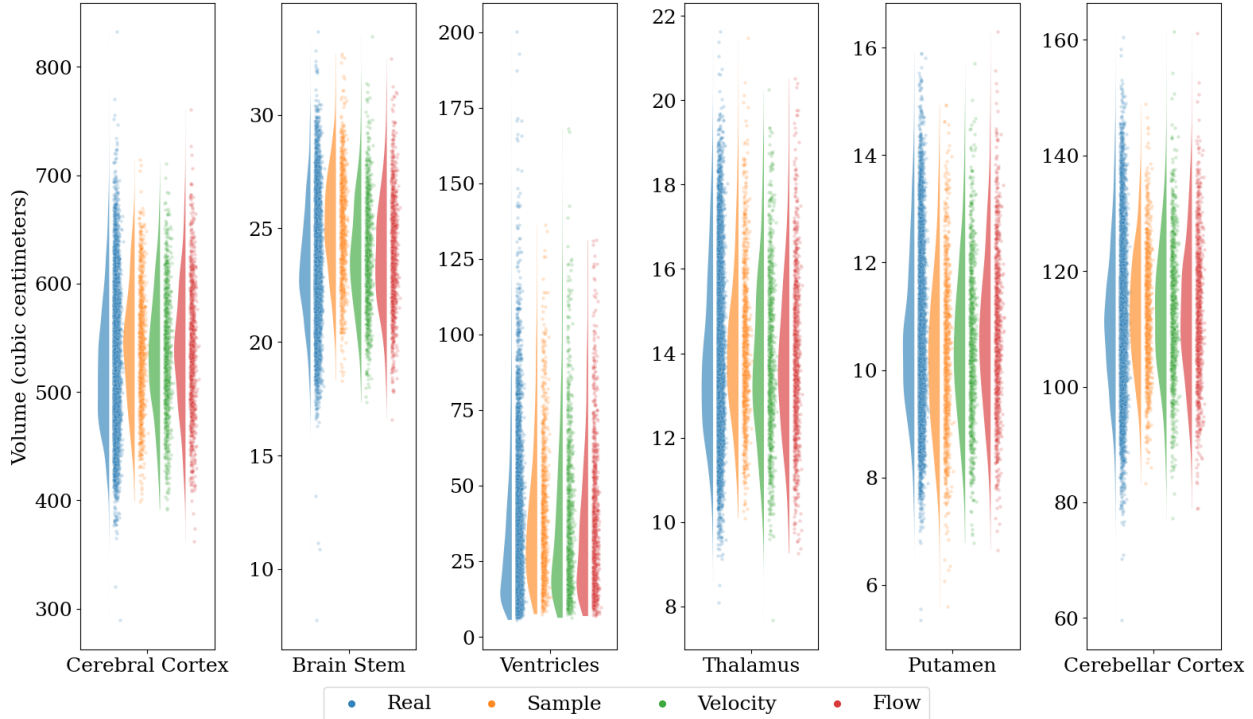


Figure 4: Raincloud plots showing the distributions of volumes as computed by SynthSeg.

Table 2: Percentage of image volumes for which there was no statistical significance in the permutation subset, where significance occurs at  $p < 0.05$ . A higher percentage here indicates that, for that permuted subset, more image volumes were statistically similar to the real data for that segmented structure.

|          | Cerebral Cortex | Brain Stem | Ventricles | Thalamus | Putamen | Cerebellar Cortex |
|----------|-----------------|------------|------------|----------|---------|-------------------|
| Sample   | 0.0%            | 0.0%       | 0.0%       | 0.0%     | 0.0%    | 0.9%              |
| Velocity | 0.0%            | 0.5%       | 0.0%       | 98.8%    | 94.1%   | 6.8%              |
| Flow     | 0.0%            | 0.0%       | 0.0%       | 94.4%    | 57.6%   | 30.4%             |

Table 3: FID between pairs of datasets. Internal is denoted as “(Int.)” and external is denoted as “(Ext.)”. Each row denotes a different real-data dataset to compare to. The first row compares to the real data from datasets used during training. The second row compares to real data from datasets withheld from training. The last row shows an average over the two.

| Reference Dataset | Training Data (Int.) | Test Data (Int.) | Test Data (Ext.) | Sample | Velocity | Flow  |
|-------------------|----------------------|------------------|------------------|--------|----------|-------|
| Training Data     | —                    | 0.11             | 10.05            | 81.82  | 43.14    | 38.53 |
| Test Data (Int.)  | 0.11                 | —                | 10.31            | 81.62  | 42.99    | 38.35 |
| Test Data (Ext.)  | 10.05                | 10.31            | —                | 94.61  | 56.02    | 52.78 |
| Average           | 5.08                 | 5.21             | 5.21             | 86.02  | 47.38    | 43.22 |

## 4 Discussion

In this study, we developed DDPMs for generating minimally preprocessed, high-resolution 3D  $T_1$  w MR brain volumes. These models generated heterogeneous data representative of the varied training datasets, including naturally occurring variations in head angle, contrast, and magnetic field inhomogeneities.

While generative models exist for 3D human brain MRI, ours are the first to be trained on data of this size and variety, at a high isotropic resolution, without skullstripping, registration, or bias-field correction, and without latent models.

Table 4: Inference time (denoted as minutes:seconds) for a single sample from any model for different NVIDIA GPUs with differing numbers of DDIM steps (Steps). The same times are true for all models. The inference time is roughly linear for the number of DDIM steps as well as for increased batch size.

| Steps | Tesla M40 | A16   | T4    | RTX 8000 | Tesla V100 | A100 | RTX 6000 ADA | H100 / H200 |
|-------|-----------|-------|-------|----------|------------|------|--------------|-------------|
| 16    | 1:32      | 0:29  | 0:19  | 0:09     | 0:08       | 0:08 | 0:06         | 0:06        |
| 32    | 3:07      | 0:59  | 0:39  | 0:18     | 0:18       | 0:17 | 0:12         | 0:11        |
| 64    | 6:16      | 1:58  | 1:16  | 0:37     | 0:36       | 0:34 | 0:24         | 0:23        |
| 128   | 12:24     | 3:57  | 2:31  | 1:15     | 1:12       | 1:08 | 0:49         | 0:47        |
| 256   | 24:42     | 7:56  | 5:10  | 2:31     | 2:24       | 2:17 | 1:39         | 1:33        |
| 512   | 49:46     | 15:54 | 10:25 | 5:02     | 4:48       | 4:34 | 3:19         | 3:07        |
| 1000  | 96:24     | 31:03 | 20:22 | 12:27    | 9:23       | 8:56 | 6:30         | 6:06        |

Despite its prevalence in applied DDPMs, we found the model using sample prediction to overall yield less realistic results. Models using velocity and flow predictions, however, achieved substantially better qualitative and quantitative results. Both velocity and flow prediction types resulted in some structures that were statistically similar to real data, and flow prediction exhibited better FID scores. Quantifying the realism of synthetic medical imaging data is still an active area of research, but in neuroimaging it is important to have accurate anatomical distributions.

There are many potential applications of our model. The primary use of a generative model is synthetic data generation, and since our model does not rely on pre-processing, the synthetic images can be used to train other models in tasks such as bias field correction, registration, and skullstripping. Routine image restoration tasks in medical imaging such as super-resolution, denoising, and inpainting can also benefit from a deep generative prior. Finally, in-line with current trends in foundation models, our model may be used alongside fine-tuning for specific tasks for which there is insufficient training data.

There are also limitations to our work. First and foremost, our model does not generate images that are qualitatively indistinguishable from real data. In small-scale in-house observer studies, our experts were always able to determine whether an image was synthetic or real given sufficient time and a holistic view of the volume. Future work will perform proper observer studies. Second, training these models is expensive. The “pay-as-you-go” cost for an 8xH200 node on Microsoft Azure is \$84.80 per hour at the time of writing. Training a single model for 136.8 hours costs approximately \$11,592; training all three models totals roughly \$34,776. However, achieving the results in this paper required multiple trials and hyperparameter tuning. We estimated a total of 25 experiments, each averaging around 200 hours, leading to a naive minimum experimental cost of \$420,000. However, at this level of compute use, alternative pricing options become more practical, such as buying hardware or reserving cloud instances (which we used). These costs are high and the necessary hardware is not widely accessible, which motivates us to openly share our trained models with the broader community. Even with frozen, trained weights, our models require 9 GiB of vRAM when using automatic mixed precision at inference time. While consumer-grade GPUs typically have this amount of vRAM, it is still not common. Finally, our models only generate structural  $T_1$ -w human brain images. In practice, this is the most common contrast to acquire at high resolution (indeed, this is how we were able to train our generative model in the first place). Future work will fine-tune and adapt these models for additional contrasts and anatomies.

In conclusion, we trained several 3D DDPMs to generate high-resolution volumetric human brain MR images. We found that DDPMs trained with velocity and flow prediction yield improved results over sample prediction for the brain generation task. We publicly released all models and code at <https://github.com/piksl-research/medforj>.

## Acknowledgments

This work is supported in part by the DoD CDMRP Award HT94252410785 (PI:Dewey), National MS Society RG-2410-44033 (PI:Dewey), the NIH under NIBIB grant R01-EB036013 (PI: J.L. Prince) and by the Congressionally Directed Medical Research Programs Grant W81XWH2010912 (PI: J.L. Prince). This material is partially supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1746891 (S.W. Remedios). GPU compute time was made possible through a grant from the Johns Hopkins Data Science and AI Institute and Johns Hopkins Research IT.

## References

- [1] Bobby Azad, Reza Azad, Sania Eskandari, Afshin Bozorgpour, Amirhossein Kazerouni, Islem Rekik, and Dorit Merhof. Foundational models in medical imaging: A comprehensive survey and future vision. *CoRR*, abs/2310.18689, 2023.
- [2] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.
- [3] Samuel W Remedios, Blake E Dewey, Aaron Carass, Dzung L Pham, and Jerry L Prince. A deep generative prior for high-resolution isotropic MR head slices. In *Medical Imaging 2023: Image Processing*, volume 12464, pages 116–122. SPIE, 2023.
- [4] Zolnamar Dorjsembe, Sodtavian Odonchimed, and Furen Xiao. Three-dimensional medical image synthesis with denoising diffusion probabilistic models. In *Medical Imaging with Deep Learning*, 2022.
- [5] Zolnamar Dorjsembe, Hsing-Kuo Pao, Sodtavian Odonchimed, and Furen Xiao. Conditional diffusion models for semantic 3D brain MRI synthesis. *IEEE Journal of Biomedical and Health Informatics*, 2024.
- [6] Walter HL Pinaya, Petru-Daniel Tudosiu, Jessica Dafflon, Pedro F Da Costa, Virginia Fernandez, Parashkev Nachev, Sebastien Ourselin, and M Jorge Cardoso. Brain imaging generation with latent diffusion models. In *MICCAI Workshop on Deep Generative Models*, pages 117–126. Springer, 2022.
- [7] Fabian Isensee, Marianne Schell, Irada Pflueger, Gianluca Brugnara, David Bonekamp, Ulf Neuberger, Antje Wick, Heinz-Peter Schlemmer, Sabine Heiland, Wolfgang Wick, Martin Bendszus, Klaus H. Maier-Hein, and Philipp Kickingereder. Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*, 40(17):4952–4964, 2019.
- [8] Kathryn A Ellis, Ashley I Bush, David Darby, Daniela De Fazio, Jonathan Foster, Peter Hudson, Nicola T Lautenschlager, Nat Lenzo, Ralph N Martins, Paul Maruff, et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer’s disease. *International psychogeriatrics*, 21(4):672–687, 2009.
- [9] Lukas Snoek, Maite M van der Miesen, Tinka Beemsterboer, Andries Van Der Leij, Annemarie Eigenhuis, and H Steven Scholte. The Amsterdam Open MRI Collection, a set of multimodal MRI datasets for individual difference analyses. *Scientific data*, 8(1):85, 2021.
- [10] Giovanni Abrahão Salum, Ary Gadelha, Pedro Mario Pan, Tais Silveira Moriyama, Ana Soledade Graeff-Martins, Ana Carina Tamanaha, Pedro Alvarenga, Fernanda Valle Krieger, Bacy Fleitlich-Bilyk, Andrea Jackowski, et al. High risk cohort study for psychiatric disorders in childhood: rationale, design, methods and preliminary results. *International journal of methods in psychiatric research*, 24(1):58–73, 2015.
- [11] Siman Liu, Yin-Shan Wang, Qing Zhang, Quan Zhou, Li-Zhi Cao, Chao Jiang, Zhe Zhang, Ning Yang, Qi Dong, Xi-Nian Zuo, et al. Chinese Color Nest Project: An accelerated longitudinal brain-mind cohort. *Developmental Cognitive Neuroscience*, 52:101020, 2021.
- [12] Andrew R Mayer, David Ruhl, Flannery Merideth, Josef Ling, Faith M Hanlon, Juan Bustillo, and Jose Canive. Functional imaging of the hemodynamic sensory gating response in schizophrenia. *Human Brain Mapping*, 34(9):2302–2312, 2013.
- [13] Joseph W Kable, M Kathleen Caulfield, Mary Falcone, Mairead McConnell, Leah Bernardo, Trishala Parthasarathi, Nicole Cooper, Rebecca Ashare, Janet Audrain-McGovern, Robert Hornik, et al. No effect of commercial cognitive training on brain activity, choice behavior, or cognitive performance. *Journal of Neuroscience*, 37(31):7390–7402, 2017.
- [14] Xi-Nian Zuo, Jeffrey S Anderson, Pierre Bellec, Rasmus M Birn, Bharat B Biswal, Janusch Blautzik, John Breitner, Randy L Buckner, Vince D Calhoun, F Xavier Castellanos, et al. An open science resource for establishing reliability and reproducibility in functional connectomics. *Scientific Data*, 1(1):1–13, 2014.

- [15] Denise C Park, Joseph P Hennessee, Evan T Smith, Micaela Y Chan, Xi Chen, Marianna Dakanali, Michelle E Farrell, Peiying Liu, Hanzhang Lu, Neil Rofsky, et al. The Dallas Lifespan Brain Study: A Comprehensive Adult Lifespan Data Set of Brain and Cognitive Aging. *Scientific Data*, 12(1):846, 2025.
- [16] Bharat B Biswal, Maarten Mennes, Xi-Nian Zuo, Suril Gohel, Clare Kelly, Steve M Smith, Christian F Beckmann, Jonathan S Adelstein, Randy L Buckner, Stan Colcombe, et al. Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, 107(10):4734–4739, 2010.
- [17] Avram J Holmes, Marisa O Hollinshead, Timothy M O’keefe, Victor I Petrov, Gabriele R Fariello, Lawrence L Wald, Bruce Fischl, Bruce R Rosen, Ross W Mair, Joshua L Roffman, et al. Brain genomics superstruct project initial data release with structural, functional, and behavioral measures. *Scientific data*, 2(1):1–16, 2015.
- [18] Lindsay M Alexander, Jasmine Escalera, Lei Ai, Charissa Andreotti, Karina Febre, Alexander Mangone, Natan Vega-Potler, Nicolas Langer, Alexis Alexander, Meagan Kovacs, et al. An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific data*, 4(1):1–26, 2017.
- [19] Anahit Babayan, Miray Erbey, Deniz Kumral, Janis D Reinelt, Andrea MF Reiter, Josefin Röbbig, H Lina Schaare, Marie Uhlig, Alfred Anwander, Pierre-Louis Bazin, et al. A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Scientific Data*, 6(1):1–21, 2019.
- [20] R Nathan Spreng, Roni Setton, Udi Alter, Benjamin N Cassidy, Bri Darboh, Elizabeth DuPre, Karin Kantarovich, Amber W Lockrow, Laetitia Mwilambwe-Tshilobo, Wen-Ming Luh, et al. Neurocognitive aging data release with behavioral, structural and multi-echo functional MRI measures. *Scientific Data*, 9(1):119, 2022.
- [21] Samuel A Nastase, Yun-Fei Liu, Hanna Hillman, Asieh Zadbood, Liat Hasenfratz, Neggin Keshavarzian, Janice Chen, Christopher J Honey, Yaara Yeshurun, Mor Regev, et al. The “Narratives” fMRI dataset for evaluating models of naturalistic language comprehension. *Scientific data*, 8(1):250, 2021.
- [22] Russell A Poldrack, Eliza Congdon, William Triplett, KJ Gorgolewski, KH Karlsgodt, JA Mumford, FW Sabb, NB Freimer, ED London, TD Cannon, et al. A phenome-wide examination of neural and cognitive function. *Scientific data*, 3(1):1–12, 2016.
- [23] Allison C Nugent, Adam G Thomas, Margaret Mahoney, Alison Gibbons, Jarrod T Smith, Antoinette J Charles, Jacob S Shaw, Jeffrey D Stout, Anna M Namyst, Arshitha Basavaraj, et al. The nimh intramural healthy volunteer dataset: A comprehensive meg, mri, and behavioral resource. *Scientific Data*, 9(1):518, 2022.
- [24] Kate Brody Nooner, Stanley J Colcombe, Russell H Tobe, Maarten Mennes, Melissa M Benedict, Alexis L Moreno, Laura J Panek, Shaquanna Brown, Stephen T Zavitz, Qingyang Li, et al. The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Frontiers in Neuroscience*, 6:152, 2012.
- [25] Pamela J. LaMontagne, Tammie LS. Benzinger, John C. Morris, Sarah Keefe, Russ Hornbeck, Chengjie Xiong, Elizabeth Grant, Jason Hassenstab, Krista Moulder, Andrei G. Vlassenko, Marcus E. Raichle, Carlos Cruchaga, and Daniel Marcus. OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer Disease. *medRxiv*, pages 2019–12, 2019.
- [26] Lauren N Koenig, Gregory S Day, Amber Salter, Sarah Keefe, Laura M Marple, Justin Long, Pamela LaMontagne, Parinaz Massoumzadeh, B Joy Snider, Manasa Kanthamneni, et al. Select Atrophied Regions in Alzheimer disease (SARA): An improved volumetric model for identifying Alzheimer disease dementia. *Neuroimage: Clinical*, 26:102248, 2020.
- [27] Theodore D Satterthwaite, John J Connolly, Kosha Ruparel, Monica E Calkins, Chad Jackson, Mark A Elliott, David R Roalf, Ryan Hopson, Karthik Prabhakaran, Meckenzie Behr, et al. The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage*, 124:1115–1119, 2016.
- [28] Lachlan T Strike, Narelle K Hansell, Baptiste Couvy-Duchesne, Paul M Thompson, Greig I de Zubicaray, Katie L McMahon, and Margaret J Wright. Genetic complexity of cortical structure: differences in genetic and environmental factors influencing cortical surface area and thickness. *Cerebral Cortex*, 29(3):952–962, 2019.
- [29] Dongtao Wei, Kaixiang Zhuang, Lei Ai, Qunlin Chen, Wenjing Yang, Wei Liu, Kangcheng Wang, Jiangzhou Sun, and Jiang Qiu. Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Scientific Data*, 5(1):1–10, 2018.
- [30] Michal Rafal Zareba, Magdalena Fafrowicz, Tadeusz Marek, Ewa Beldzik, Halszka Oginska, and Aleksandra Domagalik. Late chronotype is linked to greater cortical thickness in the left fusiform and entorhinal gyri. *Biological Rhythm Research*, 53(10):1626–1638, 2022.
- [31] Wei Liu, Dongtao Wei, Qunlin Chen, Wenjing Yang, Jie Meng, Guorong Wu, Taiyong Bi, Qinglin Zhang, Xi-Nian Zuo, and Jiang Qiu. Longitudinal test-retest neuroimaging data from healthy young adults in southwest China. *Scientific Data*, 4(1):1–9, 2017.

- [32] Ruiqi Gao, Emiel Hooeboom, Jonathan Heek, Valentin De Bortoli, Kevin P. Murphy, and Tim Salimans. Diffusion meets flow matching: Two sides of the same coin. <https://diffusionflow.github.io/>. Accessed: 2025-07-09.
- [33] M Jorge Cardoso, Wenqi Li, Richard Brown, Nic Ma, Eric Kerfoot, Yiheng Wang, Benjamin Murrey, Andriy Myronenko, Can Zhao, Dong Yang, et al. MONAI: An open-source framework for deep learning in healthcare. *arXiv preprint arXiv:2211.02701*, 2022.
- [34] Diederik Kingma, Tim Salimans, Ben Poole, and Jonathan Ho. Variational diffusion models. *Advances in Neural Information Processing Systems*, 34:21696–21707, 2021.
- [35] Diederik Kingma and Ruiqi Gao. Understanding diffusion objectives as the elbo with simple data augmentation. *Advances in Neural Information Processing Systems*, 36:65484–65516, 2023.
- [36] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising Diffusion Implicit Models. In *International Conference on Learning Representations*, 2021.
- [37] Maximilian Seitzer. pytorch-fid: FID Score for PyTorch. <https://github.com/mseitzer/pytorch-fid>, August 2020. Version 0.3.0.
- [38] Benjamin Billot, Douglas N. Greve, Oula Puonti, Axel Thielscher, Koen Van Leemput, Bruce Fischl, Adrian V. Dalca, and Juan Eugenio Iglesias. SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. *Medical Image Analysis*, 86:102789, 2023.

## A Data use statements

### A.1 A4

The A4 Study was a secondary prevention trial in preclinical Alzheimer's disease, aiming to slow cognitive decline associated with brain amyloid accumulation in clinically normal older individuals. The A4 Study was funded by a public-private-philanthropic partnership, including funding from the National Institutes of Health-National Institute on Aging, Eli Lilly and Company, Alzheimer's Association, Accelerating Medicines Partnership, GHR Foundation, an anonymous foundation, and additional private donors, with in-kind support from Avid Radiopharmaceuticals, Cogstate, Albert Einstein College of Medicine and the Foundation for Neurologic Diseases. The companion observational Longitudinal Evaluation of Amyloid Risk and Neurodegeneration (LEARN) Study was funded by the Alzheimer's Association and GHR Foundation. The A4 and LEARN Studies were led by Dr. Reisa Sperling at Brigham and Women's Hospital, Harvard Medical School, and Dr. Paul Aisen at the Alzheimer's Therapeutic Research Institute (ATRI) at the University of Southern California. The A4 and LEARN Studies were coordinated by ATRI at the University of Southern California, and the data are made available under the auspices of Alzheimer's Clinical Trial Consortium through the Global Research & Imaging Platform (GRIP). The complete A4 Study Team list is available on: <https://www.actcinfo.org/a4-study-team-lists/>. We would like to acknowledge the dedication of the study participants and their study partners who made the A4 and LEARN Studies possible.

### A.2 ABIDE

Primary support for the work by Adriana Di Martino was provided by the (NIMH K23MH087770) and the Leon Levy Foundation. Primary support for the work by Michael P. Milham and the INDI team was provided by gifts from Joseph P. Healy and the Stavros Niarchos Foundation to the Child Mind Institute, as well as by an NIMH award to MPM (NIMH R03MH096321). Funding sources for the ABIDE dataset are listed at [https://fcon\\_1000.projects.nitrc.org/indi/abide/abide\\_I.html](https://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html).

### A.3 ABIDE-II

Primary support for the work by Adriana Di Martino and her team was provided by the National Institute of Mental Health (NIMH 5R21MH107045). Primary support for the work by Michael P. Milham and his team provided by the National Institute of Mental Health (NIMH 5R21MH107045); Nathan S. Kline Institute of Psychiatric Research). Additional Support was provided by gifts from Joseph P. Healey, Phyllis Green and Randolph Cowen to the Child Mind Institute. Funding sources for the ABIDE II dataset are listed at [https://fcon\\_1000.projects.nitrc.org/indi/abide/abide\\_II.html](https://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html).

### A.4 ADNI

\* Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The original goal of ADNI was to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The current goals include validating biomarkers for clinical trials, improving the generalizability of ADNI data by increasing diversity in the participant cohort, and to provide data concerning the diagnosis and progression of Alzheimer's disease to the scientific community. For up-to-date information, see [adni.loni.usc.edu](http://adni.loni.usc.edu).

Data collection and sharing for the Alzheimer's Disease Neuroimaging Initiative (ADNI) is funded by the National Institute on Aging (National Institutes of Health Grant U19AG024904). The grantee organization is the Northern California Institute for Research and Education. In the past, ADNI has also received funding from the National Institute of Biomedical Imaging and Bioengineering, the Canadian Institutes of Health Research, and private sector contributions through the Foundation for the National Institutes of Health (FNIH) including generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.;

Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.

## **A.5 AIBL**

<sup>†</sup>Data was partially collected by the AIBL study group. AIBL study methodology has been reported previously (Ellis et al. 2009) Data used in the preparation of this article was obtained from the Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL) funded by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) which was made available at the ADNI database ([www.loni.usc.edu/ADNI](http://www.loni.usc.edu/ADNI)). The AIBL researchers contributed data but did not participate in analysis or writing of this report. AIBL researchers are listed at [www.aibl.csiro.au](http://www.aibl.csiro.au).

## **A.6 GSP**

Data were provided (in part) by the Brain Genomics Superstruct Project of Harvard University and the Massachusetts General Hospital, (Principal Investigators: Randy Buckner, Joshua Roffman, and Jordan Smoller), with support from the Center for Brain Science Neuroinformatics Research Group, the Athinoula A. Martinos Center for Biomedical Imaging, and the Center for Human Genetic Research. 20 individual investigators at Harvard and MGH generously contributed data to the overall project.

## **A.7 HABS**

<sup>‡</sup>HABS-HD MPIs: Sid E O'Bryant, Kristine Yaffe, Arthur Toga, Robert Rissman, & Leigh Johnson; and the HABS-HD Investigators: Meredith Braskie, Kevin King, James R Hall, Melissa Petersen, Raymond Palmer, Robert Barber, Yonggang Shi, Fan Zhang, Rajesh Nandy, Roderick McColl, David Mason, Bradley Christian, Nicole Philips, Stephanie Large, Joe Lee, Badri Vardarajan, Monica Rivera Mindt, Amrita Cheema, Lisa Barnes, Mark Mapstone, Annie Cohen, Amy Kind, Ozioma Okonkwo, Raul Vintimilla, Zhengyang Zhou, Michael Donohue, Rema Raman, Matthew Borzage, Michelle Mielke, Beau Ances, Ganesh Babulal, Jorge Llibre-Guerra, Carl Hill and Rocky Vig. Research reported on this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Numbers R01AG054073, R01AG058533, P41EB015922 and U19AG078109. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## **A.8 HCP1200**

Data collection and sharing for this project was provided by the Human Connectome Project (HCP; Principal Investigators: Bruce Rosen, M.D., Ph.D., Arthur W. Toga, Ph.D., Van J. Weeden, MD). HCP funding was provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute of Mental Health (NIMH), and the National Institute of Neurological Disorders and Stroke (NINDS). HCP data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California. Data used in the preparation of this work were obtained from the Human Connectome Project (HCP) database (<https://ida.loni.usc.edu/login.jsp>). The HCP project (Principal Investigators: Bruce Rosen, M.D., Ph.D., Martinos Center at Massachusetts General Hospital; Arthur W. Toga, Ph.D., University of Southern California, Van J. Weeden, MD, Martinos Center at Massachusetts General Hospital) is supported by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute of Mental Health (NIMH) and the National Institute of Neurological Disorders and Stroke (NINDS). HCP is the result of efforts of co-investigators from the University of Southern California, Martinos Center for Biomedical Imaging at Massachusetts General Hospital (MGH), Washington University, and the University of Minnesota.

## **A.9 ICBM**

Data collection and sharing for this project was provided by the International Consortium for Brain Mapping (ICBM; Principal Investigator: John Mazziotta, MD, PhD). ICBM funding was provided by the National Institute of Biomedical Imaging and BioEngineering. ICBM data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California. Data used in the preparation of this work were obtained from the International Consortium for Brain Mapping (ICBM) database ([www.loni.usc.edu/ICBM](http://www.loni.usc.edu/ICBM)). The ICBM project (Principal Investigator John Mazziotta, M.D., University of California, Los Angeles) is supported by the National Institute of Biomedical Imaging and BioEngineering. ICBM is the result of efforts of co-investigators from UCLA, Montreal Neurologic Institute, University of Texas at San Antonio, and the Institute of Medicine, Juelich/Heinrich Heine University - Germany.

## A.10 IXI

IXI data were downloaded from <https://brain-development.org/ixi-dataset>.

## A.11 MCSA

<sup>§</sup>Data used in preparation of this article were shared by the Mayo Clinic Study of Aging (MCSA). The MCSA is funded by the following sources: NIH U01 AG006786, R01 AG034676, R37 AG011378, R01 AG041851, R01 NS097495, R01 AG056366, R01 AG068206, P30 AG062677, GHR Foundation, Elsie and Marvin Dekelboum Family Foundation, Liston Award, Schuler Foundation, Alexander Foundation, Mayo Foundation for Medical Education and Research.

## A.12 NACC

The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADRCs: P30 AG062429 (PI James Brewer, MD, PhD), P30 AG066468 (PI Oscar Lopez, MD), P30 AG062421 (PI Bradley Hyman, MD, PhD), P30 AG066509 (PI Thomas Grabowski, MD), P30 AG066514 (PI Mary Sano, PhD), P30 AG066530 (PI Helena Chui, MD), P30 AG066507 (PI Marilyn Albert, PhD), P30 AG066444 (PI David Holtzman, MD), P30 AG066518 (PI Lisa Silbert, MD, MCR), P30 AG066512 (PI Thomas Wisniewski, MD), P30 AG066462 (PI Scott Small, MD), P30 AG072979 (PI David Wolk, MD), P30 AG072972 (PI Charles DeCarli, MD), P30 AG072976 (PI Andrew Saykin, PsyD), P30 AG072975 (PI Julie A. Schneider, MD, MS), P30 AG072978 (PI Ann McKee, MD), P30 AG072977 (PI Robert Vassar, PhD), P30 AG066519 (PI Frank LaFerla, PhD), P30 AG062677 (PI Ronald Petersen, MD, PhD), P30 AG079280 (PI Jessica Langbaum, PhD), P30 AG062422 (PI Gil Rabinovici, MD), P30 AG066511 (PI Allan Levey, MD, PhD), P30 AG072946 (PI Linda Van Eldik, PhD), P30 AG062715 (PI Sanjay Asthana, MD, FRCP), P30 AG072973 (PI Russell Swerdlow, MD), P30 AG066506 (PI Glenn Smith, PhD, ABPP), P30 AG066508 (PI Stephen Strittmatter, MD, PhD), P30 AG066515 (PI Victor Henderson, MD, MS), P30 AG072947 (PI Suzanne Craft, PhD), P30 AG072931 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P30 AG086401 (PI Erik Roberson, MD, PhD), P30 AG086404 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD).

## A.13 OASIS-3

OASIS-3: Longitudinal Multimodal Neuroimaging: Principal Investigators: T. Benzinger, D. Marcus, J. Morris; NIH P50 AG00561, P30 NS09857781, P01 AG026276, P01 AG003991, R01 AG043434, UL1 TR000448, R01 EB009352. AV-45 doses were provided by Avid Radiopharmaceuticals, a wholly owned subsidiary of Eli Lilly.

## A.14 OASIS-4

OASIS-4: Clinical Cohort: Principal Investigators: T. Benzinger, L. Koenig, P. LaMontagne

## A.15 PPMI

Data used in the preparation of this article was obtained on 2025-02-20 from the Parkinson's Progression Markers Initiative (PPMI) database ([www.ppmi-info.org/access-data-specimens/download-data](http://www.ppmi-info.org/access-data-specimens/download-data)), RRID:SCR\_006431. For up-to-date information on the study, visit [www.ppmi-info.org](http://www.ppmi-info.org). PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including 4D Pharma, Abbvie, AcureX, Allergan, Amathus Therapeutics, Aligning Science Across Parkinson's, AskBio, Avid Radiopharmaceuticals, BIAL, BioArctic, Biogen, Biohaven, BioLegend, BlueRock Therapeutics, Bristol-Myers Squibb, Calico Labs, Capsida Biotherapeutics, Celgene, Cerevel Therapeutics, Coave Therapeutics, DaCapo Brainscience, Denali, Edmond J. Safra Foundation, Eli Lilly, Gain Therapeutics, GE HealthCare, Genentech, GSK, Golub Capital, Handl Therapeutics, Insitro, Jazz Pharmaceuticals, Johnson & Johnson Innovative Medicine, Lundbeck, Merck, Meso Scale Discovery, Mission Therapeutics, Neurocrine Biosciences, Neuron23, Neuropore, Pfizer, Piramal, Prevail Therapeutics, Roche, Sanofi, Servier, Sun Pharma Advanced Research Company, Takeda, Teva, UCB, Vanqua Bio, Verily, Voyager Therapeutics, the Weston Family Foundation and Yumanity Therapeutics.

## A.16 SCAN

The NACC database is funded by NIA/NIH Grant U24 AG072122. SCAN is a multi-institutional project that was funded as a U24 grant (AG067418) by the National Institute on Aging in May 2020. Data collected by SCAN and shared by NACC are contributed by the NIA-funded ADRCs as follows: Arizona Alzheimer's Center - P30

AG072980 (PI: Eric Reiman, MD); R01 AG069453 (PI: Eric Reiman (contact), MD); P30 AG019610 (PI: Eric Reiman, MD); and the State of Arizona which provided additional funding supporting our center; Boston University - P30 AG013846 (PI Neil Kowall MD); Cleveland ADRC - P30 AG062428 (James Leverenz, MD); Cleveland Clinic, Las Vegas – P20AG068053; Columbia - P50 AG008702 (PI Scott Small MD); Duke/UNC ADRC – P30 AG072958; Emory University - P30AG066511 (PI Levey Allan, MD, PhD); Indiana University - R01 AG19771 (PI Andrew Saykin, PsyD); P30 AG10133 (PI Andrew Saykin, PsyD); P30 AG072976 (PI Andrew Saykin, PsyD); R01 AG061788 (PI Shannon Risacher, PhD); R01 AG053993 (PI Yu-Chien Wu, MD, PhD); U01 AG057195 (PI Liana Apostolova, MD); U19 AG063911 (PI Bradley Boeve, MD); and the Indiana University Department of Radiology and Imaging Sciences; Johns Hopkins - P30 AG066507 (PI Marilyn Albert, PhD.); Mayo Clinic - P50 AG016574 (PI Ronald Petersen MD PhD); Mount Sinai - P30 AG066514 (PI Mary Sano, PhD); R01 AG054110 (PI Trey Hedden, PhD); R01 AG053509 (PI Trey Hedden, PhD); New York University - P30AG066512-01S2 (PI Thomas Wisniewski, MD); R01AG056031 (PI Ricardo Osorio, MD); R01AG056531 (PIs Ricardo Osorio, MD; Girardin Jean-Louis, PhD); Northwestern University - P30 AG013854 (PI Robert Vassar PhD); R01 AG045571 (PI Emily Rogalski, PhD); R56 AG045571, (PI Emily Rogalski, PhD); R01 AG067781, (PI Emily Rogalski, PhD); U19 AG073153, (PI Emily Rogalski, PhD); R01 DC008552, (M.-Marsel Mesulam, MD); R01 AG077444, (PIs M.-Marsel Mesulam, MD, Emily Rogalski, PhD); R01 NS075075 (PI Emily Rogalski, PhD); R01 AG056258 (PI Emily Rogalski, PhD); Oregon Health & Science University - P30 AG066518 (PI Lisa Silbert, MD, MCR); Rush University - P30 AG010161 (PI David Bennett MD); Stanford – P30AG066515; P50 AG047366 (PI Victor Henderson MD MS); University of Alabama, Birmingham – P20; University of California, Davis - P30 AG10129 (PI Charles DeCarli, MD); P30 AG072972 (PI Charles DeCarli, MD); University of California, Irvine - P50 AG016573 (PI Frank LaFerla PhD); University of California, San Diego - P30AG062429 (PI James Brewer, MD, PhD); University of California, San Francisco - P30 AG062422 (Rabinovici, Gil D., MD); University of Kansas - P30 AG035982 (Russell Swerdlow, MD); University of Kentucky - P30 AG028283-15S1 (PIs Linda Van Eldik, PhD and Brian Gold, PhD); University of Michigan ADRC - P30AG053760 (PI Henry Paulson, MD, PhD) P30AG072931 (PI Henry Paulson, MD, PhD) Cure Alzheimer's Fund 200775 - (PI Henry Paulson, MD, PhD) U19 NS120384 (PI Charles DeCarli, MD, University of Michigan Site PI Henry Paulson, MD, PhD) R01 AG068338 (MPI Bruno Giordani, PhD, Carol Persad, PhD, Yi Murphey, PhD) S10OD026738-01 (PI Douglas Noll, PhD) R01 AG058724 (PI Benjamin Hampstead, PhD) R35 AG072262 (PI Benjamin Hampstead, PhD) W81XWH2110743 (PI Benjamin Hampstead, PhD) R01 AG073235 (PI Nancy Chiaravalloti, University of Michigan Site PI Benjamin Hampstead, PhD) 1I01RX001534 (PI Benjamin Hampstead, PhD) IRX001381 (PI Benjamin Hampstead, PhD); University of New Mexico - P20 AG068077 (Gary Rosenberg, MD); University of Pennsylvania - State of PA project 2019NF4100087335 (PI David Wolk, MD); Rooney Family Research Fund (PI David Wolk, MD); R01 AG055005 (PI David Wolk, MD); University of Pittsburgh - P50 AG005133 (PI Oscar Lopez MD); University of Southern California - P50 AG005142 (PI Helena Chui MD); University of Washington - P50 AG005136 (PI Thomas Grabowski MD); University of Wisconsin - P50 AG033514 (PI Sanjay Asthana MD FRCP); Vanderbilt University – P20 AG068082; Wake Forest - P30AG072947 (PI Suzanne Craft, PhD); Washington University, St. Louis - P01 AG03991 (PI John Morris MD); P01 AG026276 (PI John Morris MD); P20 MH071616 (PI Dan Marcus); P30 AG066444 (PI John Morris MD); P30 NS098577 (PI Dan Marcus); R01 AG021910 (PI Randy Buckner); R01 AG043434 (PI Catherine Roe); R01 EB009352 (PI Dan Marcus); UL1 TR000448 (PI Brad Evanoff); U24 RR021382 (PI Bruce Rosen); Avid Radiopharmaceuticals / Eli Lilly; Yale - P50 AG047270 (PI Stephen Strittmatter MD PhD); R01AG052560 (MPI: Christopher van Dyck, MD; Richard Carson, PhD); R01AG062276 (PI: Christopher van Dyck, MD); 1Florida - P30AG066506-03 (PI Glenn Smith, PhD); P50 AG047266 (PI Todd Golde MD PhD)