

Highly replicable multisite patterns of adolescent white matter maturation

Highlights

- ABCC releases over 24,000 processed diffusion MRI datasets from the ABCD study
- Advanced diffusion metrics are more sensitive to development than tensor measures
- Longitudinal nonlinear harmonization improves cross-scanner replicability
- The dMRI contrast quality metric captures image noise without masking age effects

Authors

Steven L. Meisler, Matthew Cieslak, Joëlle Bagautdinova, ..., Russell T. Shinohara, Damien A. Fair, Theodore D. Satterthwaite

Correspondence

steven.meisler@pennmedicine.upenn.edu (S.L.M.), theodore.satterthwaite@pennmedicine.upenn.edu (T.D.S.)

In brief

Meisler et al. present a large public resource of fully processed diffusion MRI from the ABCD study and show that advanced diffusion measures and longitudinal harmonization improve the replicability and sensitivity of developmental analyses. The resource and accompanying workflows lower barriers to rigorous studies of adolescent white matter development.

NeuroResource

Highly replicable multisite patterns of adolescent white matter maturation

Steven L. Meisler,^{1,2,3,*} Matthew Cieslak,^{1,2,3} Joëlle Bagautdinova,^{1,2,3} Timothy J. Hendrickson,^{4,5} Tanya Pandhi,^{4,5} Andrew A. Chen,⁶ Noah Hillman,^{7,8} Hamsanandini Radhakrishnan,³ Taylor Salo,^{1,2,3} Eric Feczko,^{4,5} Kimberly B. Weldon,⁴ rae McCollum,^{4,5} Begim Fayzullobekova,^{4,5} Lucille A. Moore,⁴ Lucinda M. Sisk,^{1,2,3} Christos Davatzikos,^{9,10,11} Hao Huang,^{11,12} Bárbara Avelar-Pereira,¹³ Sendy Caffarra,^{14,15,16} Kelly Chang,¹⁷ Philip A. Cook,¹¹

(Author list continued on next page)

¹The Penn Lifespan Informatics and Neuroimaging Center, University of Pennsylvania, Philadelphia, PA 19104, USA

²Lifespan Brain Institute of Penn Medicine and CHOP, University of Pennsylvania, Philadelphia, PA 19104, USA

³Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA

⁴Masonic Institute for the Developing Brain, University of Minnesota Medical School, Minneapolis, MN 55414, USA

⁵Department of Pediatrics, University of Minnesota, Minneapolis, MN 55414, USA

⁶Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC 29425, USA

⁷Penn Statistics in Imaging and Visualization Endeavor, University of Pennsylvania, Philadelphia, PA 19104, USA

⁸Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA 19104, USA

⁹Center for AI and Data Science for Integrated Diagnostics (AI2D), Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

¹⁰Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, PA 19104, USA

¹¹Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA

¹²Department of Radiology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

¹³Aging Research Center, Karolinska Institutet and Stockholm University, Stockholm 113 30, Sweden

¹⁴Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena 41125, Italy

¹⁵Graduate School of Education, Stanford University, Stanford, CA 94305, USA

¹⁶Division of Developmental-Behavioral Pediatrics, Stanford University, Stanford, CA 94304, USA

¹⁷Department of Psychology, University of Washington, Seattle, WA 98195, USA

(Affiliations continued on next page)

SUMMARY

Adolescent brain development is difficult to study at scale because diffusion MRI (dMRI) measures are sensitive to technical variability across scanning environments. The Adolescent Brain Cognitive Development (ABCD) study is the largest US-based longitudinal neuroimaging study of adolescent brain maturation. We present ABCD-Brain Imaging Data Structure (BIDS) Community Collection (ABCC) release 3.1.0, a curated resource of more than 24,000 fully processed human ABCD dMRI datasets from participants aged 8–17 years. ABCC provides processed images, nuanced image quality metrics, advanced microstructural measures, and person-specific white matter bundle tractography. Across 62 bundles, measures of diffusion restriction and non-Gaussianity—including intracellular volume fraction from neurite orientation dispersion and density imaging (NODDI), mean kurtosis tensor from diffusion kurtosis imaging (DKI), and return-to-origin probability from mean apparent propagator (MAP)-MRI—were more sensitive to development and more robust to variation in image quality than tensor-derived measures. Longitudinal harmonization of microstructural features with ComBat-GAMM markedly improved the cross-vendor generalizability of developmental effects. Together, ABCC accelerates reproducible, rigorous research on adolescent white matter development.

INTRODUCTION

Adolescence is a critical period of brain plasticity that supports cognitive development.^{1–3} However, it is also understood as a

period of vulnerability to psychopathology.^{4,5} As such, mapping the brain throughout adolescence provides insight into both typical and atypical trajectories of brain maturation. The Adolescent Brain Cognitive Development (ABCD) study⁶ is the largest

Elizabeth A. Flook,^{1,2,3} Teresa Gomez,¹⁷ Mareike Grotheer,^{18,19,20} McKenzie P. Hagen,¹⁷ Zeeshan M. Huque,²¹ Iliana I. Karipidis,^{22,23} Arielle S. Keller,^{24,25} John Kruper,¹⁷ Audrey C. Luo,^{1,2,3} Briana Macedo,^{1,2,3} Kahini Mehta,²⁶ Jamie L. Mitchell,^{16,27} Adam R. Pines,²⁸ Laura Pritschet,^{1,2,3} Amelie Rauland,^{29,30} Ethan Roy,¹⁵ Brooke L. Sevchik,^{1,2,3} Golia Shafiei,^{1,2,3} S. Parker Singleton,^{1,3} Hannah L. Stone,³¹ Kevin Y. Sun,^{1,2,3} Valerie J. Sydnor,³² Tien T. Tong,^{1,2,3} Maya Yablonski,^{15,16} Jason D. Yeatman,^{15,16,27} Ariel Rokem,^{17,33} Russell T. Shinohara,^{7,8,9} Damien A. Fair,^{4,5,34} and Theodore D. Satterthwaite^{1,2,3,35,*}

¹⁸Department of Psychology, Philipps-Universität Marburg, Marburg 35032, Germany

¹⁹Center for Mind, Brain and Behavior, Philipps-Universität Marburg, Justus-Liebig Universität Giessen, Technische Universität Darmstadt, Marburg 35032, Germany

²⁰Lise Meitner Research Group Neuroplasticity in Development and Learning, Max Planck Institute for Human Development, Berlin 14195, Germany

²¹Department of Psychology and Neuroscience, Temple University, Philadelphia, PA 19122, USA

²²Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital of Psychiatry Zurich, Zurich 8032, Switzerland

²³Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich 8057, Switzerland

²⁴Department of Psychological Sciences, University of Connecticut, Storrs, CT 06269, USA

²⁵Institute for the Brain and Cognitive Sciences, University of Connecticut, Storrs, CT 06269, USA

²⁶Department of Neuroscience, Columbia University, New York, NY 10032, USA

²⁷Department of Psychology, Stanford University, Stanford, CA 94305, USA

²⁸Department of Psychiatry and Behavioral Sciences, Stanford School of Medicine, Stanford, CA 94305, USA

²⁹Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Center Jülich, Jülich 52438, Germany

³⁰Department of Psychiatry, Psychotherapy and Psychosomatics, Faculty of Medicine, RWTH Aachen University, Aachen 52074, Germany

³¹Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

³²Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15213, USA

³³eScience Institute, University of Washington, Seattle, WA 98195, USA

³⁴Institute of Child Development, University of Minnesota, Minneapolis, MN 55455, USA

³⁵Lead contact

*Correspondence: steven.meisler@pennmedicine.upenn.edu (S.L.M.), theodore.satterthwaite@pennmedicine.upenn.edu (T.D.S.)
<https://doi.org/10.1016/j.neuron.2026.09.021>

longitudinal initiative tracking neuroimaging, behavior, cognition, and psychopathology across childhood and adolescence in the United States. Among the rich imaging data available, diffusion MRI (dMRI) is critical for characterizing the development of white matter microstructure and macrostructure.^{7–9} Prior studies using ABCD dMRI data have linked white matter features to cognitive function,^{10,11} the childhood environment,^{12,13} and early markers of psychopathology.^{14–19} Since ABCD first launched in 2015, the broader field has increasingly embraced advanced dMRI pre-processing methods,²⁰ more detailed quality control procedures,²¹ approaches to modeling dMRI beyond the diffusion tensor, and person-specific bundle tractography. However, many of these recent methodological innovations and the extensive outputs they produce—which may offer heightened sensitivity to white matter development^{22–24}—have not been generated and publicly shared as part of the official ABCD release. Applying state-of-the-art methods across large-scale datasets such as ABCD requires extensive computational resources and technical expertise, posing a significant barrier for many research groups. To accelerate rigorous and reproducible research in white matter maturation, we present the ABCD-BIDS Community Collection (ABCC)²⁵ 3.1.0 data release, which includes over 24,000 ABCD dMRI datasets with derived, analysis-ready, processed outputs (henceforth referred to as “derivatives”).

This new data resource addresses four major obstacles. First, scanner-related variability may obscure generalizable developmental effects in large multisite datasets such as ABCD. Although statistical harmonization reduces mean-level scanner differences in dMRI derivatives, most prior work has focused on

univariate intersite differences between regional measures.^{26,27} Critically, it remains unknown whether harmonization aligns or distorts biologically meaningful spatial patterns of adolescent white matter maturation across scanning environments. White matter development follows reproducible large-scale patterns across the brain,^{28–30} providing a biologically grounded benchmark for evaluating harmonization performance and microstructural model choices. However, scanner-specific factors, such as bias fields, may exhibit spatial structure that distort developmental patterns, limiting the generalizability of findings across acquisition sites. Here, we apply state-of-the-art longitudinal nonlinear harmonization and test whether reducing scanner-related variance enhances the generalizability of brain-wide developmental effects across scanning environments.

A second major challenge is that multi-shell dMRI data can be analyzed using diverse postprocessing models that yield myriad features that characterize the diffusion process. However, the relative developmental sensitivity and generalizability of these measures remain unclear. Faced with such a profusion of models and the challenges in implementing each approach, investigators typically evaluate a single or small set of measures. Historically, studies have predominantly relied on simple tensor-based measures such as fractional anisotropy (FA) and mean diffusivity (MD).³⁰ More recent studies have shown that metrics from models such as neurite orientation dispersion and density imaging (NODDI),³¹ mean apparent propagator MRI (MAP-MRI),³² and diffusion kurtosis imaging (DKI)³³ may be more sensitive to developmental processes than traditional tensor-derived measures.^{22–24} However, these comparisons have typically been made across separate studies and cohorts,

often evaluating different subsets of model features, which limits direct comparison across modeling approaches. Moreover, it remains unclear whether developmental patterns generalize across these different dMRI features or replicate consistently across scanner vendors. Here, we leverage multiple dMRI models to systematically compare the developmental sensitivity and generalizability of diverse microstructural measures in the ABCD cohort.

Third, the impact of image quality on dMRI microstructural measures remains poorly characterized. These considerations are especially consequential in developmental studies such as ABCD, where dMRI image quality can vary systematically with age^{34,35} and psychopathology.³⁶ In longitudinal multisite consortia, age, scanning environment, and image quality may also be related. For example, earlier time points for a given participant—when they are younger—may be disproportionately acquired on older scanner software versions. Such alignment can complicate statistical modeling by coupling biological and acquisition-related sources of variation. Although expert manual review is often considered the gold standard for neuroimaging quality control, manually reviewing all dMRI images at the scale of ABCD is usually impractical. In recent years, numerous automated measures of dMRI image quality have been proposed, capturing features such as head motion, contrast-to-noise ratio (CNR), *q*-space consistency, and slice-wise artifacts. However, it remains unclear which of these manual or automated quality metrics capture the most microstructural variation in the context of complex multisite longitudinal studies. We manually rated more than 3,000 ABCD dMRI images and systematically evaluated associations between dozens of measures of image quality and white matter microstructure.

Finally, we aim to lower barriers to conducting rigorous and reproducible neuroimaging research. To this end, ABCC provides fully processed, analysis-ready derivatives as both high-dimensional imaging data and tidy tabular summaries. The processing workflows used to generate these outputs—QSIprep²⁰ and QSIRecon³⁷—are all open source, transparent, BIDS compliant,^{38–40} and aligned with the NMIND criteria for rigor and usability.⁴¹ By distributing standardized derivatives generated with widely adopted community tools, ABCC reduces the need for redundant reprocessing and minimizes variability introduced by custom in-house pipelines. As such, we anticipate that the dMRI data of ABCC will accelerate robust and reproducible studies of adolescent white matter development.

RESULTS

We applied state-of-the-art open-source dMRI processing pipelines, QSIprep and QSIRecon, to longitudinal data from the ABCD study (Figure 1; see STAR Methods). Processing yielded 40 automated image quality metrics (IQMs; Table S1), as well as many macro- and microstructural features (Table S2) summarized across 67 person-specific white matter bundles (Figure S1). After quality control, 24,178 dMRI datasets across 21 sites from 10,738 participants aged 8–17 years were retained for analysis (Table 1; Table S3). In addition, 33 trained reviewers manually rated 3,101 dMRI images, and a multivariate classifier was used to extend these ratings across the full cohort. These

derivatives are distributed as part of ABCC release 3.1.0, with future releases incorporating additional ABCD datasets as they are acquired and processed.

We first characterized variation in dMRI image quality across scanner vendors (Siemens, GE, and Philips) and software versions. To account for technical differences across acquisition environments, we then applied advanced longitudinal nonlinear harmonization with longitudinal ComBat-GAMM.⁴³ Using these harmonized data, we evaluated the developmental sensitivity of diffusion microstructural measures and their relationships to image quality. Throughout these analyses, we employed generalized additive mixed models (GAMMs) to account for the longitudinal data structure and flexibly model nonlinear developmental effects while minimizing overfitting. For clarity, the main results focus on 62 consistently resolved white matter bundles and five commonly used microstructural measures. Specifically, we analyzed FA and MD from the DSI Studio tensor-fitting metric. We additionally evaluated the intracellular volume fraction (ICVF; from NODDI), return-to-origin probability (RTOP; from MAP-MRI), and mean kurtosis tensor (MKT; from DKI). These metrics span the range of diffusion modeling approaches implemented in QSIRecon. Results for the remaining microstructural measures are reported in the [supplemental information](#).

Automated measures of image quality vary by scanner vendor and software version

Scanner heterogeneity represents a major challenge for multisite neuroimaging consortia. Even with carefully harmonized acquisition protocols, images acquired across scanner vendors (Siemens, GE, and Philips) and software versions can differ systematically in quality. QSIprep produced 40 automated IQMs (see STAR Methods; Table S1), including CNR within each diffusion-weighted shell and temporal signal-to-noise ratio (tSNR) for the *b* = 0 shell. It also generated two *q*-space consistency metrics: neighboring dMRI correlation (NDC), which quantifies the similarity of neighboring diffusion directions, and dMRI contrast, which quantifies the expected difference between neighboring and orthogonal diffusion directions in *q*-space (Figure S2). To quantify scanner-driven variation and assess the effects of preprocessing, we characterized IQMs across vendors and software versions in both raw and preprocessed data.

Substantial vendor-related variation was observed across IQMs in both raw and preprocessed data (all Kruskal-Wallis tests $p < 0.001$). Vendor effects were large for CNR across shells (Figure 2A; $\eta^2 = 0.59$ – 0.67), and for *b* = 0 tSNR (Figure 2B; $\eta^2 = 0.50$). Vendor differences in NDC (Figure 2C; raw $\eta^2 = 0.39$; preprocessed $\eta^2 = 0.68$) and dMRI contrast (Figure 2D; raw $\eta^2 = 0.17$; preprocessed $\eta^2 = 0.15$) also persisted after preprocessing. Pairwise comparisons revealed systematic patterns: Siemens scanners exhibited higher CNR, tSNR, and NDC than both GE and Philips (Wilcoxon rank-sum tests; all rank-biserial $|r| \geq 0.73$, $p < 0.001$), indicative of higher image quality. Philips scanners showed markedly lower dMRI contrast than Siemens and GE ($|r| = 0.73$ – 0.80 , $p < 0.001$). Differences between Siemens and GE were comparatively smaller for both raw and preprocessed dMRI contrast (both $|r| \leq 0.11$, $p < 0.001$). As expected, preprocessing with QSIprep produced consistent within-scan improvements in IQMs calculated on both raw and

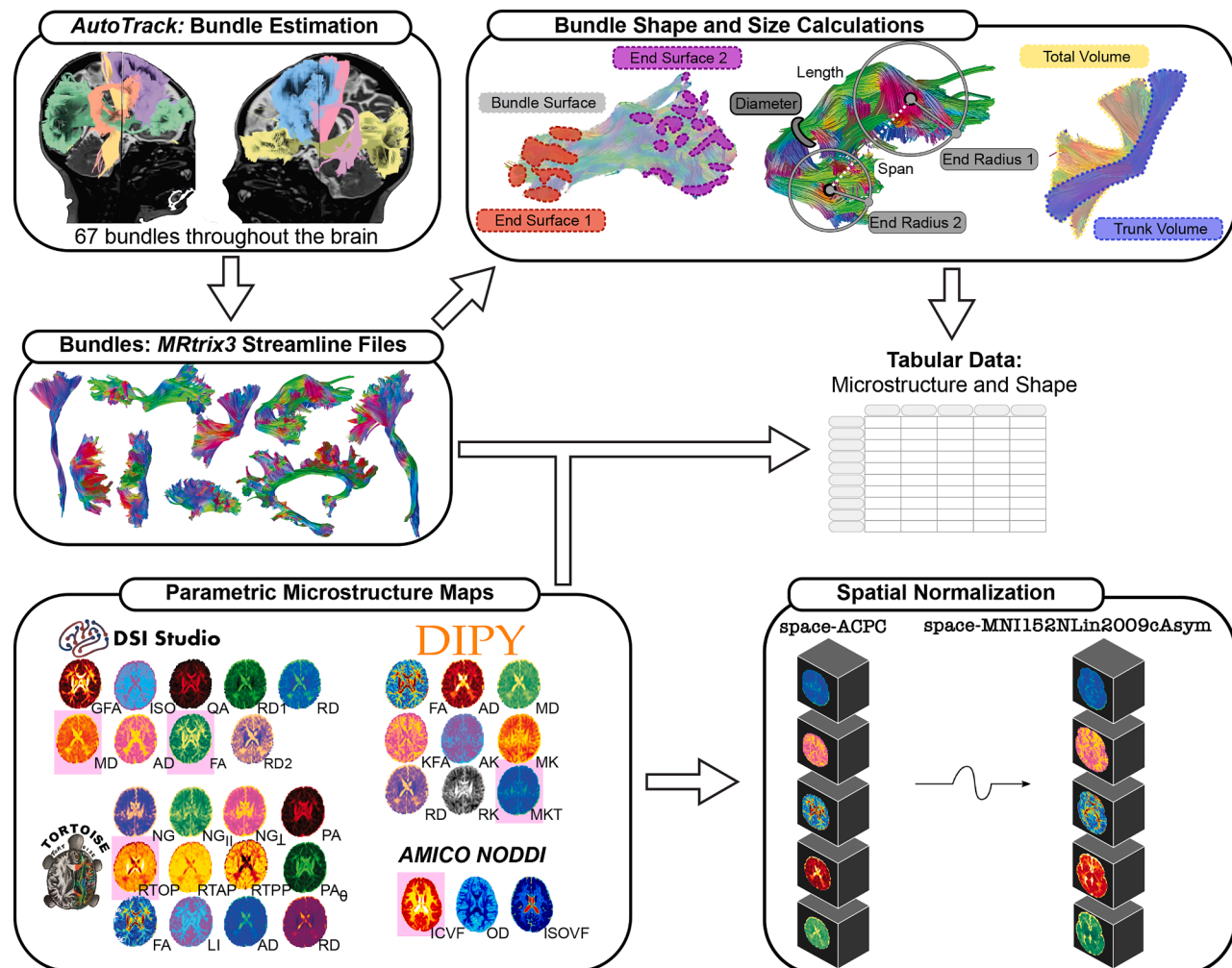


Figure 1. Postprocessing outputs from QSIRecon

QSIRecon was used to generate parametric microstructural maps, perform participant-specific bundle tractography, and compute bundle-wise summary statistics from ABCD dMRI data. Diffusion models implemented in DSI Studio, DIPY, TORTOISE, and AMICO were applied to estimate 33 microstructural maps; the five metrics highlighted in pink are those emphasized in the primary analyses. Outputs include streamline files for 67 participant-specific white matter bundles and three-dimensional NIFTI images of each microstructural map in both participant-specific and MNI152NLin2009cAsym space. Tabulated summaries of bundle-wise microstructural and macrostructural measurements are also provided. Figure adapted with permission from Cieslak et al.⁴²

preprocessed data, such as NDC and dMRI contrast (Figures 2C and 2D; paired Wilcoxon $|r| > 0.86$, $p < 0.001$).

Scanner software versions also exhibited systematic variation. Software versions tended to vary systematically with participant age (Figure 2E), reflecting scanner upgrades during the course of the study. Notably, within-vendor variation in image quality was evident across software versions (Figure 2F). Within GE, for example, preprocessed NDC differed strongly across software versions (Kruskal-Wallis test; $\eta^2 = 0.50$, $p < 0.001$), largely driven by reduced NDC in the earliest software version, DV25 (Wilcoxon rank-sum tests, all pairwise $|r| > 0.9$, $p < 0.001$). In contrast, preprocessed dMRI contrast showed substantially smaller variation across GE software versions ($\eta^2 = 2.82 \times 10^{-3}$, $p < 0.001$). Additional IQMs and batch-stratified distributions are shown in Figures S3–S10. Together, these findings demon-

strate that preprocessing substantially improves image quality, yet substantial scanner-related variation persists, underscoring acquisition heterogeneity as a fundamental challenge for large multisite dMRI studies.

Harmonization reduces mean acquisition batch effects in white matter microstructure

Although harmonization has previously been applied to ABCD dMRI data, prior efforts have largely focused on cross-sectional baseline scans.²⁷ Here, we implemented longitudinal ComBat-GAMM,⁴³ a longitudinal nonlinear harmonization approach that preserves participant-specific effects and nonlinear developmental trajectories. To quantify acquisition-related variability, we modeled each bundle-wise microstructural metric as a function of age, sex, and acquisition batch (unique scanner

Table 1. Summary of data

Session	Scanner vendor (n)			Demographics		
	Siemens	GE	Philips	Age (mean [SD])	Sex (M/F)	Total
Baseline	5,414	2,174	814	9.97 [0.63]	4,364/4,038	8,402
Year 2	4,335	1,736	637	11.99 [0.65]	3,604/3,104	6,708
Year 4	3,685	1,446	516	14.16 [0.72]	3,024/2,623	5,647
Year 6	2,067	1,043	311	16.08 [0.66]	1,796/1,625	3,421
Total	15,501	6,399	2,278	12.37 [2.28]	12,788/11,390	24,178

Counts of dMRI scans by ABCD study session and scanner vendor following application of inclusion and exclusion criteria. Demographic summaries include the mean age ($\pm$ SD) of participants at each session and the number of male (M) and female (F) participants.

device-software version combinations). Batch effect size was defined as the change in adjusted R^2 (ΔR^2_{adj}) between the full model and a reduced model excluding the batch predictor. Analyses were performed separately for unharmonized and harmonized data.

In the unharmonized data, acquisition batch explained up to 69.8% of variance in the data (Figure 3A). MKT was the most susceptible to batch effects (Figures 3B and 3C), with batch explaining 37.2% of variance on average across bundles, followed by FA (26.6%), ICVF (20.1%), MD (18.4%), and RTOP (8.2%). Cerebellar and subcortical fibers tended to be most susceptible to batch effects. Following harmonization, batch effects were effectively eliminated for all measures (Figures 3B and 3C). Batch effects for all microstructural metrics are provided in Figure S11. These results demonstrate that longitudinal harmonization can be applied at scale and substantially attenuates scanner-related differences across acquisition environments.

Advanced measures of white matter microstructure are more sensitive to development

Multi-shell dMRI enables the estimation of numerous microstructural indices beyond the diffusion tensor, yet most developmental studies have focused on a single measure or a limited subset of measures. To systematically compare their developmental sensitivity, we quantified associations between age and harmonized microstructural measures across white matter bundles. Each bundle-metric combination was modeled as a function of age and sex, and the developmental effect size was defined as the ΔR^2_{adj} between the full model and a reduced model excluding age (Figure 4A).

Across microstructural metrics, the strongest developmental effect sizes were observed for ICVF (Figure 4B), followed by MKT and RTOP (Figure 4C). ICVF explained 6%–52% of variance across bundles (mean ΔR^2_{adj} across bundles = 34.5%), MKT explained 5%–49% (mean = 32.6%), and RTOP explained 1%–49% (mean = 28.7%). In contrast, tensor-derived measures showed substantially weaker developmental sensitivity: MD explained 0%–27% of variance (mean = 12.5%), and FA explained 0%–17% (mean = 7.4%). An analogous voxel-wise analysis was performed, which converged on similar results (Figure S12). Developmental sensitivity also varied across bundle categories, with association bundles exhibiting the strongest age effect sizes, and cerebellar bundles exhibiting the weakest. Developmental effect sizes for all microstructural metrics are provided in Figure S13. Together, these results

emphasize that advanced multi-compartment and non-Gaussian metrics are more sensitive to development than conventional tensor-derived measures.

Harmonization improves the generalizability of whole-brain developmental effect sizes

Scanner effects are substantial and not uniform across the brain (Figure 3). Thus, technical variation in multisite neuroimaging may distort observed patterns of development across the brain. We therefore assessed whether longitudinal harmonization improved cross-vendor generalizability of developmental effect sizes. Within each scanner vendor, we estimated developmental effect sizes (ΔR^2_{adj}) for each bundle and microstructural metric using unharmonized and harmonized data separately. Cross-vendor correspondence was defined as the Spearman rank correlation (ρ) between vendors' vectors of bundle-wise age effect sizes.

Harmonization systematically increased cross-vendor correspondence for all metrics evaluated, often substantially (Figures 5A and S14). In harmonized data, all metrics showed mean harmonized correspondences greater than 0.91; MD showed the highest correspondence (mean harmonized ρ = 0.98 across vendor pairs; mean $\Delta\rho$ = 0.09). The most pronounced gains in generalizability following harmonization were observed for FA (mean ρ = 0.91; mean $\Delta\rho$ = 0.31) and MKT (mean ρ = 0.94; mean $\Delta\rho$ = 0.33). The largest single improvement occurred for MKT, where the GE-Philips correspondence increased from 0.47 to 0.95 ($\Delta\rho$ = 0.48; Figures 5B and 5C). These findings indicate that longitudinal nonlinear harmonization not only reduces mean vendor effects but also strengthens the reproducibility of spatial developmental patterns in large multisite studies.

Advanced diffusion models exhibit convergent developmental organization

Variation in the use of disparate microstructural models complicates synthesis across studies. Our results demonstrate that contemporary microstructural measures show greater developmental sensitivity than tensor-derived metrics (Figure 4C). However, it remains unclear whether different models of the diffusion signal capture a shared spatial pattern of development or whether tensor-derived measures simply reflect attenuated versions of a common developmental gradient.

We operationalized cross-metric correspondence as the Spearman rank correlation between bundle-wise age effects

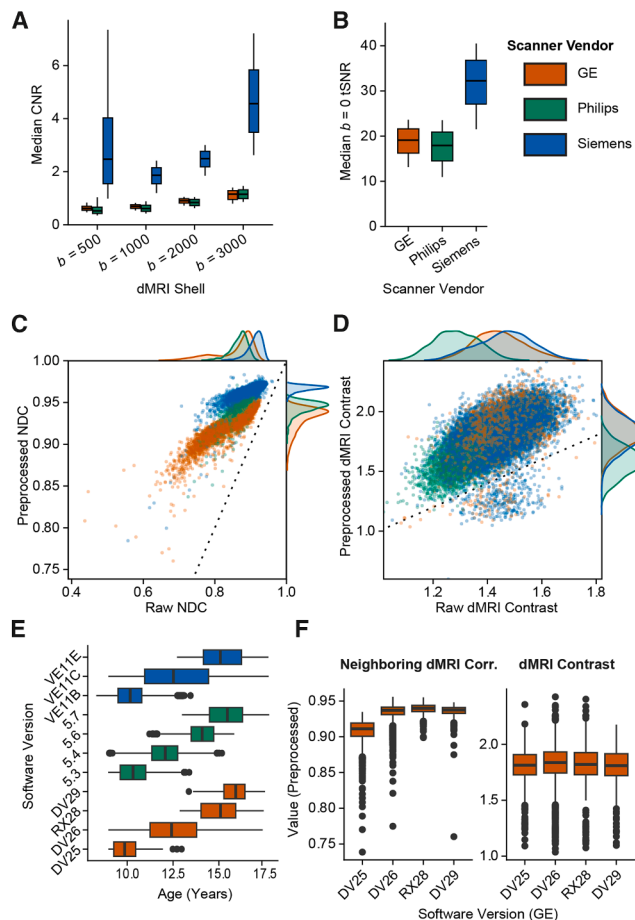


Figure 2. Automated IQMs are strongly structured by scanner vendor and software

(A) Shell-wise distributions of median CNR.
(B) Median temporal signal-to-noise ratio (tSNR) in the $b = 0$ shell.
(C and D) Neighboring dMRI correlation (NDC; left) and dMRI contrast (right) before and after preprocessing, including B_1 bias correction. The dotted line denotes the identity line; values above the line indicate improved quality following preprocessing.
(E) Distribution of ages included in each scanner software version. Boxes are colored by scanner vendor.
(F) Distribution of preprocessed NDC (left) and dMRI contrast (right) within GE software versions.

for each metric, estimated from the full harmonized dataset (Figure 5D; Figure S15). Advanced measures (ICVF, RTOP, and MKT) were highly correlated with one another ($\rho \geq 0.93$), indicating strong agreement in the pattern of developmental effects across bundles. In contrast, tensor-derived metrics were more variable; MD showed particularly weak correspondence with other measures (mean $\rho = 0.66$). The strongest association was observed between ICVF and MKT ($\rho = 0.96$; Figure 5E), whereas the weakest was between FA and MD ($\rho = 0.51$; Figure 5F). These results indicate that developmental patterns are more consistent across advanced diffusion models than across commonly used tensor-derived measures.

dMRI contrast captures the most microstructural variance among measures of data quality

dMRI data can be characterized using numerous automated IQMs as well as expert manual review. However, it remains unclear which measures meaningfully capture variation in data quality—or how sensitive different microstructural reconstruction methods are to variation in quality. QSIprep generated 40 automated IQMs indexing diverse constructs, including head motion, SNR, and q -space properties. Additionally, we trained 33 reviewers to rate a large subset of 3,101 images (see STAR Methods). From these ratings we used a multivariate classifier to predict a quality score across the entire cohort, which was analyzed alongside the other automated IQMs.

To quantify the influence of image quality, we evaluated associations between each IQM and harmonized microstructural measures. For each bundle and microstructural metric, we fit GAMMs including age, sex, harmonization batch, and each IQM separately. Conditioning on batch allowed quality-related variance to be estimated independently of batch-driven differences. Quality effect sizes were defined as the ΔR^2_{adj} between the full model and a reduced model excluding the IQM predictor.

Across the five representative microstructural metrics and 62 bundles, we identified the IQM with the largest effect size for each of the 310 bundle-metric combinations (Figure 6A; Figure S16). Preprocessed dMRI contrast (after B_1 bias correction) performed the best, accounting for the largest effect sizes in 128 combinations. In contrast, quality metrics such as mean framewise displacement and the quality classifier score rarely explained the greatest variance. Using preprocessed dMRI contrast as a representative IQM, FA was the most susceptible to quality variation (range of ΔR^2_{adj} across bundles: -0.1% to 13.8% ; mean $\Delta R^2_{\text{adj}} = 5.3\%$). In contrast, ICVF was the most robust (range: -0.1% to 2.7% ; mean $= 0.8\%$), being on average more than 6-fold less susceptible than FA (Figures 6B and 6C; Figure S16).

We further compared quality effect sizes of preprocessed dMRI contrast and mean manual rating within the subset of manually reviewed images. dMRI contrast and manual ratings were modestly related ($r = 0.27$, $p < 0.001$; Figure 6D). However, preprocessed dMRI contrast consistently showed larger quality effect sizes than manual ratings. Overall, dMRI contrast explained more microstructural variance than manual ratings in 266 of the 310 bundle-metric combinations (Figure 6E). Together, these findings indicate that dMRI contrast captures a substantial component of variability in data quality. Conversely, commonly used measures of quality such as head motion, NDC, and manual quality ratings showed markedly smaller effect sizes. Advanced microstructural metrics were more robust to variation in image quality than FA, further underscoring their suitability for studies of brain development.

Commonly used IQMs vary by acquisition batch and attenuate developmental effect sizes

In developmental neuroimaging studies, IQMs are frequently included as covariates to mitigate spurious associations. However, age, image quality, and acquisition batch may be collinear in large multisite consortia such as ABCD. We have shown this is particularly true within GE scanners in ABCD (Figures 2E and 2F).

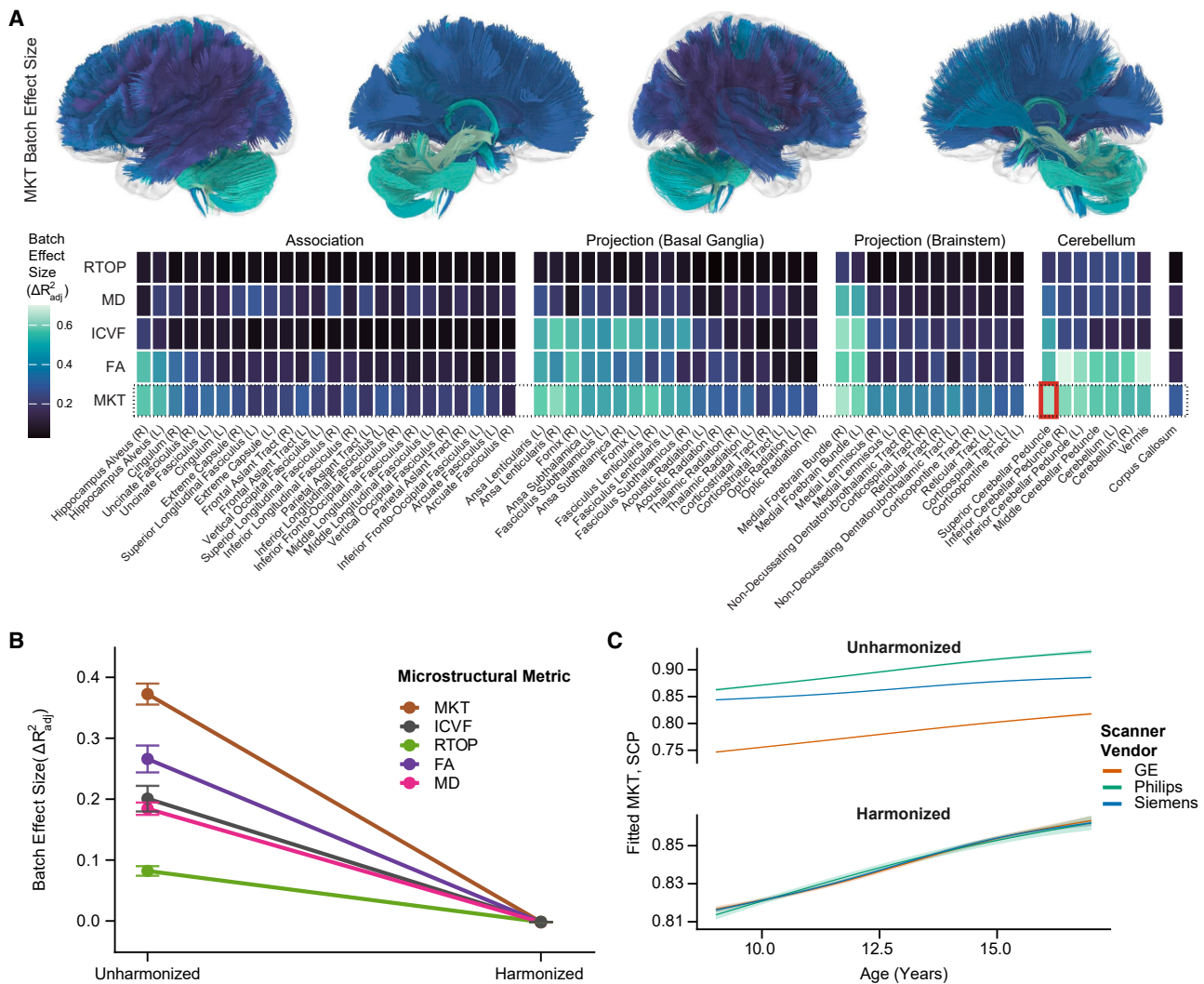


Figure 3. Harmonization reduces mean acquisition batch effects in white matter microstructure

(A) Tract plot and heatmap of batch effect sizes (ΔR^2_{adj}) in unharmonized data. Rows and columns are ordered by their mean effects, and columns are grouped by bundle category. R/L denotes right- and left-lateralized bundles. The red rectangle highlights MKT in the superior cerebellar peduncle (SCP). Tracts are colored by MKT batch effect size, as indicated by a dotted rectangle in the heatmap.

(B) Mean batch effects across bundles for each microstructural metric. Colors denote different metrics; error bars indicate the standard error of the mean.

(C) Example developmental trajectories of MKT in the SCP before and after harmonization. Lines represent GAM-fitted mean trajectories of MKT as a function of age for each vendor; shaded regions denote 95% confidence intervals of the fitted mean.

We therefore examined whether variation in dMRI IQMs reflected participant characteristics or instead primarily captured acquisition-related differences.

Several IQMs, including mean framewise displacement ($R^2 = 11.1\%$), the quality classifier score trained on manual ratings ($R^2 = 13.5\%$), and preprocessed NDC ($R^2 = 28.0\%$), showed apparent associations with age (Figures 7A and 7B). However, variance decomposition analyses revealed that these age-IQM relationships were predominantly explained by between-batch differences rather than within-batch participant variation ($\geq 82.2\%$ of the age-IQM relationship explained by batch).

We next evaluated how including these IQMs as model covariates influenced estimates of developmental effect sizes in white

matter microstructure. Bundle-wise FA age effect sizes were estimated in the harmonized data with and without each IQM included as a covariate. For all IQMs besides dMRI contrast, adding the metric systematically reduced age effect sizes across bundles (Figure 7C), consistent with attenuation of developmental effect sizes due to the collinearity between age and acquisition batch. The magnitude of this attenuation varied across bundles, resulting in a loss of correspondence with developmental effect sizes when image quality was not included as a covariate ($\rho = 0.78\text{--}0.89$).

Notably, preprocessed dMRI contrast showed minimal association with age or batch (age $R^2 = 0.6\%$, and $<0.01\%$ of the age-IQM relationships explained by batch). As such,

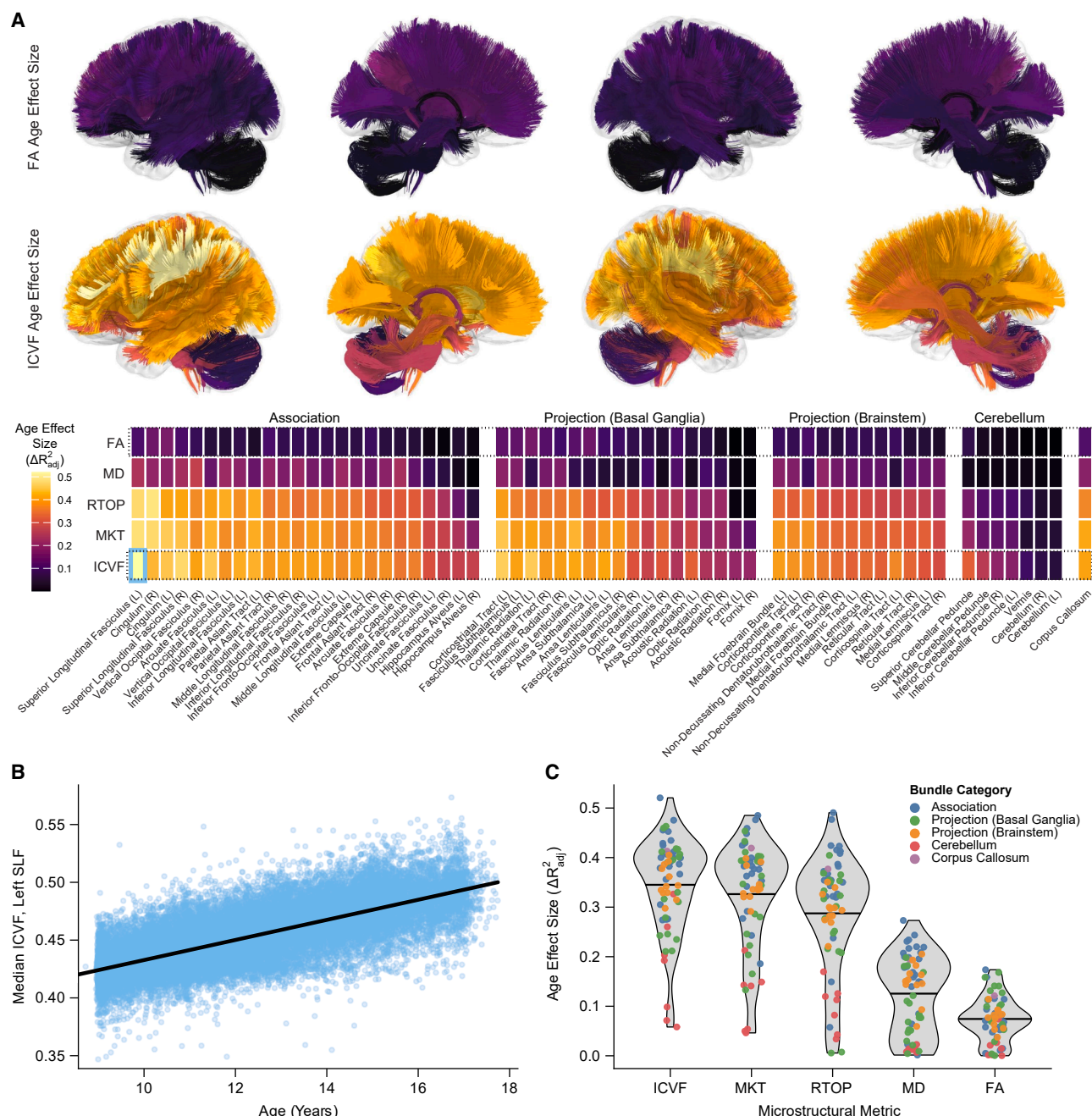


Figure 4. Advanced microstructural metrics are more sensitive to development

(A) Tract plot and heatmap of bundle-wise developmental effect sizes (ΔR^2_{adj}). Rows and columns are ordered by mean effect size, and columns are grouped by bundle category. R/L denotes right- and left-lateralized bundles. The blue rectangle highlights the largest observed effect size (ICVF in the left superior longitudinal fasciculus [SLF]). Tracts are colored by the ICVF age effect size, as indicated by a dotted rectangle in the heatmap.

(B) Age-related trajectory of ICVF in the left SLF from a GAM fit. The solid black curve shows the model expected ICVF value as a smooth function of age. The 95% confidence interval for the fitted mean trend is present but too narrow to be visible at this scale.

(C) Distribution of bundle-wise age effect sizes across microstructural metrics, ordered by decreasing mean effect size across bundles. Each point represents a bundle and is colored by bundle category. Black lines indicate the mean effect size across bundles for each metric.

developmental effect sizes were nearly identical with or without a dMRI contrast covariate ($p = 0.99$). Together, these findings indicate that, in ABCD, many commonly used diffu-

sion IQMs may largely reflect scanner differences rather than participant-specific acquisition factors, such as head motion. In some settings, including such IQMs as model

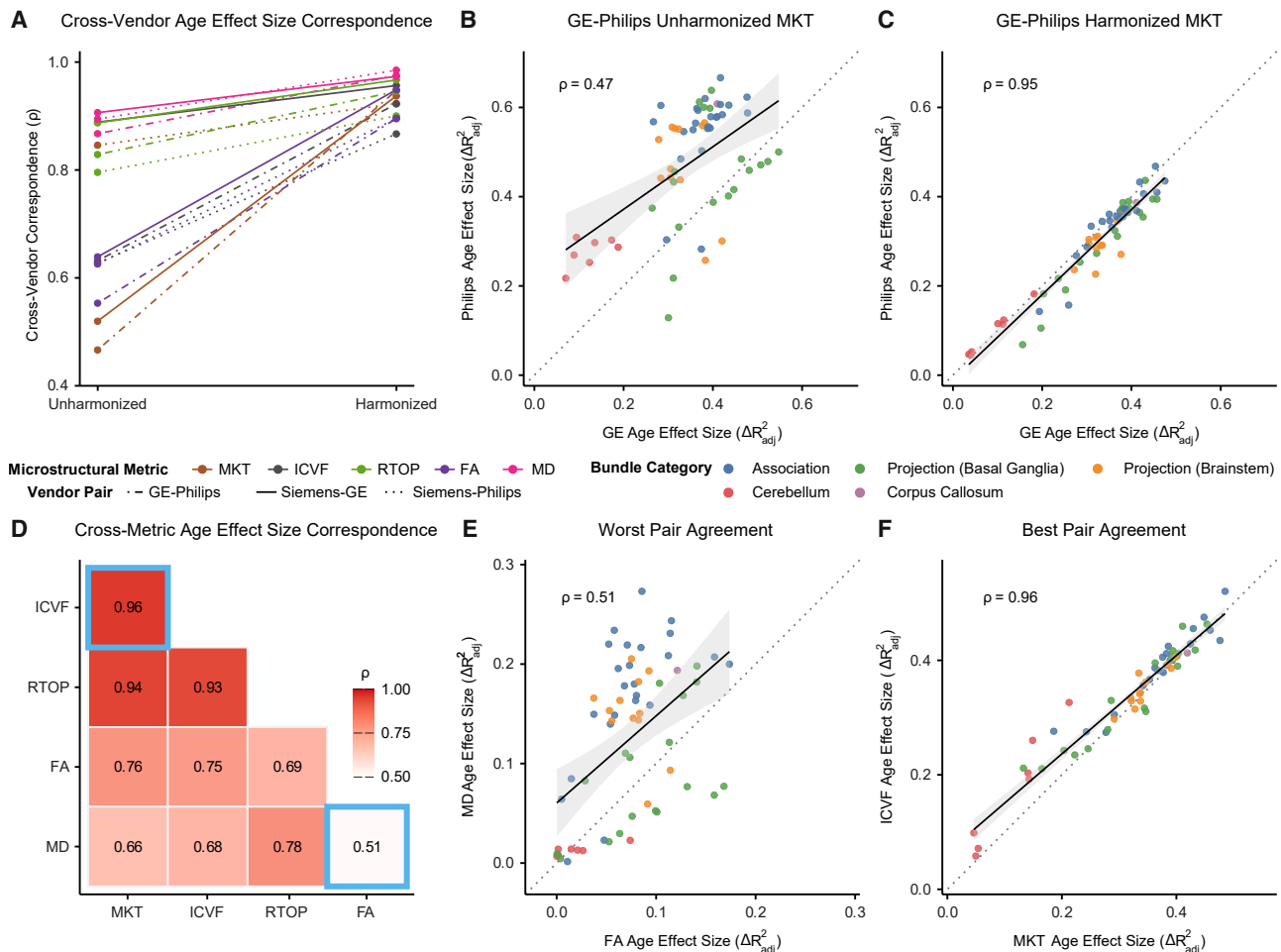


Figure 5. Developmental effect sizes generalize across vendors and microstructural measures after harmonization

(A) Cross-vendor correspondence (Spearman ρ) of bundle-wise age effect sizes in unharmonized and harmonized data. Colors denote microstructural metrics; line types indicate vendor pairs.

(B and C) Illustrative plots of the largest gain in cross-vendor correspondence. Unharmonized (left) and harmonized (right) GE-Philips correspondence for MKT. Each point represents a bundle colored by bundle category. Solid lines show linear fits with 95% confidence intervals; dotted lines denote identity.

(D) Cross-metric correspondence of developmental effect sizes estimated from harmonized data. The strongest and weakest associations are outlined in blue.

(E and F) Scatterplots illustrating the weakest (left) and strongest (right) cross-metric correspondences. Points represent bundles colored by category; solid lines show linear fits with 95% confidence intervals; dotted lines denote identity.

covariates can inadvertently distort biologically meaningful developmental signals.

DISCUSSION

Large-scale neuroimaging consortia such as ABCD provide unprecedented opportunities to characterize white matter maturation across adolescence. However, this scale also introduces important analytical challenges, including uncertainty in metric selection, scanner heterogeneity, and variation in image quality. Here, we addressed these challenges with three primary results. First, metrics from advanced diffusion models were substantially more sensitive to developmental processes than traditional tensor-derived metrics, converging on highly similar spatial patterns of maturation. Advanced measures were also more robust

to image quality variation than typically used measures such as FA. Second, scanner-related variance was substantial and spatially structured; longitudinal harmonization not only reduced mean differences across acquisition environments but also improved the cross-vendor generalizability of developmental patterns. Finally, across quality metrics, dMRI contrast captured the most variance in diffusion microstructure. Other commonly used quality metrics, such as head motion, contributed little explanatory value and in fact distorted estimates of developmental effect sizes when included as model covariates. Collectively, these findings demonstrate that practical methodological choices—including the selection of microstructural metrics, harmonization strategies, and image quality covariates—can markedly impact the sensitivity and generalizability of large-scale, multisite dMRI studies. To facilitate rigorous and reproducible developmental

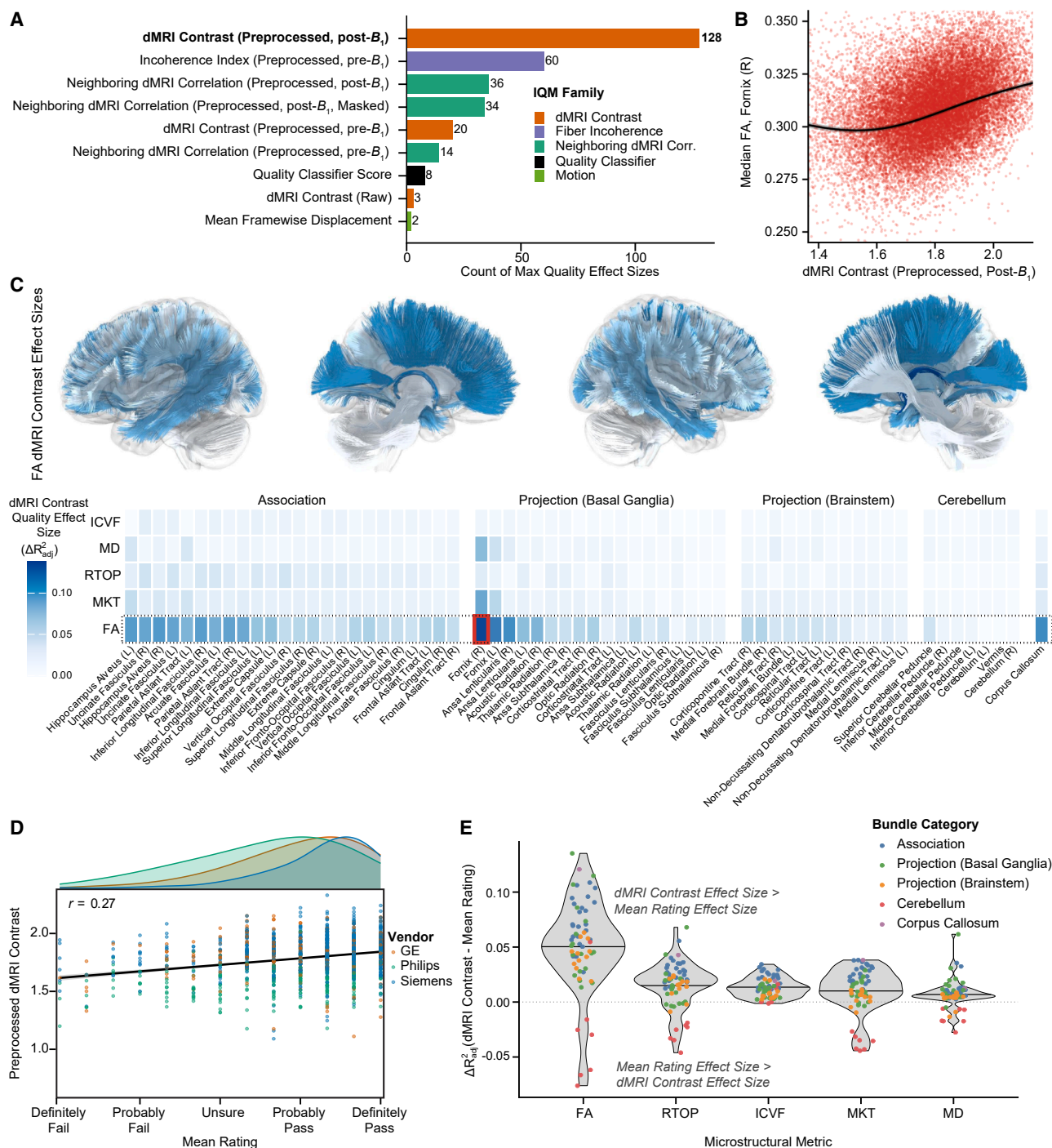


Figure 6. The dMRI contrast IQM captured the most variance in dMRI microstructure

(A) IQM winner-take-all plot. Each bar represents the number of bundle-metric combinations for which an IQM explained the greatest variance (ΔR^2_{adj}). Only IQMs with at least two instances are displayed. Bars are colored by IQM family.

(B) Illustrative plot of preprocessed dMRI contrast against FA in the right fornix. The red curve is the GAMM fit (adjusted for age, sex, and batch), with a shaded pointwise 95% confidence band.

(C) Tract plot and heatmap of bundle-wise quality effect sizes for preprocessed dMRI contrast across microstructural metrics. Rows are ordered by mean quality effect size across IQMs; columns are ordered by mean effect size and grouped by bundle category. R/L denotes right- and left-lateralized bundles. The red rectangle highlights the largest observed quality effect size (FA in the right fornix). Tracts are colored by the FA quality effect size, as indicated by a dotted rectangle in the heatmap.

(legend continued on next page)

neuroimaging, we release fully processed, analysis-ready derivatives through the ABCC.

Historically, dMRI studies of childhood and adolescent development have largely focused on tensor-derived metrics such as FA and MD, often using atlas-based analyses that obscure individual anatomical variability.³⁰ In contrast, we analyzed 33 harmonized microstructural metrics mapped to person-specific white matter tracts. We found that dMRI metrics from advanced models—particularly ICVF, MKT, and RTOP—were on average 4- to 5-fold more sensitive to developmental processes than FA. The spatial distribution of developmental effects was also broadly consistent with prior developmental literature, with stronger maturation observed in long-range association pathways than in many projection pathways, in keeping with the protracted maturation of higher-order white matter systems.^{5,28,29,44–46} Importantly, these advanced measures not only exhibited larger age-related effect sizes but also converged on highly similar spatial patterns of maturation across white matter bundles. These metrics arise from distinct mathematical models and have different biological interpretations. For example, ICVF estimates the intracellular neurite volume fraction, RTOP reflects restricted diffusion, and MKT quantifies departures from Gaussian diffusion.^{30,47–49} However, these measures likely show converging developmental patterns because they share sensitivity to restricted, non-Gaussian diffusion patterns that may not be reliably captured by tensor-based approaches.⁵⁰ Tensor modeling may be affected by multiple (potentially competing) mechanisms of white matter development.³⁰ Correspondingly, tensor-derived measures showed weaker developmental effects that had different spatiotemporal patterns than the more advanced metrics. These findings highlight the value of collecting high-quality multi-shell dMRI data and fully exploiting the rich modeling framework it enables. These results also align with prior literature demonstrating heightened age sensitivity of advanced metrics among infants,^{42,48,51} adolescents,^{23,52} and adults.^{53,54} For studies limited to diffusion tensor modeling (for example, single-shell dMRI), our results further suggest that MD—rather than FA—provides greater sensitivity to developmental change and greater robustness to variation in image quality. This observation is consistent with smaller prior studies of adolescent development.³⁰ However, developmental patterns estimated from tensor metrics should not be assumed to directly recapitulate those derived from advanced multi-shell diffusion models, and vice versa.

Technical variation across scanning environments can introduce substantial heterogeneity in dMRI data. We found that Siemens scanners generally produced higher-quality dMRI images, as indexed by metrics including CNRs and dMRI-specific measures of *q*-space consistency (NDC and dMRI contrast). Although our analyses were not designed to identify the causes of these vendor differences, several hardware and sequence im-

plementation factors likely contribute, including differences in head coil performance, gradient hardware, echo spacing, partial Fourier implementation, and simultaneous multislice reconstruction. Because these factors differ across vendor platforms and operate simultaneously, our data cannot attribute the observed differences to any individual hardware or software component. Notably, our findings are consistent with a recent longitudinal reliability study of the ABCD cohort that likewise reported higher structural and dMRI reliability for Siemens scanners than for GE or Philips systems, suggesting that these vendor differences are robust across multiple imaging measures.⁵⁵

Given the differences in MRI outputs across diverse scanning environments, the utility of large multisite consortia such as ABCD therefore critically depends on the ability to harmonize data. Without effective harmonization, acquisition-related variability can dwarf more subtle biological signals of interest.²⁶ Here, we applied an advanced statistical harmonization approach—longitudinal ComBat-GAMM—that explicitly preserves participant-specific effects and nonlinear developmental trajectories. This approach substantially mitigated scanner-related differences in dMRI microstructural measures and increased the cross-vendor generalizability of bundle-wise developmental effects. These findings indicate that scanner-related biases are not uniformly distributed across the brain but instead exhibit spatial structure that can systematically distort developmental patterns. Accordingly, harmonization should be considered even when analyses are performed within, rather than pooled across, scanner vendors—such as when findings identified within a single scanner vendor are tested for generalizability in data acquired on a different vendor. Notably, harmonization also modestly attenuated developmental effect sizes. This likely reflects the partial alignment between acquisition batch and participant age in ABCD, as scanner software versions changed over the course of the study. Consequently, no statistical harmonization approach can completely disentangle technical and biological sources of variation in this dataset. Nevertheless, the increased consistency of developmental patterns across scanner vendors suggests that harmonization preferentially reduced unwanted scanner-related variance rather than meaningful biological differences.

Our results extend prior work demonstrating that harmonization of ABCD dMRI data improves tractography reconstruction success and aligns bundle-wise microstructural measures across sites.²⁷ However, previous efforts focused on cross-sectional baseline data, excluded Philips acquisitions because of artifacts, and applied harmonization to the raw diffusion signal. In contrast, our study incorporated longitudinal data, included all scanner vendors, and applied harmonization directly to postprocessed bundle-wise features. Although harmonizing the raw diffusion signal may remove scanner effects before model fitting, nonlinear preprocessing and reconstruction algorithms can alter or reintroduce batch effects in

(D) Scatterplot showing the relationship between the average manual rating and preprocessed dMRI contrast among the subset of images that were manually reviewed. Each color represents a different scanner vendor.

(E) Differences in quality effect size between dMRI contrast and mean manual rating among the subset of manually reviewed images. A positive value reflects a higher-quality effect size for dMRI contrast compared with the effect size of the mean manual rating. Dots are colored by bundle category. Black lines represent the mean difference in quality effect sizes across bundles within the microstructural metric.

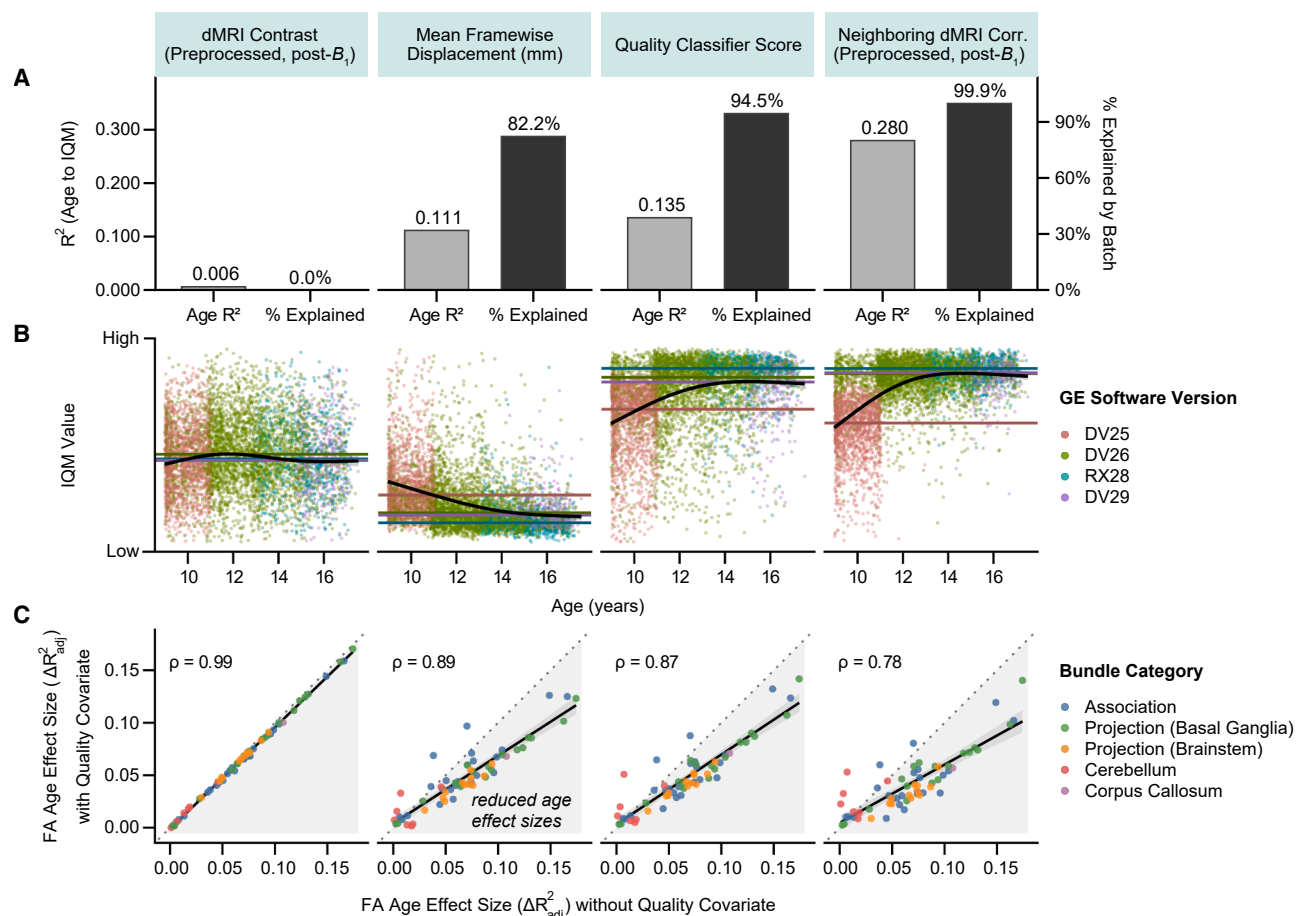


Figure 7. Batch-structured IQMs distort developmental effects in GE scanners

(A) Bar plots showing the age-to-IQM model R^2 (left) and the percentage of the age-IQM relationship explained by acquisition batch (right). Percent mediation was estimated using variance decomposition comparing models with and without batch terms.

(B) Generalized additive model (GAM) plots illustrating the relationship between age and each IQM. Each point represents a scan and is colored by the scanner software version. Black curves indicate the fitted GAM with 95% confidence intervals. Horizontal colored lines indicate the median IQM value for each software version.

(C) Comparison of bundle-wise FA age effect sizes (ΔR^2_{adj}) estimated without (x axis) and with (y axis) the inclusion of each IQM as a covariate. Points represent white matter bundles colored by their bundle category. The dashed line indicates the identity line (no change in estimated age effect size). Gray shading highlights regions where inclusion of the IQM reduces estimated developmental effect sizes. Spearman correlations (ρ) quantify the similarity of spatial developmental patterns across bundles.

derived microstructural measures. Harmonizing the features to be used in hypothesis testing therefore ensures that scanner effects are minimized. In addition, this feature-level approach substantially reduces computational burden, can be readily applied to existing processed datasets, and facilitates study-specific harmonization without requiring the complete reprocessing of the diffusion data.

In developmental neuroimaging studies, IQMs are commonly included as covariates under the assumption that data quality both influences derived microstructural measures and varies systematically with participant age.^{56,57} Our findings complicate this assumption in several important ways. First, metrics from advanced diffusion models showed relatively weak associations with image quality compared with FA, suggesting that these measures may require little or no explicit quality covariate adjust-

ment in studies of developmental change. Second, the quality metric that most strongly tracked microstructural variation—dMRI contrast—showed minimal association with age. Individual differences in dMRI contrast likely reflect the consistency of the diffusion signal across q -space, arising from factors such as residual noise, large motion spikes, and scanner-related factors. In contrast, mean framewise displacement, which remains the most ubiquitous quality covariate, decreased with age but exhibited little relationship with diffusion microstructure. These findings align with and extend recent work demonstrating that modern preprocessing pipelines substantially mitigate the influence of motion on tensor-derived metrics across the life-span.⁵⁸ Manual image assessment is often considered the gold standard for neuroimaging quality control, but it can be difficult to implement at the scale of ABCD. In our analyses, manually

derived quality ratings explained relatively little variation in diffusion microstructure relative to the variation explained by dMRI contrast. The dMRI contrast reflects nuanced q -space properties of the diffusion signal that are not readily apparent through visual inspection. As a result, automated measures such as dMRI contrast may provide a principled and scalable proxy for assessing diffusion signal quality in large datasets.

Importantly, our findings suggest that, in ABCD, apparent relationships between many IQMs and age primarily reflect acquisition-related variation rather than participant-specific acquisition factors. Scanner upgrades and software version changes occurred over the course of the study, resulting in partial alignment between participant age and scanning environment. Several common dMRI quality metrics showed associations with age that were predominantly explained by differences across, rather than within, acquisition batches. Notably, this pattern extended to head motion. Head motion, which is typically considered a participant-driven source of variation, was largely structured by acquisition batch in ABCD. Including these batch-structured metrics as covariates distorted estimates of developmental effects, likely due to collinearity between participant age and acquisition batch. These results suggest that the routine inclusion of commonly used quality covariates may not be necessary for studies of development and, in some cases, may even distort developmental inference. This consideration is especially relevant in GE scanners—which exhibited strong collinearity among age, image quality, and acquisition batch—and for FA—which showed stronger associations with image quality than metrics from advanced models. dMRI contrast may capture meaningful noise-related variation in FA, but the minimal association between dMRI contrast and age suggests that including such covariates is unlikely to meaningfully alter developmental inferences.

Our study should be interpreted in the context of several limitations. First, several commonly used dMRI derivatives are not currently produced as part of the ABCC pipeline. In particular, we did not include whole-brain interregional tractography, as such approaches are known to be prone to false-positive connections⁵⁹ and impose substantial computational and storage demands at the scale of ABCD. Additionally, QSIRecon does not yet support restriction spectrum imaging (RSI) reconstruction,⁶⁰ precluding direct comparison with the RSI data in the official ABCD dMRI release. However, alternative postprocessing workflows can be applied directly to the distributed QSIPreprocessed images, enabling future extensions of this work. Second, because optimal harmonization strategies may vary depending on the specific scientific question, we do not distribute harmonized values directly. Instead, we provide well-documented open-source code to reproduce all harmonization procedures, with the intention that these methods can be adapted to study-specific needs. Third, although longitudinal harmonization substantially reduced scanner-related variance, the acquisition batch in ABCD is partially aligned with participant age due to the longitudinal study design. As a result, scanner effects and developmental effects cannot be completely disentangled. Finally, some of our findings regarding image quality covariates may be specific to large longitudinal multisite studies such as ABCD. Accordingly, the relationships among

IQMs, diffusion measures, and phenotypes of interest should be evaluated within each dataset to guide principled selection of quality covariates.

Pre- and postprocessed dMRI data from ABCC are publicly available through release 3.1.0 via the NIH Brain Development Cohorts (NBDC) Data Hub to researchers with a valid Data Use Certification, and future releases will incorporate additional data as they are collected. Beyond the analyses presented here, our data resource provides many additional derivatives—such as measures of bundle macrostructure and volumetric microstructural maps—that support a wide range of future investigations. The end-to-end image-processing framework used here is built on widely used open-source software applicable to any BIDS-formatted dataset. As of March 2026, QSIPrep and QSIRecon have collectively been downloaded more than 58,000 times and successfully executed over 727,000 times. These tools are also the default dMRI processing pipelines for the Healthy Brain and Child Development (HBCD) study.^{42,61} This broad compatibility will facilitate combining ABCD data with other cohorts using consistent processing and harmonization strategies, enabling larger and more diverse samples for studying white matter across the lifespan. Moving forward, ABCC will accelerate efforts to understand the developing brain and variation related to cognition, environment, and psychopathology.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to the lead contact, Theodore D. Satterthwaite (theodore.satterthwaite@pennmedicine.upenn.edu).

Materials availability

This study did not generate new unique reagents or materials.

Data and code availability

The present work used raw data from the ABCD study through Release 6.0 (<https://doi.org/10.82525/jy7n-g441>). Raw ABCD data and the ABCC processed derivatives described in this article are available through the NBDC Data Hub (<https://www.nbdc-datahub.org/>) to researchers who obtain and adhere to a Data Use Certification. Additional ABCC dataset documentation may be found at https://docs.abcdstudy.org/latest/documentation/imaging/abcc_start_page.html.

Source code for QSIPrep and QSIRecon may be found on GitHub (<https://github.com/PennLINC/qsiprep> and <https://github.com/PennLINC/qsiorecon>). Container images for QSIPrep and QSIRecon are available on Docker Hub (<https://hub.docker.com/r/pennlinc/qsiprep/tags> and <https://hub.docker.com/r/pennlinc/qsiorecon/tags>) and can be executed using Docker⁶² or Singularity/Apptainer.⁶³ Study-specific analysis code and instructions for reproducing the analyses presented in this study are available at https://github.com/PennLINC/Meisler_ABCD_dMRI (archived at <https://doi.org/10.5281/zenodo.22728603>). Manual rating training material may be found at https://osf.io/89uya/overview?view_only=2559ba4e71b54c07bff70aef302540d5.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health (T32MH019112 and K99HD122318 to S.L.M.; R37MH125829 to D.A.F. and T.D.S.; 2R01MH112847 to R.T.S. and T.D.S.; 2R01MH113550 to T.D.S.; R01MH123550 to R.T.S.; F30MH138048 to K.Y.S.; RF1MH121868, RF1MH121867, RF1MH126699, R01AG060942, U19AG066567, R01EY033628, and R01EB027585 to A. Rokem; T32MH016804 and T32MH018951 to V.J.S.; R01EB031284

to H.H.; F31MH136685 to J.B.; and R01DA056499 and R25DA061823 to E.F.). S.L.M. was additionally supported by the Hartwell Foundation. G.S. was supported by a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR). A.S.K. was supported by a NARSAD Young Investigator Award from the Brain & Behavior Research Foundation. S.C. was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 837228 and PRIN PNRR P2022SMEJW. J.K. (DGE-2140004) and Z.M.H. (2038235) were supported by the National Science Foundation Graduate Research Fellowship Program. L.M.S. was supported by a National Science Foundation SBE Postdoctoral Research Fellowship (2507497). A. Rauland (269953372/GRK2150) and M.G. (EXC 3066/1 "The Adaptive Mind," project no. 533717223) were supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). M.G. is further supported by the Lise Meitner Excellence Program of the Max Planck Society and the European Union (ERC, WRAPPED, 101161197). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) study, held in the NIH Brain Development Cohorts Data Sharing Platform. This is a multisite longitudinal study designed to recruit more than 10,000 children aged 9–10 years and follow them over 10 years into early adulthood. The ABCD dataset grows and changes over time. The raw ABCD data used in this report came from ABCD release 6.0: <https://doi.org/10.82525/jy7n-g441>.

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, and U24DA041147. A full list of supporters is available at <https://abcdstudy.org/about/federal-partners/>.

AUTHOR CONTRIBUTIONS

S.L.M., M.C., and T.D.S. planned and designed the analyses; E.F., K.B.W., T.J.H., r.M., B.F., T.P., L.A.M., and D.A.F. curated, stored, and published the data; S.L.M. and M.C. processed the neuroimaging data with software developed by S.L.M., M.C., and T.S.; S.L.M. ran statistical analyses; S.L.M. and J.B. generated figures; S.L.M. and M.C. prepared training materials for manual data quality raters; A.A.C., N.H., and R.T.S. guided use of harmonization software; C.D. provided computing resources; S.L.M., M.C., H.R., B.A.-P., J.B., S.C., K.C., P.A.C., E.A.F., T.G., M.G., M.P.H., Z.M.H., I.I.K., A.S.K., J.K., A.C.L., B.M., K.M., J.L.M., A.R.P., L.P., A. Rauland, E.R., B.L.S., G.S., S.P.S., H.L.S., K.Y.S., V.J.S., T.T.T., M.Y., and A. Rokem were manual data quality raters; S.L.M. wrote the original draft; and all authors reviewed and revised the final draft.

DECLARATION OF INTERESTS

D.A.F. is a founder of Turing Medical. D.A.F. is an inventor of the FIRMM Technology 2198 (FIRMM—real-time monitoring and prediction of motion in MRI scans, exclusively licensed to Turing Medical). Any potential conflict of interest has been reviewed and managed by the University of Minnesota.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Participants
- **METHOD DETAILS**
 - Neuroimaging acquisition and curation

- Neuroimaging processing
- Manual quality rating
- Automated quality classification
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Inclusion and quality control
 - White matter feature harmonization
 - Quantifying the effect of batch on dMRI microstructure
 - Developmental sensitivity of dMRI microstructure
 - Cross-vendor and cross-metric correspondence of developmental effects
 - Sensitivity of dMRI microstructure to image quality
 - Variance decomposition of age–image quality relationships

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2026.09.021>.

Received: April 22, 2026

Revised: August 31, 2026

Accepted: September 16, 2026

REFERENCES

1. Fuhrmann, D., Knoll, L.J., and Blakemore, S.-J. (2015). Adolescence as a Sensitive Period of Brain Development. *Trends Cogn. Sci.* 19, 558–566. <https://doi.org/10.1016/j.tics.2015.07.008>.
2. Fandakova, Y., and Hartley, C.A. (2020). Mechanisms of learning and plasticity in childhood and adolescence. *Dev. Cogn. Neurosci.* 42, 100764. <https://doi.org/10.1016/j.dcn.2020.100764>.
3. Larsen, B., Sydnor, V.J., Keller, A.S., Yeo, B.T.T., and Satterthwaite, T.D. (2023). A critical period plasticity framework for the sensorimotor–association axis of cortical neurodevelopment. *Trends Neurosci.* 46, 847–862. <https://doi.org/10.1016/j.tins.2023.07.007>.
4. Crews, F., He, J., and Hodge, C. (2007). Adolescent cortical development: A critical period of vulnerability for addiction. *Pharmacol. Biochem. Behav.* 86, 189–199. <https://doi.org/10.1016/j.pbb.2006.12.001>.
5. Sydnor, V.J., Larsen, B., Bassett, D.S., Alexander-Bloch, A., Fair, D.A., Liston, C., Mackey, A.P., Milham, M.P., Pines, A., Roalf, D.R., et al. (2021). Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron* 109, 2820–2846. <https://doi.org/10.1016/j.neuron.2021.06.016>.
6. Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., Soules, M.E., Teslovich, T., Dellarco, D.V., Garavan, H., et al. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* 32, 43–54. <https://doi.org/10.1016/j.dcn.2018.03.001>.
7. Beaulieu, C., D'Arceuil, H., Hedehus, M., de Crespigny, A., Kastrup, A., and Moseley, M.E. (1999). Diffusion-weighted magnetic resonance imaging: Theory and potential applications to child neurology. *Semin. Pediatr. Neurol.* 6, 87–100. [https://doi.org/10.1016/S1071-9091\(99\)80035-7](https://doi.org/10.1016/S1071-9091(99)80035-7).
8. Schaefer, P.W., Grant, P.E., and Gonzalez, R.G. (2000). Diffusion-weighted MR Imaging of the Brain. *Radiology* 217, 331–345. <https://doi.org/10.1148/radiology.217.2.r00nv24331>.
9. Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system – a technical review. *NMR Biomed.* 15, 435–455. <https://doi.org/10.1002/nbm.782>.
10. Li, Z.A., Cai, Y., Taylor, R.L., Eisenstein, S.A., Barch, D.M., Marek, S., and Hershey, T. (2023). Associations Between Socioeconomic Status, Obesity, Cognition, and White Matter Microstructure in Children. *JAMA Netw. Open* 6, e2320276. <https://doi.org/10.1001/jamanetworkopen.2023.20276>.

11. Michel, L.C., McCormick, E.M., and Kievit, R.A. (2024). Gray and White Matter Metrics Demonstrate Distinct and Complementary Prediction of Differences in Cognitive Performance in Children: Findings from ABCD (N = 11,876). *J. Neurosci.* 44, e0465232023. <https://doi.org/10.1523/JNEUROSCI.0465-23.2023>.
12. Cotter, D.L., Ahmadi, H., Cardenas-Iniguez, C., Bottenhorn, K.L., Gauderman, W.J., McConnell, R., Berhane, K., Schwartz, J., Hackman, D.A., Chen, J.-C., et al. (2024). Exposure to multiple ambient air pollutants changes white matter microstructure during early adolescence with sex-specific differences. *Commun. Med. (Lond)* 4, 155. <https://doi.org/10.1038/s43856-024-00576-x>.
13. Roy, E., Van Rinsveld, A., Nedelec, P., Richie-Halford, A., Rauschecker, A.M., Sugrue, L.P., Rokem, A., McCandliss, B.D., and Yeatman, J.D. (2024). Differences in educational opportunity predict white matter development. *Dev. Cogn. Neurosci.* 67, 101386. <https://doi.org/10.1016/j.dcn.2024.101386>.
14. Riem, M.M.E., van Hoof, M.J., Garrett, A.S., Rombouts, S.A.R.B., van der Wee, N.J.A., van IJzendoorn, M.H., and Vermeiren, R.R.J.M. (2019). General psychopathology factor and unresolved-disorganized attachment uniquely correlated to white matter integrity using diffusion tensor imaging. *Behav. Brain Res.* 359, 1–8. <https://doi.org/10.1016/j.bbr.2018.10.014>.
15. Neumann, A., Muetzel, R.L., Lahey, B.B., Bakermans-Kranenburg, M.J., van IJzendoorn, M.H., Jaddoe, V.W., Hillegers, M.H.J., White, T., and Tiemeier, H. (2020). White Matter Microstructure and the General Psychopathology Factor in Children. *J. Am. Acad. Child Adolesc. Psychiatry* 59, 1285–1296. <https://doi.org/10.1016/j.jaac.2019.12.006>.
16. Cardenas-Iniguez, C., Moore, T.M., Kaczurkin, A.N., Meyer, F.A.C., Satterthwaite, T.D., Fair, D.A., White, T., Blok, E., Applegate, B., Thompson, L.M., et al. (2022). Direct and Indirect Associations of Widespread Individual Differences in Brain White Matter Microstructure With Executive Functioning and General and Specific Dimensions of Psychopathology in Children. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 7, 362–375. <https://doi.org/10.1016/j.bpsc.2020.11.007>.
17. Dall'Aglio, L., Xu, B., Tiemeier, H., and Muetzel, R.L. (2023). Longitudinal Associations Between White Matter Microstructure and Psychiatric Symptoms in Youth. *J. Am. Acad. Child Adolesc. Psychiatry* 62, 1326–1339. <https://doi.org/10.1016/j.jaac.2023.04.019>.
18. Retzler, C., Hallam, G., Johnson, S., and Retzler, J. (2023). Person-centred Approaches to Psychopathology in the ABCD Study: Phenotypes and Neurocognitive Correlates. *Res. Child Adolesc. Psychopathol.* 51, 1195–1212. <https://doi.org/10.1007/s10802-023-01065-w>.
19. Qiu, A., and Liu, C. (2023). Pathways link environmental and genetic factors with structural brain networks and psychopathology in youth. *Neuropsychopharmacology* 48, 1042–1051. <https://doi.org/10.1038/s41386-023-01559-7>.
20. Cieslak, M., Cook, P.A., He, X., Yeh, F.-C., Dholander, T., Adebimpe, A., Aguirre, G.K., Bassett, D.S., Betzel, R.F., Bourque, J., et al. (2021). QSIprep: an integrative platform for preprocessing and reconstructing diffusion MRI data. *Nat. Methods* 18, 775–778. <https://doi.org/10.1038/s41592-021-01185-5>.
21. Richie-Halford, A., Cieslak, M., Ai, L., Caffarra, S., Covitz, S., Franco, A.R., Karipidis, I.I., Kruper, J., Milham, M., Avelar-Pereira, B., et al. (2022). An analysis-ready and quality controlled resource for pediatric brain white-matter research. *Sci. Data* 9, 616. <https://doi.org/10.1038/s41597-022-01695-7>.
22. Kodiweera, C., Alexander, A.L., Harezlak, J., McAllister, T.W., and Wu, Y.-C. (2016). Age effects and sex differences in human brain white matter of young to middle-aged adults: A DTI, NODDI, and q-space study. *NeuroImage* 128, 180–192. <https://doi.org/10.1016/j.neuroimage.2015.12.033>.
23. Grinberg, F., Maximov, I.I., Farrher, E., Neuner, I., Amort, L., Thönneßen, H., Oberwelland, E., Konrad, K., and Shah, N.J. (2017). Diffusion kurtosis metrics as biomarkers of microstructural development: A comparative study of a group of children and a group of adults. *NeuroImage* 144, 12–22. <https://doi.org/10.1016/j.neuroimage.2016.08.033>.
24. Pines, A.R., Cieslak, M., Larsen, B., Baum, G.L., Cook, P.A., Adebimpe, A., Dávila, D.G., Elliott, M.A., Jirsaraie, R., Murtha, K., et al. (2020). Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Dev. Cogn. Neurosci.* 43, 100788. <https://doi.org/10.1016/j.dcn.2020.100788>.
25. Feczko, E., Conan, G., Marek, S., Tervo-Clemmens, B., Cordova, M., Doyle, O., Earl, E., Perrone, A., Sturgeon, D., and Klein, R. (2021). Adolescent Brain Cognitive Development (ABCD) community MRI collection and utilities. Preprint at bioRxiv. <https://doi.org/10.1101/2021.07.09.451638>.
26. Dudley, J.A., Maloney, T.C., Simon, J.O., Atluri, G., Karalunas, S.L., Altaye, M., Epstein, J.N., and Tamm, L. (2023). ABCD_Harmonizer: An Open-source Tool for Mapping and Controlling for Scanner Induced Variance in the Adolescent Brain Cognitive Development Study. *Neuroinformatics* 21, 323–337. <https://doi.org/10.1007/s12021-023-09624-8>.
27. Cetin-Karayumak, S., Zhang, F., Zurrin, R., Billah, T., Zekelman, L., Makris, N., Pieper, S., O'Donnell, L.J., and Rath, Y. (2024). Harmonized diffusion MRI data and white matter measures from the Adolescent Brain Cognitive Development Study. *Sci. Data* 11, 249. <https://doi.org/10.1038/s41597-024-03058-w>.
28. Luo, A.C., Meisler, S.L., Sydnor, V.J., Alexander-Bloch, A., Bagautdinova, J., Barch, D.M., Bassett, D.S., Davatzikos, C., Franco, A.R., Goldsmith, J., et al. (2026). Two axes of white matter development. *Nat. Commun.* 17, 1957. <https://doi.org/10.1038/s41467-026-68714-8>.
29. Sydnor, V.J., Bagautdinova, J., Larsen, B., Arcaro, M.J., Barch, D.M., Bassett, D.S., Alexander-Bloch, A.F., Cook, P.A., Covitz, S., Franco, A.R., et al. (2025). Human thalamocortical structural connectivity develops in line with a hierarchical axis of cortical plasticity. *Nat. Neurosci.* 28, 1772–1786. <https://doi.org/10.1038/s41593-025-01991-6>.
30. Lebel, C., Treit, S., and Beaulieu, C. (2019). A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR Biomed.* 32, e3778. <https://doi.org/10.1002/nbm.3778>.
31. Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., and Alexander, D.C. (2012). NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* 61, 1000–1016. <https://doi.org/10.1016/j.neuroimage.2012.03.072>.
32. Özarslan, E., Koay, C.G., Shepherd, T.M., Komlos, M.E., İrfanoğlu, M.O., Pierpaoli, C., and Basser, P.J. (2013). Mean apparent propagator (MAP) MRI: A novel diffusion imaging method for mapping tissue microstructure. *NeuroImage* 78, 16–32. <https://doi.org/10.1016/j.neuroimage.2013.04.016>.
33. Steven, A.J., Zhuo, J., and Melhem, E.R. (2014). Diffusion Kurtosis Imaging: An Emerging Technique for Evaluating the Microstructural Environment of the Brain. *Am. J. Roentgenol.* 202, W26–W33. <https://doi.org/10.2214/AJR.13.11365>.
34. Satterthwaite, T.D., Wolf, D.H., Loughhead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., and Gur, R.E. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. *NeuroImage* 60, 623–632. <https://doi.org/10.1016/j.neuroimage.2011.12.063>.
35. Yoshida, S., Oishi, K., Faria, A.V., and Mori, S. (2013). Diffusion tensor imaging of normal brain development. *Pediatr. Radiol.* 43, 15–27. <https://doi.org/10.1007/s00247-012-2496-x>.
36. Makowski, C., Lepage, M., and Evans, A.C. (2019). Head motion: the dirty little secret of neuroimaging in psychiatry. *J. Psychiatry Neurosci.* 44, 62–68. <https://doi.org/10.1503/jpn.180022>.
37. Cieslak, M., Camacho, P.B., Covitz, S., Foran, W., He, X., Humphries, J., Legarreta Gorroño, J.H., Magnussen, F., Meisler, S.L., Krause, M., et al.

- (2025). QSIRecon: Reconstruction of preprocessed q-space images (dMRI). Version 1.1.0. Zenodo. <https://doi.org/10.5281/zenodo.15213612>.
38. Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* 3, 160044. <https://doi.org/10.1038/sdata.2016.44>.
39. Gorgolewski, K.J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotà, M., Chakravarty, M.M., Churchill, N.W., Cohen, A.L., Craddock, R.C., Devenyi, G.A., et al. (2017). BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS Comput. Biol.* 13, e1005209. <https://doi.org/10.1371/journal.pcbi.1005209>.
40. Poldrack, R.A., Markiewicz, C.J., Appelhoff, S., Ashar, Y.K., Auer, T., Baillet, S., Bansal, S., Beltrachini, L., Benar, C.G., Bertazzoli, G., et al. (2024). The past, present, and future of the brain imaging data structure (BIDS). *Imaging Neurosci.* 2, 1–19. https://doi.org/10.1162/imag_a_00103.
41. Kiar, G., Clucas, J., Feczko, E., Goncalves, M., Jarecka, D., Markiewicz, C.J., Halchenko, Y.O., Hermosillo, R., Li, X., Miranda-Dominguez, O., et al. (2023). Align with the NMIND consortium for better neuroimaging. *Nat. Hum. Behav.* 7, 1027–1028. <https://doi.org/10.1038/s41562-023-01647-0>.
42. Cieslak, M., Irfanoglu, M.O., Meisler, S.L., Salo, T., Raikes, A., Cook, P.A., Chung, A.W., Lee, E.G., Li, R., Li, X., et al. (2025). Diffusion MRI Processing in the HEALTHy Brain and Child Development Study: Innovations and Applications. Preprint at bioRxiv. <https://doi.org/10.1101/2025.11.10.687672>.
43. Beer, J.C., Tustison, N.J., Cook, P.A., Davatzikos, C., Sheline, Y.I., Shinohara, R.T., and Linn, K.A.; Alzheimer's Disease Neuroimaging Initiative (2020). Longitudinal ComBat: A method for harmonizing longitudinal multi-scanner imaging data. *NeuroImage* 220, 117129. <https://doi.org/10.1016/j.neuroimage.2020.117129>.
44. Schilling, K.G., Chad, J.A., Chamberland, M., Nozais, V., Rheault, F., Archer, D., Li, M., Gao, Y., Cai, L., Del'Acqua, F., et al. (2023). White matter tract microstructure, macrostructure, and associated cortical gray matter morphology across the lifespan. *Imaging Neurosci.* 1, imag-1-00050. https://doi.org/10.1162/imag_a_00050.
45. Kim, M.E., Gao, C., Ramadass, K., Newlin, N.R., Kanakaraj, P., Bogdanov, S., Rudravarani, G., Archer, D., Hohman, T.J., Jefferson, A.L., et al. (2026). White matter micro- and macrostructure brain charts for the human lifespan. *Nature* 655, 979–989. <https://doi.org/10.1038/s41586-026-10454-2>.
46. Bagautdinova, J., Shafiei, G., Luo, A.C., Pecsok, M.K., Salo, T., Alexander-Bloch, A.F., Bassett, D.S., Gardner, M.E., Gur, R.E., Gur, R.C., et al. (2026). Anatomical white matter tracts span the cortical hierarchy to support cognitive diversity. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-026-02559-5>.
47. Reveley, C., Ye, F.Q., and Leopold, D.A. (2024). Diffusion kurtosis imaging, MAP-MRI and NODDI selectively track gray matter myelin density in the primate cerebral cortex. *Imaging Neurosci.* 2, imag-2-00368. https://doi.org/10.1162/imag_a_00368.
48. Zhao, X., Shi, J., Dai, F., Wei, L., Zhang, B., Yu, X., Wang, C., Zhu, W., and Wang, H. (2021). Brain Development From Newborn to Adolescence: Evaluation by Neurite Orientation Dispersion and Density Imaging. *Front. Hum. Neurosci.* 15, 616132. <https://doi.org/10.3389/fnhum.2021.616132>.
49. Kelm, N.D., West, K.L., Carson, R.P., Gochberg, D.F., Ess, K.C., and Does, M.D. (2016). Evaluation of diffusion kurtosis imaging in ex vivo hypomyelinated mouse brains. *NeuroImage* 124, 612–626. <https://doi.org/10.1016/j.neuroimage.2015.09.028>.
50. Sadiikov, A., Choi, H.L., Xiao, J., Cai, L.T., and Mukherjee, P. (2025). Mapping the microstructure of human cerebral cortex in vivo with diffusion MRI. *Commun. Biol.* 8, 1088. <https://doi.org/10.1038/s42003-025-08523-9>.
51. Paydar, A., Fieremans, E., Nwankwo, J.I., Lazar, M., Sheth, H.D., Adisetiyo, V., Helpert, J.A., Jensen, J.H., and Milla, S.S. (2014). Diffusional Kurtosis Imaging of the Developing Brain. *AJNR, Am. J. Neuroradiol.* 35, 808–814. <https://doi.org/10.3174/ajnr.A3764>.
52. Genc, S., Malpas, C.B., Holland, S.K., Beare, R., and Silk, T.J. (2017). Neurite density index is sensitive to age related differences in the developing brain. *NeuroImage* 148, 373–380. <https://doi.org/10.1016/j.neuroimage.2017.01.023>.
53. Taha, H.T., Chad, J.A., and Chen, J.J. (2022). DKI enhances the sensitivity and interpretability of age-related DTI patterns in the white matter of UK biobank participants. *Neurobiol. Aging* 115, 39–49. <https://doi.org/10.1016/j.neurobiolaging.2022.03.008>.
54. Bouhrara, M., Avram, A.V., Kiely, M., Trivedi, A., and Benjamini, D. (2023). Adult lifespan maturation and degeneration patterns in gray and white matter: A mean apparent propagator (MAP) MRI study. *Neurobiol. Aging* 124, 104–116. <https://doi.org/10.1016/j.neurobiolaging.2022.12.016>.
55. Pan, Y., Hong, L.E., Acheson, A., Thompson, P.M., Jahanshad, N., Zhu, A.H., Yu, J., Chen, C., Ma, T., Liu, H.-L., et al. (2024). A Site-Wise Reliability Analysis of the ABCD Diffusion Fractional Anisotropy and Cortical Thickness: Impact of Scanner Platforms. *Hum. Brain Mapp.* 45, e70070. <https://doi.org/10.1002/hbm.70070>.
56. Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., and Fischl, B. (2014). Spurious group differences due to head motion in a diffusion MRI study. *NeuroImage* 88, 79–90. <https://doi.org/10.1016/j.neuroimage.2013.11.027>.
57. Baum, G.L., Roalf, D.R., Cook, P.A., Ciric, R., Rosen, A.F.G., Xia, C., Elliott, M.A., Ruparel, K., Verma, R., Tunç, B., et al. (2018). The impact of in-scanner head motion on structural connectivity derived from diffusion MRI. *NeuroImage* 173, 275–286. <https://doi.org/10.1016/j.neuroimage.2018.02.041>.
58. Schilling, K.G., Ramadass, K., Sairanen, V., Kim, M.E., Rheault, F., Newlin, N., Nguyen, T., Barquero, L., D'archangel, M., Gao, C., et al. (2025). Head Motion in Diffusion Magnetic Resonance Imaging: Quantification, Mitigation, and Structural Associations in Large, Cross-Sectional Datasets Across the Lifespan. *Hum. Brain Mapp.* 46, e70143. <https://doi.org/10.1002/hbm.70143>.
59. Maier-Hein, K.H., Neher, P.F., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.-C., Lin, Y.-C., Ji, Q., et al. (2017). The challenge of mapping the human connectome based on diffusion tractography. *Nat. Commun.* 8, 1349. <https://doi.org/10.1038/s41467-017-01285-x>.
60. White, N.S., Leergaard, T.B., D'Arceuil, H., Bjaalie, J.G., and Dale, A.M. (2013). Probing tissue microstructure with restriction spectrum imaging: Histological and theoretical validation. *Hum. Brain Mapp.* 34, 327–346. <https://doi.org/10.1002/hbm.21454>.
61. Nelson, C.A., Frankeberger, J., and Chambers, C.D. (2024). An introduction to the HEALTHy Brain and Child Development Study (HBCD) study. *Dev. Cogn. Neurosci.* 69, 101441. <https://doi.org/10.1016/j.dcn.2024.101441>.
62. Merkel, D. (2014). Docker: lightweight linux containers for consistent development and deployment. *Linux J.* 2014, 2.
63. Kurtzer, G.M., Sochat, V., and Bauer, M.W. (2017). Singularity: Scientific containers for mobility of compute. *PLoS One* 12, e0177459. <https://doi.org/10.1371/journal.pone.0177459>.
64. Fischl, B. (2012). FreeSurfer. *NeuroImage* 62, 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>.
65. Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., van der Walt, S., Descoteaux, M., and Nimmo-Smith, I.; Dipy Contributors (2014). Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinform.* 8. <https://doi.org/10.3389/fninf.2014.00008>.

66. Irfanoglu, M.O., Nayak, A., Taylor, P., Thai, A., and Pierpaoli, C. (2025). *TORTOISEV4*: Reimagining the NIH diffusion MRI processing pipeline. *Imaging Neurosci.* 3, IMAG.a.948. <https://doi.org/10.1162/IMAG.a.948>.
67. Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., and Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* 202, 116137. <https://doi.org/10.1016/j.neuroimage.2019.116137>.
68. Yeh, F.-C. (2025). DSI Studio: an integrated tractography platform and fiber data hub for accelerating brain research. *Nat. Methods* 22, 1617–1619. <https://doi.org/10.1038/s41592-025-02762-8>.
69. Tustison, N.J., Cook, P.A., Holbrook, A.J., Johnson, H.J., Muschelli, J., Devenyi, G.A., Duda, J.T., Das, S.R., Cullen, N.C., and Gillen, D.L. (2021). The ANTsX ecosystem for quantitative biological and medical imaging. *Sci. Rep.* 11, 9068.
70. Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R.Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., and Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Dev. Cogn. Neurosci.* 32, 16–22. <https://doi.org/10.1016/j.dcn.2018.04.004>.
71. Hagler, D.J., Hatton, S.N., Cornejo, M.D., Makowski, C., Fair, D.A., Dick, A.S., Sutherland, M.T., Casey, B.J., Barch, D.M., Harms, M.P., et al. (2019). Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage* 202, 116091. <https://doi.org/10.1016/j.neuroimage.2019.116091>.
72. Tisdall, M.D., Reuter, M., Qureshi, A., Buckner, R.L., Fischl, B., and van der Kouwe, A.J.W. (2016). Prospective motion correction with volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. *NeuroImage* 127, 11–22. <https://doi.org/10.1016/j.neuroimage.2015.11.054>.
73. Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., and Ghosh, S.S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinform.* 5, 13. <https://doi.org/10.3389/fninf.2011.00013>.
74. Gorgolewski, K.J., Esteban, O., Markiewicz, C.J., Ziegler, E., Ellis, D.G., Nott, M.P., Jarecka, D., Johnson, H., Burns, C., and Manhães-Savio, A. (2018). Nipype. Zenodo. <https://zenodo.org/records/1255001>.
75. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., and Gee, J.C. (2010). N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>.
76. Avants, B.B., Tustison, N., and Song, G. (2009). Advanced normalization tools (ANTs). *Insight J.* 2, 1–35.
77. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., and Hoffmann, M. (2022). SynthStrip: skull-stripping for any brain image. *NeuroImage* 260, 119474. <https://doi.org/10.1016/j.neuroimage.2022.119474>.
78. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., and Iglesias, J.E.; ADNI (2023). SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. *Med. Image Anal.* 86, 102789. <https://doi.org/10.1016/j.media.2023.102789>.
79. Billot, B., Magdamo, C., Cheng, Y., Arnold, S.E., Das, S., and Iglesias, J.E. (2023). Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain MRI datasets. *Proc. Natl. Acad. Sci. USA* 120, e2216399120. <https://doi.org/10.1073/pnas.2216399120>.
80. Veraart, J., Novikov, D.S., Christiaens, D., Ades-Aron, B., Sijbers, J., and Fieremans, E. (2016). Denoising of diffusion MRI using random matrix theory. *NeuroImage* 142, 394–406. <https://doi.org/10.1016/j.neuroimage.2016.08.016>.
81. Irfanoglu, M.O., Nayak, A., Jenkins, J., and Pierpaoli, C. (2017). *TORTOISE v3: Improvements and new features of the NIH diffusion MRI processing pipeline*. In *Program and Proceedings of the ISMRM 25th Annual Meeting and Exhibition*.
82. Lee, H.-H., Novikov, D.S., and Fieremans, E. (2021). Removal of partial Fourier-induced Gibbs (RPG) ringing artifacts in MRI. *Magn. Reson. Med.* 86, 2733–2750. <https://doi.org/10.1002/mrm.28830>.
83. Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., and Smith, S.M. (2012). FSL. *NeuroImage* 62, 782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>.
84. Andersson, J.L.R., and Sotiropoulos, S.N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* 125, 1063–1078. <https://doi.org/10.1016/j.neuroimage.2015.10.019>.
85. Andersson, J.L.R., Graham, M.S., Zsoldos, E., and Sotiropoulos, S.N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *NeuroImage* 141, 556–572. <https://doi.org/10.1016/j.neuroimage.2016.06.058>.
86. Andersson, J.L.R., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20, 870–888. [https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7).
87. Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 59, 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>.
88. Schilling, K.G., Yeh, F.-C., Nath, V., Hansen, C., Williams, O., Resnick, S., Anderson, A.W., and Landman, B.A. (2019). A fiber coherence index for quality control of B-table orientation in diffusion MRI scans. *Magn. Reson. Imaging* 58, 82–89. <https://doi.org/10.1016/j.mri.2019.01.018>.
89. Leemans, A., and Jones, D.K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* 61, 1336–1349. <https://doi.org/10.1002/mrm.21890>.
90. Yeh, F.-C., Zaydan, I.M., Suski, V.R., Lacomis, D., Richardson, R.M., Maroon, J.C., and Barrios-Martinez, J. (2019). Differential tractography as a track-based biomarker for neuronal injury. *NeuroImage* 202, 116131. <https://doi.org/10.1016/j.neuroimage.2019.116131>.
91. Henriques, R.N., Correia, M.M., Marrale, M., Huber, E., Kruper, J., Koudoro, S., Yeatman, J.D., Garyfallidis, E., and Rokem, A. (2021). Diffusional Kurtosis Imaging in the Diffusion Imaging in Python Project. *Front. Hum. Neurosci.* 15, 675433. <https://doi.org/10.3389/fnhum.2021.675433>.
92. Basser, P.J., Mattiello, J., and LeBihan, D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson., B* 103, 247–254. <https://doi.org/10.1006/jmrb.1994.1037>.
93. Yeh, F.-C., Wedeen, V.J., and Tseng, W.-Y.I. (2010). Generalized q-sampling imaging. *IEEE Trans. Med. Imaging* 29, 1626–1635. <https://doi.org/10.1109/TMI.2010.2045126>.
94. Daducci, A., Canales-Rodríguez, E.J., Zhang, H., Dyrby, T.B., Alexander, D.C., and Thiran, J.-P. (2015). Accelerated Microstructure Imaging via Convex Optimization (AMICO) from diffusion MRI data. *NeuroImage* 105, 32–44. <https://doi.org/10.1016/j.neuroimage.2014.10.026>.
95. Guerrero, J.M., Adluru, N., Bendlin, B.B., Goldsmith, H.H., Schaefer, S.M., Davidson, R.J., Kecske, S.R., Zhang, H., and Alexander, A.L. (2019). Optimizing the intrinsic parallel diffusivity in NODDI: An extensive empirical evaluation. *PLoS One* 14, e0217118. <https://doi.org/10.1371/journal.pone.0217118>.
96. Yeh, F.-C. (2020). Shape analysis of the human association pathways. *NeuroImage* 223, 117329. <https://doi.org/10.1016/j.neuroimage.2020.117329>.
97. Yeh, F.-C., Panesar, S., Fernandes, D., Meola, A., Yoshino, M., Fernandez-Miranda, J.C., Vettel, J.M., and Verstynen, T. (2018). Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage* 178, 57–68. <https://doi.org/10.1016/j.neuroimage.2018.05.027>.

98. Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., and Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage* 103, 411–426. <https://doi.org/10.1016/j.neuroimage.2014.07.061>.
99. Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM)*, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>.
100. Zhao, C., Tapera, T.M., Bagautdinova, J., Bourque, J., Covitz, S., Gur, R.E., Gur, R.C., Larsen, B., Mehta, K., Meisler, S.L., et al. (2023). ModelArray: an R package for statistical analysis of fixel-wise data. *NeuroImage* 271, 120037. <https://doi.org/10.1016/j.neuroimage.2023.120037>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Adolescent Brain Cognitive Development Study (ABCD)	NIH Brain Development Cohorts (NBDC) https://nbdc-datashare.lassoinformatics.com/	DOI: https://doi.org/10.82525/jy7n-g441
Software and algorithms		
FreeSurfer	Fischl ⁶⁴	RRID: SCR_001847
QSIprep	Cieslak et al. ²⁰ ; https://github.com/PennLINC/qsiprep	DOI: https://doi.org/10.5281/zenodo.14187327
QSIRecon	Cieslak et al. ²⁰ ; https://github.com/PennLINC/qsiorecon	DOI: https://doi.org/10.5281/zenodo.14193615
ComBatFamily	https://github.com/andy1764/ComBatFamily	N/A
DIPY	Garyfallidis et al. ⁶⁵ ; https://github.com/dipy/dipy	RRID: SCR_000029
TORTOISE	Irfanoglu et al. ⁶⁶ ; https://github.com/QMICodeBase/TORTOISEV4	RRID: SCR_001645
MRtrix3	Tournier et al. ⁶⁷ ; https://github.com/Mrtrix3/mrtrix3	RRID: SCR_024123
DSI Studio	Yeh ⁶⁸ ; https://github.com/frankyeh/DSI-Studio	RRID: SCR_009557
ANTS	Tustison et al. ⁶⁹ ; https://github.com/antsx/ants	RRID: SCR_004757

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Participants

A full description of the recruitment procedures, priorities, and inclusion criteria for the Adolescent Brain Cognitive Development (ABCD) Study has been previously reported.⁷⁰ Participants were initially recruited between the ages of 8 and 10 years and were invited to return for MRI assessments every two years. ABCD Release 6.0 and ABCC Release 3.1.0 include participants scanned at up to four time points, corresponding to the six-year follow-up visit at ages 15–17. All participants provided informed assent, and parents or legal guardians provided written informed consent. Study procedures at most sites were approved by a central Institutional Review Board at the University of California, San Diego; several sites obtained approval from local institutional review boards. After applying all inclusion criteria (see [inclusion and quality control](#)), the final analysis sample consisted of 24,178 diffusion MRI scans from 10,738 participants (Table 1). Specifically, 2,892 participants contributed one scan, whereas 3,540, 3,018, and 1,288 participants contributed two, three, and four scans, respectively. The ABCC 3.1.0 release includes additional datasets that did not meet the inclusion criteria for the present analyses, as well as other modalities of data.²⁵ Future ABCC releases will incorporate additional ABCD data as they are acquired and processed.

METHOD DETAILS

Neuroimaging acquisition and curation

Extensive details of the ABCD neuroimaging acquisition protocols have been previously described.^{6,71} Data were collected across 21 sites in the United States using seven scanner models from three vendors (Siemens Healthineers: Prisma and Prisma Fit; GE Healthcare: Discovery MR750, SIGNA UHP, and SIGNA Premier; Philips: Achieva dStream and Ingenia; Table S3). At each study visit, participants underwent T₁-weighted (T₁w) structural imaging, diffusion MRI (dMRI), and reverse phase-encoded B₀ field map acquisitions, among other imaging modalities. T₁w images were acquired at 1 mm isotropic resolution with volumetric navigator (vNav) prospective motion correction.⁷² Multi-shell dMRI scans were collected at 1.7 mm isotropic resolution with seven b = 0 s/mm² volumes and 96 directions across four b-values (6 directions with b = 500 s/mm², 15 directions with b = 1,000 s/mm², 15 directions with b = 2,000 s/mm², and 60 directions with b = 3,000 s/mm²). For participants scanned on Philips systems, dMRI data were acquired as two separate scans (Scan 1: 3 b = 0 s/mm² volumes, 4 directions with b = 500 s/mm², 7 directions with b = 1000 s/mm², 8 directions with b = 2000 s/mm² and 29 directions with b = 3000 s/mm². Scan 2: 4 b = 0 s/mm² volumes, 2 directions with b = 500 s/mm², 8 directions with b = 1000 s/mm², 7 directions with b = 2000 s/mm² and 30 directions with b = 3000 s/mm²). The reverse phase-encoded B₀ echo-planar imaging acquisition was obtained before the diffusion scans to enable correction of susceptibility-induced geometric distortions. Raw imaging data were curated to Brain Imaging Data Structure (BIDS) format^{38,40} and were used for both the ABCC and official ABCD data releases.

Neuroimaging processing

For ABCC, T₁w, dMRI, and fieldmap images were preprocessed with the BIDS application³⁹ QSIprep²⁰ version 0.21.4, which is based on Nipype version 1.8.6.^{73,74} Data were postprocessed with QSIRecon³⁷ version 1.0.0rc2, which is based on Nipype version 1.9.1. These containerized software environments ensure computational reproducibility. The following paragraphs contain partially-edited text generated by QSIprep and QSIRecon, which is distributed with a Creative Commons Zero (CC0) license with the expressed intention of being included in manuscripts.

Anatomical preprocessing

The T₁w image was corrected for intensity non-uniformity using N4BiasFieldCorrection⁷⁵ from ANTs^{69,76} version 2.4.3, and used as an anatomical reference throughout the workflow. The anatomical reference image was reoriented into anterior commissure-posterior commissure (AC-PC) alignment via a rigid transformation extracted from a full affine registration to the MNI152NLin2009cAsym template. A full nonlinear registration to the template from the AC-PC space was estimated via symmetric nonlinear registration (SyN) using antsRegistration. Brain extraction was performed on the T₁w image using SynthStrip,⁷⁷ and automated segmentation was performed using SynthSeg^{78,79} from FreeSurfer⁶⁴ version 7.3.1.

dMRI preprocessing

MP-PCA denoising as implemented in MRtrix3's⁶⁷ dwidenoise⁸⁰ was applied with a 5-voxel window. After MP-PCA, Gibbs unringing was performed using TORTOISE's⁸¹ Gibbs.⁸² Following unringing, the mean intensity of the dMRI series was adjusted so that the mean intensity of the $b = 0$ images matched across each separate dMRI scanning sequence. FSL's⁸³ (version 6.0.8) eddy was used for head motion correction and eddy current correction.⁸⁴ eddy was configured with a q -space smoothing factor of 10, a total of 5 iterations, and 1000 voxels used to estimate hyperparameters. eddy's outlier replacement was run.⁸⁵ FSL's TOPUP⁸⁶ was used to estimate a susceptibility-induced off-resonance field based on the reverse phase-encoded B_0 reference image. The TOPUP-estimated fieldmap was incorporated into the eddy current and head motion correction interpolation. Final interpolation was performed using the Jacobian modulation method. The dMRI time series was rigidly aligned to AC-PC orientation to match the T₁w image, preserving the original 1.7 mm isotropic voxel resolution. B_1 field inhomogeneity was then corrected using dwibiascorrect from MRtrix3 with the N4⁷⁵ algorithm.

Automated image quality measures

A set of 40 automated image quality metrics (IQMs) were generated by QSIprep (Table S1), as detailed below. Vendor-wise differences in IQMs were tested with Kruskal-Wallis tests and post hoc pairwise Wilcoxon rank sum tests. A quality classifier score was also calculated after training a multivariate model to predict expert manual ratings based on automated IQMs. Pre-to-post differences in IQMs were tested with paired Wilcoxon signed-rank tests. Correlations between IQMs are reported in Figure S17.

Motion and registration

Seven IQMs were derived to index in-scanner motion and registration quality. The average and maximum framewise displacement values were recorded, following the definition described by Power and colleagues (2012).⁸⁷ Additionally, the largest values of translation and rotation were recorded, along with the maxima of their derivatives. To gauge T₁w-to-dMRI co-registration quality, the difference in spatial overlap between the anatomical brain mask and the dMRI brain mask was calculated using the Dice coefficient.

Contrast-to-noise ratio

The average, median, and standard deviation of the voxel-wise contrast-to-noise ratio was calculated within each of the four dMRI diffusion-weighted shells (i.e., excluding $b = 0$) after eddy current and motion correction, totaling 12 IQMs. Additionally, the average, median and standard deviation of the voxel-wise temporal signal to noise ratio (tSNR) of the $b = 0$ volumes was calculated. The voxel-wise values are calculated as part of FSL's eddy command within QSIprep.

Fiber coherence

The fiber coherence and incoherence indices, as described in Schilling and colleagues (2019)⁸⁸ were calculated to measure fixel-by-fixel connectivity and gauge the validity of the gradient table in the raw data and in the preprocessed dMRI after accounting for transformations from motion correction and AC-PC alignment.⁸⁹ These two indices were derived from the raw signal, as well as preprocessed image before and after B_1 bias correction, for a total of 6 IQMs.

dMRI signal similarity and dissimilarity

A total of 12 IQMs were generated to quantify expected dMRI signal similarity between adjacent slices and gradient directions, as well as expected dMRI signal dissimilarity between perpendicular gradient directions. First, the number of poor-quality slices was calculated for the raw data, as well as the preprocessed dMRI before and after B_1 bias correction for a total of three IQMs. A slice was considered low quality if the sum of its voxel-wise intensity differences from the immediately preceding and following slices exceeded twice the voxel-wise intensity difference between those two neighboring slices. Finally, the average neighboring dMRI correlations (NDC)⁹⁰ and dMRI contrasts were calculated. NDC for a given volume is quantified as the voxel-wise correlation between that volume and its nearest neighboring volume in q -space. The nearest neighbor was identified as the volume in the same b -shell with the closest angle. A value closer to 1 indicates a higher degree of expected signal similarity, denoting higher quality. dMRI contrast for a volume is quantified as NDC divided by the voxel-wise correlation between that volume and the volume with the most perpendicular direction in q -space. A higher dMRI contrast value indicates a higher degree of expected signal dissimilarity, denoting higher quality. These indices were derived on the raw and preprocessed dMRI series before and after B_1 bias correction. Additionally, NDC was calculated on both the entire 3D volume space and restricted to the brain mask; in total this yielded 9 IQMs.

dMRI postprocessing

Several dMRI reconstruction models were run as part of QSIRecon, resulting in a broad set of derived microstructural maps (Figure 1; Table S2). All maps described below were produced in participant-specific dMRI space (AC-PC orientation), as well as in MNI152NLin2009cAsym space using nearest neighbor interpolation to facilitate mass univariate analyses. Nearest-neighbor interpolation was used when warping scalar maps to standard space to avoid spatial smoothing artifacts (e.g., ringing in the corpus callosum) introduced by higher-order interpolation during nonlinear transformations. 33 of these metric maps were evaluated in this study, and 5 are focused on in the main text. Postprocessing software bundled in QSIRecon included DIPY version 1.8.0,⁶⁵ TORTOISE version 4 (5d0c688e),⁶⁶ DSI Studio version “Hou” (03476a67),⁶⁸ and AMICO version 2.0.3.

Diffusion tensor imaging and diffusion kurtosis imaging

The diffusion tensor imaging (DTI) model characterizes the 3D Gaussian diffusion process in each voxel by estimating a second-order tensor from the diffusion-weighted signal. The diffusion kurtosis imaging (DKI) model extends the tensor formulation by quantifying the degree to which water diffusion deviates from Gaussian behavior. This is achieved by additionally modeling a fourth-order kurtosis tensor that captures microstructural complexity not represented by DTI. Estimates of DTI parameters from a DKI model may be more reliable than from a standalone DTI model.⁹¹

In ABCC, the DTI model was fitted using three different software packages. The inner-shell tensor fit ($b \leq 1200$ s/mm²) was computed twice. DSI Studio used an ordinary least squares approach, while TORTOISE used a weighted linear least squares⁹² approach, in which the tensor parameter estimation is weighted in proportion to the estimated CNR. Additionally, a DKI fit was run with DIPY using the full dMRI data, producing both DKI and DTI derivatives. The parametric microstructural maps output by each software varied (Figure 1; Table S2), but the total set of DTI maps we analyzed included fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), average radial diffusivity (RD), individual radial diffusivity components (RD1 and RD2), and lattice index (LI). DKI maps included kurtosis fractional anisotropy (KFA), mean kurtosis (MK), mean kurtosis tensor (MKT), axial kurtosis (AK), and radial kurtosis (RK). Other derived DTI maps, such as the average $b = 0$ amplitude (AM) and individual tensor components (e.g., T_{xx}) are included in the data release but are not considered microstructural indices. These auxiliary metrics were not considered in the present analyses. A comparison of FA and MD results across the three tensor-fitting workflows is provided in Figure S18.

Generalized q-sampling imaging

Generalized q -sampling imaging (GQI) is a model-free reconstruction approach that estimates the diffusion orientation distribution function (ODF) directly from the q -space signal without assuming Gaussian diffusion or fitting a tensor model.⁹³ By leveraging the Fourier relationship between the diffusion signal and the spin displacement distribution, GQI recovers high-angular resolution ODFs that capture complex fiber configurations—including crossing, fanning, and branching—that are not resolvable with traditional tensor-based methods. These ODFs were used to derive quantitative indices such as quantitative anisotropy (QA), generalized fractional anisotropy (GFA), and isotropic diffusion (ISO). The helix angle (HA) was also calculated but is not analyzed in this study. In QSIRecon, GQI modeling is performed by DSI Studio.

Mean apparent propagator MRI

Mean apparent propagator MRI (MAP-MRI)³² models the full three-dimensional probability distribution of water molecule displacements, named the ensemble average propagator (EAP). MAP-MRI does not impose Gaussian assumptions and instead captures parametric features of the propagator, enabling sensitive characterization of microstructural complexity, directional dependence, and restrictions to diffusion. The method yields a suite of microstructural maps derived from the EAP that reflect different aspects of tissue organization and diffusion. MAP-MRI was run with TORTOISE and produced microstructural maps including non-Gaussianity (NG), its directional components parallel and perpendicular to the principal diffusion axis (NG_{||} and NG_⊥), propagator anisotropy (PA), the angular distance between the estimated EAP and an isotropic EAP (PA₀), as well as maps of return-to-origin (RTOP), return-to-axis (RTAP), and return-to-plane (RTPP) probabilities.

Neurite orientation dispersion and density imaging

Neurite Orientation Dispersion and Density Imaging (NODDI)³¹ is a biophysical model that characterizes microstructural features of brain tissue by estimating the relative contributions of intra-neurite, extra-neurite, and isotropic (CSF-like) compartments. The model provides indices such as intracellular volume fraction (ICVF, synonymous with neurite density index or NDI), orientation dispersion index (ODI), and isotropic volume fraction (ISO), which reflect neurite density, angular dispersion, and free-water content, respectively. We used the Accelerated Microstructure Imaging via Convex Optimization (AMICO)⁹⁴ framework to fit NODDI with fixed parallel intrinsic and isotropic diffusivities of 0.0011 and 0.003 mm²/s, respectively.⁹⁵

Tractography and tractometry

A set of 67 white matter bundles across the brain were segmented for each dMRI image using DSI Studio’s AutoTrack⁹⁶ (Figure S1). These bundles span several major anatomical categories, including association, projection, commissural, cerebellar, and brainstem pathways. This pipeline uses quantitative anisotropy and isotropic diffusion maps (both derived from GQI) to inform a non-linear registration from MNI152NLin2009aAsym space to participant-space. Bundle priors were based on a population-average white matter atlas.⁹⁷ For each bundle, deterministic tractography seeds were placed around the atlas-defined bundle region mapped to subject-space and propagated until they terminated. Streamline propagation was based on fiber orientation distribution functions (FODFs) derived from multi-shell multi-tissue constrained spherical deconvolution (MSMT-CSD),⁹⁸ as implemented in MRtrix3. The spherical harmonics from the MRtrix3 FODFs were resampled to the default direction set used by DSI Studio for cross-software compatibility. Bundle-wise averages and medians for every metric map produced by QSIRecon were calculated. For this study, we

focused on the bundle-wise median values. Additionally, bundle-wise macrostructural shape metrics,⁹⁶ such as volume and surface area, were calculated. These bundle-wise summary measures were saved as tabular data.

Manual quality rating

We trained reviewers to manually rate the quality of a subset of 3,101 ABCC preprocessed dMRI images. We used these manual ratings to train a classifier to predict quality ratings across the entire dataset. To ensure consistency, all raters completed standardized training. This involved engaging with reading materials and demonstration videos highlighting common artifacts. Reviewers also completed a pre-rating quiz, in which they were asked to all rate the same 10 images. Although there is no ground truth, an answer key (written by authors S.L.M. and M.C.) provided explanations for how expert reviewers would likely rate the images in the quiz set. The training materials may be found at https://osf.io/89uya/overview?view_only=2559ba4e71b54c07bff70aef302540d5. Each reviewer assigned every image an integer score from -2 ("definitely fail") to $+2$ ("definitely pass"), with 0 indicating uncertainty. All assessments were made using the browser-based dmripview interface,²¹ which presents the full dMRI time series alongside framewise displacement and head-translation plots, slice-wise noise timeseries, registration overlays, and color-coded fractional anisotropy maps. In general, reviewers were instructed to focus on overall *practical usability* of the dMRI. In other words, a single noisy volume might be acceptable if the remainder of the series and fractional anisotropy map were clean. However, issues that were pervasive across the entire timeseries, such as poor brain masking, severe persistent motion corruption, or systemic noise, warranted a failure. While reviewers could assess visual depictions of motion via the interface, they were blinded to all other quantitative IQMs.

A total of 33 reviewers were employed to rate images. First, a team of 23 trained reviewers evaluated images acquired on Siemens scanners. During this stage, 1,771 Siemens dMRI images from the baseline session were randomly selected so that, across all combinations of the 23 reviewers, each image received exactly three independent ratings. Each reviewer rated 231 images equally distributed across 7 quality percentile bins (indexed by the unmasked processed NDC before B_1 bias correction), and each reviewer was paired with each other reviewer 21 times to gauge inter-rater reliability. In a separate stage, 21 raters (11 of these raters overlapped with the Siemens reviewers) reviewed a total of 1,330 dMRI images, with 665 images coming from each of the GE and Philips vendors across all timepoints. Each reviewer rated 190 images, with the GE-to-Philips split being roughly half, again divided across 7 quality percentile bins. Each image was rated by 3 reviewers, and raters were paired with one another 19 times to gauge inter-rater reliability. Inter-rater disagreement for a particular image and pair of reviewers was quantified as the absolute difference in rating scores (Figure S19A). Of the 3,101 scans, 2,829 were included in the total cohort analyzed in this study. However, all 3,101 scans were used to train the quality classifier.

Automated quality classification

We trained a regression-based machine-learning model to predict manual dMRI quality ratings from the 40 image quality metrics (IQMs) produced by QSIprep. For each dMRI image, the means of the three independent manual ratings were used as the prediction targets. Following prior work on automated dMRI quality assessment by Richie-Halford, Cieslak, and colleagues,²¹ we used an eXtreme Gradient Boosting (XGBoost) model⁹⁹; however, instead of predicting a binary pass/fail label, we implemented an XGBRegressor classifier to predict the continuous mean quality rating. This approach preserved the variability in image quality and yielded a continuous prediction metric that we used directly as a covariate in subsequent statistical analyses. Images from all scanner vendors were pooled for model training.

To evaluate model performance and mitigate overfitting, we used a nested 5×5 -fold cross-validation framework. Both inner and outer folds were stratified to balance for scanner-vendor representation and the proportion of ratings above and below 0. Data in the inner folds were used to tune hyperparameters, which were optimized using a 5-fold cross validation Bayesian search procedure (200 iterations). The model configuration yielding the lowest mean cross-validated root mean squared error (RMSE) across the inner folds was selected as the best-performing model and then evaluated on the held-out outer test fold. This entire process was repeated across 1,000 random seeds, yielding 1,000 predicted quality scores per rated image. To quantify the performance of the model, we ran a Pearson correlation between the average of the 3 manual ratings and the average of the 1,000 predicted quality scores (Figure S19B).

After cross-validation performance was established, a final model was retrained on all labeled data (i.e., images with manual quality ratings). Hyperparameters were similarly optimized using a Bayesian search with 200 iterations, and the best-performing model was retained. This final trained model was then applied to all unlabeled images in the dataset to generate predicted quality scores based on their QSIprep-derived IQMs. To prevent overfitting, each labeled image was assigned a quality score equal to the average of the 1,000 predictions during cross-validation. Feature importance for each IQM was assessed using SHapley Additive exPlanations (SHAP) values (Figure S19C).

QUANTIFICATION AND STATISTICAL ANALYSIS

Inclusion and quality control

Following previously described procedures,⁷¹ a subset of ABCD images were manually reviewed by the ABCD Data Analysis, Informatics & Resource Center. These initial quality assurance checks confirmed that each dMRI scan was complete and had an

accompanying B_0 map for susceptibility distortion estimation. Scans were also flagged if derived dMRI maps (e.g., mean $b = 0$ image, FA, MD) exhibited artifacts. Any failing images were excluded before curation and processing and were not included in either the official ABCD or ABCC data releases.

ABCC release 3.1.0 includes 24,774 datasets from 10,844 participants with both QSIprep and QSIRecon outputs on the NIH Brain Development Cohorts (NBDC) Data Hub. For this study, we applied additional quality assurance protocols. A set of five bundles, including the bilateral dentatorubrothalamic tracts, bilateral corticobulbar tracts, and the anterior commissure, had relatively high failure rates of between 257 and 1,055 instances each. We did not analyze these five bundles. Furthermore, datasets were excluded for any of the following reasons: (1) They were recommended for exclusion by the ABCD fast track quality control protocol (160 datasets excluded); (2) they had missing or invalid scanner data (50 datasets excluded); (3) reconstruction failed for any of the remaining 62 bundles (300 datasets excluded); (4) they exhibited implausible tract geometry statistics (see below; 310 datasets flagged for exclusion; 77 non-overlapping datasets with previous exclusion criteria); or (5) they belonged to a harmonization batch containing fewer than 15 images (see below; 9 datasets excluded). Our final sample consisted of 24,178 datasets from 10,738 participants with complete imaging and phenotypic data.

To assess the anatomical plausibility of tractography generated by DSI Studio's AutoTrack, we conducted a series of assessments based on expected geometric properties of bundle reconstructions. We focused on four white matter bundles that are consistently identifiable and informative about reconstruction quality: the corpus callosum (CC), bilateral arcuate fasciculus (AF), bilateral corticospinal tract (CST), and bilateral inferior fronto-occipital fasciculus (IFOF). For each dataset, we first confirmed the presence of each bundle. Datasets in which any of these bundles were incomplete or had a total volume below 100 mm^3 were flagged, as such values are implausibly small. We then examined the following tract geometry metrics⁹⁶ that describe shape and consistency across hemispheres: Streamline curl which reflects how much a tract bends or twists; the length ratio which indexes the symmetry of bundle length between left and right hemispheres; and the volume ratio which compares the total spatial extent of left and right bundle homologs. Based on anatomically motivated priors, extensive review of individual-level data, and empirical distributions, we defined conservative heuristics to identify suspect data: (1) The CC should have more curl than the CST and IFOF, which are both relatively straight bundles; (2) We also would expect bilateral bundle homologs to be nearly symmetrical. Passing datasets exhibited AF length ratios < 2 , CST and IFOF length ratios < 1.25 , and all volume ratios < 5 . Datasets failing any of these criteria were marked for exclusion from all downstream analyses. In total, 310 datasets were flagged for exclusion by this process.

White matter feature harmonization

To correct for technical variation across scanning environments, we applied a harmonization procedure to the bundle-wise median values for each derived dMRI microstructural metric. We used longitudinal ComBat-GAMM, an extension of longComBat,⁴³ as implemented in the R package ComBatFamily (<https://github.com/Nhillman19/ComBatFamily> commit f8f9f3d, based off of <https://github.com/andy1764/ComBatFamily> version 0.2.1). We defined a harmonization batch as each unique scanner serial number and major software version combination with at least 15 images. This resulted in 60 unique batches (Table S3). During harmonization, we used generalized additive mixed models (GAMM) to remove the mean effect of batch while preserving the fixed effects of age and sex, including participant-specific random intercepts and age slopes to account for the longitudinal data structure. The age fixed effect was modeled as a penalized spline (maximum basis dimension $k = 4$). Harmonization for each microstructural metric was performed jointly across all bundles to take advantage of the empirical Bayesian estimation of average batch differences.

Quantifying the effect of batch on dMRI microstructure

We quantified the variance attributable to harmonization batch before and after harmonization. For each bundle and microstructural metric, we fit generalized additive mixed models (GAMMs) relating the median bundle value to age, sex, and acquisition batch, with person-specific random intercepts and age slopes to account for the longitudinal data structure. Age was modeled as a penalized spline (maximum basis dimension $k = 4$). The batch effect size for each bundle-metric combination was defined as the change in adjusted R^2 (ΔR^2_{adj}) between the GAM component of the full model and a reduced model excluding the batch term. This procedure was performed separately for unharmonized and harmonized data.

Developmental sensitivity of dMRI microstructure

We quantified developmental sensitivity of diffusion microstructural measures during childhood and adolescence. For each bundle and microstructural metric, we fit GAMMs relating the median bundle value to age while controlling for sex, with person-specific random intercepts and age slopes to account for the longitudinal data structure. Age was modeled as a penalized spline (maximum basis dimension $k = 4$). The age effect size for each bundle-metric combination was defined as the ΔR^2_{adj} between the GAM component of the full model and a reduced model excluding the age term. Additional models were also fit including individual image quality metrics (IQMs) as covariates to evaluate how controlling for image quality influenced estimates of developmental effects.

To complement the bundle-wise analyses, we performed a mass univariate voxel-wise analysis of developmental effects across all white matter voxels. All analyses were conducted using the ModelArray framework (version 0.1.5),¹⁰⁰ which enables efficient parallelized mass univariate modeling of large neuroimaging datasets. We analyzed the QSIRecon-derived microstructural maps in MNI152NLin2009cAsym standard space. As in the primary bundle-wise analyses, each microstructural map was harmonized using longitudinal ComBat-GAMM, preserving the fixed effects of age and sex while including participant-specific random intercepts and

age slopes to account for the longitudinal data structure. The age effect was modeled as a penalized spline (maximum basis dimension $k = 4$). Following harmonization, generalized additive mixed models (GAMMs) were fit independently at each voxel for each microstructural metric, relating voxel values to age while controlling for sex and including participant-specific random intercepts and age slopes. Developmental sensitivity at each voxel was quantified as the change in adjusted R^2 (ΔR^2_{adj}) between the full model and a reduced model excluding the age term.

Cross-vendor and cross-metric correspondence of developmental effects

We tested whether harmonization improved the spatial correspondence of developmental effects between scanner vendors. Within each scanner vendor, we fit GAMMs relating median bundle values to age and sex, with random effects for participants. For Philips scanners, too few participants (549 out of 1,259) had longitudinal data to reliably estimate random slopes, so models included only person-specific random intercepts. For Siemens and GE scanners, both random intercepts and age slopes were included. Age was modeled as a penalized spline (maximum basis dimension $k = 4$). For each vendor, microstructural metric, and bundle, we calculated the developmental age effect size as the ΔR^2_{adj} between the GAM component of the full model and a reduced model excluding the age term. For each metric and pair of vendors (Siemens–Philips, Siemens–GE, and GE–Philips), we then calculated Spearman rank correlations between the vendors' vectors of bundle-wise age effect sizes. This procedure was performed separately on unharmonized and harmonized data.

We also characterized the correspondence of developmental patterns across microstructural metrics. Using the age effect sizes estimated from the harmonized data pooled across vendors, we calculated Spearman rank correlations between each pair of metrics' vectors of bundle-wise age effect sizes.

Sensitivity of dMRI microstructure to image quality

We evaluated the susceptibility of derived diffusion microstructural metrics to variation in image quality. For each bundle and microstructural metric, we fit GAMMs relating the harmonized median bundle value to each image quality metric (IQM), while controlling for age, sex, and harmonization batch, with person-specific random intercepts and age slopes to account for the longitudinal data structure. Conditioning on harmonization batch ensured that quality-related variance was estimated independently of acquisition-related differences across scanners and software versions. Age and the IQM were modeled as penalized splines (maximum basis dimension $k = 4$). Our analysis space included 33 *QSIRecon*-derived microstructural metrics as dependent variables and 41 image quality metrics as predictors, comprising the 40 automated IQMs generated by QSIprep and the quality classifier score derived from manual ratings. For each model, we additionally fit a reduced model excluding the IQM term. The quality effect size for each bundle-metric combination was defined as the ΔR^2_{adj} between the GAM components of the full and reduced models.

For the subset of 2,829 manually rated images, we additionally calculated quality effect sizes for the mean manual rating and preprocessed dMRI contrast. Because the manually rated subset contained limited longitudinal observations, these analyses were performed using generalized additive models without random effects.

Variance decomposition of age–image quality relationships

To determine whether apparent age–IQM associations reflected biological developmental effects or instead arose from alignment between participant age and acquisition batch, we performed variance-decomposition analyses within a generalized additive modeling framework. For each IQM, we fit three models: (i) an age-only model ($IQM \sim s(\text{age})$), (ii) a batch-only model ($IQM \sim \text{batch}$), and (iii) a combined model ($IQM \sim s(\text{age}) + \text{batch}$). We quantified the proportion of age-related variance explained by batch by comparing the variance explained in the age-only model with the age-related variance retained after conditioning on batch ($R^2_{\text{full}} - R^2_{\text{batch}}$). Finally, to assess the downstream impact of quality adjustment on developmental inference, we calculated Spearman rank correlations between bundle-wise age effect sizes estimated from models with versus without added image quality covariates.