NeuroImage 144 (2017) 255–258

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

# Harvard Aging Brain Study: Dataset and accessibility

Alexander Dagley [a,d,*], Molly LaPoint [a,d], Willem Huijbers [a,b,d,g], Trey Hedden [a,c,d,e], Donald G. McLaren [a,c,d], Jasmeer P. Chatwal [a,b,d], Kathryn V. Papp [a,b,c], Rebecca E. Amariglio [a,b,c], Deborah Blacker [a,c,h,i], Dorene M. Rentz [a,b,c], Keith A. Johnson [a,b,c,f], Reisa A. Sperling [a,b,c], Aaron P. Schultz [a,d]

[a] Harvard Aging Brain Study, Department of Neurology, Massachusetts General Hospital, Boston, MA 02114, USA
[b] Center for Alzheimer Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115, USA
[c] Harvard Medical School, Boston, MA 02115, USA
[d] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129, USA
[e] Harvard University, Department of Psychology, Center for Brain Science, Cambridge, MA 02138, USA
[f] Department of Radiology, Massachusetts General Hospital, Boston, MA, USA
[g] The German Center for Neurodegenerative Diseases (DZNE), Population Health Sciences, Bonn, Germany
[h] Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA
[i] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Cambridge, MA, USA

ARTICLE INFO

*Article history:*
Accepted 24 March 2015
Available online 3 April 2015

ABSTRACT

The Harvard Aging Brain Study is sharing its data with the global research community. The longitudinal dataset consists of a 284-subject cohort with the following modalities acquired: demographics, clinical assessment, comprehensive neuropsychological testing, clinical biomarkers, and neuroimaging. To promote more extensive analyses, imaging data was designed to be compatible with other publicly available datasets. A cloud-based system enables access to interested researchers with blinded data available contingent upon completion of a data usage agreement and administrative approval. Data collection is ongoing and currently in its fifth year.



## Introduction

The Harvard Aging Brain Study (HABS NIH-P01AG036694) is a longitudinal observational study designed to further our understanding of differentiating "normal" aging from preclinical Alzheimer's dementia (AD). Longitudinal data collection in HABS is ongoing and now in its fifth year. Table 1 highlights demographics for the baseline HABS cohort.

From study initiation, we planned to make the HABS dataset available to the global research community, with the hope that sharing this richly characterized cohort beyond our immediate lab and established collaborators would accelerate advances in our understanding of "normal" aging and preclinical AD (Sperling et al., 2011). With this in mind, we intentionally utilized MRI and PET acquisitions that are compatible with other publicly available datasets, including the Alzheimer's Disease Neuroimaging Initiative (ADNI), http://adni.loni.usc.edu/; Australian Imaging, Biomarker & Lifestyle Flagship Study of Aging (AIBL), http://aibl.csiro.au; and the Genome Superstruct Project (GSP), https://thedata.harvard.edu/dvn/dv/GSP. Recent studies have combined MRI and PET data from HABS with ADNI and AIBL (Mormino et al., 2014) and HABS with GSP (Schultz et al., 2014) to increase statistical power. The HABS dataset can augment existing data sharing initiatives, and we hope this will provide new insights into aging and preclinical AD.

The first data freeze (v1.00), which includes baseline clinical and neuropsychological assessments, regional PiB-PET measures, and regional structural MRI measurements, was released on August, 2014. We plan to release additional data in early 2016, including a richer set of neuropsychological variables and raw imaging data including structural MRI, resting state fMRI, task-based fMRI, PiB-PET, and FDG-PET scans. Additional information regarding HABS as well as information about data requests and what data is currently being shared is available at the HABS website: http://www.nmr.mgh.harvard.edu/lab/harvardagingbrain. This website will be updated with additional data, information, and announcements over time. Our goal here is to provide for potential users a citable introduction to the available data, its structure, and our quality control methods.

### *The Harvard Aging Brain Study database*

The HABS data sharing system is designed specifically to make HABS data available to researchers from around the world. This involves an easily accessible system consisting of an online data request form and a simple excel spreadsheet containing basic phenotypic information with blinded subject identifiers for study participants. This spreadsheet is then distributed to approved users and serves as the link to all other measurements and modalities. We were able to design this relatively

* Corresponding author at: Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129, USA.
*E-mail address:* adagley@partners.org (A. Dagley).

http://dx.doi.org/10.1016/j.neuroimage.2015.03.069
1053-8119/

**Table 1**
Demographics for HABS cohort.

| Label | |
|---|---|
| *n* | 284 |
| Age | 62–90; 73.67 ± 6.13 years |
| Sex | 167 females; 117 males |
| Years of education | 15.81 ± 3.04 |
| VIQ | 120.77 ± 9.24 |
| Ethnicity | 231 white; 45 African American; 6 Asian; 1 Native American |

simple system because the repository covers a single study with a predefined set of visits and measurements associated with each subject.

The first HABS data release (v1.0) contains only baseline data, but future releases will include longitudinal data, and will be made available after each data collection wave has been completed and data is curated. Currently, a subset of the neuropsychological, clinical, and imaging data for baseline visits have been made available (specific details are available on the HABS website). ROI values for structural MRI and PiB-PET are available in spreadsheet form as are a number of neuropsychological tests, clinical assessments, and demographic information. Over the course of 2015 we plan to prepare the raw imaging data for sharing, with the goal of making the image data available by early 2016. Following this we will begin to release longitudinal data.

Only data from HABS participants will be included; however, data from affiliated studies using HABS participants may be included if participants have consented to their data being shared as part of the affiliated study. We will not augment the dataset with data from other samples, although we expect that the HABS dataset will be used in concert with ADNI, ABIL, and GSP datasets among others. We plan to host the data in-perpetuity or until such time that we find a centralized repository for the data to be stored and shared. Given the strong movement towards greater sharing of neuroimaging datasets, we look forward to new tools and storage methods that will be developed in the future.

### *Blinding process*

The spreadsheet contains blinded subject identifiers and collection dates that can be used to match the spreadsheet-based information to future data releases as well as raw imaging data. Data collection dates are blinded by adding a randomized number of days to each date. Each subject has a single random integer, which is used for all dates of that subject. This blinds the collection date but fully preserves relative date differences between any collection dates and enables longitudinal analyses.

### *Data access*

Access to these data requires the completion of an online data request form, which includes user registration and acceptance of the terms of a data use agreement. The data use agreement focuses on three primary requirements: 1) agreement to abide by human subject research and data sharing policies, 2) agreement to credit the Harvard Aging Brain Study as the source of the data, and 3) agreement to not attempt in any way to obtain the identities of the subjects. The data use agreement and the data request form can be accessed on the HABS website.[1] Once the data request is submitted, it is sent to the HABS data managers for review and approval by the data committee. Two weeks will be asked to process the request. The approval process is primarily designed to allow us to track who is utilizing the HABS data and ensure that the proposed purpose of the data request is consistent with the data use agreement. Evaluation of the proposed research hypothesis and aims will not factor into the decision. Our goal is for this dataset to be an open resource for the entire research community. When approved, a download link to the excel file, containing information for the full cohort, will be sent to the requester. The registration process also enables us to compile e-mail addresses that can be used to notify users of updates and changes to the dataset.

## *Harvard Aging Brain Study methods*

### *Data structure*

As a longitudinal project, the HABS dataset consists of multiple study visit dates each year, necessitating a temporal grouping scheme for storage and subsequent analysis. A StudyArc variable was created to distinguish time points, with HAB_1.0 representing baseline, HAB_2.0 representing year 2, HAB_3.0 representing year 3, etc. HAB_1.0 includes visits 1 to 7. Table 2 depicts the breakdown of required visits and testing done at each visit.

At thirty-six months (HAB_4.0), all study procedures are repeated, including the full range of neuroimaging modalities. We are currently seeking funding to continue to follow subjects out to 60 months from baseline, with an additional round of neuroimaging at 60 months.

### *Recruitment criteria*

Inclusion criteria included: 65 years of age or older (4 exceptions were made to help meet diversity targets; minimum age at entry was 62 years), a score of 0 on the Clinical Dementia Rating Scale, a score of greater than 25 on the Mini-Mental State Examination, scores above age and education-adjusted cutoffs on the 30-Minute Delayed Recall of the Logical Memory Story A (Wechsler, 1987, ADNI based cut-offs; http://www.adni-info.org/), and a score of less than 11 on the Geriatric Depression Scale. Exclusion criteria included: history of alcoholism, drug abuse, head trauma, or current serious medical/psychiatric illness. The criteria utilized for inclusion/exclusion ensures the HABS cohort is composed of cognitively normal, healthy older individuals at study enrollment. As of March 2015, 6% of the cohort has progressed to mild cognitive impairment, and we expect additional individuals to progress over the next 5 years.

### *Clinical assessments, neuropsychological testing, and clinical biomarkers*

HABS collects a range of demographic information (e.g., age, sex, education, race), clinical assessments (e.g., CDR, MMSE, GDS, estimated IQ), neuropsychological testing (e.g., processing speed, executive function, memory), blood work (fasting blood draw for creatinine, cholesterol, HDL, LDL, glucose), APOE genotyping, and physiological measures (e.g., height, weight, blood pressure). Lumbar puncture was performed on a subset of subjects for CSF derived measures of aß 1–42, total tau, and phospho-tau. A full list of currently available data can be found on the HABS website (this will be updated as new measures and modalities become available).

### *Imaging modalities*

Subjects undergo a thorough imaging protocol consisting of the following modalities: PiB-PET (C-11 Pittsburgh Compound-B; dynamic 0–60 min), FDG-PET (F-18 fluorodeoxyglucose; 45–75 min), 3D T1-weighted structural MRI, task fMRI, resting state fMRI, diffusion tensor imaging (DTI), 3D T2-weighted FLAIR, susceptibility weighted imaging (SWI), and pulsed arterial spin labeling (pASL). All PET data is collected on a Siemens ECAT HR+ PET scanner at Massachusetts General Hospital (MGH), in Boston, MA. All MR data is collected at the MGH-Martinos Center for Biomedical Imaging in Charlestown, MA on one of two matched 3 T Siemens Tim Trio 3 T scanners with a 12-channel phased-array head coil. Complete PET and MRI acquisition protocols can be found on the HABS website.[2]

[1] https://www.nmr.mgh.harvard.edu/lab/harvard-aging-brain-study/public-data-releases.

[2] https://www.nmr.mgh.harvard.edu/lab/harvardagingbrain/data.