Harmonized diffusion MRI data and white matter measures from the baseline Adolescent Brain Cognitive Development (ABCD) study
(Dated: 04/04/2023)

Suheyla Cetin-Karayumak, Ph.D. [a], Fan Zhang, Ph.D.[b], Tashrif Billah [a], Leo Zekelman [d, e], Nikos Makris, M.D., Ph.D.[a,f], Steve Pieper, Ph.D.[c], Lauren J. O'Donnell, Ph.D.[b], Yogesh Rathi, Ph.D. [a,b]

[a] Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.
[b] Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.
[c] Isomics, Inc., Cambridge, Massachusetts, USA.
[d] Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.
[e] Program in Speech and Hearing Bioscience and Technology, Division of Medical Sciences, Harvard University, Boston, Massachusetts, USA.
[f] Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA.

Correspondence:
Suheyla Cetin-Karayumak, Ph.D.
Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School
399 Revolution Drive, STE 1155
Somerville, MA 02145
t: +1 857-3205098
e: skarayumak@bwh.harvard.edu

Abstract:

The Adolescent Brain Cognitive Development (ABCD) study is a groundbreaking effort aimed at providing a comprehensive understanding of adolescent brain development. With data collected from over 10,000 children across 21 sites, this study promises to unlock key insights into the cognitive, behavioral, and neuroimaging data that underpins this critical period of development. However, the potential use of diffusion MRI (dMRI) data in this study is hindered by site-specific scanner and protocol variability, which precludes naive pooling of the data from across sites for analysis. In response to this challenge, we have developed an aggregated database of harmonized dMRI data, comprising imaging data from 9345 subjects and their white matter tracts, as well as tract-specific microstructural measures. This resource has been made publicly available via the NIMH Data Archive (NDA). The generation of this database required enormous computational effort, taking ~50,000 CPU hours to effectively harmonize the dMRI data, run whole brain tractography, and perform white matter parcellation. The data available through the NDA includes: harmonized dMRI data, several dMRI derived measures (e.g., fractional anisotropy) for 73 anatomically well-defined white matter fiber bundles in all 9345 subjects and streamlines corresponding to each fiber bundle for visualization. In this work, we further demonstrate the efficacy of our dMRI data harmonization algorithm on the ABCD dataset using multiple experiments, allowing researchers to perform their own analyses on the harmonized data. The data resources made available through this work will allow for new structural connectivity studies of neurodevelopment in children and adolescents at an unprecedented scale.

1) Introduction:

The Adolescent Brain Cognitive Development (ABCD) study® is a landmark research project focused on child health and brain development (Jernigan, Brown, and Dowling 2018). The ABCD Research Consortium consists of 21 research sites across the US and has collected one of the largest multi-domain datasets, including neuroimaging, behavioral, cognitive, and genetic data, from over 10,000 children from middle childhood into early adulthood. Using this extensive dataset, the ABCD study aims to provide the opportunity to understand the factors that shape brain and cognitive development during this crucial period of life. To ensure uniformity in scanning protocols, the ABCD study has taken extensive measures to maintain imaging acquisition parameters and protocols consistent across all 21 participating sites (Casey et al. 2018). Despite these efforts, data acquired from different sites and vendors (Siemens, Philips, and GE) introduce significant site effects (Hagler et al. 2019), due to scanner-specific MR acquisition sequences and data reconstruction algorithms. In fact, given the vast array of design choices and vendor-specific decisions beyond the users' control, it is almost always challenging to completely standardize the different scanners in multi-site neuroimaging studies such as the ABCD study, leading to inter-scanner differences in the acquired data.

Diffusion-weighted magnetic resonance imaging (dMRI) is a non-invasive imaging technique that can detect microstructural changes and reveal white matter connectivity in the brain. Studies have shown that even small scanner-related differences in the dMRI data collected from various scanners can result in significant measurement biases in the dMRI

measures of white matter connectivity and microstructure (Vollmar et al. 2010; H. Mirzaalian et al. 2016; Cetin Karayumak et al. 2019; S. Cetin-Karayumak, Kubicki, and Rathi 2018; Cetin Karayumak S*, Billah T*, Bouix S, Rathi Y 2019; Ning et al. 2020; Tax et al. 2019). Even when scanners from the same manufacturer are used across different sites, the measurement biases in the dMRI data can be substantial. For example, Thieleking et al. (Thieleking et al. 2021) showed that the difference in the fractional anisotropy (FA) from the same healthy participants (N = 121) scanned on two 3T Siemens Magnetom scanners could be up to 33 times larger than the effect seen in healthy aging. In addition, Schilling et al. (Schilling et al. 2021) demonstrated the significant variability in connectivity measures across different scanners (e.g., the density of streamlines in white matter fiber connections/tracts), which also resulted in significant differences in the related microstructural measures. Moreover, multiple studies have highlighted that scanner-related variability is highly non-linear across various tissues and white matter regions/tracts (Tax et al. 2019; Ning et al. 2020; Cetin Karayumak et al. 2019; Bayer et al. 2022; Botvinik-Nezer et al. 2020; Schilling et al. 2021). Thus, many studies underscore the significance of eliminating bias due to scanner or site from dMRI data at the voxel level before running joint analysis in multi-site studies such as the ABCD study (H. Mirzaalian et al. 2016; Hengameh Mirzaalian et al. 2018; Cetin Karayumak et al. 2019; Pinto et al. 2020; Hengameh Mirzaalian et al. 2015; Fortin et al. 2017; Huynh et al. 2019; Moyer et al. 2020; Hagler et al. 2019).

Enabling pooled large-scale analysis of multi-site datasets represents an outstanding opportunity for improving our understanding of the human brain in health and disease. However, such research requires large sample sizes, which can only be obtained by appropriate pooling and removal of scanner-specific effects from dMRI data acquired from multiple scanners. *"Harmonization"* is a way to mitigate the measurement differences attributed to the scanner-, protocol-, or other site-related effects (H. Mirzaalian et al. 2016; Hengameh Mirzaalian et al. 2018; Cetin Karayumak et al. 2019; Pinto et al. 2020; Hengameh Mirzaalian et al. 2015; Fortin et al. 2017; Huynh et al. 2019; Moyer et al. 2020). The goal of harmonizing dMRI data is to preserve variability that is purely related to biology or disease and remove variability caused by intrinsic or acquisition-related factors of the scanners, which can conceal the desired effect. Various data harmonization methods have been proposed in the literature (Kochunov et al. 2014; Thompson et al. 2014, 2020; G. Zhang et al. 2017; Fortin et al. 2017; Johnson, Li, and Rabinovic 2007; S. Cetin-Karayumak, Kubicki, and Rathi 2018). A majority of these approaches are based on adding statistical covariates to remove site effects. These methods are commonly used in multiple fields, from genetics to functional MRI and recently in dMRI (Kochunov et al. 2014; Thompson et al. 2014, 2020; G. Zhang et al. 2017; Fortin et al. 2017; Johnson, Li, and Rabinovic 2007). These methods typically involve fitting a dMRI model (such as diffusion tensor imaging) to obtain the desired dMRI measures (e.g., FA) for each subject at each site, followed by regression modeling where the site is added as a linear covariate to minimize inter-site effects. These approaches are limited for several reasons. First, they assume linear site effects in the microstructural measures (such as FA, mean diffusivity, kurtosis, etc.). However, this assumption has been challenged by many studies demonstrating nonlinear site effects (Suheyla Cetin-Karayumak, Stegmayer, et al. 2020; Cetin Karayumak et al. 2019; Botvinik-Nezer et al. 2020; De Luca et al., n.d.). Second, white matter connectivity (i.e., tractography) can be affected by scanner biases (Schilling et al. 2021), which raises doubts about the effectiveness of data

harmonization after running whole-brain tractography. Hence, harmonizing the structural connectivity matrices (derived from tractography of unharmonized data) can be inadequate and ineffective in removing scanner effects across dMRI data.

On the other hand, harmonizing dMRI data obtained directly from the scanner has several advantages. As demonstrated in several of our earlier works (Cetin Karayumak et al. 2019; H. Mirzaalian et al. 2016), nonlinear and voxel-wise site effects can be removed allowing robust and unbiased tractography estimation across sites. Further, any dMRI microstructural model (e.g. single or multi-tensor, standard model of diffusion, NODDI, etc. (Özarslan et al. 2013; H. Zhang et al. 2012; Jelescu et al. 2016)) can be used without having to worry about the confounds due to scanner-related biases during the model estimation process. The dMRI community has recognized the value of performing harmonization directly on the scanner data and has organized multiple community challenges to determine the best-performing algorithm as part of the Medical Image Computing and Computer-assisted Intervention (MICCAI) conference (Ning et al. 2020; Tax et al. 2019). In all metrics, our algorithm was chosen as the best-performing method (Ning et al. 2020; Tax et al. 2019)). The effectiveness of our approach has also been demonstrated through several studies in schizophrenia and other disorders (Cetin Karayumak et al. 2019; Suheyla Cetin-Karayumak, Di Biase, et al. 2020; Seitz-Holland et al. 2021; Seitz et al. 2021; Biase et al. 2021; De Luca et al. 2022; Seitz-Holland et al. 2022; Suheyla Cetin-Karayumak et al., n.d.; Elad et al. 2021; Ye et al. 2021; Lv et al. 2020; Hegde et al. 2021; Kelly et al. 2021; Levitt et al., n.d.; S. Cetin-Karayumak, Kubicki, and Rathi 2018).

Furthermore, the analysis of very large dMRI data sets represents a significant computational burden for many neuroscience researchers and laboratories. Of particular neuroscientific interest is the quantitative hypothesis-driven study of the brain's major white matter fiber tracts (F. Zhang et al. 2022). To enable the study of the fiber tracts across very large dMRI datasets, it is critical to define anatomical fiber pathways (e.g. the arcuate fasciculus) consistently across all subjects, irrespective of their age, gender, or disease indications. It is also critical to automate the extraction of these fiber pathways and the measurement of their tissue microstructure. Several approaches have demonstrated the advantages of automated extraction of white matter fiber tracts (F. Zhang, Wu, et al. 2018; Wasserthal, Neher, and Maier-Hein 2018; F. Zhang, Cetin Karayumak, et al. 2020; Schilling et al. 2020; Yendiki et al. 2011). In this study, we apply our robust multi-fiber tractography method, which allows the measurement of fiber-specific microstructural properties (Malcolm, Shenton, and Rathi 2010; Reddy and Rathi 2016). We then perform automated tract extraction using a well-established fiber clustering pipeline (O'Donnell and Westin 2007; O'Donnell et al. 2012a) together with a neuroanatomically curated white matter atlas (F. Zhang, Wu, et al. 2018). We have previously demonstrated that this framework consistently parcellates white matter tracts across the lifespan (F. Zhang, Wu, et al. 2018) with high test-retest reproducibility (F. Zhang et al. 2019).

The aim of this study is to create a database of harmonized dMRI data and tract-specific microstructure measures to enable novel scientific investigations across thousands of subjects from the ABCD study. To produce this novel database, which includes data from 9345 subjects, we applied our advanced dMRI harmonization, tractography, and white matter parcellation computational pipeline. First, our robust dMRI data harmonization algorithm was applied to remove scanner-specific biases from the multi-site dMRI data. Harmonizing the raw data

acquired directly from the scanner in this way enables laboratories to then perform any desired analyses to test their neuroscientific hypotheses of interest. Second, robust multi-fiber tractography was computed in the whole brain of all 9345 subjects for highly consistent tracing of white matter connections. Third, 73 subject-specific anatomical white matter tracts were extracted from the tractography of each subject in an automated fashion, and their fiber-specific microstructural properties were quantified. The release of this fiber tract microstructure data will enable laboratories to directly test neuroscientific hypotheses of interest without needing to perform computationally intensive processing. This paper includes several experiments that demonstrate the efficacy of dMRI data harmonization and the success of white matter tract identification on the ABCD dataset. Overall, the release of the proposed database represents an unparalleled opportunity for researchers to investigate neurodevelopmental brain changes that were previously challenging to investigate using unharmonized data or smaller sample sizes. The database includes harmonized dMRI data, tractography, extracted white matter tracts, and microstructure measures from the ABCD study. The database is now available through the NIMH Data Archive repository, allowing open access to the neuroimaging and neuroscience community.

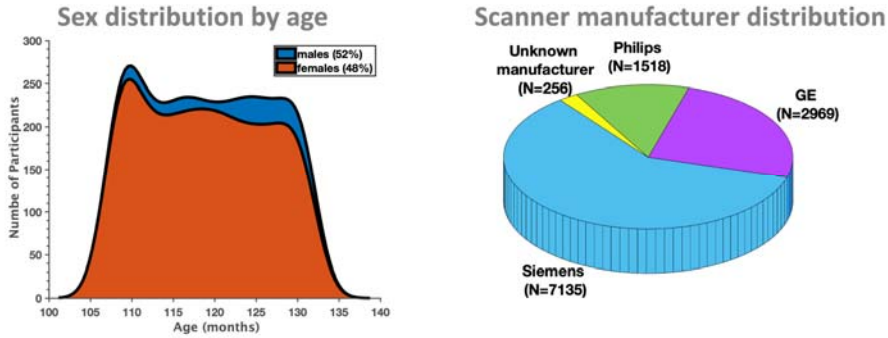
2) Materials and Methods:

2.1 The ABCD study database.

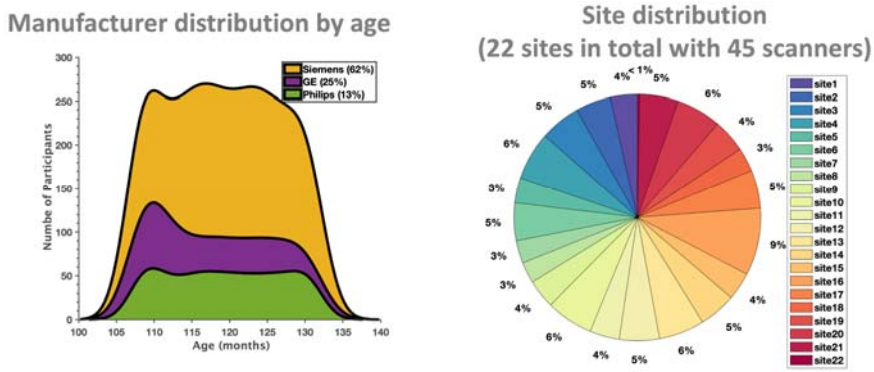
The ABCD study® has collected an extensive dataset of 11,878 participants across 21 sites with a baseline age of 8-11 years to gain a better understanding of neurodevelopment between childhood and early adulthood (Figure 1a). The dataset provides a longitudinally collected wealth of measured attributes, such as neuroimaging, cognitive, biospecimen, behavioral, youth self-report and parent self-report metrics, and environmental data. The ABCD dataset can be publicly accessed by researchers through a data use agreement with the NDA. Notably, the neuroimaging data includes dMRI, structural MRI (both T1- and T2-weighted), and functional MRI data, all of which were collected every two years for almost all participants (48% of whom are female). The participants were scanned using Siemens, Philips, and GE scanners with similar acquisition parameters across 21 sites.

The acquired MRI data was minimally preprocessed by the ABCD study Data Analysis, Informatics, and Resource Center group. It is important to note that the ABCD study extensively screened participants for their history of psychiatric disorders, neurological problems, traumatic brain injury (TBI), seizure, low gestational weights, and more. Participants who did not meet the criteria for MRI scanning were excluded from the study, such as those with a history of severe TBI, abnormal MRI scans, or seizures. Detailed information about data recruitment, acquisition, and MRI data preprocessing steps can be found in (Hagler et al. 2019) and (Hatton 2020). We note that the preprocessing steps of the ABCD study did not include harmonization or tractography analysis of the dMRI data. Therefore, this study is focused on processing multi-site dMRI data from the baseline session using our harmonization and connectivity frameworks, as well as making the processed data of the ABCD study available to the NDA community.

a) The ABCD database (N=11,878**)**



b) After quality control (QC) applied to the dMRI data by the ABCD study team, data from the unknown manufacturer was removed (N=10,866).



c) Additional dMRI data QC steps applied by our team (**N=9345** is the final number of participants used in this study).

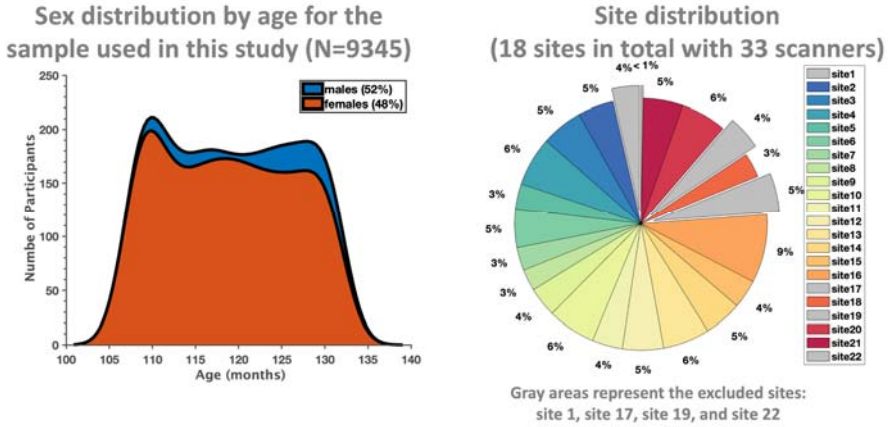


Figure 1. The ABCD study database: (a) The original database before quality control (QC) procedures applied to the dMRI data (N=11,878); (b) After QC procedures are applied to the dMRI data by the ABCD study team (N=10,866 for the quality passed dMRI scans); (c) Additional QC steps are applied to the dMRI data in this study, where 4 sites with 12 scanners are removed from the sample (N=9345 for the final dMRI scans used in the

subsequent processing steps in this study).

2.2 The details on dMRI data acquisition and minimal preprocessing.

DMRI scans were acquired with 1.7x1.7x1.7 mm^3 resolution using multiband EPI with a slice acceleration factor of 3 (Casey et al. 2018). Siemens and GE scanners collected dMRI data using a single acquisition with 96 diffusion gradient directions, seven b=0 volumes, and four b-values (6 directions with b=500 s/mm^2, 15 directions with b=1000 s/mm^2, 15 directions with b=2000 s/mm^2, and 60 directions with b=3000 s/mm^2). The Philips scanner used two acquisitions to collect the dMRI data. Philips scan 1 parameters were: 3 directions with b=0 s/mm^2, 4 directions with b=500 s/mm^2, 7 directions with b=1000 s/mm^2, 8 directions with b=2000 s/mm^2 and 29 directions with b=3000 s/mm^2. Philips scan 2 parameters were: 4 directions with b=0 s/mm^2, 2 directions with b=500 s/mm^2, 8 directions with b=1000 s/mm^2, 7 directions with b=2000 s/mm^2 and 30 directions with b=3000s/mm^2 (Table 1). The minimal preprocessing steps were consistently applied to the dMRI data of each study site by the ABCD study, which included: eddy and motion correction, b0 inhomogeneity correction, gradient unwarp, and resampling to isotropic resolution (1.7 mm^3) (Hagler et al. 2019; Hatton 2020). While these preprocessing steps are well established and known to improve data quality, they do not correct for biases due to multiple acquisition sites and scanners, which will be addressed as part of this study.

<u>The ABCD study dMRI database of baseline scans:</u>

Scanners	Number of subjects	B-values and the number of gradients/volumes
Siemens	7135	*The full protocol consists of 6 b=500, 15 b=1000, 15 b=2000, and 60 b=3000 volumes.*
GE	2969	*The full protocol consists of 6 b=500, 15 b=1000, 15 b=2000, and 60 b=3000 volumes.*
Philips	1518	*Includes two acquisitions: 1) 4 b=500, 7 b=1000, 8 b=2000 and 30 b=3000 volumes; 2) 2 b=500, 8 b=1000. 7 b=2000 and 30 b=3000 volumes.*
Unknown	256	-
Total	11,878	

Table 1. The table includes the data about the scanner manufacturer, inlcuding the number of subjects scanned by the Siemens/GE/Philips scanner, and the acquisition information. We note that additional 256 subjects, who do not have scanner information available, were marked as unknown in this table.

2.2 Scanner-Software Upgrades in each ABCD Study Site.

Several studies have revealed that even slight modification to the scanner software can result in bias in the acquired dMRI data, similar to a "site-effect," even when the acquisition protocol remains unchanged (Medawar et al. 2021; Bayer et al. 2022). However, it is often difficult to predict the bias in the data due to such software updates. Therefore, scanner software (version) updates should ideally be avoided entirely during the course of a study to avoid introducing bias

in the dMRI data. However, this is often not feasible in practice. Similar to many large-scale multi-site neuroimaging studies, several scanners within the ABCD study underwent software upgrades during the acquisition of the baseline data. Therefore, we treated each upgrade as an independent scanner. Hence, after counting each software upgrade as a separate "scanner," the ABCD study included 45 total scanners. Supplementary Table 1 provides details about the scanners and related upgrades.

2.3 DMRI Data Quality Control.

The ABCD study applied several automated quality control (QC) procedures, such as head motion statistics and the detection of signal drops to remove low quality dMRI scans (Hagler et al. 2019; Hatton 2020). The trained technicians at the sites of the ABCD study also did manual checks by reviewing the dMRI scans for poor image quality and various imaging artifacts. The scans that did not pass the QC checks were marked as unacceptable data. Refer to Figure 1b for the distribution of the subjects (per site) that passed the required dMRI data quality.

Additionally, we applied our own automated and manual QC steps to ensure the data was of good quality for harmonization and connectivity analysis (i.e., whole brain tractography and white matter parcellation). This involved using the average whole-brain FA as well as regional FA to screen for outliers, as described in our previous study (Cetin Karayumak et al. 2019). We also conducted manual QC checks on a randomly selected subset of subjects from each scanner. Our findings showed that the dMRI data from most GE and Siemens sites had good quality and was suitable for harmonization and tractography analysis. However, the minimally processed dMRI data from the Philips scanners were found to have certain artifacts, such as ringing and motion artifacts as well as excessive smoothing (potentially during pre-processing steps). Additionally, the Philips scans were obtained in two separate acquisitions with two files, which required additional processing for motion and eddy current correction. Due to these factors, data from the Philips scanners that included 3 sites with 12 scanners were excluded from the current study. Our final sample consisted of dMRI data from 9345 participants from Siemens and GE scanners, which were obtained from 18 sites and 33 scanners. Refer to Figure 1c for the distribution of subjects and sites used in this work (i.e., for harmonization and tractography analysis). Supplementary Table 2 provides more details about the sites and scanners used in this work.

2.4 Overview of the proposed pipeline to process the dMRI Data (N=9345).

This section describes the major computational steps used to create a database of harmonized and processed dMRI data derived from the ABCD study (see Figure 2). First, brain masks were generated to isolate only the brain tissue and exclude the skull. This was crucial to focus solely on the brain region, making it an indispensable prerequisite for many processing pipelines in neuroimaging studies including ours. To this end, we applied our deep learning-based brain masking tool[1] that robustly extracted the brain from the field-of-view volume in dMRI data of 9345 subjects, as shown in Figure A.1. The details of this method and performance evaluation are described under Appendix A. We then applied the following computational steps

[1] Our brain masking tool is publicly available in GitHub: https://github.com/pnlbwh/CNN-Diffusion-MRIBrain-Segmentation

respectively: a) dMRI data harmonization, b) whole-brain tractography, c) subject-specific white matter parcellation, d) extracting dMRI measures.

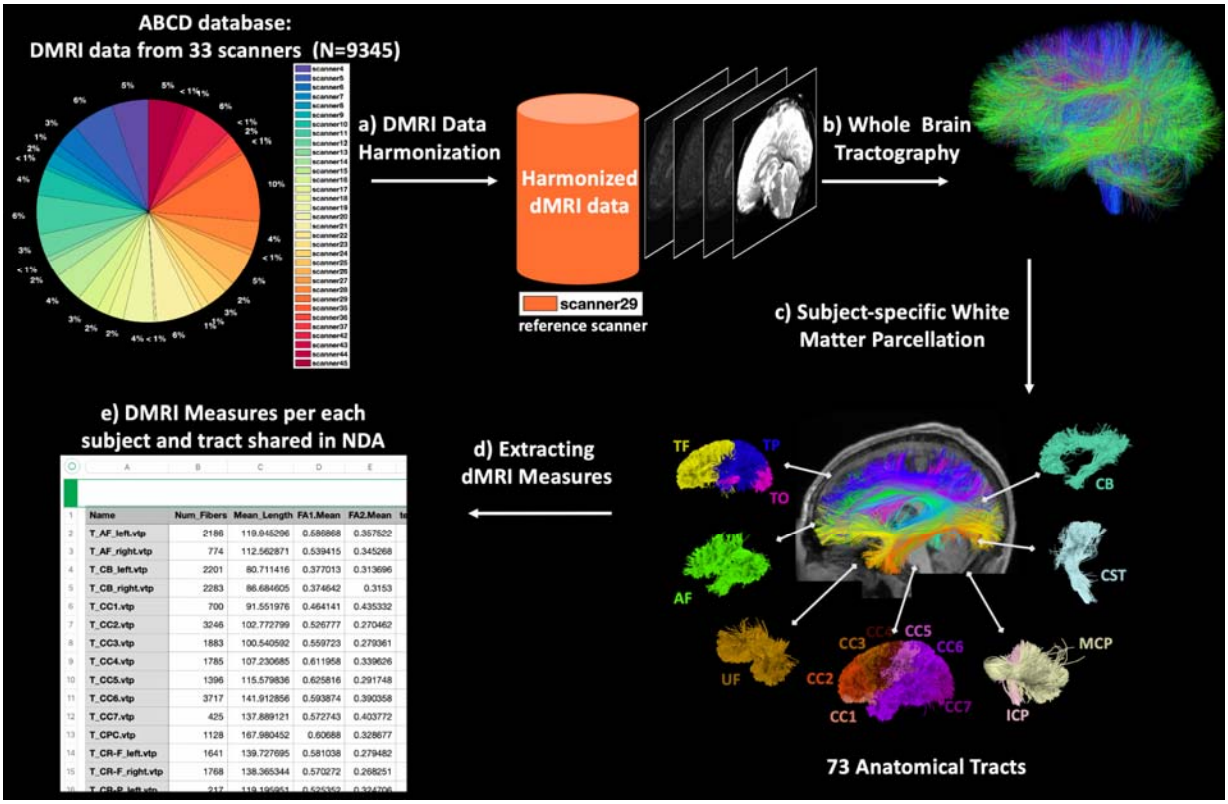


Figure 2. Pipeline overview. (a) DMRI data are harmonized to the reference scanner (scanner 29) to remove possible site/scanner-specific effects. (b) Whole-brain tractography was computed for each subject using the UKF tractography algorithm. (c) WM parcellation was performed using the WMA package in conjunction with an anatomical white matter atlas, resulting in parcellation of 73 anatomically well-defined tracts. (d) Multiple dMRI measures were extracted for each parcellated tract.

(a) DMRI Data Harmonization. We applied our retrospective harmonization algorithm[2] to remove the scanner-related differences across 33 scanners. Our harmonization approach is based on rotation invariant spherical harmonics (RISH) features and works at the signal level to remove the scanner-related biases across datasets while accounting for non-linearities in the dMRI data (which can vary by region and tissue). We note that this method was selected as the best-performing method in a harmonization community challenge (Ning et al. 2020).

The harmonization algorithm proceeds by selecting one scanner as the reference and with all other scanners harmonized to this reference site. Scanner 29 from site 16 (a Siemens Prisma scanner) was selected as the reference due to its large sample size (N=??) and high data quality. The dMRI data from the remaining 32 scanners were harmonized to this reference

[2] Our multi-site diffusion MRI data harmonization algorithm is available in GitHub: https://github.com/pnlbwh/dMRIharmonization

scanner (for a full list, see Supp. Table 2). While scanner 29 was chosen as the reference in this study, we have previously shown that choosing a different scanner as the reference does not affect the performance of harmonization (Cetin Karayumak et al. 2019).

Harmonization of the dMRI data is carried out in two steps. The first step, referred to as template creation, involves matching approximately 35 subjects from each scanner with subjects from the reference scanner based on age, sex, and IQ. This matching process is conducted at a group level to eliminate differences between scanners while preserving inter-subject biological variability. To ensure we capture only scanner related effects, subjects with a history of epilepsy, stroke, tumor, hemorrhage, or psychiatric disorders were not included in this step. The full exclusion criteria can be found in the supplement. The matched set of 35 subjects is then used to learn a nonlinear mapping between dMRI data from different scanners. This inter-scanner mapping is voxel-specific and uses the fiber-orientation independent RISH feature representation to learn a specific mapping for each RISH feature. These RISH features capture different microstructural tissue properties that are independent of the orientation (Reisert et al. 2017) and provide the ability to reconstruct the harmonized dMRI signal. In the second step, *harmonizing dMRI data*, the learned mappings (from the template creation step) across RISH features were applied to the remaining dMRI data of the corresponding scanner to be harmonized with the reference. Further details can be found in (Cetin Karayumak et al. 2019).

For harmonizing this large-scale dataset (N = 9345), we used the AWS EC2 platform. Specifically, r5d.4xlarge EC2 instances (128GiB of memory, 16 vCPUs) were used in template generation, where it took ~48 hours to generate each template. We then used r5d.large EC2 instances (16GiB of memory, 2 vCPUs) in the harmonization step, where harmonization of each subject took ~15 mins. The whole pipeline was tailored for parallel execution in 32 instances. In total, it took about 3872 CPU hours (~162 CPU days) to complete the processing of all dMRI data (N=9345) in this step.

(b) Whole Brain Tractography. *Tractography* was performed using our well-validated multi-tensor unscented Kalman filter (UKF) fiber tracking algorithm (Malcolm, Shenton, and Rathi 2010; Reddy and Rathi 2016; *Ukftractography* n.d.; Fillard et al. 2011) on the b=3000 s/mm^2 shell. UKF tractography accounts for crossing fibers, and can consistently trace fibers across various populations (F. Zhang, Wu, et al. 2018) as well as disease states (e.g. tumors) (Olszewski et al. 2017; F. Zhang, Savadjiev, et al. 2018; Hamoda et al. 2019). We successfully applied tractography on 9345 subjects from the harmonized ABCD dataset. For each harmonized dMRI data, ~450k fiber streamlines were generated to create a whole brain tractography dataset.

The UKF tractography was run on the AWS EC2 platform. Specifically, r5d.large EC2 instances (16GiB of memory, 2 vCPUs) were used, where tractography took ~4.5 hours per subject. A total of ~500 instances were used, with each instance running tractography for about 20 subjects. In total, it took about ~42052 CPU hours/~1752 CPU days to complete whole brain tractography.

(c) Subject-specific White Matter Parcellation. Tractography parcellation was performed using our WhiteMatterAnalysis (WMA) (*Whitematteranalysis* n.d.) fiber clustering pipeline (O'Donnell and Westin 2007; O'Donnell et al. 2012b) in conjunction with our anatomical white

matter atlas (F. Zhang, Wu, et al. 2018). This pipeline produces a fine-scale whole-brain tractography parcellation into 800 fiber clusters and a coarse-scale anatomical tract parcellation, including 57 major deep WM tracts and 16 categories of superficial fiber clusters organized according to the brain lobes they connect. Refer to Table 2 for a list of anatomical tracts. Figure 2 gives a visualization of example tracts from 10 randomly selected subjects from 10 different scanners.

We used the AWS r5d.large EC2 instances (16GiB of memory, 2 vCPUs) for the computation, where it took ~1.6 hours per subject. A total of ~500 instances were used, with each instance running white matter parcellation for about 20 subjects. In total, it took about ~14952 CPU hours/~623 CPU days to complete the white matter parcellation.

Tract category (number of tracts)	Tract name (L - left; R - right; C - commissural)
Association tracts (24)	arcuate fasciculus (AF) – LR cingulum bundle (CB) – LR external capsule (EC) – LR extreme capsule (EmC) – LR inferior longitudinal fasciculus (ILF) – LR inferior occipito-frontal fasciculus (IoFF) – LR middle longitudinal fasciculus (MdLF) – LR posterior limb of internal capsule (PLIC) – LR superior longitudinal fasciculus I (SLF I) – LR superior longitudinal fasciculus II (SLF II) – LR superior longitudinal fasciculus II (SLF III) – LR uncinate fasciculus (UF) – LR
Cerebellar tracts (8)	cortico-ponto-cerebellar (CPC) – C inferior cerebellar peduncle (ICP) – LR middle cerebellar peduncle (MCP) – C intracerebellar input and Purkinje tract (CBLM-I&P) – LR intracerebellar parallel tract (CBLM-PaT) – LR
Commissural tracts (7)	corpus callosum 1 (CC 1) – C corpus callosum 2 (CC 2) – C corpus callosum 3 (CC 3) – C corpus callosum 4 (CC 4) – C corpus callosum 5 (CC 5) – C corpus callosum 6 (CC 6) – C corpus callosum 7 (CC 7) – C
Projection tracts (18)	corticospinal tract (CST) – LR corona-radiata-frontal (CR-F) – LR corona-radiata-parietal (CR-P) – LR striato-frontal (SF) – LR striato-occipital (SO) – LR striato-parietal (SP) – LR thalamo-frontal (TF) – LR

	thalamo-occipital (TO) – LR thalamo-parietal (TP) – LR
Superficial tracts (16)	superficial-frontal (Sup-F) – LR superficial-frontal-parietal (Sup-FP) – LR superficial-occipital (Sup-O) – LR superficial-occipital-temporal (Sup-OT) – LR superficial-parietal (Sup-P) – LR superficial-parietal-occipital (Sup-PO) – LR superficial-parietal-temporal (Sup-PT) – LR superficial-temporal (Sup-T) – LR

Table 2. A total of 73 anatomical fiber tracts are identified for each subject, including 24 association tracts, 8 cerebellar tracts, 7 commissural tracts, 18 projection tracts, and 16 categories of short and medium range superficial fiber clusters organized according to the brain lobes that they connect.

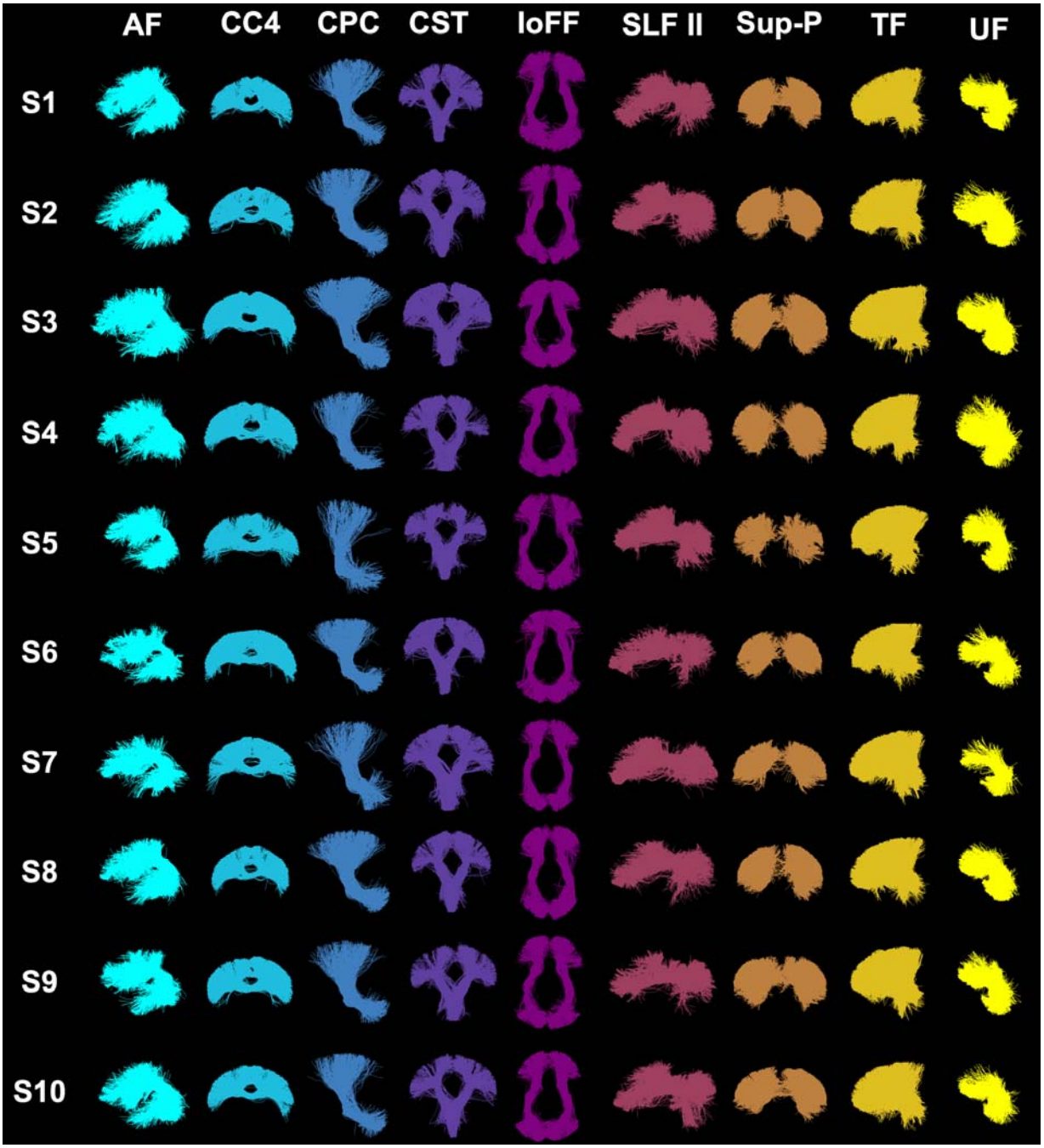


Figure 3. Visualization of example anatomical tracts. 10 randomly selected subjects (from 10 different sites) are used. In general, the parcellated anatomical tracts are visually highly similar across the different subjects. Refer to Table 2 for the definition of the acronyms for tract names.

(d) Extracting DMRI Measures. Multiple dMRI measures were extracted from each parcellated cluster and anatomical tract using SlicerDMRI (Norton et al. 2017; F. Zhang, Noh, et al. 2020)

for all subjects. These include the widely used dMRI measures: FA, mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD), number of streamlines (NoS), number of streamline points (NoF), and streamline length. The tensors associated at each streamline location were also saved in the corresponding VTK file which can be accessed from the NDA. These extracted dMRI measures are available via the NDA repository. Additional measures of interest can also be computed from the tensors by downloading the appropriate VTK file corresponding to a particular anatomical tract or fiber cluster.

2.6 Demonstration of the efficacy of harmonization on the ABCD datasets.

In this section, we conducted experiments to demonstrate that we accurately harmonized the baseline dMRI data in the ABCD study.

(a) Experiment 1: Tractography analysis. We first assessed the effects of dMRI harmonization on white matter parcellation (Section 2.4.c) and the extracted dMRI measures (Section 2.4.d) at the tract-level. Specifically, for each tract per scanner, we measured the identification rate (F. Zhang, Wu, et al. 2018; F. Zhang, Cetin Karayumak, et al. 2020) to quantify the quality of the tracts detected before and after harmonization. Here, a tract is considered to be successfully identified if there are at least 10 streamlines inside the tract (F. Zhang, Wu, et al. 2018; F. Zhang, Cetin Karayumak, et al. 2020; F. Zhang, Xie, et al. 2020), and its identification rate is defined as the percentage of successfully identified tracts across all analyzed subjects. In three randomly selected sites, we repeated this experiment on the original dMRI data to be able to compare the identification rate before and after harmonization. In addition, for each tract, we also computed the mean FA across all subjects for each dataset before and after harmonization and compared it to the reference data (to which all datasets were harmonized). We repeated the entire analysis at the cluster-level as well to demonstrate the effects of dMRI harmonization on the clusters.

(b) Experiment 2: Regions-of-interest (ROI) analysis. We further evaluated the performance of the harmonization on multi-shell dMRI measures. To this end, we computed several multi-shell dMRI measures such as Return-To-Origin Probability (RTOP), Return to Axis Probability (RTAP), Return to the Plane Probability (RTPP) and Mean Squared Displacement (MSD) (Özarslan et al. 2013; Ning, Westin, and Rathi 2015; Garyfallidis et al. 2014) in the original, harmonized and reference dMRI datasets using all b-value shells in the dMRI data. We note that for computational efficiency, we compared these measures across the 35 matched subjects that were used in the template creation step. The average of the multi-shell dMRI measures were calculated over the whole brain white matter skeleton and 42 white matter regions using the Illinois Institute of Technology's brain ROI atlas in the standard MNI space (Varentsova, Zhang, and Arfanakis 2014; Qi and Arfanakis 2021). To evaluate the performance of the harmonization, we compared the original and harmonized datasets to the reference dataset using unpaired t-tests.

To further verify the effectiveness of harmonization, we conducted a test using a set of 35 subjects that were not included in the process of creating the template for harmonization. These subjects were selected from the dataset that was being harmonized to the reference data, and were chosen based on their similarity in age, sex, and IQ to the reference data, in

order to minimize any biological differences that could affect the results. This test on the unseen dataset was carried out to replicate the ROI analysis results in an independent dataset that was not used during the template creation (i.e., learning process).

3) Results:

3.1 Sharing the dMRI measures and processed dMRI data in NDA.

The generated imaging files and corresponding dMRI measures for each tract are shared in NDA through two separate submissions.

First, a total of 804 different derived measures, including FA, MD, NoS, etc. for each of the 73 white matter fiber bundles (bilateral, commissural, superficial and cerebellar) were uploaded on the NDA for 9345 subjects. Refer to Table 3 for the list of shared dMRI measures. These measures will be made available by the NDA during their next data release.

Second, we shared the harmonized and processed imaging data, including the harmonized dMRI data, brain masks, whole brain tractography, anatomical fiber tracts (both full-size and downsampled VTK files for tract visualization), and white matter clusters with the NDA community. Refer to Figure 4 for the shared imaging data.

Shared DMRI Measures in NDA per each subject and tract
The total number of streamlines (TotalNoS) in the whole brain. This is the sum of the number of streamlines of all clusters in WhiteMatterClusters.
The below dMRI measures are provided for each tract:
The number of fiber streamline points (NoP)
The number of fiber streamlines (NoS)
The mean length of all fiber streamlines (LenS) (in mm)
The mean tract fractional anisotropy (FA) using tensor 1 (Ten1)
The mean tract axial diffusivity (AD) using tensor 1 (Ten1)
The mean tract mean diffusivity (MD) using tensor 1 (Ten1)
The mean tract radial diffusivity (RD) using tensor 1 (Ten1)
The mean tract fractional anisotropy (FA) using tensor 2 (Ten2)
The mean tract axial diffusivity (AD) using tensor 2 (Ten2)
The mean tract mean diffusivity (MD) using tensor 2 (Ten2)
The mean tract radial diffusivity (RD) using tensor 2 (Ten2)
Table 3. A total of 804 different derived measures, including FA, MD, NoS, etc. for each of the 73 white matter fiber bundles (bilateral, commissural, superficial and cerebellar) were uploaded

on the NDA for 9345 subjects.

Shared processed imaging data file structure in NDA:

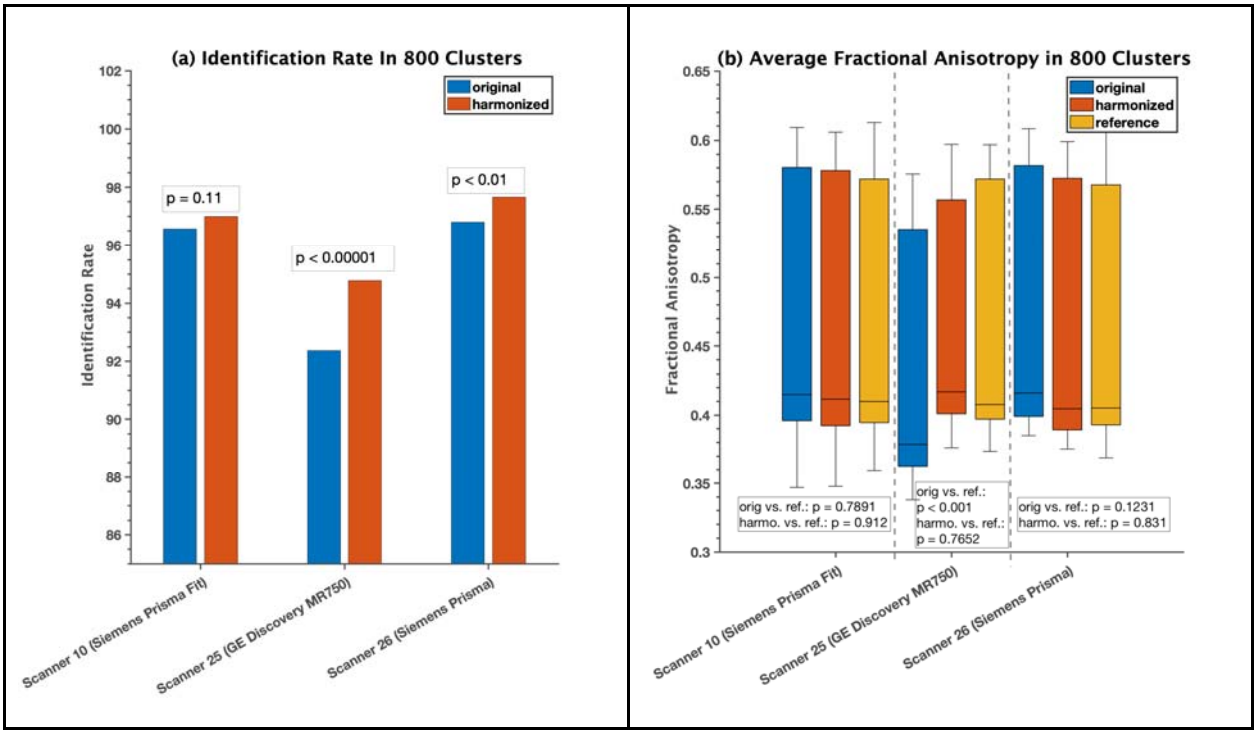
- ${GUID}/ses-baselineYear1Arm1/
  - dwi/
    - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_dwi.json -> json file includes paths to the paths of all files under the dwi folder
    - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_dwi.nii.gz
    - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_dwi.bval
    - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_dwi.bvec
    - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_dwi_mask.nii.gz
  - Tractography/
    - WholeBrainTractography/
      - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_wholebrain.json -> json file includes the path of whole brain tractography file (harmonized_${GUID}_ses-baselineYear1Arm1_run-01_wholebrain.vtk)
      - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_wholebrain.vtk
    - WhiteMatterClustering/
      - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_clusters.json ->json file includes paths to all 800 clusters
      - tracts_commissural/
        - cluster_00001.vtp
        - ...
        - cluster_00800.vtp
      - tracts_left_hemisphere/
        - cluster_00001.vtp
        - ...
        - cluster_00800.vtp
      - tracts_right_hemisphere/
        - cluster_00001.vtp
        - ...
        - cluster_00800.vtp
      - diffusion_measurements_commissural.csv
      - diffusion_measurements_left_hemisphere.csv
      - diffusion_measurements_left_hemisphere.csv
      - diffusion_measurements_right_hemisphere.csv
    - AnatomicalTracts/
      - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_tracts.json -> json file includes paths to all 73 white matter tracts
      - diffusion_measurements_anatomical_tracts.csv
      - T_AF_left.vtp
      - ...
    - AnatomicalTracts_downsampled/ -> Provided for visualization purposes
      - harmonized_${GUID}_ses-baselineYear1Arm1_run-01_tracts_down.json -> json file includes paths to the 73 white matter tracts (downsampled)
      - T_AF_left.vtp
      - ...

Figure 4. Harmonized and processed dMRI data from the ABCD study (N=9345) that have been shared through NDA.

3.2 Validation on harmonized dMRI measures.

The baseline dMRI data of 9345 subjects from the ABCD study underwent harmonization, which brought the dMRI data of all 32 scanners from 18 different sites in line with the reference dataset (dataset #29 from site 16). The harmonization demonstrated consistent results across all dMRI measures in two separate experiments. The outcomes of each of these experiments are detailed in the following sections.

In Experiment 1, the impact of harmonization on dMRI measures obtained through tractography analysis were evaluated. The study focused on three randomly selected sites that used different scanners (Siemens Prisma, Siemens Prisma-fit, GE) to investigate the effect of harmonization on white matter parcellation and extracted dMRI measures. The analysis included both cluster- and tract-level evaluations. Figure 4 illustrates the effect of harmonization on white matter parcellation. Harmonization was found to improve the identification rates in both cluster- and tract-level parcellations (Figure 4 (a) and (c)) for all three sample scanners, except for scanner-a, where both the original and harmonized data had a 100% tract identification rate. Furthermore, the mean cluster and tract FA values (Figure 4 (b) and (d)) were found to be closer to the reference data after harmonization. Across all 9345 harmonized dMRI data from 21 sites in the ABCD study, high identification rates were observed in both tract- and cluster-level parcellations, at 99.9% and 97.5%, respectively.



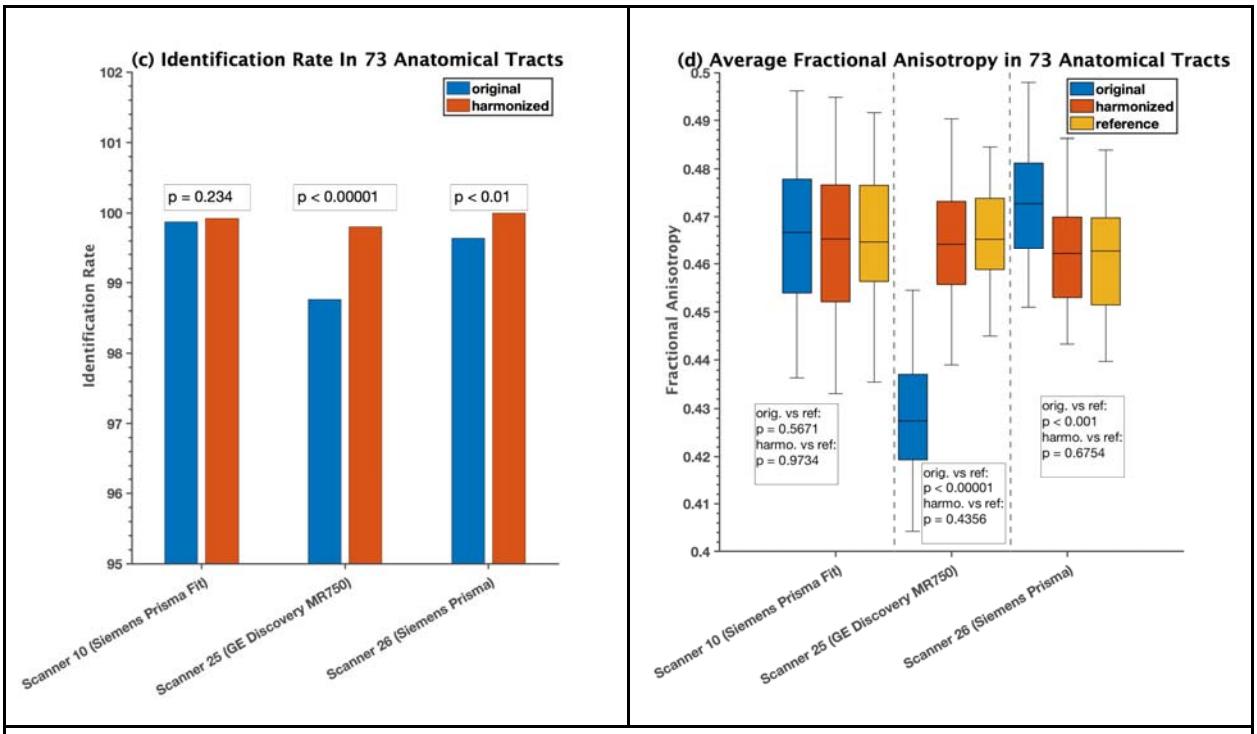


Figure 5. Experiment 1, section 2.6.a, investigated the effects of dMRI harmonization on white matter parcellation in three randomly selected scanners. For this analysis, thirty-five (35) subjects from each scanner, who were part of the learning process of harmonization, were selected for comparisons. These subjects were matched in terms of age, sex, and IQ to the reference scanner (scanner 29). The cluster-level analysis is presented in (a) and (b), while the tract-level analysis is shown in (c) and (d). In (a) and (c), a paired t-test was conducted for each scanner to determine whether there was a significant improvement in the identification rate after harmonization. In (b) and (d), two unpaired t-tests were performed for each scanner to compare the average FA between (i) the original and reference scanners and (ii) the harmonized and reference scanners. The resulting p-values were presented on the plots.

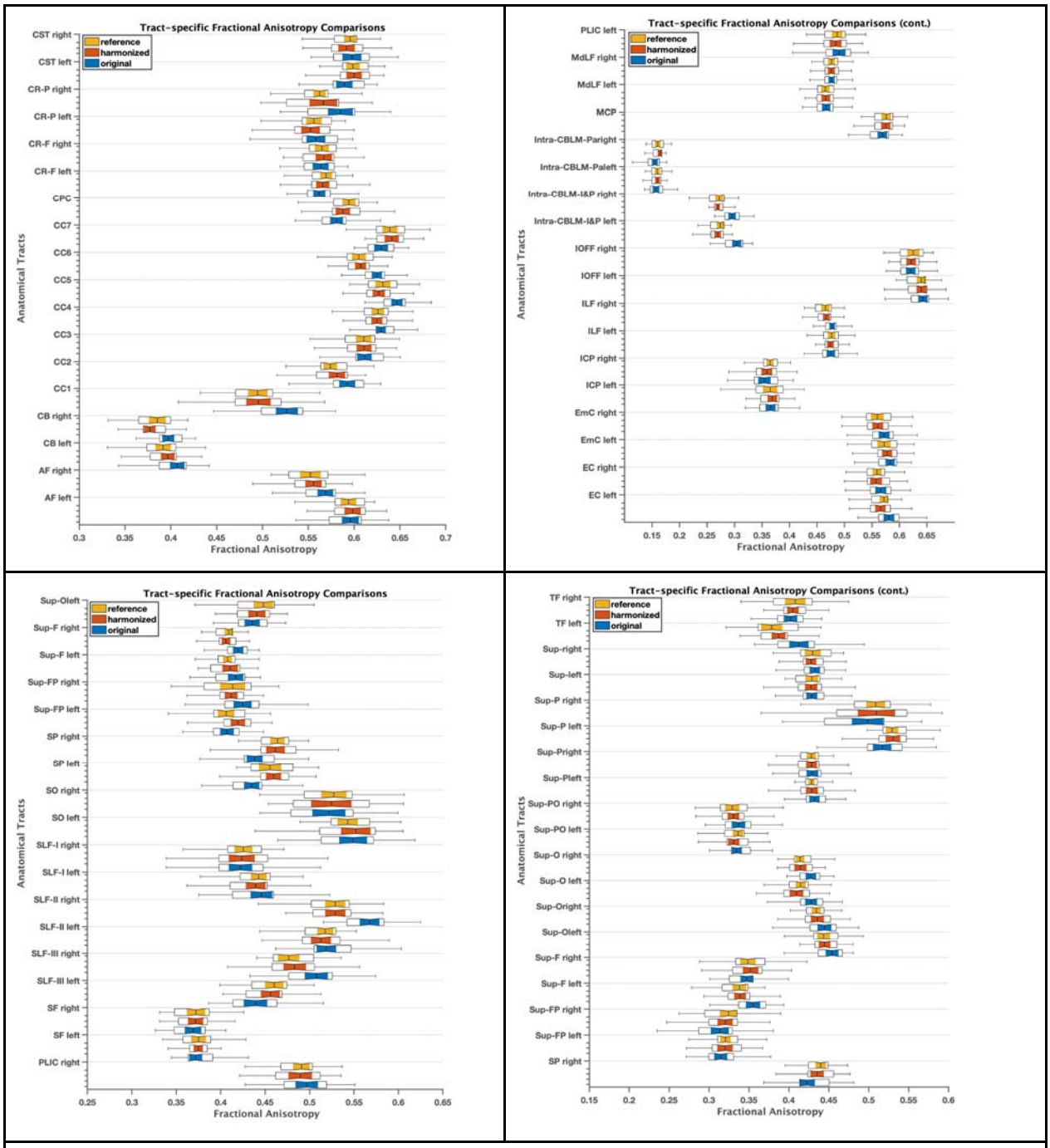


Figure 6. Experiment 1, section 2.6.a, demonstrated the effects of dMRI harmonization on 73 white matter tracts by comparing the average FA of reference, original, and harmonized datasets. Table 2 provides the definition of the acronyms for the tract names. For this analysis, 35 subjects from scanner 10 (Siemens Prisma fit) were selected for comparisons. These subjects were used as part of the learning process of harmonization and were matched in terms of age, sex, and IQ to the reference scanner (scanner 29). Any significant differences observed prior to harmonization ($p<0.01$) were eliminated after harmonization ($p>0.1$).

In Experiment 2, the impact of harmonization on the multi-shell dMRI measures (RTOP, RTAP, RTPP, and MSD) was evaluated. The outcome of this experiment is presented in Figure 7, which demonstrates a comparison of the average multi-shell dMRI measures for different white matter ROIs across eight distinct scanners. The comparison was made between the original data (before harmonization), harmonized data, and reference data. The results indicated significant differences ($p<0.01$) in the multi-shell dMRI measures between the original and reference datasets. However, after harmonization, these differences were reduced and were statistically diminished ($p>0.1$).

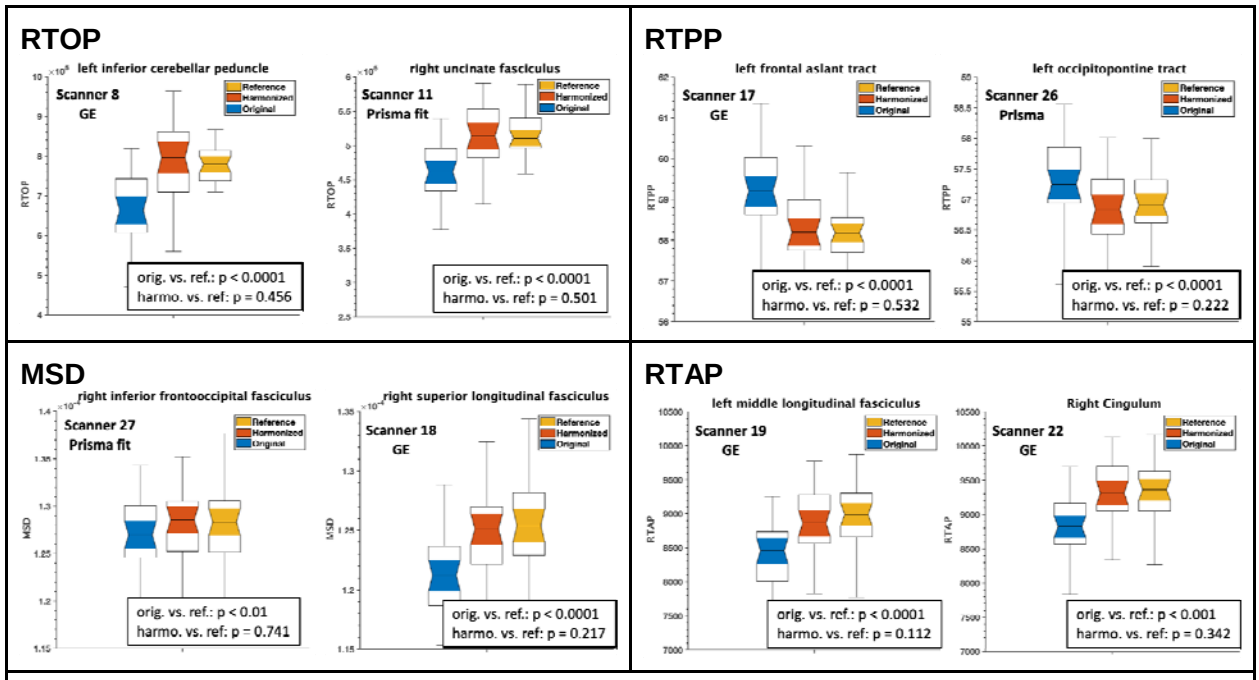


Figure 7. Experiment 2, section 2.6.b, demonstrated the effects of dMRI harmonization on ROI analysis and multi-shell dMRI measures, namely RTOP, RTPP, MSD, RTAP, in eight randomly selected scanners and white matter ROIs. **Thirty-five (35) subjects from each scanner were chosen to create a template for harmonization**, and they were matched in terms of age, sex, and IQ to the reference scanner (scanner 29). Using these subjects from each scanner, two unpaired t-tests were conducted to compare each multi-shell dMRI measure between (i) the original and reference scanners (orig. vs. ref.) and (ii) the harmonized and reference scanners (harmo. vs. ref.). The resulting p-values were presented on the plots.

To further validate the performance of harmonization, we repeated experiment 2 on the dMRI data of new and unseen subjects (i.e., these were not included in the learning process of harmonization/template creation). These subjects were selected from eight scanners and matched with the reference dataset in terms of age, sex, and IQ. The results of the experiment are demonstrated in Figure 8, which compares the average RTOP, RTAP, RTPP, and MSD measures of the several white matter ROIs between the reference, original, and harmonized

datasets. Once again, harmonization eliminated any existing scanner-related differences between the original and reference datasets.

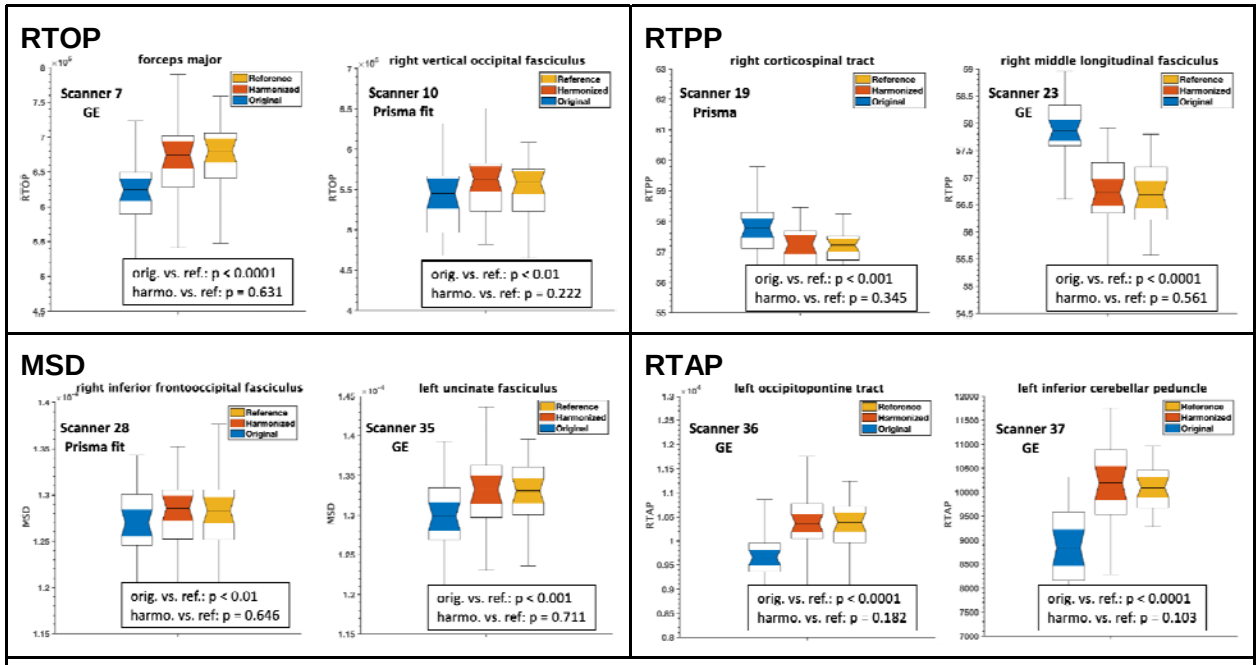


Figure 8. Experiment 2, section 2.6.b, investigated the effects of dMRI harmonization on ROI analysis and multi-shell dMRI measures, namely RTOP, RTPP, MSD, RTAP, in eight randomly selected scanners and white matter ROIs. For this analysis, **an unseen set of thirty-five (35) subjects who were not involved in creating the template** and were still matched in terms of age, sex, and IQ to the reference scanner (scanner 29) were selected. Using these subjects from each scanner, two unpaired t-tests were conducted to compare each multi-shell dMRI measure between (i) the original and reference scanners (orig. vs. ref.) and (ii) the harmonized and reference scanners (harmo. vs. ref.). The resulting p-values were presented on the plots.

4) Discussion:

Our study aimed to harmonize and process the baseline dMRI data from the ABCD study comprehensively. The ultimate goal was to create a harmonized and processed dMRI database for the ABCD study, which would enable researchers to analyze the entire dMRI data together. Our previous success in harmonizing a large schizophrenia database (Suheyla Cetin-Karayumak, Di Biase, et al. 2020) and in the harmonization community challenge (Ning et al. 2020) inspired us to undertake this effort. After eliminating the poor quality dMRI scans through automated and manual QC protocols, we applied the following computational procedures using our processing pipeline on the remaining quality passed dMRI data acquired across 18 sites with 33 scanners using Siemens and GE manufacturers (N=9345). 1) harmonization of multi-site dMRI data at the signal level to remove any statistical differences in measurements caused by the scanners; 2) Whole brain tractography on the harmonized dMRI data to generate approximately 450k fiber streamlines per subject; 3) White matter analysis to identify white

matter clusters and anatomical tracts, including 800 fiber clusters, 73 white matter tracts (57 major deep white matter tracts, and 16 superficial white matter tracts); 4) Extraction of dMRI-related measures from each of the parcellated white matter clusters and tracts.

We evaluated the effectiveness of our dMRI data harmonization by conducting two experiments. In experiment 1, our results showed that the dMRI harmonization improved the accuracy of white matter parcellation and led to a higher identification rate, as well as a closer alignment with the reference data in terms of FA. Additionally, the results showed that the WM parcellation was robust across all harmonized datasets, with over 99.9% of the anatomical tracts being accurately identified. In experiment 2, our results demonstrated how well the harmonization process removed global and local scanner-related differences across the datasets in multi-shell measures (RTOP, RTPP, RTAP, MSD). Further, we demonstrated that the harmonization performance was consistent on unseen new datasets, which were not part of the template creation.

The measures of the white matter anatomical tracts extracted from the harmonized and processed dMRI data is accessible through the NDA. This includes a total of 808 different derived measures, including FA, MD, NoF, etc., for each of the white matter fiber bundles (78 bilateral, commissural, superficial, and cerebellar) per each subject. We have also shared the processed imaging data, including the harmonized dMRI data, brain masks, whole-brain tractography, white matter clusters, and white matter tracts in NDA. Leveraging the harmonized and processed dMRI data will allow for pooled large-scale data analysis as if the data came from the same scanner, which would significantly increase statistical power of all neuroimaging studies using this comprehensive dataset, the ABCD study. Moreover, the harmonized dMRI data is agnostic to the use of any type of diffusion model (Panagiotaki et al. 2012) and thus can be used to derive other measures of interest (e.g. NODDI (H. Zhang et al. 2012), WMTI (Fieremans, Jensen, and Helpern 2011), structural disturbance (Ning et al. 2017)).

This study has several limitations that are necessary to note. The primary limitation is that we had to exclude the dMRI scans acquired on the Philips scanner as it did not pass our quality checks. Although the excluded scans make up a small portion of the overall dMRI data (less than 13% of all dMRI scans), these subjects can still be distinct enough to affect the statistical significance of studies investigating specific biology- or disease-related characteristics. Additionally, all dMRI data included in this sample comes from only baseline sessions. On the other hand, the ABCD study has been collecting dMRI data for each subject every two years. Hence, our future work will focus on harmonizing the dMRI data from following years to provide the ability to analyze the longitudinal dMRI data (e.g., characterize long-term white matter changes).

In conclusion, we presented our efforts to create a harmonized database on the multi-site dMRI data from the ABCD study. The creation of this database involved a significant computational undertaking, which required approximately 50,000 CPU hours to successfully harmonize the dMRI data, conduct tractography for the entire brain, and execute white matter parcellation. The results of our efforts have been made available to the scientific community via the NDA. The harmonized dMRI database from the baseline session includes the output of the harmonization, brain masking and whole brain tractography, as well as white matter tracts, white matter clusters, and derived dMRI-related measures. We believe that this harmonized database, which pools together data from multiple sites and scanners without significant scanner bias, will

significantly enhance the statistical power of research based on the ABCD study. This will provide the ability to run more advanced statistical and machine learning analyses aimed at uncovering the neurodevelopmental changes in the white matter of adolescents.

Acknowledgments:

We gratefully acknowledge funding provided by the following National Institutes of Health (NIH) grant: R01 MH119222 (PIs: Dr. Yogesh Rathi, Dr. Lauren J. O'Donnell). We also acknowledge funding provided by the Brigham and Women's Hospital Program for Interdisciplinary Neurosciences through a gift from Lawrence and Tina Rand and the Brain and Behavior Research Foundation NARSAD Young Investigator Award (PI: Dr. Suheyla Cetin-Karayumak).

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

Disclosures:

The authors declare that they have nothing to disclose.

References:

A., Roy, Conjeti S., Navab N., and C. Wachinger. 2019. "QuickNAT: A Fully Convolutional Network for Quick and Accurate Segmentation of Neuroanatomy." *NeuroImage* 186 (February): 713–27.

Bayer, Johanna M. M., Paul M. Thompson, Christopher R. K. Ching, Mengting Liu, Andrew Chen, Alana C. Panzenhagen, Neda Jahanshad, Andre Marquand, Lianne Schmaal, and Philipp G. Sämann. 2022. "Site Effects How-to and When: An Overview of Retrospective Techniques to Accommodate Site Effects in Multi-Site Neuroimaging Analyses." *Frontiers in Neurology* 13 (October): 923988.

Biase, Maria A. Di, Maria A. Di Biase, Andrew Zalesky, Suheyla Cetin-Karayumak, Yogesh Rathi, Jinglei Lv, Danny Boerrigter, et al. 2021. "Large-Scale Evidence for an Association Between Peripheral Inflammation and White Matter Free Water in Schizophrenia and

Healthy Individuals." *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbaa134.

Botvinik-Nezer, Rotem, Felix Holzmeister, Colin F. Camerer, Anna Dreber, Juergen Huber, Magnus Johannesson, Michael Kirchler, et al. 2020. "Variability in the Analysis of a Single Neuroimaging Dataset by Many Teams." *Nature* 582 (7810): 84–88.

Casey, B. J., Tariq Cannonier, May I. Conley, Alexandra O. Cohen, Deanna M. Barch, Mary M. Heitzeg, Mary E. Soules, et al. 2018. "The Adolescent Brain Cognitive Development (ABCD) Study: Imaging Acquisition across 21 Sites." *Developmental Cognitive Neuroscience* 32 (August): 43–54.

Cetin Karayumak S*, Billah T*, Bouix S, Rathi Y. 2019. "Multi-Site Diffusion MRI Harmonization." https://doi.org/10.5281/zenodo.2584275.

Cetin-Karayumak, S., M. Kubicki, and Y. Rathi. 2018. "Harmonizing Diffusion MRI Data across Magnetic Field Strengths." *Medical Image Computing and Computer-Assisted Intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*. https://link.springer.com/chapter/10.1007/978-3-030-00931-1_14.

Cetin Karayumak, Suheyla, Sylvain Bouix, Lipeng Ning, Anthony James, Tim Crow, Martha Shenton, Marek Kubicki, and Yogesh Rathi. 2019. "Retrospective Harmonization of Multi-Site Diffusion MRI Data Acquired with Different Acquisition Parameters." *NeuroImage* 184 (January): 180–200.

Cetin-Karayumak, Suheyla, Maria A. Di Biase, Natalia Chunga, Benjamin Reid, Nathaniel Somes, Amanda E. Lyall, Sinead Kelly, et al. 2020. "White Matter Abnormalities across the Lifespan of Schizophrenia: A Harmonized Multi-Site Diffusion MRI Study." *Molecular Psychiatry* 25 (12): 3208–19.

Cetin-Karayumak, Suheyla, Amanda Lyall, Maria Di Biase, Johanna Seitz-Holland, Fan Zhang, Sinead Kelly, Doron Elad, et al. n.d. "Characterization of Extracellular Free Water Pathologies in Schizophrenia Using Multi-Site Diffusion MRI Harmonization." https://doi.org/10.21203/rs.3.rs-1742219/v1.

Cetin-Karayumak, Suheyla, Katharina Stegmayer, Sebastian Walther, Philip R. Szeszko, Tim Crow, Anthony James, Matcheri Keshavan, Marek Kubicki, and Yogesh Rathi. 2020. "Exploring the Limits of ComBat Method for Multi-Site Diffusion MRI Harmonization." *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2020.11.20.390120.

De Luca, Alberto, Suheyla Cetin Karayumak, Alexander Leemans, Yogesh Rathi, Stephan Swinnen, Jolien Gooijers, Amanda Clauwaert, et al. 2022. "Cross-Site Harmonization of Multi-Shell Diffusion MRI Measures Based on Rotational Invariant Spherical Harmonics (RISH)." *NeuroImage* 259 (October): 119439.

———. n.d. "Harmonization of Diffusion Kurtosis Imaging Metrics with Rotational Invariant Spherical Harmonics (RISH) to Remove Cross-Site Biases." https://archive.ismrm.org/2021/1724.html.

Dey, Raunak, and Yi Hong. 2018. "CompNet: Complementary Segmentation Network for Brain MRI Extraction." *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. https://doi.org/10.1007/978-3-030-00931-1_72.

Elad, Doron, Suheyla Cetin□Karayumak, Fan Zhang, Kang Ik K. Cho, Amanda E. Lyall, Johanna Seitz□Holland, Rami Ben□Ari, et al. 2021. "Improving the Predictive Potential of Diffusion MRI in Schizophrenia Using Normative models—Towards Subject□level Classification." *Human Brain Mapping*. https://doi.org/10.1002/hbm.25574.

Fieremans, Els, Jens H. Jensen, and Joseph A. Helpern. 2011. “White Matter Characterization with Diffusional Kurtosis Imaging.” *NeuroImage* 58 (1): 177–88.

Fillard, Pierre, Maxime Descoteaux, Alvina Goh, Sylvain Gouttard, Ben Jeurissen, James Malcolm, Alonso Ramirez-Manzanares, et al. 2011. “Quantitative Evaluation of 10 Tractography Algorithms on a Realistic Diffusion MR Phantom.” *NeuroImage* 56 (1): 220–34.

Fortin, Jean-Philippe, Drew Parker, Birkan Tunç, Takanori Watanabe, Mark A. Elliott, Kosha Ruparel, David R. Roalf, et al. 2017. “Harmonization of Multi-Site Diffusion Tensor Imaging Data.” *NeuroImage* 161 (November): 149–70.

Garyfallidis, Eleftherios, Matthew Brett, Bagrat Amirbekian, Ariel Rokem, Stefan van der Walt, Maxime Descoteaux, Ian Nimmo-Smith, and Dipy Contributors. 2014. “Dipy, a Library for the Analysis of Diffusion MRI Data.” *Frontiers in Neuroinformatics* 8 (February): 8.

Hagler, Donald J., Jr, Seann Hatton, M. Daniela Cornejo, Carolina Makowski, Damien A. Fair, Anthony Steven Dick, Matthew T. Sutherland, et al. 2019. “Image Processing and Analysis Methods for the Adolescent Brain Cognitive Development Study.” *NeuroImage* 202 (November): 116091.

Hamoda, Hesham M., A. T. Makhlouf, J. Fitzsimmons, Y. Rathi, N. Makris, R. I. Mesholam-Gately, J. D. Wojcik, et al. 2019. “Abnormalities in Thalamo-Cortical Connections in Patients with First-Episode Schizophrenia: A Two-Tensor Tractography Study.” *Brain Imaging and Behavior* 13 (2): 472–81.

Hatton, Sean. 2020. “Preview of the Adolescent Brain Cognitive Development (ABCD) Study Release 3.0.” *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2020.02.303.

Hegde, Rachal, Paulo Lizano, Sinead Kelly, Olivia Lutz, Suheyla Cetin Karayumak, Ofer Pasternak, Marek Kubicki, Shaun Eack, and Matcheri Keshavan. 2021. “Cognitive Enhancement Therapy Might Halt Progressive White Matter Microstructural Changes in Early Course Schizophrenia.” *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2021.02.534.

Huynh, Khoi Minh, Geng Chen, Ye Wu, Dinggang Shen, and Pew-Thian Yap. 2019. “Multi-Site Harmonization of Diffusion MRI Data via Method of Moments.” *IEEE Transactions on Medical Imaging* 38 (7): 1599–1609.

Jelescu, Ileana O., Jelle Veraart, Els Fieremans, and Dmitry S. Novikov. 2016. “Degeneracy in Model Parameter Estimation for Multi-Compartmental Diffusion in Neuronal Tissue.” *NMR in Biomedicine* 29 (1): 33–47.

Jenkinson, Mark, Christian F. Beckmann, Timothy E. J. Behrens, Mark W. Woolrich, and Stephen M. Smith. 2012. “FSL.” *NeuroImage* 62 (2): 782–90.

Jernigan, Terry L., Sandra A. Brown, and Gayathri J. Dowling. 2018. “The Adolescent Brain Cognitive Development Study.” *Journal of Research on Adolescence: The Official Journal of the Society for Research on Adolescence* 28 (1): 154–56.

Johnson, W. Evan, Cheng Li, and Ariel Rabinovic. 2007. “Adjusting Batch Effects in Microarray Expression Data Using Empirical Bayes Methods.” *Biostatistics* 8 (1): 118–27.

Kelly, S., S. Guimond, O. Pasternak, and O. Lutz. 2021. “White Matter Microstructure across Brain-Based Biotypes for Psychosis–findings from the Bipolar-Schizophrenia Network for Intermediate Phenotypes.” *Psychiatry Research*. https://www.sciencedirect.com/science/article/pii/S0925492720302067.

Kochunov, Peter, Neda Jahanshad, Emma Sprooten, Thomas E. Nichols, René C. Mandl, Laura Almasy, Tom Booth, et al. 2014. “Multi-Site Study of Additive Genetic Effects on Fractional Anisotropy of Cerebral White Matter: Comparing Meta and Megaanalytical Approaches for Data Pooling.” *NeuroImage* 95 (July): 136–50.

Levitt, James, Fan Zhang, Mark Vangel, Paul Nestor, Yogesh Rathi, Suheyla Cetin-Karayumak, Marek Kubicki, et al. n.d. “The Organization of Frontostriatal Brain Wiring in Non-Affective Early Psychosis Compared with Healthy Subjects Using a Novel Diffusion Imaging Fiber Cluster Analysis.” https://doi.org/10.21203/rs.3.rs-2062646/v1.

Lv, Jinglei, Maria Di Biase, Robin F. H. Cash, Luca Cocchi, Vanessa L. Cropley, Paul Klauser, Ye Tian, et al. 2020. “Individual Deviations from Normative Models of Brain Structure in a Large Cross-Sectional Schizophrenia Cohort.” *Molecular Psychiatry*, September. https://doi.org/10.1038/s41380-020-00882-5.

Malcolm, James G., Martha E. Shenton, and Yogesh Rathi. 2010. “Filtered Multitensor Tractography.” *IEEE Transactions on Medical Imaging* 29 (9): 1664–75.

Medawar, Evelyn, Ronja Thieleking, Iryna Manuilova, Maria Paerisch, Arno Villringer, A. Veronica Witte, and Frauke Beyer. 2021. “Estimating the Effect of a Scanner Upgrade on Measures of Grey Matter Structure for Longitudinal Designs.” *PloS One* 16 (10): e0239021.

Mirzaalian, Hengameh, Lipeng Ning, Peter Savadjiev, Ofer Pasternak, Sylvain Bouix, Oleg Michailovich, Sarina Karmacharya, et al. 2018. “Multi-Site Harmonization of Diffusion MRI Data in a Registration Framework.” *Brain Imaging and Behavior* 12 (1): 284–95.

Mirzaalian, Hengameh, Amicie de Pierrefeu, Peter Savadjiev, Ofer Pasternak, Sylvain Bouix, Marek Kubicki, Carl-Fredrik Westin, Martha E. Shenton, and Yogesh Rathi. 2015. “Harmonizing Diffusion MRI Data Across Multiple Sites and Scanners.” In *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015*, edited by Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro Frangi, 9349:12–19. Lecture Notes in Computer Science. Cham: Springer International Publishing.

Mirzaalian, H., L. Ning, P. Savadjiev, O. Pasternak, S. Bouix, O. Michailovich, G. Grant, et al. 2016. “Inter-Site and Inter-Scanner Diffusion MRI Data Harmonization.” *NeuroImage* 135 (July): 311–23.

Moyer, Daniel, Greg Ver Steeg, Chantal M. W. Tax, and Paul M. Thompson. 2020. “Scanner Invariant Representations for Diffusion MRI Harmonization.” *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 84 (4): 2174–89.

Ning, Lipeng, Elisenda Bonet-Carne, Francesco Grussu, Farshid Sepehrband, Enrico Kaden, Jelle Veraart, Stefano B. Blumberg, et al. 2020. “Cross-Scanner and Cross-Protocol Multi-Shell Diffusion MRI Data Harmonization: Algorithms and Results.” *NeuroImage* 221 (November): 117128.

Ning, Lipeng, Evren Özarslan, Carl-Fredrik Westin, and Yogesh Rathi. 2017. “Precise Inference and Characterization of Structural Organization (PICASO) of Tissue from Molecular Diffusion.” *NeuroImage* 146 (February): 452–73.

Ning, Lipeng, Carl-Fredrik Westin, and Yogesh Rathi. 2015. “Estimating Diffusion Propagator and Its Moments Using Directional Radial Basis Functions.” *IEEE Transactions on Medical Imaging* 34 (10): 2058–78.

Norton, Isaiah, Walid Ibn Essayed, Fan Zhang, Sonia Pujol, Alex Yarmarkovich, Alexandra J.

Golby, Gordon Kindlmann, et al. 2017. “SlicerDMRI: Open Source Diffusion MRI Software for Brain Cancer Research.” *Cancer Research* 77 (21): e101–3.

O’Donnell, Lauren J., William M. Wells 3rd, Alexandra J. Golby, and Carl-Fredrik Westin. 2012a. “Unbiased Groupwise Registration of White Matter Tractography.” *Medical Image Computing and Computer-Assisted Intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 15 (Pt 3): 123–30.

O’Donnell, Lauren J., William M. Wells III, Alexandra J. Golby, and Carl-Fredrik Westin. 2012b. “Unbiased Groupwise Registration of White Matter Tractography.” In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, edited by Nicholas Ayache, Hervé Delingette, Polina Golland, and Kensaku Mori, 123–30. Lecture Notes in Computer Science 7512. Springer Berlin Heidelberg.

O’Donnell, Lauren J., and Carl-Fredrik Westin. 2007. “Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas.” *IEEE Transactions on Medical Imaging* 26 (11): 1562–75.

Olszewski, Amy K., Zora Kikinis, Christie S. Gonzalez, Ioana L. Coman, Nikolaos Makris, Xue Gong, Yogesh Rathi, et al. 2017. “The Social Brain Network in 22q11.2 Deletion Syndrome: A Diffusion Tensor Imaging Study.” *Behavioral and Brain Functions: BBF* 13 (1): 4.

Özarslan, Evren, Cheng Guan Koay, Timothy M. Shepherd, Michal E. Komlosh, M. Okan İrfanoğlu, Carlo Pierpaoli, and Peter J. Basser. 2013. “Mean Apparent Propagator (MAP) MRI: A Novel Diffusion Imaging Method for Mapping Tissue Microstructure.” *NeuroImage* 78 (September): 16–32.

Panagiotaki, Eleftheria, Torben Schneider, Bernard Siow, Matt G. Hall, Mark F. Lythgoe, and Daniel C. Alexander. 2012. “Compartment Models of the Diffusion MR Signal in Brain White Matter: A Taxonomy and Comparison.” *NeuroImage* 59 (3): 2241–54.

Pinto, Maíra Siqueira, Roberto Paolella, Thibo Billiet, Pieter Van Dyck, Pieter-Jan Guns, Ben Jeurissen, Annemie Ribbens, Arnold J. den Dekker, and Jan Sijbers. 2020. “Harmonization of Brain Diffusion MRI: Concepts and Methods.” *Frontiers in Neuroscience* 14. https://doi.org/10.3389/fnins.2020.00396.

Qi, Xiaoxiao, and Konstantinos Arfanakis. 2021. “Regionconnect: Rapidly Extracting Standardized Brain Connectivity Information in Voxel-Wise Neuroimaging Studies.” *NeuroImage* 225 (January): 117462.

Reddy, Chinthala P., and Yogesh Rathi. 2016. “Joint Multi-Fiber NODDI Parameter Estimation and Tractography Using the Unscented Information Filter.” *Frontiers in Neuroscience* 10 (April): 166.

Reisert, Marco, Elias Kellner, Bibek Dhital, Jürgen Hennig, and Valerij G. Kiselev. 2017. “Disentangling Micro from Mesostructure by Diffusion MRI: A Bayesian Approach.” *NeuroImage* 147 (February): 964–75.

Schilling, Kurt G., François Rheault, Laurent Petit, Colin B. Hansen, Vishwesh Nath, Fang-Cheng Yeh, Gabriel Girard, et al. 2020. “Tractography Dissection Variability: What Happens When 42 Groups Dissect 14 White Matter Bundles on the Same Dataset?” *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2020.10.07.321083.

Schilling, Kurt G., Chantal M. W. Tax, Francois Rheault, Colin Hansen, Qi Yang, Fang-Cheng Yeh, Leon Cai, Adam W. Anderson, and Bennett A. Landman. 2021. “Fiber Tractography Bundle Segmentation Depends on Scanner Effects, Vendor Effects, Acquisition Resolution,

Diffusion Sampling Scheme, Diffusion Sensitization, and Bundle Segmentation Workflow." *NeuroImage* 242 (November): 118451.

Seitz-Holland, Johanna, Suheyla Cetin-Karayumak, Joanne D. Wojcik, Amanda Lyall, James Levitt, Martha E. Shenton, Ofer Pasternak, et al. 2021. "Elucidating the Relationship between White Matter Structure, Demographic, and Clinical Variables in Schizophrenia-a Multicenter Harmonized Diffusion Tensor Imaging Study." *Molecular Psychiatry*, January. https://doi.org/10.1038/s41380-021-01018-z.

Seitz-Holland, Johanna, Joanne D. Wojcik, Suheyla Cetin-Karayumak, Amanda E. Lyall, Ofer Pasternak, Yogesh Rathi, Mark Vangel, et al. 2022. "Cognitive Deficits, Clinical Variables, and White Matter Microstructure in Schizophrenia: A Multisite Harmonization Study." *Molecular Psychiatry* 27 (9): 3719–30.

Seitz, Johanna, Suheyla Cetin-Karayumak, Amanda Lyall, Ofer Pasternak, Madhura Baxi, Mark Vangel, Godfrey Pearlson, et al. 2021. "Investigating Sexual Dimorphism of Human White Matter in a Harmonized, Multisite Diffusion Magnetic Resonance Imaging Study." *Cerebral Cortex* 31 (1): 201–12.

Smith, Stephen M. 2002. "Fast Robust Automated Brain Extraction." *Human Brain Mapping* 17 (3): 143–55.

Tax, Chantal Mw, Francesco Grussu, Enrico Kaden, Lipeng Ning, Umesh Rudrapatna, C. John Evans, Samuel St-Jean, et al. 2019. "Cross-Scanner and Cross-Protocol Diffusion MRI Data Harmonisation: A Benchmark Database and Evaluation of Algorithms." *NeuroImage* 195 (July): 285–99.

Thieleking, Ronja, Rui Zhang, Maria Paerisch, Kerstin Wirkner, Alfred Anwander, Frauke Beyer, Arno Villringer, and A. Veronica Witte. 2021. "Same Brain, Different Look?—The Impact of Scanner, Sequence and Preprocessing on Diffusion Imaging Outcome Parameters." *Journal of Clinical Medicine*. https://doi.org/10.3390/jcm10214987.

Thompson, Paul M., Neda Jahanshad, Christopher R. K. Ching, Lauren E. Salminen, Sophia I. Thomopoulos, Joanna Bright, Bernhard T. Baune, et al. 2020. "ENIGMA and Global Neuroscience: A Decade of Large-Scale Studies of the Brain in Health and Disease across More than 40 Countries." *Translational Psychiatry* 10 (1): 100.

Thompson, Paul M., Jason L. Stein, Sarah E. Medland, Derrek P. Hibar, Alejandro Arias Vasquez, Miguel E. Renteria, Roberto Toro, et al. 2014. "The ENIGMA Consortium: Large-Scale Collaborative Analyses of Neuroimaging and Genetic Data." *Brain Imaging and Behavior* 8 (2): 153–82.

Ukftractography. n.d. Github. Accessed November 5, 2021. https://github.com/pnlbwh/ukftractography.

Varentsova, Anna, Shengwei Zhang, and Konstantinos Arfanakis. 2014. "Development of a High Angular Resolution Diffusion Imaging Human Brain Template." *NeuroImage* 91 (May): 177–86.

Vollmar, Christian, Jonathan O'Muircheartaigh, Gareth J. Barker, Mark R. Symms, Pamela Thompson, Veena Kumari, John S. Duncan, Mark P. Richardson, and Matthias J. Koepp. 2010. "Identical, but Not the Same: Intra-Site and Inter-Site Reproducibility of Fractional Anisotropy Measures on Two 3.0T Scanners." *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2010.03.046.

Wasserthal, Jakob, Peter Neher, and Klaus H. Maier-Hein. 2018. "TractSeg - Fast and Accurate

White Matter Tract Segmentation." *NeuroImage* 183 (December): 239–53.

Whitematteranalysis. n.d. Github. Accessed November 5, 2021. https://github.com/SlicerDMRI/whitematteranalysis.

Ye, Hua, Andrew Zalesky, Jinglei Lv, Samantha M. Loi, Suheyla Cetin-Karayumak, Yogesh Rathi, Ye Tian, Christos Pantelis, and Maria A. Di Biase. 2021. "Network Analysis of Symptom Comorbidity in Schizophrenia: Relationship to Illness Course and Brain White Matter Microstructure." *Schizophrenia Bulletin*, March. https://doi.org/10.1093/schbul/sbab015.

Yendiki, Anastasia, Patricia Panneck, Priti Srinivasan, Allison Stevens, Lilla Zöllei, Jean Augustinack, Ruopeng Wang, et al. 2011. "Automated Probabilistic Reconstruction of White-Matter Pathways in Health and Disease Using an Atlas of the Underlying Anatomy." *Frontiers in Neuroinformatics* 5 (October): 23.

Zhang, Fan, Suheyla Cetin Karayumak, Nico Hoffmann, Yogesh Rathi, Alexandra J. Golby, and Lauren J. O'Donnell. 2020. "Deep White Matter Analysis (DeepWMA): Fast and Consistent Tractography Segmentation." *Medical Image Analysis* 65 (October): 101761.

Zhang, Fan, Alessandro Daducci, Yong He, Simona Schiavi, Caio Seguin, Robert E. Smith, Chun-Hung Yeh, Tengda Zhao, and Lauren J. O'Donnell. 2022. "Quantitative Mapping of the Brain's Structural Connectivity Using Diffusion MRI Tractography: A Review." *NeuroImage* 249 (April): 118870.

Zhang, Fan, Thomas Noh, Parikshit Juvekar, Sarah F. Frisken, Laura Rigolo, Isaiah Norton, Tina Kapur, et al. 2020. "SlicerDMRI: Diffusion MRI and Tractography Research Software for Brain Cancer Surgery Planning and Visualization." *JCO Clin Cancer Inform*.

Zhang, Fan, Peter Savadjiev, Weidong Cai, Yang Song, Yogesh Rathi, Birkan Tunç, Drew Parker, et al. 2018. "Whole Brain White Matter Connectivity Analysis Using Machine Learning: An Application to Autism." *NeuroImage* 172 (May): 826–37.

Zhang, Fan, Ye Wu, Isaiah Norton, Yogesh Rathi, Alexandra J. Golby, and Lauren J. O'Donnell. 2019. "Test-Retest Reproducibility of White Matter Parcellation Using Diffusion MRI Tractography Fiber Clustering." *Human Brain Mapping* 40 (10): 3041–57.

Zhang, Fan, Ye Wu, Isaiah Norton, Laura Rigolo, Yogesh Rathi, Nikos Makris, and Lauren J. O'Donnell. 2018. "An Anatomically Curated Fiber Clustering White Matter Atlas for Consistent White Matter Tract Parcellation across the Lifespan." *NeuroImage* 179 (October): 429–47.

Zhang, Fan, Guoqiang Xie, Laura Leung, Michael A. Mooney, Lorenz Epprecht, Isaiah Norton, Yogesh Rathi, et al. 2020. "Creation of a Novel Trigeminal Tractography Atlas for Automated Trigeminal Nerve Identification." *NeuroImage* 220 (June): 117063.

Zhang, Guohao, Peter Kochunov, Elliot Hong, Sinead Kelly, Christopher Whelan, Neda Jahanshad, Paul Thompson, and Jian Chen. 2017. "ENIGMA-Viewer: Interactive Visualization Strategies for Conveying Effect Sizes in Meta-Analysis." *BMC Bioinformatics* 18 (Suppl 6): 253.

Zhang, Hui, Torben Schneider, Claudia A. Wheeler-Kingshott, and Daniel C. Alexander. 2012. "NODDI: Practical in Vivo Neurite Orientation Dispersion and Density Imaging of the Human Brain." *NeuroImage* 61 (4): 1000–1016.

Appendix:

A. Convolutional Neural Network (CNN) DMRI Brain Segmentation.

The CNN architecture in this method is based on the study of (Dey and Hong 2018), which was trained to skull-strip the T1-weighted images. In our case, however, we trained the network on baseline (b=0 s/mm^2) images of the dMRI data to separate the brain from non-brain regions in the image. To further improve the segmentation, we modified the CNN architecture by integrating a multi-view aggregation step (A. et al. 2019). This step combined the results from models trained on 2D slices along three primary axes: coronal, sagittal, and axial. The final brain mask is obtained by combining the probability maps from all three networks.

The brain masking process included a pre-processing step where we registered the b0 images to the standard MNI space along with performing data normalization (scaling the b0 values between 0 and 1) to improve the performance of the model in the training step. The deep learning models were trained using manually curated masks of dMRI data from 1500 subjects collected from 10 different datasets, including 200 subjects from the ABCD study. We employed the Adam optimizer with a learning rate of 1e-3 during the training process, and all networks were run for up to 10 epochs. We implemented the deep learning algorithm in Keras-Tensorflow 2.2.4. Finally, as a post-processing step, we transformed the resulting brain mask from the MNI space back to the subject space. These steps are summarized in Figure A.2.

To demonstrate the effectiveness of our method with respect to existing tools, we compared our deep learning brain masking method with the most commonly used approach, Brain Extraction Tool (BET) (Smith 2002). We evaluated approximately 400 manually corrected masks from the ABCD study with the output of our deep learning network, resulting in a Dice overlap coefficient of 0.99 and a Jaccard index of 0.987. We note that none of these ~400 subjects were part of the training dataset. In contrast, BET produced lower performance on the same dataset, with a Dice overlap coefficient of 0.957 and a Jaccard index of 0.945.

Finally, we deployed this brain masking tool on the dMRI data of 9345 subjects from the ABCD study using Amazon Web Services (AWS) Elastic Compute Cloud (EC2) g4dn.2xlarge instances (32 GiB CPU memory, NVIDIA T4 GPU with 32 GiB GPU memory, 8 vCPUs), where the entire brain masking process took about 1 min and 30 seconds for each subject. Refer to Figure A.1 for the demonstration of the dMRI data brain masks on 5 randomly picked brains.

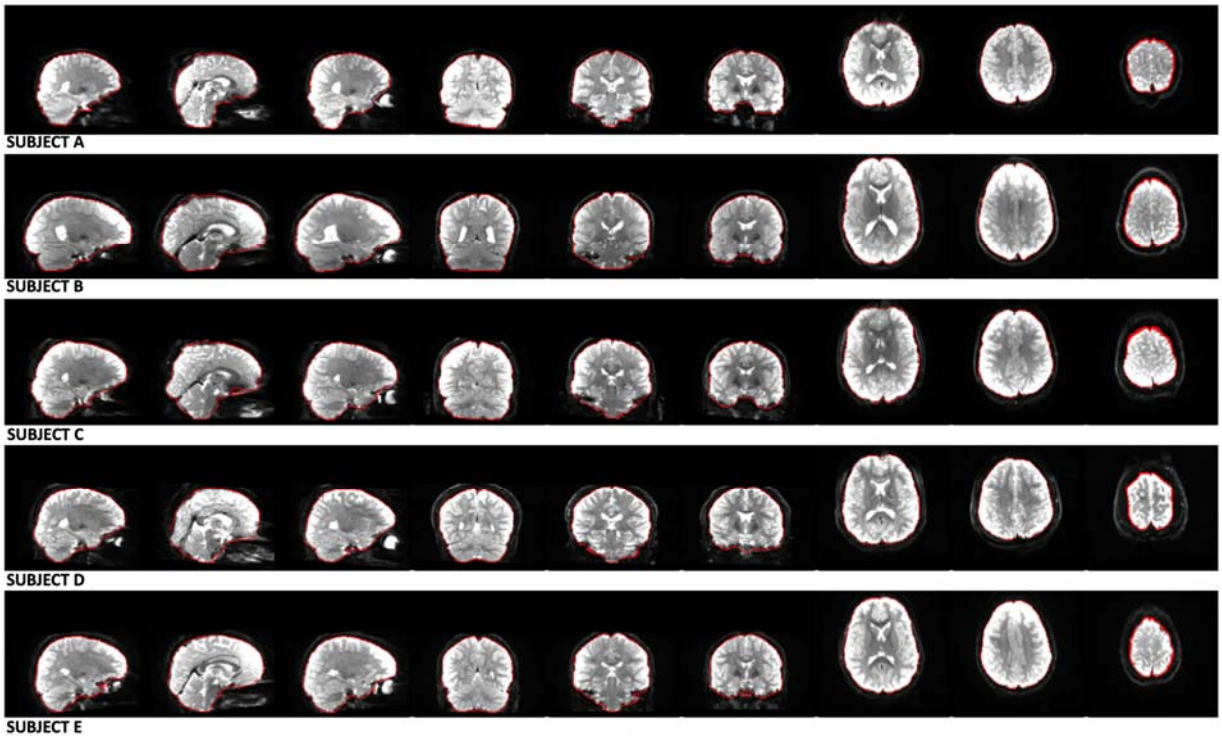


Figure A.1. Our deep learning CNN method (https://github.com/pnlbwh/CNN-Diffusion-MRIBrain-Segmentation) was used to perform dMRI Brain Segmentation on the ABCD study's dMRI data. The results of the segmentation are denoted by the red outlines, which successfully delineate the brain of five randomly selected subjects (Subject A, B, C, D, E) from the ABCD study. We employed fsl's slicesdir (Jenkinson et al. 2012) for visualization purposes, and each row of the visualization represents a different subject's brain, while each column displays the various brain slices of that subject.

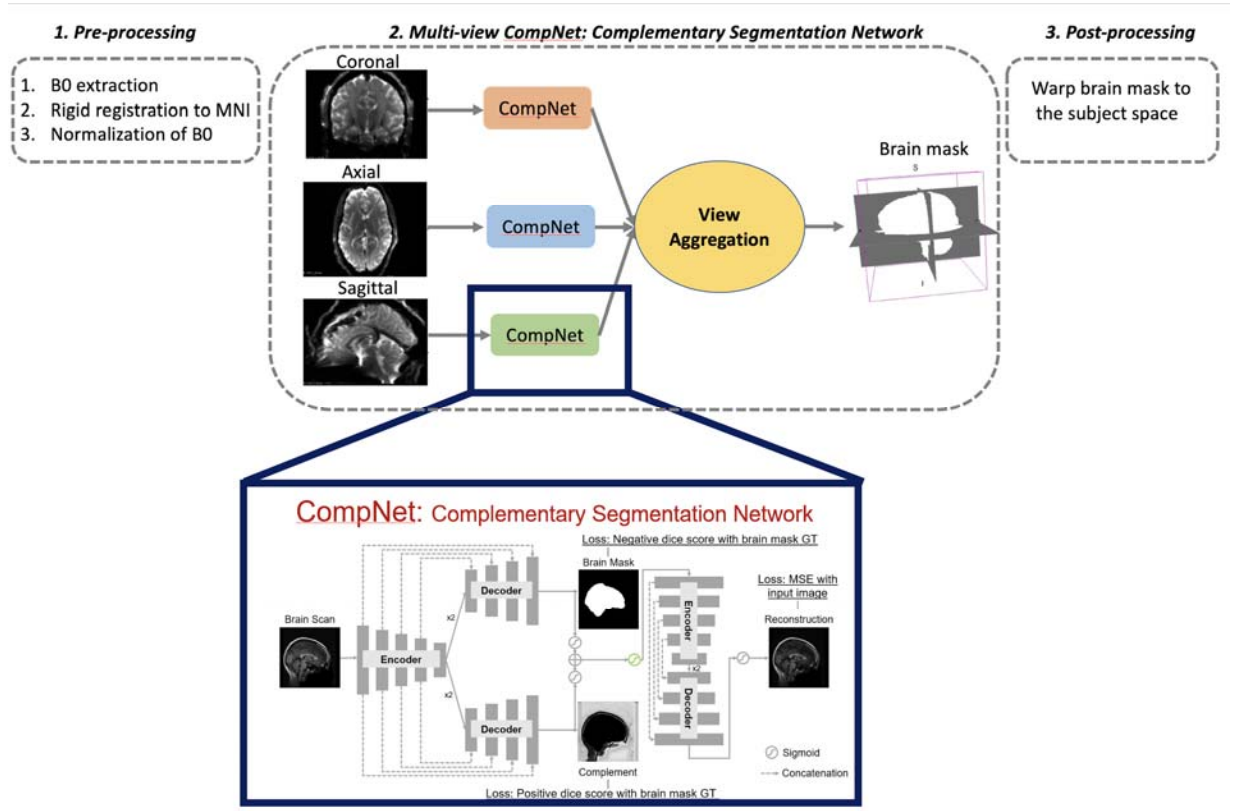


Figure A.2. - CNN DMRI Brain Segmentation: *(1) Pre-processing; (2) Multi-view CompNet; (3) Post-processing.*

Multi-view CompNet includes three branches: (i) Segmentation Branch - learns the brain region; (ii) Complementary Branch - learns complement of the brain region; (iii) Reconstruction Branch - provides direct feedback to the segmentation and complementary branches and expects reasonable reconstructions. Segmentation and reconstruction branches include a series of encoder and decoder networks with a kernel of size 3 × 3. The number of convolutional filters in the encoder starts from 32, followed by 64, 128, 256, and 512, while the number in the decoder starts from 256, followed by 128, 64, and 32.