



## OPEN ACCESS

## EDITED BY

Tomáš Paus,  
Université de Montréal, Canada

## REVIEWED BY

John Breitner,  
McGill University, Canada  
Alexa Pichet Binette,  
Lund University, Sweden

## \*CORRESPONDENCE

C. Elizabeth Shaaban  
beth.shaaban@pitt.edu

## SPECIALTY SECTION

This article was submitted to  
Population Neuroimaging,  
a section of the journal  
Frontiers in Neuroimaging

RECEIVED 25 June 2022

ACCEPTED 05 September 2022

PUBLISHED 26 September 2022

## CITATION

Shaaban CE, Tudorascu DL,  
Glymour MM, Cohen AD, Thurston RC,  
Snyder HM, Hohman TJ, Mukherjee S,  
Yu L and Snitz BE (2022) A guide for  
researchers seeking training in  
retrospective data harmonization for  
population neuroscience studies of  
Alzheimer's disease and related  
dementias.

*Front. Neuroimaging* 1:978350.  
doi: 10.3389/fnimg.2022.978350

## COPYRIGHT

© 2022 Shaaban, Tudorascu, Glymour,  
Cohen, Thurston, Snyder, Hohman,  
Mukherjee, Yu and Snitz. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# A guide for researchers seeking training in retrospective data harmonization for population neuroscience studies of Alzheimer's disease and related dementias

C. Elizabeth Shaaban<sup>1\*</sup>, Dana L. Tudorascu<sup>2</sup>,  
M. Maria Glymour<sup>3</sup>, Ann D. Cohen<sup>2</sup>, Rebecca C. Thurston<sup>1,2</sup>,  
Heather M. Snyder<sup>4</sup>, Timothy J. Hohman<sup>5</sup>,  
Shubhabrata Mukherjee<sup>6</sup>, Lan Yu<sup>7</sup> and Beth E. Snitz<sup>8</sup>

<sup>1</sup>Department of Epidemiology, School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States, <sup>2</sup>Department of Psychiatry, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States, <sup>3</sup>Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, San Francisco, CA, United States, <sup>4</sup>Medical and Scientific Relations, Alzheimer's Association, Chicago, IL, United States, <sup>5</sup>Vanderbilt Memory and Alzheimer's Center, Vanderbilt University Medical Center, Nashville, TN, United States, <sup>6</sup>Department of Medicine, University of Washington, Seattle, WA, United States, <sup>7</sup>Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States, <sup>8</sup>Department of Neurology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

Due to needs surrounding rigor and reproducibility, subgroup specific disease knowledge, and questions of external validity, data harmonization is an essential tool in population neuroscience of Alzheimer's disease and related dementias (ADRD). Systematic harmonization of data elements is necessary to pool information from heterogeneous samples, and such pooling allows more expansive evaluations of health disparities, more precise effect estimates, and more opportunities to discover effective prevention or treatment strategies. The key goal of this Tutorial in Population Neuroimaging Curriculum, Instruction, and Pedagogy article is to guide researchers in creating a customized population neuroscience of ADRD harmonization training plan to fit their needs or those of their mentees. We provide brief guidance for retrospective data harmonization of multiple data types in this area, including: (1) clinical and demographic, (2) neuropsychological, and (3) neuroimaging data. Core competencies and skills are reviewed, and resources are provided to fill gaps in training as well as data needs. We close with an example study in which harmonization is a critical tool. While several aspects of this tutorial focus specifically on ADRD, the concepts and resources are likely to benefit population neuroscientists working in a range of research areas.

## KEYWORDS

neurocognitive testing, neuropsychological assessment, cognitive aging, pooling data, neuroimaging, sample size, external validity

## Introduction: Background and rationale

Neuroscience studies, including those focused on Alzheimer's disease and related dementias (ADRD), are often marked by small sample size and highly selective participation. Consequently, many studies do not represent diversity with respect to race/ethnicity, age, comorbid conditions, education, income, or geographic factors, limiting the population relevance of the research. Further, study participation may be influenced by complex combinations of these or other variables such as gender/sex or genetic characteristics.

These selection processes create critical limitations to the quality of the neuroscientific evidence base. Work in the Adolescent Brain Cognitive Development Study ( $N = 11,878$ ) has demonstrated that small sample sizes reduce reproducibility of study findings. Investigators showed that brain-behavior correlations in brain-wide association studies can differ, not only in strength, but worryingly in directionality. Across multiple draws of small sample size the direction of associations may be reversed, but results are more reproducible at large sample sizes (Marek et al., 2022). This has negative implications for both reproducibility and pooled and meta-analyses. Small, highly selected samples also reduce statistical power to test for subgroup effects (e.g., women vs. men; *APOE4* carriers vs. non-carriers; those with vs. without cardiovascular risk factors). Evaluating effects within subgroups, and estimating differences in effects across groups, is essential to determine how to tailor interventions to prevent neurocognitive decline and identify drivers of brain health inequalities.

Increasing sample size is important but cannot fully rectify these limitations; attention to other features of rigorous and robust research designs is needed. First, to obtain larger sample sizes, researchers may wish to combine multiple data sets. In population neuroscience studies, attempts to pool neuroimaging data collected across scanners, sites, and cohorts will be subject to “scanner effects” (Fortin et al., 2018)—technical sources of variance. Scanner effects have been reported in multiple neuroimaging measures derived from MRI and PET (Fortin et al., 2016, 2017, 2018; Yu et al., 2018; Pomponio et al., 2020). Even in multi-center studies using good principles of study design for prospective harmonization, scanner or site-related factors can explain large proportions of variance in the neuroimaging measure (Shinohara et al., 2017). Scanner or site effects, unless corrected, introduce error variance, reducing power to detect effects of interest such as sex interactions (Leek and Storey, 2007). Furthermore, scanner differences likely become associated with study population and compositional differences, rendering this research vulnerable to potential confounding bias. Thus, not accounting analytically for scanner effects can be consequential.

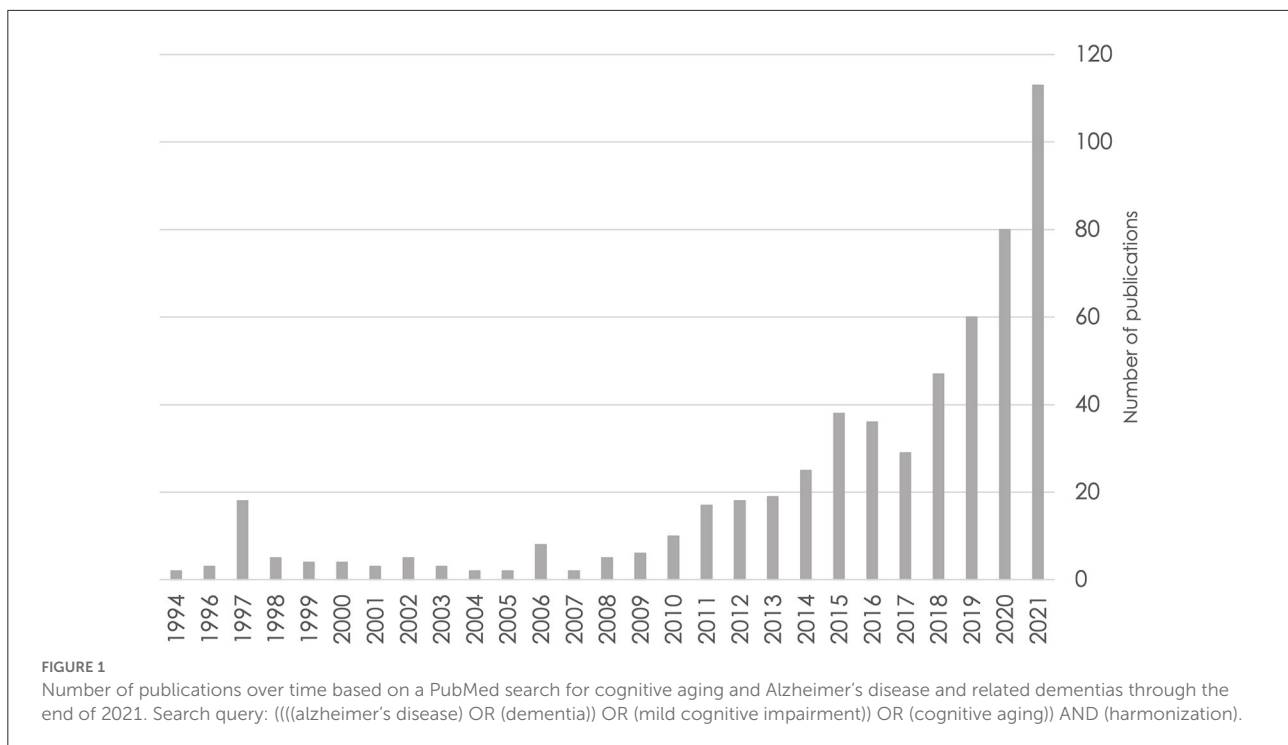
Second, when there is effect modification of an association between an exposure and an outcome, external validity will be impacted by differing prevalence of effect modifiers in the sample vs. the target population (Cole and Stuart, 2010; Keyes and Westreich, 2019). An illustration of this phenomenon can be seen in a comparison of associations in the highly selected Alzheimer's Disease Neuroimaging Initiative (ADNI) study and the community-based Atherosclerosis Risk in Communities (ARIC) study (Gianattasio et al., 2021). Compared to ARIC participants, ADNI participants were more likely to be male, *APOE4+*, married, more highly educated, to have mild cognitive impairment (MCI) or dementia, and less likely to be Black or have a history of hypertension (Gianattasio et al., 2021). Approximately 1/3 of associations varied significantly by study, and some effect size differences were very large (e.g., odds ratio for association of *APOE4* with  $(A\beta)^+$  = 8.6 in ADNI but 2.8 in ARIC), likely due to differences in prevalence of effect modifiers between the two studies (Gianattasio et al., 2021).

Data pooling underlies the ability to address each of the limitations described above, but data cannot be pooled without careful harmonization. As data sharing of numerous neuroscience datasets is becoming more common, the number of publications using harmonization to study brain health is increasing (see Figure 1 for an example from the cognitive aging and ADRD literature). Thus, harmonization is now an essential skill for population neuroscientists.

Harmonization “refers to all efforts to combine data from different sources and provide users with a comparable view of data from different studies” (Data Sharing for Demographic Research, 2022). The aim is to synthesize data to render it similar enough to either be (1) combined for pooled data analysis or (2) analyzed in parallel in the same manner and compared (e.g., reproducing an analysis carried out in one study sample within a different study sample). A “stringent” approach to harmonization involves multiple studies agreeing in advance of data collection to use the same assessments and protocols to prospectively collect the same data (Fortier et al., 2011). On the other hand, a more “flexible” harmonization approach allows for differing assessments and study protocols (Fortier et al., 2011). This approach may be carried out prospectively, but also allows for retrospective harmonization. We focus our comments in this paper specifically on retrospective data harmonization so that early career researcher (ECR) population neuroscientists and others new to the field develop training to make efficient and accurate use of existing data. The audience and objectives of this article are detailed next.

## Audience, environment, objectives, and outcomes

Approaches to retrospective harmonization vary, and there exist few sources of integrated guidance addressing the



varied data types that population neuroscientists commonly use. We aim to address how new researchers can obtain training in this area with this introductory level Tutorial in Population Neuroimaging Curriculum, Instruction, and Pedagogy article. The article is directed to ECRs (students, postdocs, and early career faculty) and others new to the field as well as faculty teaching related courses and mentoring trainees. This information would be especially relevant for researchers writing government and foundation funded training grants. While the example (section Example research plan incorporating population neuroscience of AD/DRD harmonization) and many data sources in Table 2 are AD/DRD-specific, population neuroscientists across the life course and health and disease states are likely to benefit from this tutorial. The learning approach and environment are highly self-directed and based upon a mentored academic model in which the trainee works with mentors to identify gaps in knowledge and training elements to fill those needs and build their skills.

The objectives of this article are to (1) illustrate for researchers, reviewers, and funders the need for population neuroscience data harmonization (Introduction); (2) describe core competencies and skills necessary for harmonization methods of the data types that population neuroscientists should have expertise in; (3) assist readers in identifying their own training gaps and list a selection of relevant learning resources; and (4) use an example AD/DRD research question to further examine considerations in harmonization of demographic/clinical, neuropsychological, and neuroimaging data. Overall, the expected learning outcome of this article

is the creation of a customized population neuroscience of AD/DRD (Ganguli et al., 2018) harmonization learning plan to fit readers' needs or those of their mentees. Evaluative feedback on the developed training plan can be carried out through an iterative process of mentor feedback and revision, while the final evaluation for those writing training grants will be in the form of peer review and feedback on the grant to the investigator. Key elements for positive evaluation are the extent to which the training resources selected by the researcher map onto knowledge gaps and the extent to which the training is integrated into and necessary to answer the researcher's scientific questions.

## Brief introduction to harmonization and core competencies and skills

Detailed guidance regarding overall retrospective data harmonization is provided in the Maelstrom Research Guidelines (Fortier et al., 2017). In this section, we provide a brief introduction to the Maelstrom best practices and related core competencies and skills researchers need to carry out rigorous harmonization in population neuroscience studies. Core competencies and training resources to address researcher knowledge gaps are outlined in Table 1. We recommend that researchers use this table and work with their mentors to identify which core competencies are training gaps (Table 1, left column) and design their own training plan based on the resources provided (Table 1, right column). This should be an iterative process of design and drafting, mentor feedback, and training

TABLE 1 Core competencies/gaps in knowledge and specific training resources in harmonization of population neuroscience studies of ADRD.

Core competency/gap in knowledge	Training resource(s)
<b>Overall harmonization skills</b>	
Multidisciplinary collaboration skills	<ul style="list-style-type: none"> <li>● Scientific leadership and development courses; refer to researcher's own institution offerings on these topics               <ul style="list-style-type: none"> <li>○ Example at the University of Pittsburgh: <a href="https://www.oacd.health.pitt.edu/micro-credential-postdocs">https://www.oacd.health.pitt.edu/micro-credential-postdocs</a></li> <li>Note: Although it is aimed at postdocs, faculty may also enroll</li> </ul> </li> <li>● Mentor(s) with multidisciplinary project leadership experience               <ul style="list-style-type: none"> <li>○ Discuss during mentorship meetings</li> </ul> </li> <li>● Mentors or consultants with specific area expertise               <ul style="list-style-type: none"> <li>○ One-on-one and/or lab meetings</li> <li>○ Directed readings</li> </ul> </li> </ul>
Pre-statistical harmonization and detailed documentation of study design, variables, and variable transformation	<ul style="list-style-type: none"> <li>● Maelstrom retrospective harmonization guidelines (Fortier et al., 2017)</li> <li>● Common data element resources on how variables can be mapped to a harmonized variable               <ul style="list-style-type: none"> <li>○ NIH's common data element (CDE) repository (<a href="https://cde.nlm.nih.gov/">https://cde.nlm.nih.gov/</a>)</li> <li>○ Gateway to Global Aging Data site's (<a href="https://g2aging.org/documentation">https://g2aging.org/documentation</a>) data documentation</li> </ul> </li> <li>● Considerations for harmonization, pooled study design, and analyses (Lesko et al., 2018)</li> </ul>
Reproducible statistical coding, analysis, and power calculations	<ul style="list-style-type: none"> <li>● Statistical coding and analysis coursework at researcher's own institution</li> <li>● GitHub for promoting code reproducibility (<a href="https://github.com/">https://github.com/</a>)</li> <li>● Learn about incorporating code review (Vable et al., 2021)</li> <li>● Mentor(s) with biostatistics expertise</li> <li>● Online coursework</li> <li>● LinkedIn Learning (see if your institution has an institutional subscription)</li> <li>● Statistical Horizons (<a href="https://statisticalhorizons.com/">https://statisticalhorizons.com/</a>) and Code Horizons (<a href="https://codehorizons.com/">https://codehorizons.com/</a>)</li> <li>● Neuroimaging analysis in R through Neuroconductor (<a href="https://neuroconductor.org/courses">https://neuroconductor.org/courses</a>)</li> </ul>
<b>Demographic and clinical variable harmonization skills</b>	
Domain expertise	<ul style="list-style-type: none"> <li>● Mentors or consultants with domain expertise               <ul style="list-style-type: none"> <li>○ One-on-one and/or lab meetings</li> <li>○ Directed readings</li> </ul> </li> </ul>
<b>Neurocognitive assessment harmonization skills</b>	
Domain expertise	<ul style="list-style-type: none"> <li>● A mentor or consultant who is a neuropsychologist               <ul style="list-style-type: none"> <li>○ One-on-one and/or lab meetings</li> <li>○ Directed readings</li> </ul> </li> <li>● Observation of neuropsychological test administration and scoring</li> <li>● Review of test protocols, materials, and stimuli</li> <li>● Suggested readings               <ul style="list-style-type: none"> <li>○ Common tests selected for data sharing in AD research, characteristics, considerations (Bellio et al., 2020)</li> </ul> </li> </ul>
Methods expertise such as standardization, equipercentile equating, multiple imputation, factor analysis, and item response theory-based approaches	<ul style="list-style-type: none"> <li>● A mentor or consultant who has psychometric methods expertise               <ul style="list-style-type: none"> <li>○ One-on-one and/or lab meetings</li> <li>○ Directed readings</li> </ul> </li> <li>● Advanced Psychometric Methods in Cognitive Aging Research (<math>\Psi</math>MCA: <a href="https://psymca.org/">https://psymca.org/</a>) Annual conference and workgroups. Workgroups are application-based admission.</li> <li>● Suggested readings               <ul style="list-style-type: none"> <li>○ Methods to harmonize and combine neuropsychological assessment data for meta-analysis (Griffith et al., 2015)</li> <li>○ Cautions about sum and mean score approaches (standardization) (McNeish and Wolf, 2020)</li> <li>○ Multiple imputation approach to harmonization in AIBL and ADNI (Shishegar et al., 2021)</li> <li>○ Equipercentile equating based approach in the NACC Uniform Data Set neuropsychological test battery (Monsell et al., 2016)</li> <li>○ Detailed paper with workflow on IRT-based neuropsychological data harmonization and co-calibration in studies of cognitive aging and ADRD (Mukherjee et al., 2022)</li> </ul> </li> </ul>

(Continued)

TABLE 1 (Continued)

Core competency/Gap in knowledge	Training resource(s)
<b>Neuroimaging harmonization skills</b>	<ul style="list-style-type: none"> <li>○ IRT-based harmonization of neuropsychological data for an analysis of genetics in late-onset AD subgroups across five studies (Mukherjee et al., 2020)</li> <li>○ Example cross-national harmonization (US and India) (Vonk et al., 2022)</li> <li>○ Cross-national harmonization with brief cognitive assessments, with good discussion of assumptions, alternatives (Kobayashi et al., 2021)</li> </ul>
Image processing skills	<ul style="list-style-type: none"> <li>● A mentor or consultant who has neuroimaging harmonization expertise               <ul style="list-style-type: none"> <li>○ One-on-one and/or lab meetings</li> <li>○ Directed readings</li> </ul> </li> <li>● FreeSurfer neuroimaging processing course: <a href="https://surfer.nmr.mgh.harvard.edu/fswiki/CourseDescription">https://surfer.nmr.mgh.harvard.edu/fswiki/CourseDescription</a></li> <li>● Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL) course and online resources: <a href="https://open.win.ox.ac.uk/pages/fslcourse/website/">https://open.win.ox.ac.uk/pages/fslcourse/website/</a></li> <li>● NIPY: Neuroimaging analysis using Python <a href="https://nipy.org/#">https://nipy.org/#</a></li> <li>● Neurohackademy Lectures: <a href="https://neurohackademy.org/course_type/lectures/">https://neurohackademy.org/course_type/lectures/</a></li> <li>● PET Pharmacokinetics Course: A 3-day course which provides an overview of principles involved in PET kinetic modeling and analysis. The course includes lectures, interactive discussions, and hands-on computer exercises. It runs every other year before NeuroReceptor Mapping (NRM) and before Brain on the intervening years</li> <li>● Rotations in your local MRI and/or PET imaging center</li> </ul>
Pre- and post-statistical harmonization	<ul style="list-style-type: none"> <li>● See <i>Statistical coding and analysis</i> under <i>Overall harmonization skills</i> above</li> <li>● Suggested readings               <ul style="list-style-type: none"> <li>○ Original ComBat paper for gene expression microarray data (Johnson et al., 2007)</li> <li>○ MRI, cross-sectional regression-based harmonization                   <ul style="list-style-type: none"> <li>○ White Stripe image-intensity normalization (Shinohara et al., 2014)</li> <li>○ RAVEL (Removal of Artificial Voxel Effect by Linear regression) (Fortin et al., 2016)</li> <li>○ ComBat for cortical thickness (Fortin et al., 2018)</li> <li>○ ComBat for diffusion tensor imaging (Fortin et al., 2017)</li> <li>○ ComBat for fMRI (Yu et al., 2018)</li> <li>○ ComBat combined with generalized additive models (ComBat-GAM) to address harmonization across a wide age range (Pomponio et al., 2020)</li> <li>○ CovBat to address site effects in covariance (in addition to the more typical site effects in mean and variance) (Chen et al., 2022)</li> </ul> </li> <li>○ MRI, longitudinal regression-based harmonization                   <ul style="list-style-type: none"> <li>○ ComBat for longitudinal cortical thickness (Beer et al., 2020)</li> </ul> </li> <li>○ Machine learning-based harmonization                   <ul style="list-style-type: none"> <li>○ MRI harmonization via MISPEL (Multi-scanner Image harmonization via Structure Preserving Embedding Learning) when more than two scanners are used (Torbati et al., 2021b)</li> <li>○ DeepHarmony addresses MRI contrast differences across two scanners (Dewey et al., 2019)</li> <li>○ mica addresses MRI contrast differences across more than two scanners (Wrobel et al., 2020)</li> </ul> </li> <li>○ PET harmonization                   <ul style="list-style-type: none"> <li>○ Standardization with Centiloids for PET amyloid imaging (Klunk et al., 2015; Royse et al., 2021)</li> </ul> </li> </ul> </li> </ul>

(Continued)

TABLE 1 (Continued)

Core competency/Gap in knowledge	Training resource(s)
	<ul style="list-style-type: none"> <li>○ Non-linear distributional mapping (NoDiM) to address potential non-linearities in amyloid PET tracer measurement scales (Properzi et al., 2019)</li> <li>○ Impact of RAVEL on MRI and PET outcomes (Minhas et al., 2020)</li> <li>○ New and combined pipelines</li> <li>○ Combining RAVEL and ComBat to harmonize across different scanner strengths and remove both variation due to varying imaging intensity and other scanner effects (Torbati et al., 2021a)</li> </ul>

This table should be used by the researcher seeking training to identify the core competencies which are knowledge gaps for them in the left column. The researcher should then select one or more relevant training resources in the right column to address the training needs.

plan revision. This is a key component in career development award proposals. Data resources for ADRD studies are provided in Table 2. These may be sources of data for use in researcher harmonization studies, additional training information, and in some cases, small grants may be available. Some of the data sources listed are highly selected samples/cohorts [e.g., ADNI and the National Alzheimer's Coordinating Center (NACC)], and population neuroscientists are encouraged to evaluate external validity more formally when these samples are used.

Population neuroscience entails a convergence of expertise in epidemiology and neuroscience/neuroimaging (Paus, 2010; Falk et al., 2013). If harmonization is to be undertaken in large epidemiological cohorts with neuroimaging, the population neuroscientist must be skilled in harmonizing demographic and clinical data as well as neuropsychological and neuroimaging data. Next, we provide an introduction to harmonization, core competencies, and necessary skills both overall and for the specific data types population neuroscientists will encounter.

## Overall harmonization

Retrospective population neuroscience harmonization requires expertise across a range of disciplines. As such, multidisciplinary collaboration skills are critical. Pre-statistical harmonization ensures rigorous, high quality research results, and includes selection of appropriate studies and variables to incorporate in the harmonization. This involves creating the DataSchema—the list of variables needed to answer the specific study question—and assessing these variables in each study for harmonization potential (Fortier et al., 2017). The DataSchema includes the key predictors(s), outcome(s), confounders, and effect modifiers of interest. The research question, population studied, and necessary data on exposures, outcomes, and other key variables should dictate which studies are selected for harmonization. Detailed documentation of each contributing study's design characteristics and variables will be needed for

study selection and assessment of variable compatibility with the DataSchema and harmonization potential. This process will require expert input. Specifically, the following study design information should be documented:

- Is the study population representative or volunteer based? Who is the target population for the sample if the population is representative, and if not, who is in the sample?
- What was involved in study participation? Were there different modalities or degrees of participation (e.g., home visit vs. clinic visit)? What were predictors of participation, if known?
- How were measurements conducted? Are there alternative sources of information about people who did not complete the measurement?

Next, data will be requested from contributing studies and transformations of available variables into a common data format will be applied with statistical analysis software. All decisions regarding transformations will need to be documented. If it is unclear how multiple variables can be mapped onto a final harmonized version, referring to NIH's common data element (CDE) repository (<https://cde.nlm.nih.gov/>) or the Gateway to Global Aging Data site's (<https://g2aging.org/documentation>) data documentation may give helpful starting points. Throughout the harmonization and analysis process, we recommend the use of GitHub (<https://github.com/>) paired with code review by another team member (Vable et al., 2021) to promote transparent, reproducible statistical analysis. More details about how to incorporate these features into a cognitive data harmonization workflow can be found in Mukherjee et al. (2022), but are applicable to all data types reviewed in this Tutorial.

Statistical coding and analysis skills will be needed when: running power or sample size calculations to confirm the study is appropriately powered; assessing variable distributions, missingness, harmonized data quality, and representativeness; transforming variables; using imputation

TABLE 2 Data resources for harmonization of population neuroscience studies of ADRD.

Study or consortium and description	Location
<b>Sources of multiple studies that could be harmonized together</b>	
<b>ADDI: Alzheimer's Disease Data Initiative.</b> A source for data sharing and funding with data analysis tools.	<ul style="list-style-type: none"> <li>• <a href="https://www.alzheimersdata.org/">https://www.alzheimersdata.org/</a></li> </ul>
<b>AD Knowledge Portal.</b> An open access data repository established as part of the AMP-AD program (Accelerating Medicines Partnership in Alzheimer's Disease). Shares National Institute on Aging (NIA)-funded translational data in the cognitive aging-to-ADRD spectrum.	<ul style="list-style-type: none"> <li>• <a href="https://adknowledgeportal.synapse.org/">https://adknowledgeportal.synapse.org/</a></li> </ul>
<b>DPUK: Dementias Platform UK.</b> A partnership of public and private organizations based at Oxford University. A source of shared data from 42 cohorts and more than 3 million study participants through their data portal.	<ul style="list-style-type: none"> <li>• <a href="https://www.dementiasplatform.uk/">https://www.dementiasplatform.uk/</a></li> <li>• Data portal: <a href="https://www.dementiasplatform.uk/research-hub/data-portal">https://www.dementiasplatform.uk/research-hub/data-portal</a></li> <li>• DPUK Paper: (Bauermeister et al., 2020)</li> </ul>
<b>GAINN: Global Alzheimer's Association Interactive Network.</b> A source of information on shared datasets with tools for cohort discovery (which cohorts could be used to answer your scientific question of interest).	<ul style="list-style-type: none"> <li>• <a href="https://gaaain.org/">https://gaaain.org/</a></li> <li>• GAINN Paper: (Ashish et al., 2016)</li> </ul>
<b>Human Connectome Project (HCP).</b> A source of data across 20 studies of brain connectomics.	<ul style="list-style-type: none"> <li>• <a href="https://www.humanconnectome.org/">https://www.humanconnectome.org/</a></li> </ul>
<b>Rush Alzheimer's Disease Center.</b> A source of data across multiple Rush cohort studies. Includes: clinical evaluations, cognitive testing, laboratory tests, neuroimaging, etc.	<ul style="list-style-type: none"> <li>• <a href="https://www.radc.rush.edu/">https://www.radc.rush.edu/</a></li> <li>• Religious Orders Study (ROS) Paper: (Bennett et al., 2012a)</li> <li>• Rush Memory and Aging Project (MAP) Papers: (Bennett et al., 2005, 2012b)</li> <li>• Rush Minority Aging Research Study (MARS) Paper: (Barnes et al., 2012)</li> </ul>
<b>Sources of single studies that could be used in a new retrospective data harmonization across studies</b>	
<b>ADNI: Alzheimer's Disease Neuroimaging Initiative.</b> Multicenter study with data from participants across the AD spectrum. Clinical, neuropsychological, and neuroimaging data.	<ul style="list-style-type: none"> <li>• <a href="https://adni.loni.usc.edu/">https://adni.loni.usc.edu/</a></li> <li>• ADNI Papers: (Weiner et al., 2010, 2015, 2017)</li> </ul>
<b>Cambridge Centre for Ageing Neuroscience (Cam-CAN).</b> A study of successful cognitive aging across a wide age range (18–87) with demographic, physiological, neuropsychological, and multimodal neuroimaging data. The neuroimaging study is a sub-study of a larger, population-based study.	<ul style="list-style-type: none"> <li>• <a href="https://camcan-archive.mrc-cbu.cam.ac.uk/dataaccess/">https://camcan-archive.mrc-cbu.cam.ac.uk/dataaccess/</a></li> <li>• Cam-CAN data paper: (Taylor et al., 2017)</li> </ul>
<b>Harvard Aging Brain Study (HABS).</b> A longitudinal study of healthy cognitive aging vs. preclinical AD with clinical, neuropsychological, and multimodal neuroimaging data.	<ul style="list-style-type: none"> <li>• <a href="https://habs.mgh.harvard.edu/researchers/">https://habs.mgh.harvard.edu/researchers/</a></li> <li>• HABS data paper: (Dagley et al., 2017)</li> </ul>
<b>Health and Retirement Study (U.S.).</b> A population representative survey on aging. May be useful for external validity assessment. Includes data on demographics, health (including cognitive data), healthcare services, work and employment, economic status, family structure, and social network retirement.	<ul style="list-style-type: none"> <li>• <a href="https://hrs.isr.umich.edu/about">https://hrs.isr.umich.edu/about</a></li> </ul>
<b>NACC: National Alzheimer's Coordinating Center.</b> Multicenter study with data from more than 45,000 participants from the United States National Institute on Aging funded Alzheimer's Disease Research Centers. Includes resources for trainees in their Research Education Component information (2 <sup>nd</sup> link).	<ul style="list-style-type: none"> <li>• <a href="https://naccdata.org/">https://naccdata.org/</a></li> <li>• Research Education Component (REC) training resources: <a href="https://naccdata.org/adrc-resources/rec-home">https://naccdata.org/adrc-resources/rec-home</a></li> <li>• NACC Papers: (Beekly et al., 2004, 2007)</li> </ul>
<b>OASIS-3.</b> A longitudinal study of cognitive aging to dementia spectrum from the Washington University Knight Alzheimer Disease Research Center. As such, some of this data is likely part of the NACC dataset (see above). Includes clinical, neuropsychological, and multimodal neuroimaging data.	<ul style="list-style-type: none"> <li>• <a href="https://www.oasis-brains.org/">https://www.oasis-brains.org/</a></li> <li>• OASIS-3 data paper: (LaMontagne et al., 2019)</li> </ul>
<b>PResymptomatic Evaluation of Experimental or Novel Treatments for AD (Prevent AD).</b> A longitudinal study of cognitively unimpaired people with a family history of AD. Includes clinical, neuropsychological, biofluid, and neuroimaging data.	<ul style="list-style-type: none"> <li>• Neuroimaging data: <a href="https://openpreventad.loris.ca/">https://openpreventad.loris.ca/</a></li> <li>• Other data: <a href="https://registeredpreventad.loris.ca/">https://registeredpreventad.loris.ca/</a></li> <li>• Prevent AD data paper: (Tremblay-Mercier et al., 2021)</li> </ul>
<b>Projects with retrospective data harmonization actively in progress or data that has already been harmonized; may also have additional data for new harmonization across studies</b>	
<b>CCC: Cross-Cohort Collaboration Consortium.</b> A collaborative across multiple observational cohort studies.	<ul style="list-style-type: none"> <li>• <a href="https://chs-nhlbi.org/node/6539">https://chs-nhlbi.org/node/6539</a></li> </ul>

(Continued)

TABLE 2 (Continued)

Study or consortium and description	Location
<p><b>CHARGE: Cohorts for Heart and Aging Research in Genomic Epidemiology.</b> A source of genomic, risk factor, subclinical disease, and cardiovascular events data across multiple cohorts and more than 50,000 study participants. The NeuroCHARGE work group heads up collaborations with other consortia using cognitive, neuroimaging, and clinical neurological data.</p>	<ul style="list-style-type: none"> <li>• <a href="https://web.chargeconsortium.com/">https://web.chargeconsortium.com/</a></li> <li>• CHARGE Paper: (Psaty et al., 2009)</li> </ul>
<p><b>COSMIC: Cohort Studies of Memory in an International Consortium.</b> A source of data from 47 population-based cohort studies of cognitive aging from 35 countries (~150,000 study participants).</p>	<ul style="list-style-type: none"> <li>• <a href="https://cheba.unsw.edu.au/consortia/cosmic">https://cheba.unsw.edu.au/consortia/cosmic</a></li> <li>• COSMIC Paper: (Sachdev et al., 2013)</li> </ul>
<p><b>ENIGMA: Enhancing Neuroimaging Genetics Through Meta Analysis.</b> An international collaboration of studies evaluating genetics in multiple brain-related conditions. May be a source of data, training materials.</p>	<ul style="list-style-type: none"> <li>• <a href="https://enigma.ini.usc.edu/">https://enigma.ini.usc.edu/</a></li> <li>• Videos including training presentations: <a href="https://bit.ly/3HzDiw">https://bit.ly/3HzDiw</a></li> <li>• ENIGMA Paper: (Stein et al., 2012)</li> <li>• <a href="https://g2aging.org/">https://g2aging.org/</a></li> <li>• The U.S. Health and Retirement Study is one contributing study included on this site.</li> </ul>
<p><b>Gateway to Global Aging Data.</b> A source of cohort and data documentation, questionnaires, and harmonized longitudinal data from the Health and Retirement Studies from around the world (more than 40 countries). Includes data on demographics, health (including cognitive data), healthcare services, work and employment, economic status, family structure, and social network retirement.</p>	<ul style="list-style-type: none"> <li>• <a href="https://dss.niagads.org/">https://dss.niagads.org/</a></li> <li>• Studies included: <a href="https://dss.niagads.org/studies/">https://dss.niagads.org/studies/</a></li> <li>• Future summaries of endophenotype harmonization: <a href="https://www.vmacdata.org/">https://www.vmacdata.org/</a></li> <li>• Future home of harmonized images: <a href="https://loni.usc.edu/">https://loni.usc.edu/</a></li> </ul>
<p><b>National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS).</b> Harmonized genomic and clinical data from 30<sup>+</sup> cohort studies of Alzheimer's disease. Endophenotype harmonization (e.g., cognition, brain MRI, amyloid PET imaging, autopsy measures of neuropathology, vascular risk factors, and fluid biomarkers) was recently funded and will be released in phases over the next 5 years. Through U24 AG074855, <i>Alzheimer's Disease Sequencing Project Phenotype Harmonization Consortium</i></p>	

or latent variable-based harmonization approaches for neuropsychological data; weighting to address selective participation; and carrying out primary and sensitivity analyses to test the major harmonized variables being used in the analyses.

Accurate, transparent reporting is needed when (1) reporting back to original contributing studies, (2) publishing harmonized study results, and (3) providing harmonized data to future users.

Finally, researchers will need to learn effective project leadership and respectful partnership with stakeholders while requesting and working on data and detailed documentation relating to the project (Fortier et al., 2017; Lesko et al., 2018). Additional guidance regarding decisions relevant to these skills in harmonization and pooled analyses may be found in Lesko et al. (2018).

## Clinical and demographic variable harmonization

Clinical and demographic variable harmonization requires domain expertise regarding how to define and collapse across categories. Because flexible, retrospective harmonization requires that data be combined in a way that supports “inferential equivalence” (Fortier et al., 2011), researchers will

need to assess which variables cannot be combined due to compromised measurement validity. For example, consider alternative approaches to assessing prevalent hypertension, a common and nominally straightforward risk factor:

- Self-reported response to “Has your doctor ever told you that you have high blood pressure or hypertension?”
- Selecting hypertension from a list when instructed “Have you ever been diagnosed with any of the following conditions? Please select all that apply.”
- Hypertension recorded in medical records before or after practice guidelines changed in 2017.
- Hypertension based on study measurements of blood pressure.

Can hypertension measured in these heterogeneous ways across studies be conceptualized as the same variable? What is the sensitivity and specificity of each measure for prevalent hypertension and how will the misclassification in each study impact findings? Are there possible pre-processing steps to make the measures more comparable, or bias corrections to reduce the impact of misclassification analytically? This will also require content knowledge about the construct being measured, review of the literature on measurement characteristics of each approach, and good documentation from the studies to know how the measurement was conducted.



## Neuropsychological assessment harmonization

Neuropsychological assessment harmonization requires expertise regarding cognitive domains and processes, knowledge of testing protocols and standards, as well as relevant analytic competencies such as descriptive statistics, data visualization, and variable transformations. Different harmonization strategies include linear transformations/standardization (such as z-scores), equipercentile equating, multiple imputation approaches, and psychometric and latent variable techniques including factor analysis and item response theory (IRT).

Standardization methods, when used to enable data pooling, impose strict assumptions and are only appropriate if the contributing data sets were all representative of the same population, or if populations are known to have the same distribution of the neuropsychological scores. This is rarely plausible because of the strong influence of cognitive function on study participation. Standardization methods can also create circularity when comparing studies. For example, if the impairment definition is based on the within-study distribution of cognition, the prevalence of impairment will be identical for all studies.

Z-scores are a commonly used approach in ADRD research. However, there are several cautions regarding their use. While averaging all z-scored tests within a cognitive domain is an often-used approach to obtain a domain score, this simply puts test scores on the same scale, but has not harmonized them absent confirmation that the tests equivalently measure the underlying domain of interest. Aside from making the distributional assumption mentioned above, simple average domain z-scores assume equal test contribution within domains (e.g., a memory domain z-score comprised of four test z-scores assumes that each test makes up 25% of the memory domain). Some of these problems of sum and mean scores have been recently reviewed (McNeish and Wolf, 2020), and a recent paper illustrates some of these principles by moving from a preclinical Alzheimer cognitive composite (PACC) z-score to a harmonized PACC using IRT (Hampton et al., 2022). A final caution on z-scores is that many neuropsychological assessment batteries change over time, and z-scoring cannot provide inferential equivalence under this circumstance.

Equipercentile equating preserves rank across two assessments, determines the score on one assessment that is equated to the score on the other, and the equated score can then be imputed as the value for the assessment of interest. An example application in ADRD research addressed changing neuropsychological batteries in the NACC Uniform Data Set (UDS), equating the Mini-Mental State Examination (MMSE, Folstein et al., 1975) with the Montreal Cognitive Assessment (MoCA, Nasreddine et al., 2005; Monsell et al., 2016). This approach is only applied when an a priori level of correlation between the two assessments is achieved (here, a correlation

coefficient of  $\geq 0.6$ ). The sample can be divided into a training set to develop the equating and a test set to test the accuracy.

Multiple imputation and IRT do not rely on the population distributional assumption of standardization, but instead rely on the availability of at least one, but ideally many, items that are identical across samples. Multiple imputation relies on the assumption that the associations between items are identical across studies. In multiple imputation approaches, an assessment which was not completed in one study but was completed in others is treated as systematically missing in the stacked dataset. The missing data is then imputed. One recent approach in ADRD has used a random forest model to learn the association between the neuropsychological assessment of interest (non-missing) with all other variables in the dataset and then imputes the missing values of the assessment of interest based on that structure (Shishegar et al., 2021). A starting value for the missing data is preselected and entered into the model. The model then outputs an estimate for the missing data and those initial and new values are compared. This iterative process stops when a predetermined difference between the initial and new values meets a predetermined, sufficiently low threshold.

Item response theory offers methods to assess the assumption that tests (referred to as test *items* in latent variable modeling) are equivalent across studies and estimate latent variables even under modest violations of that assumption. Item response theory models are thus more flexible and rigorous because they build in methods to assess harmonization validity. However, IRT methods still rely on the availability of at least some truly equivalent items. We describe latent variable approaches in detail in Section Neuropsychological assessment harmonization. This approach has been recently detailed in ADRD research in the following publications (Mukherjee et al., 2020, 2022). Suggested readings on these topics and cross-national harmonization approaches are listed in Table 1.

## Neuroimaging harmonization

Population neuroscientists in ADRD research should know that several neuroimaging harmonization approaches exist. For example, standardization approaches of interest include binarization and the Centiloid scale. Many studies make use of binarization to determine positivity or negativity on some biomarker of interest, e.g., A $\beta$ , tau, and neurodegeneration. The limitations of this approach are that it does not deal with processing pipeline differences or scanner effects, loses information from the continuous version of the variable, and allows for only coarse longitudinal change tracking (Lesko et al., 2018). The Centiloid scale is a standardization approach to put different amyloid PET tracers on the same scale, allowing data pooling across tracers (Klunk et al., 2015). The Centiloid scale is framed by 0–100; the 0 anchor represents high-certainty amyloid negative cases, i.e., amyloid level in the brains of healthy

young controls ( $\leq 45$  years of age), and the 100 anchor reflects amyloid level in the brains of typical AD dementia patients. Because these anchor points are averages, the full range of the scale can run from below 0 to  $> 100$ . Transformation equations and amyloid positivity cut points across tracers for cross-sectional and longitudinal analyses have been suggested (Royse et al., 2021). Other approaches such as non-linear distributional mapping (NoDiM) do not assume linearities in amyloid PET tracer measurement scales (Properzi et al., 2019).

There are regression-based statistical harmonization approaches which can be applied either pre- or post-image processing. These methods include RAVEL (Removal of Artificial Voxel Effect by Linear regression, Fortin et al., 2016) and ComBat [combatting batch effects when combining batches of gene expression microarray data (Fortin et al., 2018), with its original use in gene expression data]. Also important is the ability to combine multiple approaches when building neuroimaging and data processing pipelines, such as our own pipelines from MRI pre-processing harmonization using RAVEL to inform PET quantification (Minhas et al., 2020) and approaches incorporating RAVEL and ComBat in the same pipeline to address both MRI image intensity and other scanner effects (Torbati et al., 2021a). We review RAVEL and ComBat in more detail in Section Neuroimaging harmonization.

Finally, machine learning-based neuroimaging harmonization approaches include MISPEL (Multi-scanner Image harmonization via Structure Preserving Embedding Learning), an approach to MRI harmonization developed for use with more than two scanners (Torbati et al., 2021b). DeepHarmony and mica are harmonization approaches that address MRI contrast when two scanners (Dewey et al., 2019) or more are used (Wrobel et al., 2020).

Researchers are advised to note that varying imaging processing softwares and versions are another source of unwanted noise when pooling neuroimaging data (Tudorascu et al., 2016). The strongest approach for dealing with this problem would be to process all images with the same program and version pipeline, and program and version should always be reported. Suggested readings on harmonization approaches dealing with cross-sectional and longitudinal MRI, PET, and new and combined pipelines are provided in Table 1. To illustrate an example of retrospective data harmonization in a population neuroscience of ADRD study, next we describe the workflow of a planned study on sex differences in the AD biomarker cascade.

## Example research plan incorporating population neuroscience of ADRD harmonization

### Background and study aim

Men and women may differ in pathways to AD with critical implications for personalized interventions. Women are

consistently found to have more tau accumulation in the brain than men after accounting for age (Filon et al., 2016; Hohman et al., 2018; Oveisgharan et al., 2018; Buckley et al., 2019a,b, 2020; Luchsinger et al., 2020; Edwards et al., 2021; Palta et al., 2021), with few studies reporting no differences or reverse directionality (Morris et al., 2010; Altmann et al., 2014; Buckley et al., 2019a; Ziontz et al., 2019). Some studies find women also have more brain A $\beta$  than men (Barnes et al., 2005; Jack et al., 2015; Hohman et al., 2018; Liesinger et al., 2018; Oveisgharan et al., 2018; Sundermann et al., 2018; Buckley et al., 2019b; Luchsinger et al., 2020; Rahman et al., 2020; Edwards et al., 2021; Palta et al., 2021), though others do not (Morris et al., 2010; Mielke et al., 2012; Altmann et al., 2014; Filon et al., 2016; Buckley et al., 2018; Sperling et al., 2020; Edwards et al., 2021; Yan et al., 2021). In addition, the relationship between A $\beta$  and tau may vary by sex (effect modification), with this relationship being stronger in women than men (Buckley et al., 2019b, 2020).

In addition to A $\beta$  and tau, cerebral small vessel disease (cSVD), pathology of the small arteries, veins, and capillaries of the brain (Pantoni, 2010; Wardlaw et al., 2013, 2015), may be an important part of the pathway to AD (Kester et al., 2014; McAleese et al., 2015; Tosto et al., 2015; Lee et al., 2016, 2018; Dobbie et al., 2019; Greenberg et al., 2020), and late-life women have a greater burden and risk of cSVD than age-matched men (Longstreth et al., 1998; Uehara et al., 1999; Vermeer et al., 2002; van Dijk et al., 2008; Nyquist et al., 2014). Among those with clinical AD dementia and mixed pathology on postmortem exam, women are likelier than men to have mixed AD and cerebrovascular pathology (Barnes et al., 2019). Several pieces of evidence implicate vascular damage as an important part of AD pathophysiology. First, cerebrovascular disease and AD share risk factors (Dichgans and Zietemann, 2012; Jorgensen et al., 2018; Shaaban et al., 2019). Second, cerebrovascular dysfunction has been shown early in the transition from cognitively unimpaired (CU) to impaired, preceding A $\beta$  deposition (Iturria-Medina et al., 2016). Finally, postmortem clinical-pathologic samples demonstrates that pure AD pathology in AD dementia cases is rare ( $\sim 4\%$ ), while 87% have co-occurring vascular pathology (Kapasi et al., 2017). Therefore, a research and public health focus on modifiable vascular contributors to AD dementia is imperative (Gorelick et al., 2011; Snyder et al., 2015; Corriveau et al., 2016). Understanding the role of cSVD in sex differences in the AD pathophysiological cascade could reveal intervention targets and markers of target engagement which could be used to reduce AD dementia.

One specific aim of this study is to quantify sex differences in the cSVD-AD pathway. We will also explore associations of sex-related factors (e.g., pregnancy history, menopause, hormone use; signs and symptoms of hypogonadism) with cSVD, A $\beta$ , and tau. In this study, we will harmonize data from five longitudinal cohort studies: PiB Normal Aging (Aizenstein et al., 2008), Heart SCORE A and B (neuroimaging sub-studies of the parent study, Heart Strategies Concentrating on Risk Evaluation; Snitz et al., 2020), MYHAT-NI (a neuroimaging sub-study of

the parent study, Monongahela-Youghiogheny Healthy Aging Team; Sullivan et al., 2020), and Human Connectome Project (HCP)-Pitt (Cohen et al., 2021).

## Why is harmonization needed?

First, to detect sex differences and explore sex-specific relationships, we need a larger sample size than available in any individual study. Harmonization will allow us to standardize and pool data across the contributing studies and conduct joint analyses in the larger sample. Second, we would like to enhance the external validity of our estimates. Although all contributing studies have been carried out at the University of Pittsburgh and draw from the local southwestern Pennsylvania population, they were volunteer-based and not population-representative. The selection factors that led to women being included in a study sample likely differed from those operating in men; failing to account for these potentially gives a misleading picture of sex differences. Furthermore, cardiovascular risk factors and common comorbidities of aging (1) are important to consider in the cSVD-AD pathway; (2) are far more common in the population than in highly selected studies; and (3) may vary in prevalence by sex. Since our long-term goal is to improve brain health on the population level, addressing these threats to external validity is crucial to understanding whether sex differences exist at the level of various populations of interest.

Harmonization will allow us to standardize. Weighting and other methods will allow us to adjust the estimates from our study sample to those we should find in the local population. This is further described below in Section Harmonization for external validity analyses.

## A scientific caution

We caution the reader to be thoughtful about the capacity, within harmonization approaches, to remove differences in measures due to certain variables. For example, neuroimaging harmonization approaches can regress out differences due to sex. However, our primary scientific interest is in sex differences and sex-specific pathways, and therefore we wish to preserve the variance in our outcomes that is attributable to sex. The neuroimaging harmonization approaches we describe below can accommodate this if specified in the model. Such decisions must be made while designing the harmonization plan and are fully dependent on each specific scientific question.

## How will harmonization be carried out?

Harmonization will follow the procedures laid out in Sections Brief introduction to harmonization and core competencies and skills and Overall harmonization above,

following Maelstrom guidance (Fortier et al., 2017). We will develop the DataSchema and assess for harmonization potential of the variables we have in the contributing studies by data type.

## Clinical and demographic variable harmonization

We will pool data including demographics, cardiovascular risk factors/common comorbidities of aging (e.g., hypertension, diabetes, congestive heart failure, obesity, smoking, and physical activity) and cognitive status based on content area expertise on the study team. Coding of these characteristics across cohorts will be documented and transformed as needed to develop a harmonized dataset. All cohort studies except MYHAT-NI adjudicate cognitive status yearly with a consensus conference modeled on the University of Pittsburgh Alzheimer's Disease Research Center. Neurologists, psychiatrists, neuropsychologists, and other clinicians review medical history, medications, neurologic and psychiatric exams, neuropsychological testing, and neuroimaging. In MYHAT-NI, cognitive status is initially based on the Clinical Dementia Rating (CDR) (Morris, 1993) scale: cognitively unimpaired (CU), CDR = 0; MCI, CDR = 0.5; dementia, CDR  $\geq$  1. Etiologic diagnosis of all incident dementia cases in MYHAT-NI is then determined by a "virtual consensus conference" (Lee et al., 2020), with inter-disciplinary experts reviewing clinical data online and making etiologic diagnostic ratings.

## Neuropsychological assessment harmonization

All contributing studies administer detailed neuropsychological assessments yearly. We wish to conduct a rigorous harmonization to allow data pooling across contributing studies and the possibility of comparisons with other studies in the future. In our case, all contributing studies are recruited from the same overall geographic population, with many aspects of shared language and culture, and multiple equivalent neuropsychological test items across studies. Furthermore, our contributing studies are longitudinal, and some are longstanding with potential for changing test batteries over time. Item response theory best meets our needs for inferential equivalence across studies and over changing test batteries, and the data meet the requirements for IRT.

We will follow the rigorous IRT-based approach recently detailed in a cognitive harmonization workflow paper (Mukherjee et al., 2022) and initially developed for a genetics of late-onset AD across five studies (Mukherjee et al., 2020). We refer readers to Mukherjee et al. (2022) for detailed methods. This approach has been calibrated across the full spectrum of cognitive diagnoses from CU to AD dementia. Briefly, first, test administration and scoring procedures across contributing studies are fully documented to understand potentially

important differences. Test variables are assigned to cognitive domains (memory, executive function, language, visuospatial abilities) by neuropsychologists. Next, we will assess data distributions, recoding reverse coded items, and missingness as part of the data quality control step. All transformations will be documented. We will confirm that the tests load onto their respective cognitive domain factors and obtain the best fitting model using confirmatory factor analysis. The scores will be co-calibrated to other studies with overlapping measurements, such as the Adult Changes in Thought study, NACC, the Framingham Study, and ADNI. Co-calibrations can be daisy-chained together, so even studies with no overlapping measurements with our study may still be compared, although each step introduces uncertainty in the measurement quality.

## Neuroimaging harmonization

The inter-visit interval for neuroimaging is 24 months in all contributing studies. MRI has been obtained across all studies using 3T Siemens scanners (1 TIM TRIO and 3 Prismas).  $^{11}\text{C}$ -PiB is used in all cohorts for A $\beta$  PET, and 4/5 studies collect tau PET using  $^{18}\text{F}$ -AV-1451. All contributing study PET scans were conducted on either a Siemens Biograph mCT PET/CT or a Siemens/CTI ECAT HR+ PET. Our neuroimaging harmonization approach includes both pre- and post-processing regression-based harmonization methods, described below. The harmonized neuroimaging values are then used as the outcomes of interest in the overall substantive statistical analyses.

### Pre-processing neuroimaging harmonization

The images at all visits for the combined pool of participants ( $N = 870$  cross-sectional and 645 longitudinal) will be pre-processed together as follows. We will use RAVEL (Fortin et al., 2016) applied to the T1 and FLAIR MRI scans. RAVEL incorporates two key steps that differ from typical MRI pre-processing pipelines. First, following typical segmentation into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), transformation into a common anatomical space, and skull stripping, White Stripe image-intensity normalization (Shinohara et al., 2014) is applied. This is important because MRI intensity units are not standard across scanners, and this step removes variation due simply to arbitrary unit differences between visits and scanners. This approach z-scores voxel intensity based on the mean and standard deviation of intensity in normal appearing white matter (NAWM). Because of its large size, NAWM is less susceptible to partial volume effects and represents biologically healthy tissue. This processing is rapid—typically under 5 s per scan on a laptop (Shinohara et al., 2014). Second, a control region of interest (ROI) is identified (here, CSF) where image intensity should not vary as a function of AD or other biological variables of interest. Any variance seen in this ROI represents non-biological differences (e.g., scanner effects); this variance is regressed out in voxel-level

linear regression. One advantage of RAVEL over other image intensity normalization methods is that it maintains variance due to biological factors of interest to study (e.g. sex, age, hypertension) by including them in the voxel-level regression (Fortin et al., 2016). This is critical as our primary interest is in examining sex differences. RAVEL is implemented in R statistical analysis software (R). After RAVEL is applied, the MRI is warped back to person-specific space, and in our pipeline, FreeSurfer software is applied to perform MR bias correction, automated ROI parcellation, and tissue segmentation. These harmonized MRIs will then be used for analysis of cSVD and to derive the PET ROIs according to our RAVEL to PET pipeline (Minhas et al., 2020).

### Post-processing neuroimaging harmonization

RAVEL harmonization may be more important for MRI markers than PET markers (Minhas et al., 2020). If an alternative PET harmonization approach is needed, we will apply post-processing statistical harmonization methods using ComBat (Fortin et al., 2018). In this case, regression-based harmonization is performed using PET SUVR based on FreeSurfer ROIs which have not been harmonized with RAVEL. Similar to the second step of RAVEL, it removes scanner effects while maintaining participant characteristic-related variance of interest when these variables are added to the harmonization model. ComBat is computationally efficient to use (Fortin et al., 2018) and can remove non-biological sources of variance when harmonized data acquisition protocols were not used (Fortin et al., 2018). It has been demonstrated to be effective when applied to multiple neuroimaging measures including GM volume, cortical thickness, diffusion tensor imaging, and fMRI (Fortin et al., 2017, 2018; Yu et al., 2018; Pomponio et al., 2020) and is implemented in R (Fortin, 2020).

## Harmonization for external validity analyses

We will leverage availability of data from a local county-wide population-representative sample from the Behavioral Risk Factor Surveillance System (BRFSS; <https://www.cdc.gov/brfss/>). Carrying out these analyses requires that our overall ADRD dataset and the BRFSS dataset are stacked, thus requiring harmonization. We can then use variables that were measured in both the BRFSS and in our contributing ADRD studies to correct for selection (e.g., demographics and cardiovascular risk factors/common comorbidities of aging). For example, if men in the ADRD studies are much more likely to be married than men in the BRFSS, whereas for women marital status does not differ substantially, we can upweight unmarried men in our ADRD studies so the joint distribution of sex and marital status in our analyzed data matches the joint distribution in the population. We will use the two main G-methods for these external validity analyses: (1) inverse probability weighting (IPW) for sampling and (2) G-computation (Bareinboim and Pearl, 2016; Lesko

et al., 2017; Westreich et al., 2017). This will allow us to adjust the estimates from our study sample and make them more generalizable to the target population.

## Conclusion

Studies of ADRD have proliferated and data sharing has increased and will be an NIH requirement as of January 2023 (see NIH Policy for Data Management and Sharing: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>). While several data resources listed in Table 2 provide data already harmonized and ready to use in analysis, many of the data resources listed are opportunities to discover and request original data only. Investigators requesting data will very often need to harmonize the data themselves, yet without access to ready guidance as to how to carry out and report the retrospective harmonization according to best practices in the field, especially across the multiple types of data ADRD population neuroscientists work with. This is a recipe for an “anything goes” approach, and it has been shown that harmonization and reporting practices vary widely across studies (Fortier et al., 2017). We hope this Tutorial will help begin to fill this gap. Finally, we recommend that new and existing investigators help develop further best practices and training materials for our field to standardize and enhance rigor across approaches.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

## Author contributions

CS conceived this article and wrote the first draft. DT, MG, AC, RT, HS, TH, SM, LY, and BS contributed to the first draft and reviewed critically the entire manuscript. All authors contributed to the article and approved the submitted version.

## References

- Aizenstein, H. J., Nebes, R. D., Saxton, J. A., Price, J. C., Mathis, C. A., Tsopelas, N. D., et al. (2008). Frequent amyloid deposition without significant cognitive impairment among the elderly. *Arch. Neurol.* 65, 1509–1517. doi: 10.1001/archneur.65.11.1509
- Altmann, A., Tian, L., Henderson, V. W., and Greicius, M. D. (2014). Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann. Neurol.* 75, 563–573. doi: 10.1002/ana.24135
- Ashish, N., Bhatt, P., and Toga, A. W. (2016). Global data sharing in Alzheimer disease research. *Alzheimer Dis. Assoc. Disord.* 30, 160–168. doi: 10.1097/wad.0000000000000121

## Funding

CS was supported by Award Number K01AG071849 from the National Institute on Aging of the US National Institutes of Health for work in this publication.

## Acknowledgments

We thank Dr. Paul Crane for discussions regarding methods for harmonization of neuropsychological test data.

## Conflict of interest

Author RT is a consultant for Bayer, Astellas, and Happify. Author HS is a full-time employee of the Alzheimer's Association. Author TH is a member of the Scientific Advisory Board for Vivid Genomics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Author disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

- Bareinboim, E., and Pearl, J. (2016). Causal inference and the data-fusion problem. *Proc. Natl. Acad. Sci. USA.* 113, 7345–7352. doi: 10.1073/pnas.1510507113

- Barnes, L. L., Lamar, M., and Schneider, J. A. (2019). Sex differences in mixed neuropathologies in community-dwelling older adults. *Brain Res.* 1719, 11–16. doi: 10.1016/j.brainres.2019.05.028

- Barnes, L. L., Shah, R. C., Aggarwal, N. T., Bennett, D. A., and Schneider, J. A. (2012). The Minority Aging Research Study: ongoing efforts to obtain brain donation in African Americans without dementia. *Curr. Alzheimer Res.* 9, 734–745. doi: 10.2174/156720512801322627

- Barnes, L. L., Wilson, R. S., Bienias, J. L., Schneider, J. A., Evans, D. A., and Bennett, D. A. (2005). Sex differences in the clinical manifestations of Alzheimer disease pathology. *Arch. Gen. Psychiatry* 62, 685–691. doi: 10.1001/archpsyc.62.6.685
- Bauermeister, S., Orton, C., Thompson, S., Barker, R. A., Bauermeister, J. R., Ben-Shlomo, Y., et al. (2020). The Dementias Platform UK (DPUK) data portal. *Eur. J. Epidemiol.* 35, 601–611. doi: 10.1007/s10654-020-00633-4
- Beekly, D. L., Ramos, E. M., Lee, W. W., Deitrich, W. D., Jacka, M. E., Wu, J., et al. (2007). The National Alzheimer's Coordinating Center (NACC) database: the Uniform Data Set. *Alzheimer Dis. Assoc. Disord.* 21, 249–258. doi: 10.1097/WAD.0b013e318142774e
- Beekly, D. L., Ramos, E. M., van Belle, G., Deitrich, W., Clark, A. D., Jacka, M. E., et al. (2004). The National Alzheimer's Coordinating Center (NACC) Database: an Alzheimer disease database. *Alzheimer Dis. Assoc. Disord.* 18, 270–277.
- Beer, J. C., Tustison, N. J., Cook, P. A., Davatzikos, C., Sheline, Y. I., Shinohara, R. T., et al. (2020). Longitudinal ComBat: a method for harmonizing longitudinal multi-scanner imaging data. *Neuroimage* 220, 117129. doi: 10.1016/j.neuroimage.2020.117129
- Bellio, M., Oxtoby, N. P., Walker, Z., Henley, S., Ribbens, A., Blandford, A., et al. (2020). Analyzing large Alzheimer's disease cognitive datasets: considerations and challenges. *Alzheimers Dement. (Amst.)* 12, e12135. doi: 10.1002/dad2.12135
- Bennett, D. A., Schneider, J. A., Arvanitakis, Z., and Wilson, R. S. (2012a). Overview and findings from the religious orders study. *Curr. Alzheimer Res.* 9, 628–645. doi: 10.2174/156720512801322573
- Bennett, D. A., Schneider, J. A., Buchman, A. S., Barnes, L. L., Boyle, P. A., and Wilson, R. S. (2012b). Overview and findings from the Rush Memory and Aging Project. *Curr. Alzheimer Res.* 9, 646–663. doi: 10.2174/156720512801322663
- Bennett, D. A., Schneider, J. A., Buchman, A. S., Mendes de Leon, C., Bienias, J. L., and Wilson, R. S. (2005). The Rush Memory and Aging Project: study design and baseline characteristics of the study cohort. *Neuroepidemiology* 25, 163–175. doi: 10.1159/000087446
- Buckley, R. F., Mormino, E. C., Amariglio, R. E., Properzi, M. J., Rabin, J. S., Lim, Y. Y., et al. (2018). Sex, amyloid, and APOE epsilon4 and risk of cognitive decline in preclinical Alzheimer's disease: findings from three well-characterized cohorts. *Alzheimers Dement.* 14, 1193–1203. doi: 10.1016/j.jalz.2018.04.010
- Buckley, R. F., Mormino, E. C., Chhatwal, J., Schultz, A. P., Rabin, J. S., Rentz, D. M., et al. (2019a). Associations between baseline amyloid, sex, and APOE on subsequent tau accumulation in cerebrospinal fluid. *Neurobiol. Aging* 78, 178–185. doi: 10.1016/j.neurobiolaging.2019.02.019
- Buckley, R. F., Mormino, E. C., Rabin, J. S., Hohman, T. J., Landau, S., Hanseeuw, B. J., et al. (2019b). Sex differences in the association of global amyloid and regional tau deposition measured by positron emission tomography in clinically normal older adults. *JAMA Neurol.* 76, 542–551. doi: 10.1001/jamaneurol.2018.4693
- Buckley, R. F., Scott, M. R., Jacobs, H. I. L., Schultz, A. P., Properzi, M. J., Amariglio, R. E., et al. (2020). Sex mediates relationships between regional tau pathology and cognitive decline. *Ann. Neurol.* 88, 921–932. doi: 10.1002/ana.25878
- Chen, A. A., Beer, J. C., Tustison, N. J., Cook, P. A., Shinohara, R. T., and Shou, H. (2022). Mitigating site effects in covariance for machine learning in neuroimaging data. *Hum. Brain Mapp.* 43, 1179–1195. doi: 10.1002/hbm.25688
- Cohen, A. D., Bruña, R., Chang, Y. F., Cheng, Y., Doman, J., Huppert, T., et al. (2021). Connectomics in brain aging and dementia - the background and design of a study of a connectome related to human disease. *Front. Aging Neurosci.* 13, 669490. doi: 10.3389/fnagi.2021.669490
- Cole, S. R., and Stuart, E. A. (2010). Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 trial. *Amer. J. Epidemiol.* 172, 107–115. doi: 10.1093/aje/kwq084
- Corriveau, R. A., Bosetti, F., Emr, M., Gladman, J. T., Koenig, J. I., Moy, C. S., et al. (2016). The science of Vascular Contributions to Cognitive Impairment and Dementia (VCID): a framework for advancing research priorities in the cerebrovascular biology of cognitive decline. *Cell Mol. Neurobiol.* 36, 281–288. doi: 10.1007/s10571-016-0334-7
- Dagley, A., LaPoint, M., Huijbers, W., Hedden, T., McLaren, D. G., Chatwal, J. P., et al. (2017). Harvard aging brain study: dataset and accessibility. *Neuroimage* 144(Pt B), 255–258. doi: 10.1016/j.neuroimage.2015.03.069
- Data Sharing for Demographic Research (2022). *Data Harmonization*. Available Online at: <https://www.icpsr.umich.edu/web/pages/DSDR/harmonization.html> (accessed June 19, 2022).
- Debette, S., Schilling, S., Duperron, M. G., Larsson, S. C., and Markus, H. S. (2019). Clinical significance of magnetic resonance imaging markers of vascular brain injury: a systematic review and meta-analysis. *JAMA Neurol.* 76, 81–94. doi: 10.1001/jamaneurol.2018.3122
- Dewey, B. E., Zhao, C., Reinhold, J. C., Carass, A., Fitzgerald, K. C., Sotirchos, E. S., et al. (2019). DeepHarmony: a deep learning approach to contrast harmonization across scanner changes. *Magn. Reson. Imaging* 64, 160–170. doi: 10.1016/j.mri.2019.05.041
- Dichgans, M., and Zietemann, V. (2012). Prevention of vascular cognitive impairment. *Stroke* 43, 3137–3146. doi: 10.1161/strokeaha.112.651778
- Edwards, L., La Joie, R., Iaccarino, L., Strom, A., Baker, S. L., Casaletto, K. B., et al. (2021). Multimodal neuroimaging of sex differences in cognitively impaired patients on the Alzheimer's continuum: greater tau-PET retention in females. *Neurobiol. Aging* 105, 86–98. doi: 10.1016/j.neurobiolaging.2021.04.003
- Falk, E. B., Hyde, L. W., Mitchell, C., Faul, J., Gonzalez, R., Heitzeg, M. M., et al. (2013). What is a representative brain? Neuroscience meets population science. *Proc. Natl. Acad. Sci. USA.* 110, 17615–17622. doi: 10.1073/pnas.1310134110
- Filon, J. R., Intorcica, A. J., Sue, L. I., Vazquez Arreola, E., Wilson, J., Davis, K. J., et al. (2016). Gender differences in Alzheimer disease: brain atrophy, histopathology burden, and cognition. *J. Neuropathol. Exp. Neurol.* 75, 748–754. doi: 10.1093/jnen/nlw047
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Fortier, I., Doiron, D., Burton, P., and Raina, P. (2011). Invited commentary: consolidating data harmonization—how to obtain quality and applicability? *Am. J. Epidemiol.* 174, 261–264; author reply 265–266. doi: 10.1093/aje/kwr194
- Fortier, I., Raina, P., Van den Heuvel, E. R., Griffith, L. E., Craig, C., Saliba, M., et al. (2017). Maelstrom research guidelines for rigorous retrospective data harmonization. *Int. J. Epidemiol.* 46, 103–105. doi: 10.1093/ije/dyw075
- Fortin, J. P. (2020). *ComBat Harmonization: Harmonization of Multi-Site Imaging Data With ComBat*. Available online at: <https://github.com/jfortin1/ComBatHarmonization> (accessed April 30, 2020).
- Fortin, J. P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., et al. (2018). Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* 167, 104–120. doi: 10.1016/j.neuroimage.2017.11.024
- Fortin, J. P., Parker, D., Tunc, B., Watanabe, T., Elliott, M. A., Ruparel, K., et al. (2017). Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* 161, 149–170. doi: 10.1016/j.neuroimage.2017.08.047
- Fortin, J. P., Sweeney, E. M., Muschelli, J., Crainiceanu, C. M., and Shinohara, R. T. (2016). Removing inter-subject technical variability in magnetic resonance imaging studies. *Neuroimage* 132, 198–212. doi: 10.1016/j.neuroimage.2016.02.036
- Ganguli, M., Albanese, E., Seshadri, S., Bennett, D. A., Lyketsos, C., Kukull, W. A., et al. (2018). Population neuroscience: dementia epidemiology serving precision medicine and population health. *Alzheimer Dis. Assoc. Disord.* 32, 1–9. doi: 10.1097/wad.0000000000000237
- Gianattasio, K. Z., Bennett, E. E., Wei, J., Mehrotra, M. L., Mosley, T., Gottesman, R. F., et al. (2021). Generalizability of findings from a clinical sample to a community-based sample: a comparison of ADNI and ARIC. *Alzheimers Dement.* 17, 1265–1276. doi: 10.1002/alz.12293
- Gorelick, P. B., Scuteri, A., Black, S. E., Decarli, C., Greenberg, S. M., Iadecola, C., et al. (2011). Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American heart association/American stroke association. *Stroke* 42, 2672–2713. doi: 10.1161/STR.0b013e3182299496
- Greenberg, S. M., Bacskaï, B. J., Hernandez-Guillamon, M., Pruzin, J., Sperling, R., and van Veluw, S. J. (2020). Cerebral amyloid angiopathy and Alzheimer disease - one peptide, two pathways. *Nat. Rev. Neurol.* 16, 30–42. doi: 10.1038/s41582-019-0281-2
- Griffith, L. E., van den Heuvel, E., Fortier, I., Sohler, N., Hofer, S. M., Payette, H., et al. (2015). Statistical approaches to harmonize data on cognitive measures in systematic reviews are rarely reported. *J. Clin. Epidemiol.* 68, 154–162. doi: 10.1016/j.jclinepi.2014.09.003
- Hampton, O. L., Mukherjee, S., Properzi, M. J., Schultz, A. P., Crane, P. K., Gibbons, L. E., et al. (2022). Harmonizing the preclinical Alzheimer cognitive composite for multicohort studies. *Neuropsychology*. doi: 10.1037/neu0000833. [Epub ahead of print].
- Hohman, T. J., Dumitrescu, L., Barnes, L. L., Thambisetty, M., Beecham, G., Kunkle, B., et al. (2018). Sex-specific association of apolipoprotein E with cerebrospinal fluid levels of tau. *JAMA Neurol.* 75, 989–998. doi: 10.1001/jamaneurol.2018.0821
- Iturria-Medina, Y., Sotero, R. C., Toussaint, P. J., Mateos-Pérez, J. M., and Evans, A. C. (2016). Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat. Commun.* 7, 11934. doi: 10.1038/ncomms11934

- Jack, C. R. Jr., Wiste, H. J., Weigand, S. D., Knopman, D. S., Vemuri, P., Mielke, M. M., et al. (2015). Age, sex, and APOE epsilon4 effects on memory, brain structure, and beta-amyloid across the adult life span. *JAMA Neurol.* 72, 511–519. doi: 10.1001/jamaneurol.2014.4821
- Johnson, W. E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118–127. doi: 10.1093/biostatistics/kxj037
- Jorgensen, D. R., Shaaban, C. E., Wiley, C. A., Gianaros, P. J., Mettenberg, J., and Rosano, C. (2018). A population neuroscience approach to the study of cerebral small vessel disease in midlife and late life: an invited review. *Am. J. Physiol. Heart Circ. Physiol.* 314, H1117–H1136. doi: 10.1152/ajpheart.00535.2017
- Kapasi, A., DeCarli, C., and Schneider, J. A. (2017). Impact of multiple pathologies on the threshold for clinically overt dementia. *Acta Neuropathol.* 134, 171–186. doi: 10.1007/s00401-017-1717-7
- Kester, M. I., Goos, J. D., Teunissen, C. E., Benedictus, M. R., Bouwman, F. H., Wattjes, M. P., et al. (2014). Associations between cerebral small-vessel disease and Alzheimer disease pathology as measured by cerebrospinal fluid biomarkers. *JAMA Neurol.* 71, 855–862. doi: 10.1001/jamaneurol.2014.754
- Keyes, K. M., and Westreich, D. (2019). UK Biobank, big data, and the consequences of non-representativeness. *Lancet* 393, 1297. doi: 10.1016/s0140-6736(18)33067-8
- Klunk, W. E., Koeppe, R. A., Price, J. C., Benzinger, T. L., Devous, M. D. Sr., Jagust, W. J., et al. (2015). The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement.* 11, 1.e4–15.e4. doi: 10.1016/j.jalz.2014.07.003
- Kobayashi, L. C., Gross, A. L., Gibbons, L. E., Tommet, D., Sanders, R. E., Choi, S. E., et al. (2021). You say tomato, i say radish: can brief cognitive assessments in the U.S. health retirement study be harmonized with its international partner studies? *J. Gerontol. B Psychol. Sci. Soc. Sci.* 76, 1767–1776. doi: 10.1093/geronb/gbaa205
- LaMontagne, P. J., Benzinger, T. L., Morris, J. C., Keefe, S., Hornbeck, R., Xiong, C., et al. (2019). OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *medRxiv* doi: 10.1101/2019.12.13.19014902
- Lee, J., Ganguli, M., Weerman, A., Chien, S., Lee, D. Y., Varghese, M., et al. (2020). Online clinical consensus diagnosis of dementia: development and validation. *J. Am. Geriatr. Soc.* 68(Suppl 3), S54–S59. doi: 10.1111/jgs.16736
- Lee, S., Viqar, F., Zimmerman, M. E., Narkhede, A., Tosto, G., Benzinger, T. L., et al. (2016). White matter hyperintensities are a core feature of Alzheimer's disease: evidence from the dominantly inherited Alzheimer network. *Ann. Neurol.* 79, 929–939. doi: 10.1002/ana.24647
- Lee, S., Zimmerman, M. E., Narkhede, A., Nasrabad, S. E., Tosto, G., Meier, I. B., et al. (2018). White matter hyperintensities and the mediating role of cerebral amyloid angiopathy in dominantly-inherited Alzheimer's disease. *PLoS ONE* 13, e0195838. doi: 10.1371/journal.pone.0195838
- Leek, J. T., and Storey, J. D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* 3, 1724–1735. doi: 10.1371/journal.pgen.0030161
- Lesko, C. R., Buchanan, A. L., Westreich, D., Edwards, J. K., Hudgens, M. G., and Cole, S. R. (2017). Generalizing study results: a potential outcomes perspective. *Epidemiology* 28, 553–561. doi: 10.1097/ede.0000000000000664
- Lesko, C. R., Jacobson, L. P., Althoff, K. N., Abraham, A. G., Gange, S. J., Moore, R. D., et al. (2018). Collaborative, pooled and harmonized study designs for epidemiologic research: challenges and opportunities. *Int. J. Epidemiol.* 47, 654–668. doi: 10.1093/ije/dyx283
- Liesinger, A. M., Graff-Radford, N. R., Duara, R., Carter, R. E., Hanna Al-Shaikh, F. S., Koga, S., et al. (2018). Sex and age interact to determine clinicopathologic differences in Alzheimer's disease. *Acta Neuropathol.* 136, 873–885. doi: 10.1007/s00401-018-1908-x
- Longstreth, W. T. Jr., Bernick, C., Manolio, T. A., Bryan, N., Jungreis, C. A., and Price, T. R. (1998). Lacunar infarcts defined by magnetic resonance imaging of 3660 elderly people: the Cardiovascular Health Study. *Arch. Neurol.* 55, 1217–1225.
- Luchsinger, J. A., Palta, P., Rippon, B., Soto, L., Ceballos, F., Pardo, M., et al. (2020). Sex differences in *in vivo* Alzheimer's disease neuropathology in late middle-aged hispanics. *J. Alzheimers Dis.* 74, 1243–1252. doi: 10.3233/jad-191183
- Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., et al. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature* 603, 654–660. doi: 10.1038/s41586-022-04492-9
- McAleese, K. E., Firbank, M., Dey, M., Colloby, S. J., Walker, L., Johnson, M., et al. (2015). Cortical tau load is associated with white matter hyperintensities. *Acta Neuropathol. Commun.* 3, 60. doi: 10.1186/s40478-015-0240-0
- McNeish, D., and Wolf, M. G. (2020). Thinking twice about sum scores. *Behav. Res. Methods* 52, 2287–2305. doi: 10.3758/s13428-020-01398-0
- Mielke, M. M., Wiste, H. J., Weigand, S. D., Knopman, D. S., Lowe, V. J., Roberts, R. O., et al. (2012). Indicators of amyloid burden in a population-based study of cognitively normal elderly. *Neurology* 79, 1570–1577. doi: 10.1212/WNL.0b013e31826e2696
- Minhas, D. S., Yang, Z., Muschelli, J., Laymon, C. M., Mettenberg, J. M., Zammit, M. D., et al. (2020). "Statistical methods for processing neuroimaging data from two different sites with a down syndrome population application," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (Lisbon: Springer), 367–379.
- Monsell, S. E., Dodge, H. H., Zhou, X. H., Bu, Y., Besser, L. M., Mock, C., et al. (2016). Results from the NACC Uniform Data Set neuropsychological battery crosswalk study. *Alzheimer Dis. Assoc. Disord.* 30, 134–139. doi: 10.1097/wad.0000000000000111
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414.
- Morris, J. C., Roe, C. M., Xiong, C., Fagan, A. M., Goate, A. M., Holtzman, D. M., et al. (2010). APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. *Ann. Neurol.* 67, 122–131. doi: 10.1002/ana.21843
- Mukherjee, S., Choi, S.-E., Lee, M., Scollard, P., Trittschuh, E. H., Mez, J., et al. (2022). Cognitive domain harmonization and co-calibration in studies of older adults. *Neuropsychology*. doi: 10.1037/neu0000835
- Mukherjee, S., Mez, J., Trittschuh, E. H., Saykin, A. J., Gibbons, L. E., Fardo, D. W., et al. (2020). Genetic data and cognitively defined late-onset Alzheimer's disease subgroups. *Mol. Psychiatry* 25, 2942–2951. doi: 10.1038/s41380-018-0298-8
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x
- Nyquist, P. A., Bilgel, M. S., Gottesman, R., Yanek, L. R., Moy, T. F., Becker, L. C., et al. (2014). Extreme deep white matter hyperintensity volumes are associated with African American race. *Cerebrovasc. Dis.* 37, 244–250. doi: 10.1159/000358117
- Oveisgharan, S., Arvanitakis, Z., Yu, L., Farfel, J., Schneider, J. A., and Bennett, D. A. (2018). Sex differences in Alzheimer's disease and common neuropathologies of aging. *Acta Neuropathol.* 136, 887–900. doi: 10.1007/s00401-018-1920-1
- Palta, P., Rippon, B., Tahmi, M., Pardo, M., Johnson, A., Tomljanovic, Z., et al. (2021). Sex differences in *in vivo* tau neuropathology in a multiethnic sample of late middle-aged adults. *Neurobiol. Aging* 103, 109–116. doi: 10.1016/j.neurobiolaging.2021.03.007
- Pantoni, L. (2010). Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol.* 9, 689–701. doi: 10.1016/s1474-4422(10)70104-6
- Paus, T. (2010). Population neuroscience: why and how. *Hum. Brain Mapp.* 31, 891–903. doi: 10.1002/hbm.21069
- Pomponio, R., Erus, G., Habes, M., Doshi, J., Srinivasan, D., Mamourian, E., et al. (2020). Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *Neuroimage* 208, 116450. doi: 10.1016/j.neuroimage.2019.116450
- Properzi, M. J., Buckley, R. F., Chhatwal, J. P., Donohue, M. C., Lois, C., Mormino, E. C., et al. (2019). Nonlinear Distributional Mapping (NoDiM) for harmonization across amyloid-PET radiotracers. *Neuroimage* 186, 446–454. doi: 10.1016/j.neuroimage.2018.11.019
- Psaty, B. M., O'Donnell, C. J., Gudnason, V., Lunetta, K. L., Folsom, A. R., Rotter, J. I., et al. (2009). Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ. Cardiovasc. Genet.* 2, 73–80. doi: 10.1161/circgenetics.108.829747
- Rahman, A., Schelbaum, E., Hoffman, K., Diaz, I., Hristov, H., Andrews, R., et al. (2020). Sex-driven modifiers of Alzheimer risk: a multimodality brain imaging study. *Neurology* 95, e166–e178. doi: 10.1212/wnl.00000000000009781
- Royle, S. K., Minhas, D. S., Lopresti, B. J., Murphy, A., Ward, T., Koeppe, R. A., et al. (2021). Validation of amyloid PET positivity thresholds in centiloids: a multisite PET study approach. *Alzheimers Res. Ther.* 13, 99. doi: 10.1186/s13195-021-00836-1
- Sachdev, P. S., Lipnicki, D. M., Kochan, N. A., Crawford, J. D., Rockwood, K., Xiao, S., et al. (2013). COSMIC (Cohort Studies of Memory in an International Consortium): an international consortium to identify risk and protective factors and biomarkers of cognitive ageing and dementia in diverse ethnic and sociocultural groups. *BMC Neurol.* 13, 165. doi: 10.1186/1471-2377-13-165
- Shaaban, C. E., Jia, Y., Chang, C. H., and Ganguli, M. (2019). Independent and joint effects of vascular and cardiometabolic risk factor pairs for risk of all-cause dementia: a prospective population-based study. *Int. Psychogeriatr.* 31, 1421–1432. doi: 10.1017/S1041610219001066

- Shinohara, R. T., Oh, J., Nair, G., Calabresi, P. A., Davatzikos, C., Doshi, J., et al. (2017). Volumetric analysis from a harmonized multisite brain MRI study of a single subject with multiple sclerosis. *Am. J. Neuroradiol.* 38, 1501–1509. doi: 10.3174/ajnr.A5254
- Shinohara, R. T., Sweeney, E. M., Goldsmith, J., Shiee, N., Mateen, F. J., Calabresi, P. A., et al. (2014). Statistical normalization techniques for magnetic resonance imaging. *Neuroimage Clin.* 6, 9–19. doi: 10.1016/j.nicl.2014.08.008
- Shishegar, R., Cox, T., Rolls, D., Bourgeat, P., Doré, V., Lamb, F., et al. (2021). Using imputation to provide harmonized longitudinal measures of cognition across AIBL and ADNI. *Sci. Rep.* 11, 23788. doi: 10.1038/s41598-021-02827-6
- Snitz, B. E., Tudorascu, D. L., Yu, Z., Campbell, E., Lopresti, B. J., Laymon, C. M., et al. (2020). Associations between NIH toolbox cognition battery and *in vivo* brain amyloid and tau pathology in non-demented older adults. *Alzheimers Dement.* 12, e12018. doi: 10.1002/dad2.12018
- Snyder, H. M., Corriveau, R. A., Craft, S., Faber, J. E., Greenberg, S. M., Knopman, D., et al. (2015). Vascular contributions to cognitive impairment and dementia including Alzheimer's disease. *Alzheimers Dement.* 11, 710–717. doi: 10.1016/j.jalz.2014.10.008
- Sperling, R. A., Donohue, M. C., Raman, R., Sun, C. K., Yaari, R., Holdridge, K., et al. (2020). Association of factors with elevated amyloid burden in clinically normal older individuals. *JAMA Neurol.* 77, 735–745. doi: 10.1001/jamaneurol.2020.0387
- Stein, J. L., Medland, S. E., Vasquez, A. A., Hibar, D. P., Senstad, R. E., Winkler, A. M., et al. (2012). Identification of common variants associated with human hippocampal and intracranial volumes. *Nat. Genet.* 44, 552–561. doi: 10.1038/ng.2250
- Sullivan, K., Liu, A., Chang, C., Cohen, A., Lopresti, B., Minhas, D., et al. (2020). Alzheimer's disease pathology in a community-based sample of older adults without dementia: the MYHAT neuroimaging study. *Brain Imaging Behav.* 15, 1355–1363. doi: 10.1007/s11682-020-00334-2
- Sundermann, E. E., Tran, M., Maki, P. M., and Bondi, M. W. (2018). Sex differences in the association between apolipoprotein E epsilon4 allele and Alzheimer's disease markers. *Alzheimers Dement.* 10, 438–447. doi: 10.1016/j.dadm.2018.06.004
- Taylor, J. R., Williams, N., Cusack, R., Auer, T., Shafto, M. A., Dixon, M., et al. (2017). The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* 144(Pt B), 262–269. doi: 10.1016/j.neuroimage.2015.09.018
- Torbati, M. E., Minhas, D. S., Ahmad, G., O'Connor, E. E., Muschelli, J., Laymon, C. M., et al. (2021a). A multi-scanner neuroimaging data harmonization using RAVEL and ComBat. *NeuroImage* 245, 118703. doi: 10.1016/j.neuroimage.2021.118703
- Torbati, M. E., Tudorascu, D. L., Minhas, D. S., Maillard, P., DeCarli, C. S., and Hwang, S. J. (2021b). Multi-scanner harmonization of paired neuroimaging data via Structure Preserving Embedding Learning. *IEEE Int. Conf. Comput. Vis. Workshops* 2021, 3277–3286. doi: 10.1109/iccvw54120.2021.00367
- Tosto, G., Zimmerman, M. E., Hamilton, J. L., Carmichael, O. T., and Brickman, A. M. (2015). The effect of white matter hyperintensities on neurodegeneration in mild cognitive impairment. *Alzheimers Dement.* 11, 1510–1519. doi: 10.1016/j.jalz.2015.05.014
- Tremblay-Mercier, J., Madjar, C., Das, S., Pichet Binette, A., and Dyke, S. O. M., Étienne, P., et al. (2021). Open science datasets from PREVENT-AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease. *Neuroimage Clin.* 31, 102733. doi: 10.1016/j.nicl.2021.102733
- Tudorascu, D. L., Karim, H. T., Maronge, J. M., Alhilali, L., Fakhran, S., Aizenstein, H. J., et al. (2016). Reproducibility and bias in healthy brain segmentation: comparison of two popular neuroimaging platforms. *Front. Neurosci.* 10, 503. doi: 10.3389/fnins.2016.00503
- Uehara, T., Tabuchi, M., and Mori, E. (1999). Risk factors for silent cerebral infarcts in subcortical white matter and basal ganglia. *Stroke* 30, 378–382.
- Vable, A. M., Diehl, S. F., and Glymour, M. M. (2021). Code review as a simple trick to enhance reproducibility, accelerate learning, and improve the quality of your team's research. *Am. J. Epidemiol.* 190, 2172–2177. doi: 10.1093/aje/kwab092
- van Dijk, E. J., Prins, N. D., Vrooman, H. A., Hofman, A., Koudstaal, P. J., and Breteler, M. M. (2008). Progression of cerebral small vessel disease in relation to risk factors and cognitive consequences: Rotterdam Scan study. *Stroke* 39, 2712–2719. doi: 10.1161/strokeaha.107.513176
- Vermeer, S. E., Koudstaal, P. J., Oudkerk, M., Hofman, A., and Breteler, M. M. (2002). Prevalence and risk factors of silent brain infarcts in the population-based Rotterdam Scan Study. *Stroke* 33, 21–25. doi: 10.1161/hs10102.101629
- Vonk, J. M. J., Gross, A. L., Zammit, A. R., Bertola, L., Avila, J. F., Jutten, R. J., et al. (2022). Cross-national harmonization of cognitive measures across HRS HCAP (USA) and LASI-DAD (India). *PLoS ONE* 17, e0264166. doi: 10.1371/journal.pone.0264166
- Wardlaw, J. M., Smith, E. E., Biessels, G. J., Cordonnier, C., Fazekas, F., Frayne, R., et al. (2013). Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* 12, 822–838. doi: 10.1016/S1474-4422(13)70124-8
- Wardlaw, J. M., Valdes Hernandez, M. C., and Munoz-Maniega, S. (2015). What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. *J. Am. Heart Assoc.* 4:e01140. doi: 10.1161/jaha.114.001140
- Weiner, M. W., Aisen, P. S., Jack, C. R. Jr., Jagust, W. J., Trojanowski, J. Q., Shaw, L., et al. (2010). The Alzheimer's Disease Neuroimaging Initiative: progress report and future plans. *Alzheimers Dement.* 6, 202–211.e7. doi: 10.1016/j.jalz.2010.03.007
- Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Cedarbaum, J., et al. (2015). 2014 Update of the Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. *Alzheimers Dement.* 11, e1–120. doi: 10.1016/j.jalz.2014.11.001
- Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., et al. (2017). The Alzheimer's Disease Neuroimaging Initiative 3: continued innovation for clinical trial improvement. *Alzheimers Dement.* 13, 561–571. doi: 10.1016/j.jalz.2016.10.006
- Westreich, D., Edwards, J. K., Lesko, C. R., Stuart, E., and Cole, S. R. (2017). Transportability of trial results using inverse odds of sampling weights. *Am. J. Epidemiol.* 186, 1010–1014. doi: 10.1093/aje/kwx164
- Wrobel, J., Martin, M. L., Bakshi, R., Calabresi, P. A., Elliot, M., Roalf, D., et al. (2020). Intensity warping for multisite MRI harmonization. *Neuroimage* 223, 117242. doi: 10.1016/j.neuroimage.2020.117242
- Yan, S., Zheng, C., Paranjpe, M. D., Li, Y., Li, W., Wang, X., et al. (2021). Sex modifies APOE ε4 dose effect on brain tau deposition in cognitively impaired individuals. *Brain* 144, 3201–3211. doi: 10.1093/brain/awab160
- Yu, M., Linn, K. A., Cook, P. A., Phillips, M. L., McInnis, M., Fava, M., et al. (2018). Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Hum. Brain Mapp.* 39, 4213–4227. doi: 10.1002/hbm.24241
- Ziontz, J., Bilgel, M., Shafer, A. T., Moghekar, A., Elkins, W., Helpfrey, J., et al. (2019). Tau pathology in cognitively normal older adults. *Alzheimers Dement. (Amst.)* 11, 637–645. doi: 10.1016/j.dadm.2019.07.007