

Reproducible Brain Charts: An open data resource for mapping brain development and its associations with mental health

Highlights

- RBC facilitates robust, reproducible research in developmental neuroscience
- RBC provides harmonized phenotypes and uses uniform, reproducible processing pipelines
- RBC includes raw and fully processed imaging data as well as extensive QC measures
- All RBC data are openly shared without a data use agreement via INDI

Authors

Golia Shafiei, Nathalia B. Esper, Mauricio S. Hoffmann, ..., Giovanni A. Salum, Michael P. Milham, Theodore D. Satterthwaite

Correspondence

sattertt@pennmedicine.upenn.edu

In brief

Shafiei et al. introduce Reproducible Brain Charts (RBC), an open, reproducible developmental neuroimaging resource combining data from five large cohorts of youth. They provide harmonized psychiatric phenotypes, curated raw data, and uniformly processed derived imaging data. All data and processing pipelines are openly shared without a data use agreement.

NeuroResource

Reproducible Brain Charts: An open data resource for mapping brain development and its associations with mental health

Golia Shafiei,^{1,2,3,25} Nathalia B. Esper,^{4,25} Mauricio S. Hoffmann,^{5,6,7,8} Lei Ai,⁴ Andrew A. Chen,⁹ Jon Cluce,⁴ Sydney Covitz,¹⁰ Steven Giavasis,⁴ Connor Lane,⁴ Kahini Mehta,¹¹ Tyler M. Moore,^{2,3} Taylor Salo,^{1,2,3} Tinashe M. Tapera,¹² Monica E. Calkins,^{2,3} Stanley Colcombe,^{13,14} Christos Davatzikos,^{15,16}

(Author list continued on next page)

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

²Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

³Lifespan Brain Institute (LiBI) of Penn Medicine and CHOP, University of Pennsylvania, Philadelphia, PA, USA

⁴Child Mind Institute, New York, NY, USA

⁵Department of Neuropsychiatry, Universidade Federal de Santa Maria (UFSM), Santa Maria, Brazil

⁶Graduate Program in Psychiatry and Behavioral Sciences, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

⁷National Institute of Developmental Psychiatry & National Center for Innovation and Research in Mental Health, São Paulo, Brazil

⁸Care Policy and Evaluation Centre, London School of Economics and Political Science, London, UK

⁹Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC, USA

¹⁰Department of Bioengineering, Stanford University, Stanford, CA, USA

¹¹Department of Neuroscience, Columbia University, New York, NY, USA

¹²Khoury College of Computer Sciences, Northeastern University, Boston, MA, USA

¹³Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute, Orangeburg, NY, USA

¹⁴Department of Psychiatry, NYU Grossman School of Medicine, New York, NY, USA

¹⁵Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA

¹⁶Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, PA, USA

¹⁷Department of Psychiatry, Federal University of São Paulo (UNIFESP), São Paulo, Brazil

¹⁸Department of Psychology, University of Washington, Seattle, WA, USA

¹⁹eScience Institute, University of Washington, Seattle, WA, USA

(Affiliations continued on next page)

SUMMARY

Mental disorders are increasingly understood as disorders of brain development. Large and heterogeneous samples are required to define generalizable links between brain development and psychopathology. To this end, we introduce Reproducible Brain Charts (RBC), an open resource that integrates data from 5 large studies of brain development in youth from three continents ($N = 6,346$). Bifactor models were used to create harmonized psychiatric phenotypes, capturing major dimensions of psychopathology. Following rigorous quality assurance, neuroimaging data were carefully curated and processed using consistent pipelines in a reproducible manner. Initial analyses of RBC emphasize the benefit of careful quality assurance and data harmonization in delineating developmental effects and associations with psychopathology. Critically, all RBC data—including harmonized psychiatric phenotypes, unprocessed images, and fully processed imaging derivatives—are openly shared without a data use agreement via the International Neuroimaging Data-sharing Initiative. Together, RBC facilitates large-scale, reproducible, and generalizable research in developmental and psychiatric neuroscience.

INTRODUCTION

Mental disorders are increasingly understood as disorders of brain development.^{1–3} Neuroimaging studies of brain develop-

ment have the potential to track healthy brain maturation and identify deviations linked to psychopathology. However, large and diverse samples are necessary to detect reliable patterns of neurodevelopment and identify generalizable links to

Raquel E. Gur,^{2,3} Ruben C. Gur,^{2,3} Pedro M. Pan,¹⁷ Andrea P. Jackowski,¹⁷ Ariel Rokem,^{18,19} Luis A. Rohde,⁶ Russell T. Shinohara,^{16,20} Nim Tottenham,²¹ Xi-Nian Zuo,²² Matthew Cieslak,^{1,2,3} Alexandre R. Franco,^{4,13,14} Gregory Kiar,⁴ Giovanni A. Salum,^{4,6,7,23,24} Michael P. Milham,^{4,13,25} and Theodore D. Satterthwaite^{1,2,3,16,25,26,*}

²⁰Penn Statistics in Imaging and Visualization Center, Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA, USA

²¹Department of Psychology, Columbia University, New York, NY, USA

²²Developmental Population Neuroscience Research Center, IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

²³ADHD Outpatient Program & Developmental Psychiatry Program, Hospital de Clinicas de Porto Alegre, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil

²⁴Medical Council UNIFAJ & UNIMAX, São Paulo, Brazil

²⁵These authors contributed equally

²⁶Lead contact

*Correspondence: sattertt@pennmedicine.upenn.edu

<https://doi.org/10.1016/j.neuron.2025.08.026>

psychopathology.^{4–12} Multiple independent open science initiatives have facilitated this by sharing data publicly.^{7,8,13–19} While it is possible to aggregate data across independent studies,^{13,20–22} it is not necessarily a straightforward process due to variation in neuroimaging and psychiatric phenotyping protocols used.^{23,24} Obstacles in combining disparate data lead most investigators to use only a fraction of the data available.^{7,25,26} To address these challenges, we introduce the Reproducible Brain Charts (RBC) initiative: a large-scale, open data resource for the developing brain and psychiatry.

RBC addresses five major obstacles. First, there is abundant evidence that large, high-quality samples are essential for defining reliable brain-behavior associations.^{5–9,27–29} This is particularly challenging for studies of brain development, where large samples are also necessary to define a normative growth curve.³ To yield generalizable results, samples must not only be large but also diverse.^{5,8,12,19,30,31} Basic dimensions of diversity include age, sex, socioeconomic status, clinical diagnosis, race, and genetic ancestry.^{5,8} Accordingly, large-scale samples must encompass such variability to identify generalizable patterns of neurodevelopment and their links to mental health. In response, RBC has assembled a diverse dataset from five major developmental cohorts across three continents, spanning various recruitment strategies and thereby enriching both psychopathology and demographic diversity. This model serves as a foundational starting point for future expansions and the inclusion of data from additional studies and cohorts.

Second, combining psychiatric phenotypic data across large-scale studies presents multiple challenges.^{32–36} An initial challenge is obvious: different studies often employ disparate assessment tools to measure the same construct. Moreover, even when the same assessment is used, important psychometric properties may vary across populations. Aggregated data thus require careful harmonization of both measures and response properties. RBC addresses this by mapping differing tools to a common framework using a bifactor model.^{37–44} Bifactor models provide a robust solution by extracting a general factor (the “*p* factor”) that captures shared variance across symptoms and distinct domain-specific factors (e.g., internalizing or externalizing symptoms) that remain independent of the general factor. This approach effectively summarizes correlated psychiatric symptoms and diagnoses, aligning with dimensional

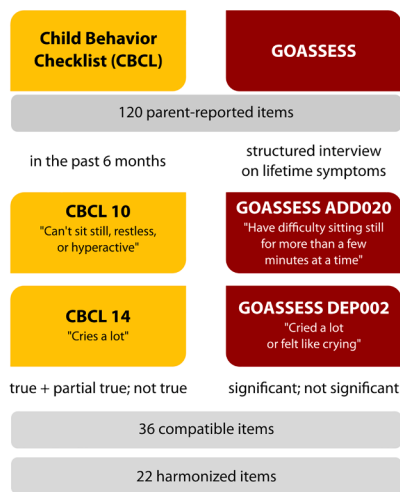
models like Hierarchical Taxonomy of Psychopathology (HiTOP)^{40,41} and Research Domain Criteria (RDoC).⁴⁵

A third major challenge is that both image acquisition parameters and image processing procedures vary considerably across large-scale studies. Even within multi-site studies with harmonized protocols, variation in scanners and protocol adherence introduces significant technical variability.^{25,30} Furthermore, many large-scale studies do not publicly release fully processed data. When they do, different studies utilize discrepant image processing pipelines, further complicating data aggregation and the reproducible integration of findings across studies.^{46–51} To address this, we processed all RBC data with standard tools like FreeSurfer⁵² and the Configurable Pipeline for Analysis of Connectomes (C-PAC⁵³). C-PAC’s highly configurable workflow allowed us to uniformly apply multiple versions of image processing across the entire data resource. Notably, all image processing steps were executed within the FAIRlybig framework,⁵⁴ maintaining a detailed audit trail via DataLad.⁵⁵ Fully documented and traceable data curation and processing enable researchers to rerun and adapt workflows for their analyses, facilitating harmonization and integration across datasets.⁵⁶ Following such reproducible image processing, we further reduced acquisition-related variation using statistical techniques adapted from computational genomics to harmonize imaging features.^{57–63}

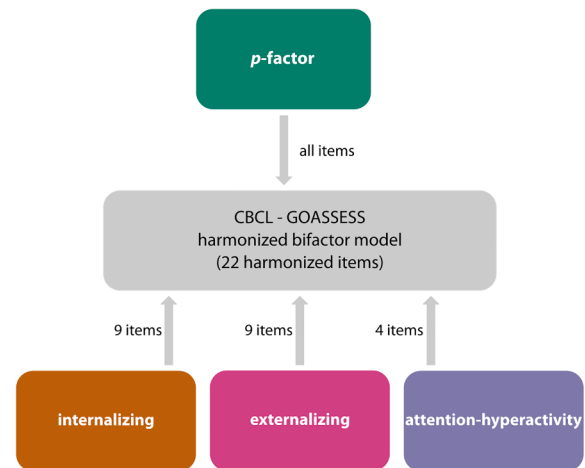
Fourth, variation in data quality remains a very important confound in neuroimaging research, with in-scanner motion profoundly impacting imaging features such as functional connectivity. This challenge is particularly acute for studies of brain development and psychiatry, as younger children and individuals with significant symptoms tend to move more during scanning sessions.^{20,64} Without meticulous quality control (QC), the substantial effects of data quality can easily overshadow the more subtle variations associated with brain development or psychopathology, leading to spurious associations that may be misinterpreted as biologically meaningful.^{23,64–68} To address this, RBC provides an extensive array of QC metrics alongside specific suggestions for exclusion criteria, bolstering the validity of the integrated data.

Fifth and finally, the mechanics of data access remain a major obstacle for investigators who hope to exploit large-scale data resources. Administratively cumbersome data use agreements

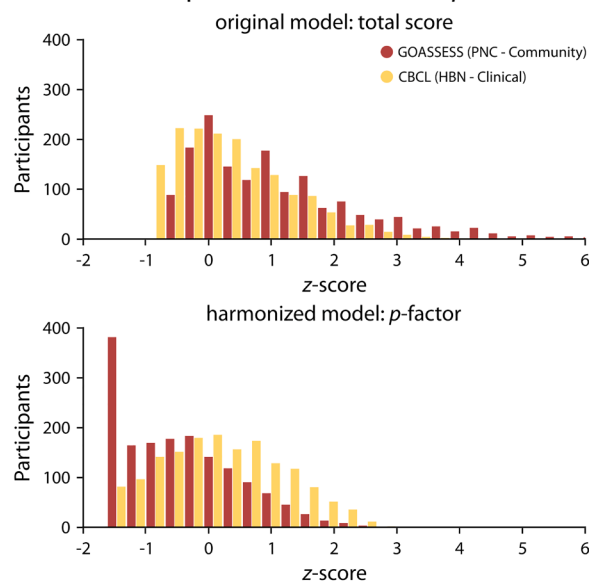
A Potential problem: different measures



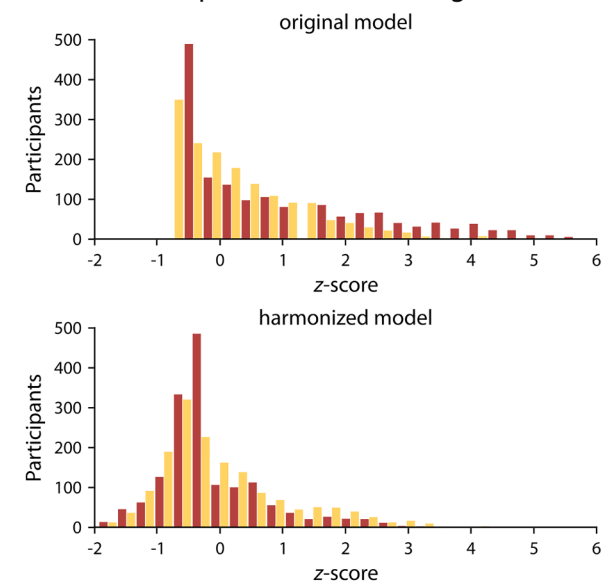
B Potential solution: harmonized bifactor model



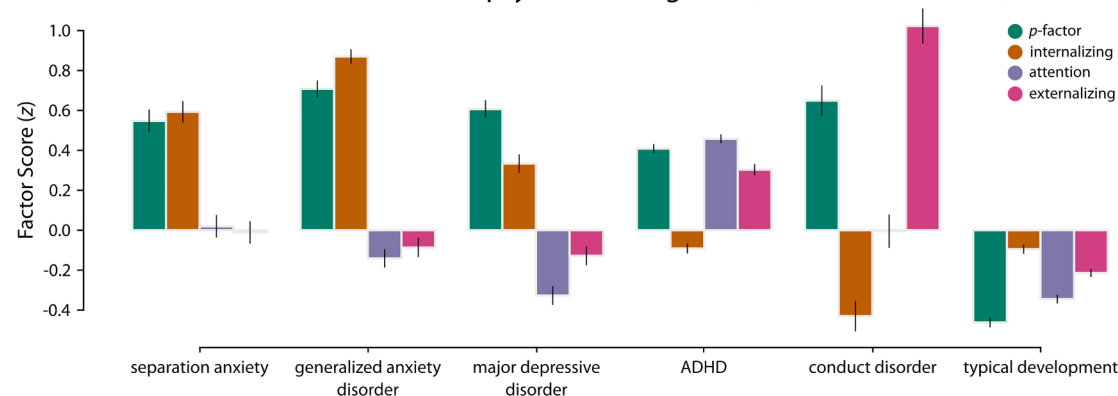
C Model comparison: total score & p-factor



D Model comparison: externalizing score



E Harmonized factor scores based on psychiatric diagnosis (combined RBC data)



(legend on next page)

(DUAs) often delay researchers and diminish the returns on public investments.^{22,27,30,69–72} In contrast, the fully de-identified data in RBC is released as a completely open resource. This open-access approach allows all harmonized psychiatric phenotypes and imaging data to be freely shared via the International Neuroimaging Data-sharing Initiative (INDI²⁴) without the need for a DUA, thereby accelerating research and maximizing the impact of public investments.

RESULTS

RBC aggregates five diverse neurodevelopmental datasets

RBC integrates demographics, psychiatric phenotyping, and both T1-weighted structural and resting-state and task functional Magnetic Resonance Imaging (MRI) data from five diverse and prominent developmental cohorts (see Table S1; total $N = 6,346$). Specific studies include the following: (1) Brazilian High Risk Cohort (BHRC⁷³); (2) Developmental Chinese Color Nest Project (CCNP^{74,75}); (3) Healthy Brain Network (HBN¹⁷); (4) Nathan Kline Institute-Rockland Sample (NKI¹⁸); (5) Philadelphia Neurodevelopmental Cohort (PNC^{16,76}). The vast majority of data included in RBC is from childhood, adolescence, and early adulthood (age range: 5–22 years old), with the exception of the NKI dataset that also includes participants from across the lifespan (up to 85 years old). All datasets have a relatively balanced sex distribution (about 45% female across all datasets). In addition, RBC provides information about participant race and ethnicity, handedness, body mass index (BMI), participant education, parental education, and psychopathology (see below).

RBC provides harmonized phenotypic data across studies

Selection of phenotypic instruments predated the RBC, and as such, reflects the design of each study. Specifically, the Child Behavior Checklist (CBCL) was used in BHRC, Chinese Color Nest Project (CCNP), Healthy Brain Network (HBN), and NKI while GOASSESS was used in Philadelphia Neurodevelopmental Cohort (PNC). The CBCL is a 120-item parent-report assessment of emotional and behavioral phenotypes.⁷⁷ PNC did not include the CBCL but rather assessed psychopathology using a highly structured psychiatric screening interview (GOASSESS).⁷⁸

In RBC, we sought to derive major dimensions of psychopathology that both harmonized differences between samples assessed with the same instrument (i.e., the CBCL) as well as harmonizing differences between instruments (i.e., the CBCL

vs. GOASSESS; see Figure 1A). To this end, we modeled psychopathology using a bifactor model.^{37–39} Specifically, bifactor models yield a factor that represents overall psychopathology (also known as the “ p factor”) and orthogonal factors for specific symptom domains. We first evaluated 12 bifactor models based on previous literature to identify the model that best harmonized phenotypic data, minimizing between-study differences in CBCL.⁴⁴ Furthermore, we extensively evaluated methods for harmonizing GOASSESS with CBCL using CFA.^{42,43} Ultimately, we developed a harmonized model that included a general psychopathology factor as well as specific factors for internalizing, externalizing, and attention symptoms⁴³ (Figure 1B; also see Tables S2–S3 for model fits and factor structure for harmonized CBCL–GOASSESS model). Participant scores for general psychopathology and domain-specific factors are publicly shared in RBC.

An illustrative example of the impact of harmonization on psychiatric phenotypes in RBC is apparent when comparing the overall psychopathology factor (p factor) and the externalizing factor in PNC and HBN (Figures 1C and 1D). As noted above, PNC and HBN used different instruments to assess psychiatric phenotypes (CBCL vs. GOASSESS). They also differed in their sample recruitment strategy: while PNC is a community-based sample, HBN primarily consists of help-seeking youth with significant mental health symptoms. As such, we expect higher levels of psychopathology in HBN than PNC. Notably, before harmonization, this difference was not apparent (Figure 1C; original psychopathology score: $t = -8.4e-15$, $p = 0.9$). Following harmonization, higher levels of psychopathology were evident in HBN as expected (Figure 1C; harmonized psychopathology score: $t = 22.9$, $p = 1.3e-110$). These harmonized factor scores aligned with clinical diagnostic categories defined by the DSM (Figure 1E). For example, conduct disorder was marked by elevated externalizing factor and p factor values, whereas major depression and generalized anxiety were marked by higher p factor and internalizing factor scores. In contrast, all general and specific factors were low in typically developing youth. Together, this harmonization process allows for direct pooling of phenotypic data across the diverse studies included in RBC.

Neuroimaging data are curated in a fully reproducible manner in RBC

We aggregated structural and functional neuroimaging data from multiple independently collected, large-scale data resources in RBC, while addressing challenges due to variations in imaging data acquisition (Tables S4–S5). Imaging metadata, described by the Brain Imaging Data Structure (BIDS⁷⁹), can be used to

Figure 1. RBC provides harmonized phenotypic data

(A) Phenotypic data were separately collected by studies included in RBC using different instruments (i.e., CBCL vs. GOASSESS). To harmonize phenotypic data, expert 1-to-1 semantic item-matching was used to identify 36 compatible items across studies, which were then reduced to 22 harmonized items.

(B) The McElroy bifactor model was used on harmonized items, yielding a general psychopathology factor (i.e., p factor) and orthogonal factors that represent specific domains (i.e., externalizing, internalizing, and attention).

(C and D) The impact of the RBC phenotypic data harmonization is demonstrated for PNC (GOASSESS questionnaire) and HBN (CBCL checklist), depicting original total scores (i.e., normalized sum scores based on full item sets) and harmonized factors for p factor (C) and externalizing factor (D). Expected cohort differences (e.g., higher levels of psychopathology in HBN) becomes more evident in harmonized data.

(E) Factor scores aligned with clinical diagnostic categories. For example, higher externalizing and p factors were observed in conduct disorder, while higher internalizing and p factors were observed in major depression and anxiety disorders. All general and specific factors were low in typical development.

automatically configure processing workflows (e.g., “BIDS-apps”⁷⁹). Such automatic configuration facilitates processing of large datasets that may have been collected using different data acquisition protocols. However, given that automatic configuration of processing workflows relies on metadata, any inaccuracies in metadata can lead to pipelines that are reproducible but wrong. This reliance makes careful curation of metadata essential as any issues at the curation step will influence all subsequent analysis. Yet, data curation is typically an ad-hoc process, often involving poorly-recorded manual intervention that is not reproducible, compromising the chain of reproducibility from the start.

To overcome these challenges, we converted all datasets to the BIDS format and curated the data using the reproducible workflows provided by the Curation of BIDS (CuBIDS⁸⁰) package. CuBIDS summarizes the heterogeneity in image acquisition and facilitates metadata curation. Moreover, CuBIDS uses DataLad⁵⁵ to ensure reproducibility throughout the curation process. Each study in RBC was curated with CuBIDS, yielding summary tables that describe and summarize heterogeneity in image acquisition (see [STAR Methods](#) for details). For example, PNC had 3 different CuBIDS parameter groups for structural images (i.e., T1-weighted MRI scans), separating T1 images to a main group with the majority of scans ($n = 1597$) and 2 variant groups with only a few scans each (e.g., $n = 3$ and $n = 1$). The CuBIDS summary files also indicate the source of the variance in image acquisition parameters. For example, the sources of the variance in PNC T1 images were obliquity for one variant group and slightly different echo and repetition times for the other variant group. Notably, functional MRI data were generally more heterogeneous, where the main source of the variance was the number of volumes acquired during fMRI scans. In all cases, using CuBIDS to curate the RBC data bolstered our confidence in the metadata used by BIDS-apps to configure image processing pipelines.

RBC provides fully processed structural and functional MRI data

Differences in data processing pipelines across studies present significant barriers to aggregating data from multiple resources. To support cross-study analyses, we used uniform processing pipelines and maintained a comprehensive audit trail in RBC. Specifically, following data curation, we used consistent image processing pipelines across studies to generate a standardized set of commonly used measures of brain structure and function.^{15,19,56,81,82} To ensure reproducibility and transparency, we adopted the “FAIRly-big” framework,⁵⁴ which enabled all data preparation and analyses to be accompanied by a detailed audit trail in DataLad.⁵⁵ This audit trail not only allows for tracking all steps in the processing pipeline but also provides a robust mechanism to rerun the pipelines, preserving methodological integrity. While the use of random numbers in imaging pipelines, such as FreeSurfer,⁵² sMRIprep,⁸³ and C-PAC,⁵³ can lead to slight variations in derived outputs (e.g., zip files with differing shasums), the pipeline commands are fully rerunnable, and their execution history is carefully recorded. This ensures that the entire process remains transparent and auditable, even if certain imaging derivatives are not strictly identical upon re-execution.

In RBC, we provide fully processed data including an extensive set of commonly used structural and functional data derivatives ([Figure 2](#)). Structural derivatives include surface area, cortical thickness, gray matter volume, folding and curvature indices, as well as summary whole-brain measures such as total intracranial volume. Functional data derivatives include preprocessed time series and functional connectivity matrices with multiple edge weights (e.g., pairwise Pearson and partial correlations between regional time series; [Figure 2B](#)). In addition, RBC provides regional measures such as Regional Homogeneity (ReHo),⁸⁴ Amplitude of Low Frequency Fluctuation (ALFF⁸⁵), and fractional ALFF (fALFF⁸⁶) ([Figure 2C](#)). Both structural and functional data are available in parcellated format using 16 anatomical, functional, and multimodal atlases (see [STAR Methods](#) for further details).

RBC data are accompanied by harmonized measures of QC

Data quality is one of the most important confounding factors in neuroimaging research. Image quality—driven mainly by in-scanner motion—affects both measures of brain structure^{68,87,88} and functional connectivity.^{23,64–66} Impact of data quality is even more pronounced in studies of development and psychopathology, where younger individuals and those with psychiatric symptoms tend to have higher in-scanner motion.^{20,64,89} Inconsistent quality control (QC) criteria complicate cross-study comparability, yielding different samples and discrepant results from the same data. Thus, harmonized QC metrics that can be used for sample construction and model covariates (e.g., continuous QC metrics) are required to account for the impact of data quality.

To ensure consistent sample selection and comparable cross-study analyses, we generated harmonized measures of neuroimaging data QC in RBC. These measures are accompanied by specific QC guidelines that allow for consistent quality assurance. Summary information on the number of participants with structural and functional scans before and after applying RBC’s recommended QC is available in [Figure S1](#). Overall, approximately 90% of RBC data had adequate structural and functional QC (range: 75% for HBN to 99% for CCNP). Study- and modality-specific data on how RBC’s recommended QC affects sample size are detailed in [Table S6](#). As expected, image quality varied by study, with lower quality in studies with younger participants and greater psychopathology (e.g., HBN).

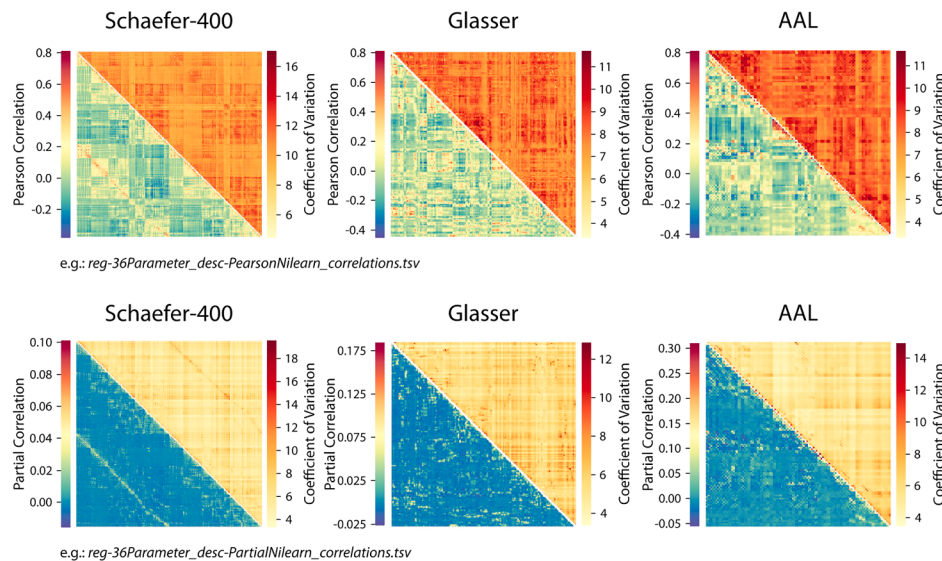
Structural MRI

RBC provides both harmonized QC information for sMRI based on expert manual ratings as well as automated, quantitative indices of image quality. Specifically, every structural image was manually evaluated by 2–5 expert raters in multiple phases using Swipes for Science, a web application for binary image classification⁹⁰ ([Figure 3A](#)). Expert raters assigned ratings of “Pass” or “Fail” to four slices per each image. Each image was assigned a score per rater by averaging across slices, and then an overall QC score of Pass, “Artifact,” or Fail was assigned to the image based on the average rating across raters. Specifically, Pass or Fail labels were assigned to a given image if all raters agreed about its quality (i.e., all raters assigned the

A Functional MRI processing workflow



B Functional connectivity matrices



C Regional measures

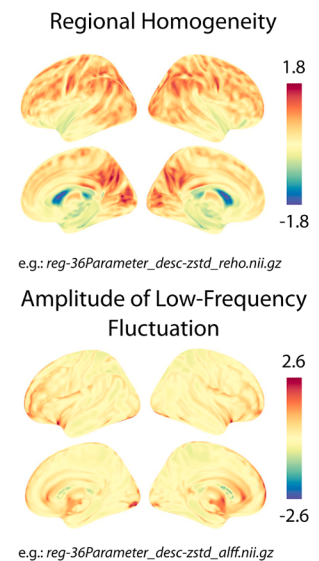


Figure 2. RBC provides processed neuroimaging derivatives

(A) RBC integrates structural and functional neuroimaging data from 5 neurodevelopmental datasets.

(B and C) Following careful data curation and consistent data preprocessing, neuroimaging data derivatives were estimated and publicly released in RBC. A few examples of group-average functional data derivatives are depicted: (B) functional connectivity matrices for 3 atlases with multiple edge weights (Pearson and partial correlations between processed regional time series) and their corresponding coefficients of variation (ratio between standard deviation and average absolute correlation value); (C) regional properties such as Regional Homogeneity (ReHo) and amplitude of low frequency fluctuation (ALFF).

same label of either Pass or Fail to the image). The intermediate artifact label was assigned to an image when there was disagreement between raters. The reliability of ratings was assessed across raters using a subset of images, using ratings for 1,896 participants (totaling 7,584 slice ratings). We found a high level of agreement among raters (71.5% agreement; average Cohen's kappa = 0.71; Figure 3B). The majority of structural scans across studies were labeled as Pass and were considered of adequate quality (82.6% on average across studies; Figure 3C). However, there was heterogeneity in data quality across datasets: HBN had the most artifact and fail scans compared with the other four studies, consistent with the fact that HBN includes younger help-seeking individuals with higher levels of psychopathology (with ADHD being the most frequent diagnosis). Furthermore, we found a positive trend between the average manual ratings and participant age (age range: 5–21 years old; Spearman rho = 0.28; Figure 3E), consistent with lower in-scanner motion and improved image quality in older individuals.

In addition to the manual expert ratings, we calculated the Euler number for each structural scan as part of the FreeSurfer processing pipeline. The Euler number reflects the topological

complexity of the reconstructed cortical surface: a lower Euler number indicates more defects in surface reconstructions (i.e., holes in the reconstructed cortical surface), and a ceiling of 0 represents the highest possible quality by this metric. Previous reports have demonstrated that the Euler number is an accurate, fully automated measure of data quality, capturing variation in data quality within the coarse categories provided by manual ratings.^{88,91–93} Consistent with prior work, we found that the Euler number was positively associated with average manual expert ratings in RBC (Spearman rho = 0.63; Figure 3D). We recommend including Euler number as a covariate in secondary analysis of RBC data in addition to excluding data that do not meet RBC's data quality criteria based on the categorical QC labels.

Functional MRI

fMRI derivatives in RBC are accompanied by extensive measures of QC, including multiple indices of both in-scanner motion and image registration quality. To facilitate consistent QC across datasets, we generated a summary functional QC score. Specifically, fMRI runs with low in-scanner motion (i.e., median framewise displacement [FD] ≤ 0.2) and high image normalization quality (i.e., normalized cross correlation ≥ 0.8) were considered of adequate quality.

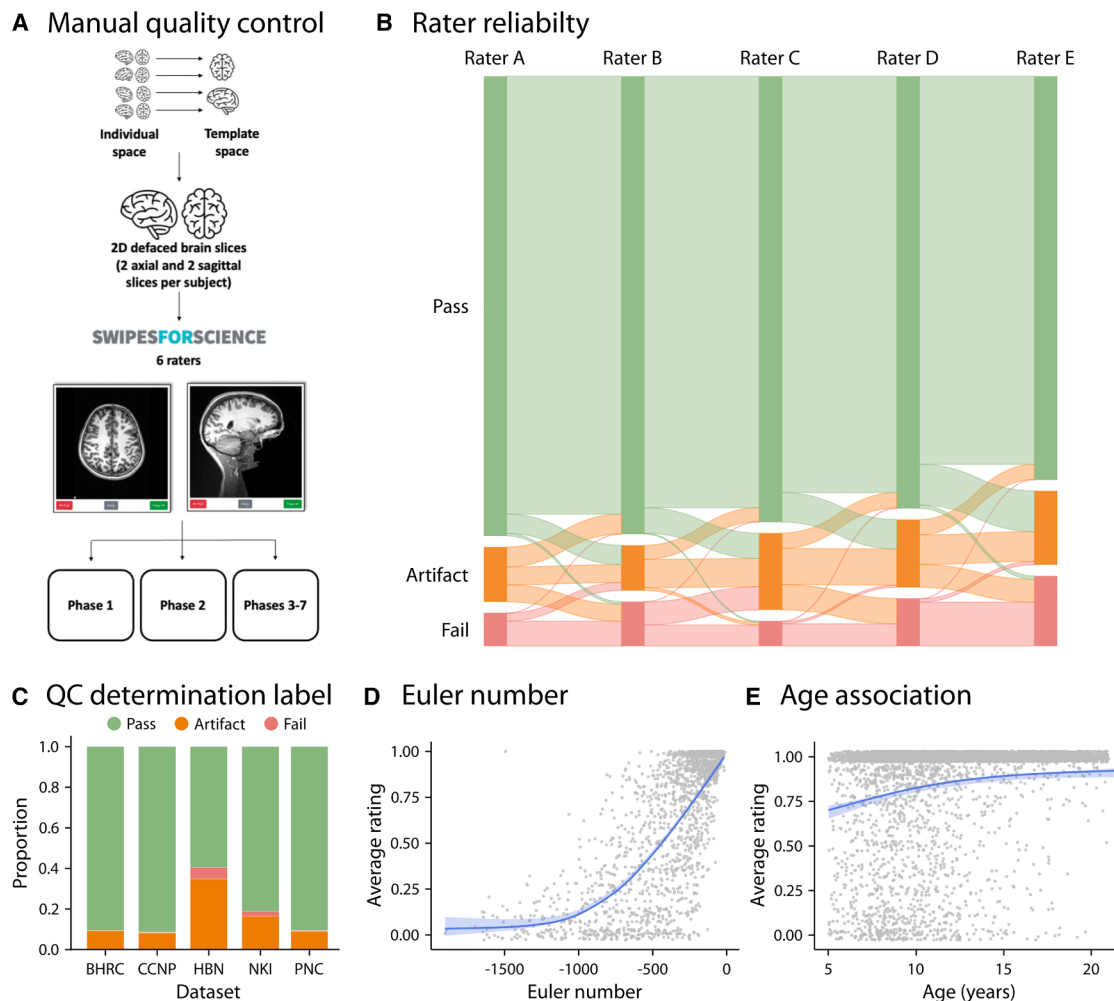


Figure 3. RBC data are accompanied by harmonized structural QC metrics

Harmonized structural QC metrics were estimated based on expert manual ratings and automated methods.

(A) Each structural image was manually evaluated by 2–5 expert raters in 7 phases using Swipes for Science. Expert raters assigned pass or fail labels to 4 two-dimensional slices per image. An overall QC score of pass, artifact, or fail was then assigned to each image: scans were labeled pass or fail if all raters were in agreement about their quality while artifact was assigned when there was any disagreement.

(B) Comparing final QC labels to expert ratings demonstrated a high level of agreement among raters (71.5% agreement; average pairwise Cohen's kappa = 0.71).

(C) Proportion of data with pass, artifact, and fail labels are depicted for each study. HBN had the most artifact and fail scans compared with the other four studies.

(D) A continuous automated measure of data quality, Euler number, was estimated as part of the FreeSurfer processing pipeline. As expected, Euler number was positively associated with the average expert ratings (Spearman rho = 0.63).

(E) Average expert ratings were positively associated with age, indicating higher in-scanner motion, hence, lower image quality in younger individuals (Spearman rho = 0.28).

Using this approach, a relatively small proportion of participants per study were excluded due to poor data quality (9.7% excluded across studies). As expected, due to younger participants and higher prevalence of psychopathology, HBN had the largest number of individuals with increased in-scanner motion (Figures 4A and 4B). In-scanner motion decreased with age (Spearman rho = −0.22; Figure 4C). Furthermore, median FD was positively associated with higher *p* factor (Spearman rho = 0.15; Figure 4D), suggesting that individuals with higher overall psychopathology tend to have higher in-scanner motion. Finally, median FD during fMRI was negatively related to manual ratings

of structural image quality (Spearman rho = −0.45; Figure 4E), indicating that individuals with higher in-scanner motion during functional scans also had lower structural scan quality. Taken together, these extensive structural and functional QC metrics allow for consistent quality assurance of imaging data in RBC. We strongly encourage researchers to use our recommended quality assurance ratings when working with RBC dataset. However, we also provide all data—even the scans that do not pass QC—to help accelerate progress on image QC research (see STAR Methods for details on different versions of publicly released RBC data).

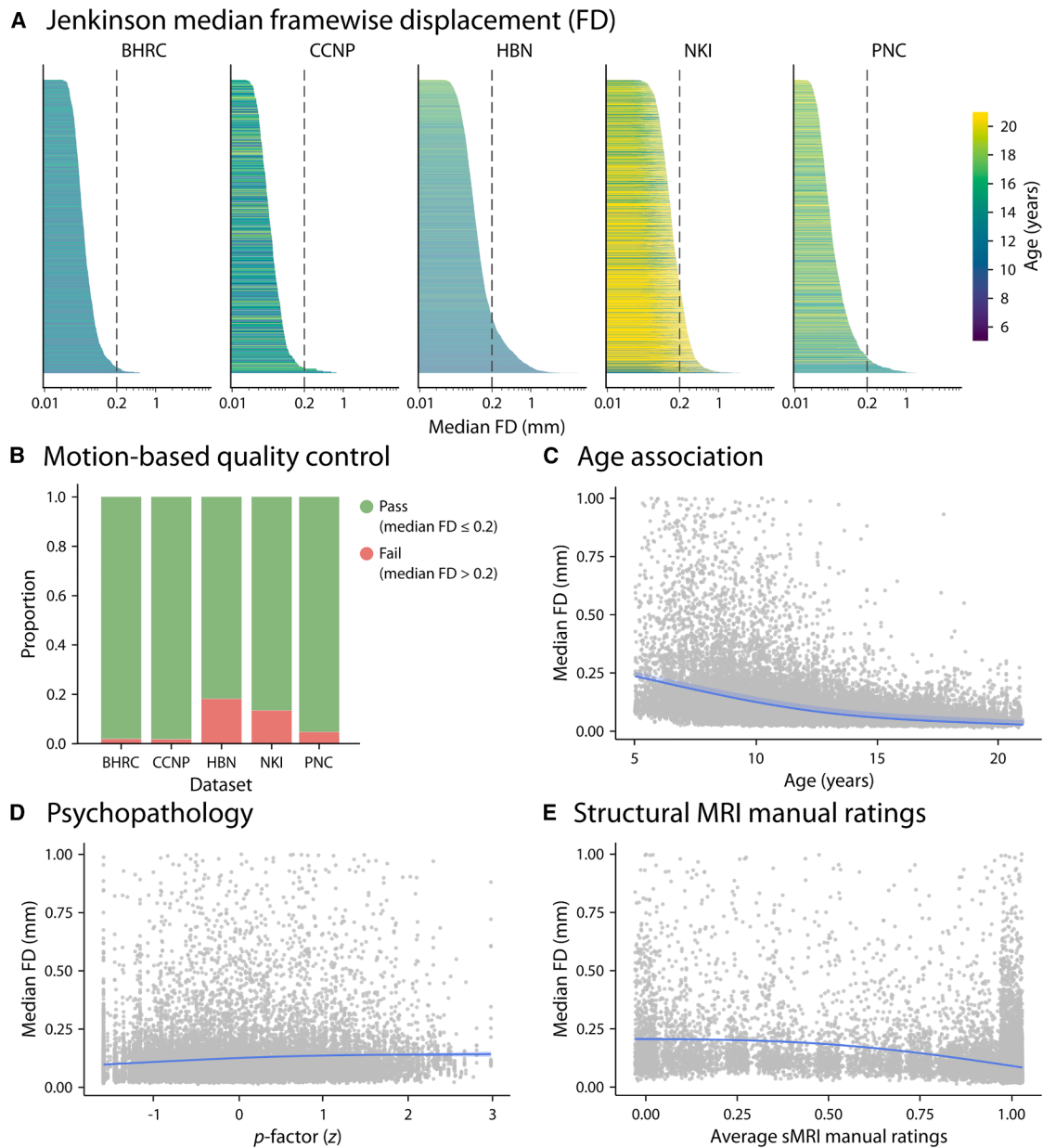


Figure 4. RBC data are accompanied by harmonized functional QC metrics

An extensive list of harmonized QC metrics were generated for functional data, including summary measures of in-scanner motion, image registration, and normalization quality indices.

(A) Median Framewise Displacement (FD) is depicted as an example functional QC metric (bar plots colored by participant age). Dashed line indicates RBC's recommended threshold for median FD, demonstrating that only a small proportion of individuals had high in-scanner motion.

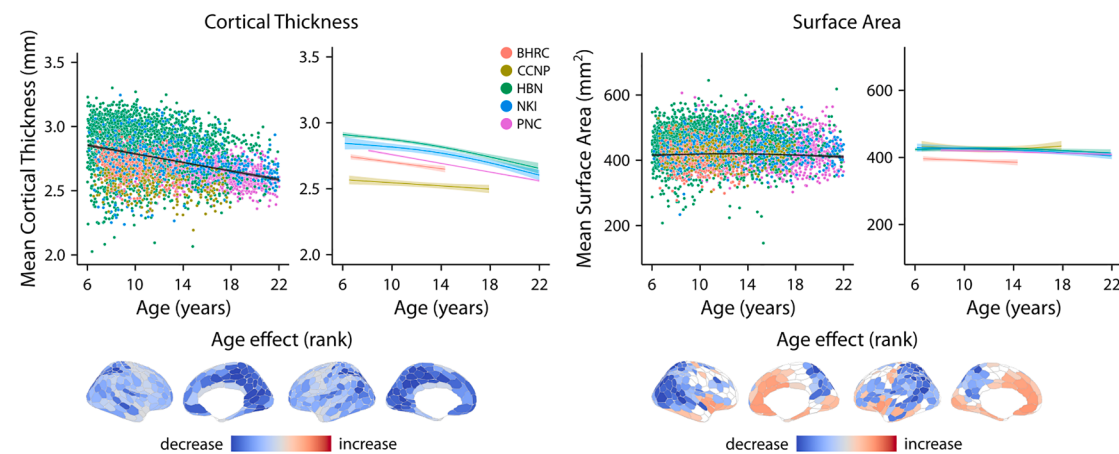
(B) As demonstrated in (A), the majority of RBC functional data pass the in-scanner motion threshold of 0.2. Consistent with structural QC (Figure 3), HBN had the largest number of individuals with high in-scanner motion.

(C) Median FD was negatively associated with participant age, suggesting higher motion in younger individuals (Spearman $\rho = -0.22$).

(D) In-scanner motion displayed a positive trend with p factor, suggesting that individuals with higher overall psychopathology tend to have higher in-scanner motion (Spearman $\rho = 0.15$).

(E) Median FD in functional data was negatively associated with average manual ratings in structural data, indicating that individuals with higher in-scanner motion during functional scans also had lower structural scan quality (Spearman $\rho = -0.45$).

A Combined RBC data



B Combined RBC data with QC and harmonization

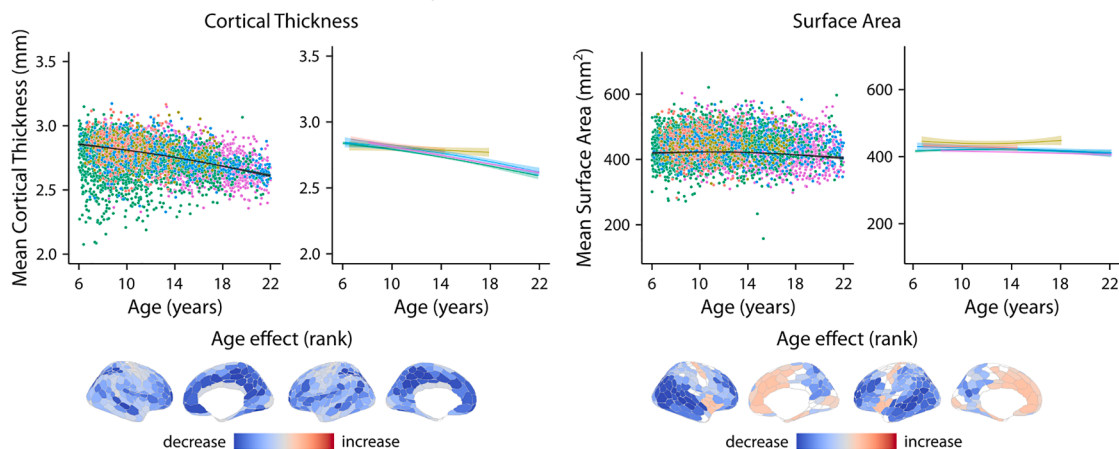


Figure 5. Structural data derivatives are associated with age in youth

Generalized additive models (GAMs) were used to examine the relationship between participant age and structural data derivatives, including cortical thickness (CT) and surface area (SA).

(A) We found an overall decrease in mean CT during development in aggregated RBC data without QC or harmonization. However, developmental effects varied markedly between studies as demonstrated by study-specific model fits (color-coded curves). We also examined regional age effects using separate models for each brain region. Age effects were quantified using ranked partial R^2 and are depicted on the cortical surface after correcting for multiple comparisons (FDR corrected).

(B) To assess QC and harmonization effects, we repeated the analysis after excluding data with Fail structural QC and harmonizing neuroimaging data using CovBat-GAM. Following QC and harmonization, the developmental effects on mean CT and SA became more similar between studies as evident by study-specific fits. Similar to before, results demonstrated an overall decrease in mean CT with development. Regional analyses identified heterogeneous developmental effects on the cortex.

Neuroimaging features are associated with age and psychopathology in youth

To illustrate the utility of RBC, we examined how neuroimaging features were related to participant age and overall psychopathology. To ensure sensitivity to both linear and non-linear relationships, we used Generalized Additive Models (GAMs), while controlling for covariates such as sex and data quality.^{60,94,95}

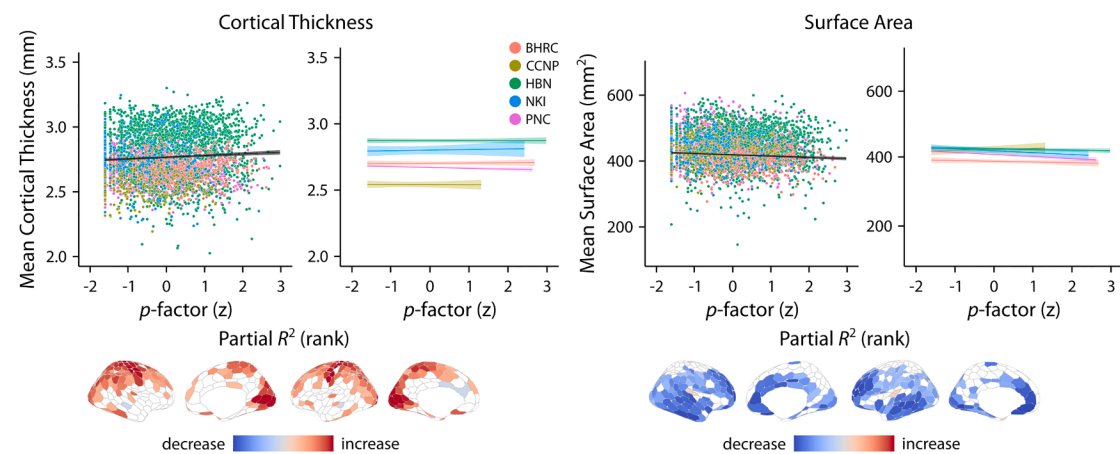
Prior to statistical analyses, imaging measures were harmonized using CovBat-GAM while protecting the effects of model covariates such as age (as a smooth term), sex, data quality, and the p factor (as linear terms)^{57–61} (see Figure S2 for site variation before and after harmonization). Sample sizes varied

($N = 3,847$ to $4,827$) depending on data modality (i.e., structural vs. functional features), analysis type (i.e., age vs. psychopathology analysis), and whether QC was implemented (Table S7).

Structural neuroimaging features are associated with development and psychopathology

We first assessed developmental associations with brain structure using structural MRI derivatives, including cortical thickness (CT), surface area (SA), and gray matter volume (GMV). The data were parcellated into 400 regions using the Schaefer-400 atlas.⁹⁶ As a baseline, we evaluated age effects in combined RBC data without QC implementation or imaging data harmonization (Figure 5A). We found that developmental effects varied

A Combined RBC data



B Combined RBC data with QC and harmonization

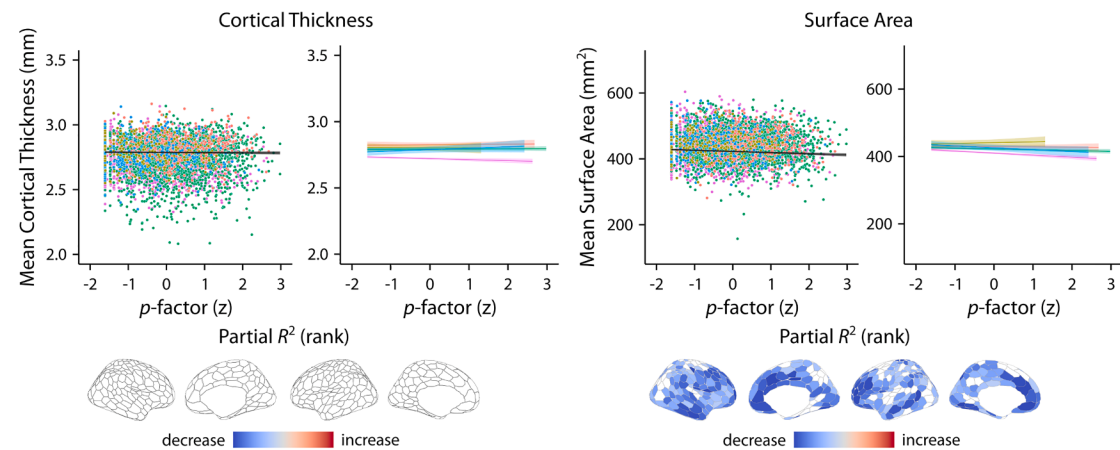


Figure 6. Structural data derivatives are associated with psychopathology in youth

(A) GAM analyses using aggregated RBC data without QC or harmonization identified an overall increase in mean CT with increased p factor and decreasing trend in mean SA with increased p factor. Similar to age-effects (Figure 5), whole-brain associations between brain structure and psychopathology varied between studies. Regional analyses identified significant associations between structural features and p factor in a subset of brain regions.

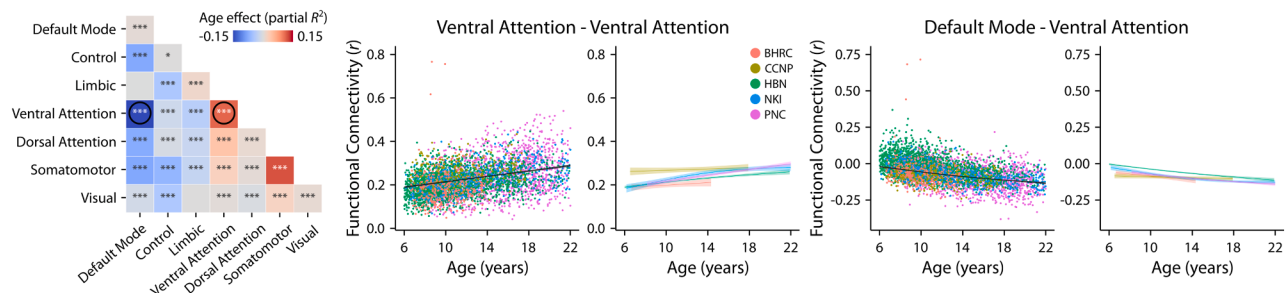
(B) Following QC and harmonization, whole-brain association converged between studies, demonstrating no relationship between mean CT and p factor and a decreasing trend in mean SA with increased p factor. Regional analyses identified no significant associations between regional CT and psychopathology and significant decreasing patterns in SA with increased psychopathology. Note that CT was significantly associated with p factor in a subset of regions before QC and harmonization. This suggests that including low-quality data or combining multi-site datasets without data harmonization may result in spurious associations.

markedly between studies. For example, mean CT displayed a sharp decline with age in BHRC and PNC whereas these effects were more gradual in CCNP, HBN, and NKI. However, following QC and harmonization, the developmental patterns converged across studies and sites (Figure 5B; see Figure S3A for developmental patterns across sites). Mean CT declined markedly with age; GMV and SA showed more gradual changes (CT and SA in Figure 5 and GMV in Figure S4). Regional analyses identified heterogeneous developmental effects in CT and SA across the cortex. CT decreased with age in most cortical regions, with the most prominent decrease observed in medial parietal, lateral and medial prefrontal, and temporal regions. SA displayed a combination of increases and decreases with age, such that medial frontal and parts of visual and motor cortices displayed an increase in SA whereas lateral parietal, temporal and prefrontal

cortices displayed a decrease in SA during development (Figure 5B). Regional variations were consistent with the original analyses after controlling for global effects by including mean CT, SA, or GMV in addition to other covariates (i.e., sex and Euler number; Figure S5). In addition, we repeated the analyses after implementing QC but before harmonizing imaging data (Figure S6A). We found that although QC removed data with low quality, data harmonization was required to effectively account for site and study differences.

We next examined the relationship between brain structure and psychopathology (p factor) using the aggregated data (Figure 6). Similar to age-related findings, whole-brain associations between structural features and p factor converged across studies and sites following QC and data harmonization (Figure 6B; see Figure S3B for results across sites). Although regional analyses identified

A Combined RBC data



B Combined RBC data with QC and harmonization

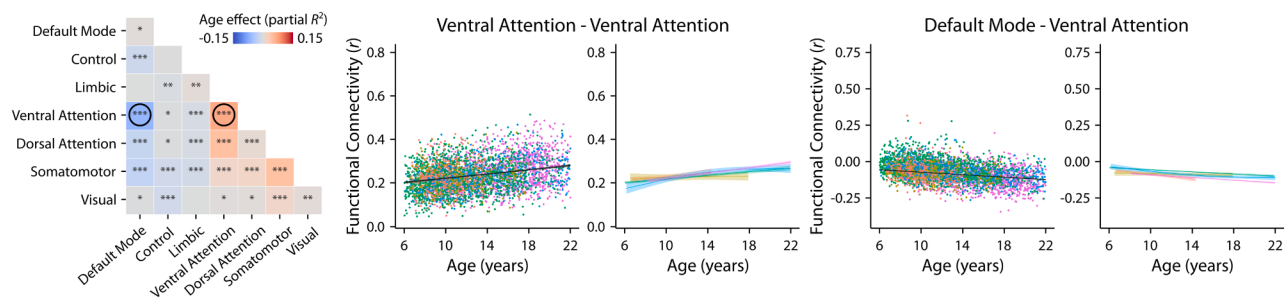


Figure 7. Functional data derivatives are associated with age in youth

(A) Aggregated data without QC or harmonization were used to model age-related variations in network-level functional connectivity.

Results are depicted in a matrix (left), where the diagonal values correspond to within-network age-effects (partial R^2) and off-diagonal values correspond to between-network age effects. Asterisks indicate statistical significance after correcting for multiple comparisons (FDR-corrected p values: * indicates $0.001 < p < 0.05$; ** indicates $0.0001 < p < 0.001$; *** indicates $p < 0.0001$). Overall, we found an increase in within-network connectivity with development while between-network age-effects were more heterogeneous. Example results (marked via circles in the matrix) are shown for within-network connectivity in the ventral attention network and between-network connectivity between the default mode and ventral attention networks. Study-specific model fits varied between studies, especially for within-network connectivity.

(B) We repeated the analyses following QC and harmonization. Within- and between-network age effects displayed similar associations as before (with variations in effect size and significance). However, study-specific model fits converged and displayed consistent patterns across studies.

significant association with p factor in both CT and SA in combined RBC data without QC and harmonization (Figure 6A), only SA was significantly associated with p factor following QC and harmonization (Figure 6B). The results demonstrated that SA decreased with overall psychopathology, with the most prominent decrease in medial and lateral prefrontal cortices. We also repeated the analyses after implementing QC but before harmonizing imaging data (Figure S6B). This analysis suggested the presence of significant associations between CT and p factor without data harmonization. However, these effects largely disappeared after harmonization with CovBat-GAM. Together, these findings underscore the importance of QC and harmonization when combining heterogeneous data. Importantly, including low-quality data in the analysis or combining multi-site datasets without data harmonization may result in spurious associations between psychopathology and brain structure.

Functional neuroimaging features are associated with development and psychopathology

To assess the developmental variations in brain function, we investigated how within- and between-network functional connectivity were associated with age (Figure 7) and overall psychopathology (Figure 8). We evaluated functional connectivity matrices with 400 cortical regions,⁹⁶ each assigned to one of the canonical 7

Yeo-Krienen networks.⁹⁷ Similar to the findings for structural data, age-related variations in network-level connectivity converged to similar patterns across studies and sites following structural and functional QC and data harmonization (Figure 7B compared with Figure 7A; see Figure S3C for results across sites). Following QC and harmonization, we observed an overall increase in within-network connectivity during development for all resting-state networks. However, the amount of age-related increase in within-network connectivity varied between networks with the ventral attention network demonstrating the largest increase in connectivity during development (Figure 7B; Partial $R^2 = 0.067$, $p_{FDR} < 0.0001$). In contrast, between-network functional connectivity decreased with age in most networks. However, these age effects were heterogeneous (Figure 7B). For example, connectivity between ventral attention and default mode networks significantly decreased with age (Figure 7B; Partial $R^2 = -0.065$, $p_{FDR} < 0.0001$), while connectivity between ventral attention and dorsal attention networks increased with development (Figure 7B; Partial $R^2 = 0.046$, $p_{FDR} < 0.0001$).

Finally, we assessed the relationship between p factor and network-level functional connectivity (Figure 8). Without QC or harmonization, we found significant associations between p factor and within- and between-network functional connectivity in

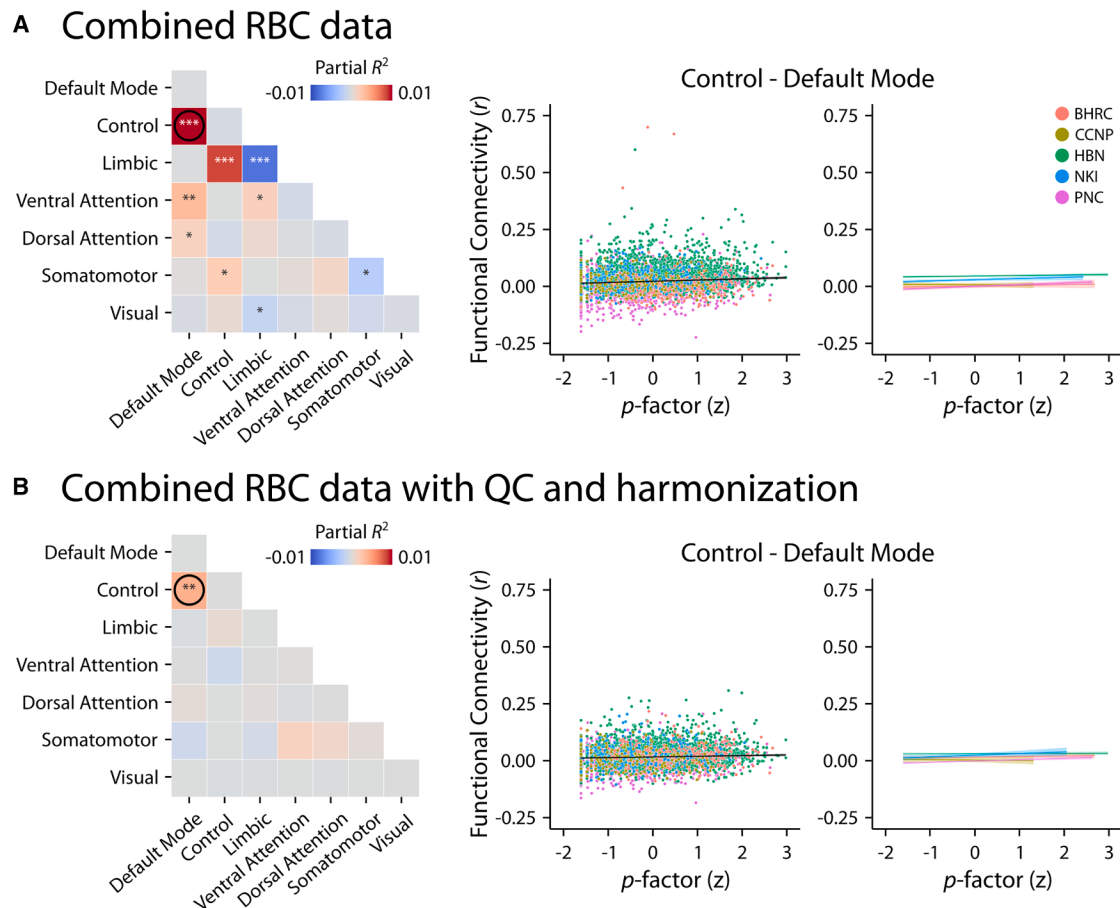


Figure 8. Functional data derivatives are associated with psychopathology in youth

(A) GAM analysis with aggregated data without QC or harmonization identified significant associations between network-level connectivity and p factor in multiple networks.

Results are depicted in a matrix (left), where the diagonal values correspond to within-network age-effects (partial R^2) and off-diagonal values correspond to between-network age-effects. Asterisks indicate statistical significance of the findings after FDR correction (corrected p values: * indicates $0.001 < p < 0.05$; ** indicates $0.0001 < p < 0.001$; *** indicates $p < 0.0001$). An example relationship (marked via a circle in the matrix) between psychopathology and between-network connectivity is shown for the default mode and frontoparietal control networks.

(B) Following QC and harmonization, we only identified a significant positive association between p factor and between-network connectivity in the default mode and frontoparietal control networks.

multiple networks (Figure 8A). However, following QC and harmonization, we observed significantly increased connectivity only between default mode and frontoparietal control networks with increasing psychopathology (Figure 8B; Partial $R^2 = 0.004$, $p_{FDR} = 0.0002$; see Figure S3D for results across sites). As for sMRI, we repeated the analyses after implementing QC but before harmonizing imaging data (Figure S6C for age effects and Figure S6D for psychopathology). Together, these analyses showcase how RBC data can be used and highlight the importance of rigorous QC and data harmonization.

Evidence for sex differences in brain development

Lastly, we examined the effects of biological sex. Consistent with previous literature,^{98,99} we found widespread significant sex differences in structural features, even after controlling for global effects (Figures S7A and S7B). As in previous reports,^{100–102} we found significant sex differences in network-level functional connectivity

(Figure S7C). Specifically, we found that females had greater within-network connectivity in the default mode network (Partial $R^2 = 0.005$, $p_{FDR} < 0.0001$). In contrast, we found that males had greater between-network connectivity between the default mode network and the ventral attention (Partial $R^2 = 0.009$, $p_{FDR} < 0.0001$), dorsal attention (Partial $R^2 = 0.006$, $p_{FDR} < 0.0001$), visual (Partial $R^2 = 0.004$, $p_{FDR} < 0.0001$), and somatomotor (Partial $R^2 = 0.002$, $p_{FDR} = 0.01$) networks. These findings indicate a less segregated default mode network in males. Interactions between biological sex and age are summarized in Figures S8.

DISCUSSION

Developmental and psychiatric neuroimaging research often faces significant obstacles: limited sample sizes, variability in

image acquisition and psychiatric phenotyping, and inconsistent data analysis workflows.^{5,7,9,25,26,30} These collectively hinder the generalizability of results, and act as a brake on scientific progress. RBC addresses these challenges by integrating and harmonizing data from over 6,000 participants across five major neurodevelopmental cohorts. We employed advanced harmonization techniques to overcome variability in psychiatric phenotyping and imaging protocols, ensuring that data from different studies can be meaningfully combined with confidence. Furthermore, neuroimaging data with uniform processing are complemented by standardized QC measures to ensure rigor. Initial results defining consistent patterns of brain development underscore the degree to which RBC facilitates robust and generalizable studies of the developing brain.

RBC is a response to the proliferation of large-scale studies of brain development.^{13–18,74,75} RBC builds on previous data aggregation and harmonization efforts such as the International Neuroimaging Data-sharing Initiative (INDI), the Autism Brain Imaging Data Exchange (ABIDE) Preprocessed, and the ADHD-200 Preprocessed,^{13,22,56,103,104} while distinguishing itself through a strong emphasis on harmonization, reproducibility, and data quality. However, this emphasis posed unique and sometimes unforeseen challenges; RBC reflects 6 years of sustained methodological efforts. Next, we detail what was required as well as lessons learned across several key domains.

Concise, harmonized psychiatric phenotyping

Harmonizing psychiatric phenotypes remains a major challenge for the field; RBC was no exception. The harmonized dimensions of psychopathology released with RBC were only possible after a series of detailed methodological studies, which evaluated a broad range of bifactor models to harmonize constructs.^{42–44} The resulting factor scores provide a parsimonious summary of mental health data, disambiguating general and specific dimensions of psychopathology.^{37–39} This approach aligns with the dimensional and hierarchical characterization of psychopathology emphasized in HiTOP.^{40,41} It should be noted that recent studies have raised important concerns regarding the alignment of bifactor models of psychopathology with established theoretical frameworks.¹⁰⁵ Despite potential shortcomings, the validity of the general factor of psychopathology (i.e., *p* factor) derived from these models is supported by extensive evidence.^{44,106–110} Importantly, bifactor models also offer pragmatic advantages to researchers, such as parsing inter-item covariance into orthogonal scores (i.e., latent dimensions) that can be used simultaneously in hypothesis testing. When dealing with the inevitable comorbidity of psychopathology symptoms, extracting a general factor to explain them is often the most practical solution. Here, we exploited bifactor modeling's strength in facilitating cross-study comparisons despite the different sampling strategies of the component studies.⁴²

Reproducible image curation

The BIDS format has been a boon for the field.^{26,79,111} However, many studies—like several included in RBC—have not previously been released in BIDS. Furthermore, BIDS metadata may be inaccurate or missing. This is a critical and under-recognized problem as widely-used image processing pipelines (e.g.,

BIDS-apps) automatically configure workflows based on imaging metadata—leading to processed data that may be inaccurate (e.g., reproducible but wrong). Typically, metadata is corrected manually in an ad-hoc fashion, compromising reproducibility before image processing even begins. This was an unanticipated challenge in RBC—large datasets that required significant metadata curation. To address this, we created CuBIDS,⁸⁰ which allows for reproducible BIDS curation. Notably, the detailed summary of metadata provided by CuBIDS also revealed more significant protocol variation in each of the component studies than originally anticipated. Such protocol variation is usually unacknowledged; studies typically report only the intended protocol. Moving forward, the practical consequences of such variation merit evaluation and likely suggest the need for multi-level harmonization methods that can address variation in image acquisition both within and across studies.¹¹²

Uniform image processing and reproducible workflows

In RBC, we followed the example of prior data aggregations—such as the ABIDE and ADHD-200 Preprocessed^{56,104}—and released fully processed data. We started by leveraging containerized, open-source pipelines—such as FreeSurfer,⁵² sMRI-Prep,⁸³ and C-PAC⁵³—to ensure uniform processing of structural and functional MRI data. The need for such uniform processing was underscored after our initial benchmarking study revealed that even seemingly innocuous analytic choices (e.g., template version) can introduce significant variability.⁵¹ Both that study and prior work revealed that the use of global signal regression (GSR) has a substantial impact on derived features; the impact of GSR on findings remains one of the most common and time-consuming questions addressed in peer review.^{113–119} We took advantage of the exceptional configurability of C-PAC to execute two high-performance denoising pipelines both with and without GSR.

The use of containerized processing pipelines is a major asset for reproducibility but does not on its own allow for full audit trail. In RBC, we adopted the “FAIRly-big” workflow⁵⁴: a framework that uses DataLad⁵⁵ to ensure reproducible processing of large datasets. This framework aligns with the FAIR principles (findability, accessibility, interoperability, and reusability)¹²⁰ by enabling open, modular, and reusable workflows. However, while we successfully applied the FAIRly-big workflow in RBC, it was not straightforward: it had a steep learning curve despite our team's significant technical expertise. This experience led our team to develop dedicated software—the BIDS App Bootstrap (BABS¹²¹)—that automates the application of FAIRly-big. While we did not use BABS software for this initial release of RBC, moving forward we anticipate that it will significantly lower the barriers to researchers adopting FAIRly-big and DataLad in their own work.

QC

One critical yet often-overlooked factor in cross-study reproducibility is the role of QC and QC-based sample selection. Even when identical processing pipelines and QC metrics are used, differences in inclusion criteria can result in different samples included in data analysis, potentially leading to divergent findings from the same dataset. To address these challenges, we

provide extensive harmonized QC metrics and recommended QC guidelines for using RBC data. These include both categorical QC measures that can be used for sample selection as well as continuous measures of data quality that can be used as model covariates.

While RBC provides harmonized QC metrics and standardized guidelines, our analyses highlight significant remaining challenges. For example, we observed that participants who are younger or had higher levels of psychopathology tended to have poorer data quality. This raises important concerns about how QC procedures may inadvertently reduce generalizability to populations of the most interest. Ultimately, while our QC efforts confirm many findings from prior studies,^{17,18} they also underscore the trade-offs inherent in balancing data quality with inclusivity.

Convergent developmental findings and ongoing challenges in translational psychiatry

As an initial evaluation of the utility of RBC, we examined the associations of structural and functional imaging features with age. Aligning with a large body of prior work,^{3,122–127} we found evidence for a decline in CT, SA, and GMV in all datasets following careful QC and harmonization. Consistent with a rich literature from lifespan network neuroscience,^{95,128–140} we also found evidence for a decline in between-network functional connectivity paired with increases in within-network connectivity. These findings suggest that resting-state functional networks become more segregated and specialized during development.^{95,137,140–143} Notably, these results were far less consistent prior to harmonization and QC, emphasizing how variation in data quality and acquisition parameters may obscure even robust effects of brain development.

In contrast to the highly consistent developmental findings, the existing literature on associations between major dimensions of psychopathology and imaging features have been more varied.^{144–152} These inconsistencies may be due to small sample sizes, limited reliability, and biological heterogeneity.^{9,28,153–156} Echoing at least part of the existing literature,^{144,145,147,148,152,157} we found that reduced SA and GMV were linked to higher overall psychopathology. Additionally, we found that higher overall psychopathology was associated with greater connectivity between the default mode and frontoparietal control networks, suggesting a loss of normative developmental network segregation.^{128,139,152,158–162} These results are not surprising in that they are consistent with multiple published reports—this consistency and the large sample used bolsters confidence in our results. However, in contrast to the larger effects of brain development, it should be emphasized that the effect sizes of associations with psychopathology were small. Such small effect sizes remain a major challenge for the field, highlighting the need for more sensitive imaging measures, better methods for characterizing psychopathology, and approaches for parsing heterogeneity in links between brain and behavior.^{28,147,155,156,163–166}

Notably, our analyses of brain development and links to psychopathology also highlighted the importance of QC and data harmonization. For example, our initial analysis of combined RBC data without QC or neuroimaging data harmonization identified significant associations between CT and psychopathology

in multiple cortical regions. However, those associations were no longer significant following QC and data harmonization. We observed similar effects in the fMRI data. These findings underscore that large samples are necessary but not sufficient—high-quality, harmonized data are also required.

Limitations and future directions

There are several methodological and technical limitations that must be considered when using RBC. First, RBC provides extensive demographics and phenotypic data (e.g., *p* factor) that can be accessed and used without any restrictions. However, more sensitive phenotypes—such as certain behavioral, clinical, and cognitive measures—are not released as part of the RBC dataset due to specific data privacy and DUA protocols set by the component studies. We note that for HBN and NKI datasets—two of the five studies included in RBC, comprising a large portion of the total data ($n = 3,940$; 62% of total participants)—granular phenotypic data are already publicly available via their respective study portals. Second, the majority of the data included in RBC are cross-sectional, with the exception of a smaller subset of individuals who have follow-up longitudinal data. Future work is required to provide public large longitudinal datasets to systematically study neurodevelopmental trajectories of brain and behavior data within the same individuals over time. Third, while RBC provides a relatively diverse data resource and incorporates data from five prominent neurodevelopmental datasets spanning three different continents, it is not fully representative of diverse populations. Fourth, parent or caregiver reports were the primary source of psychopathology measures in RBC. While using parent-reported measures of psychopathology provides valuable insights and is considered a reasonable approach, these reports reflect only one perspective and often diverge from youth self-reports.^{167,168} Finally, RBC's neuroimaging data consist of structural and functional MRI data. Future efforts are required to provide large-scale neurodevelopmental datasets that include multiple neuroimaging modalities, such as diffusion-weighted MRI (DWI) studies of brain microstructure and arterial spin-labeled (ASL) MRI measures of cerebral perfusion.

Moving forward, RBC's utility will be amplified by its adherence to FAIR principles—findability, accessibility, interoperability, and reusability. RBC data are easily *findable* and readily *accessible*: all raw and fully processed RBC data are publicly shared via the International Neuroimaging Data-sharing Initiative (INDI) and are accessible without any DUA requirements. By openly releasing de-identified data and removing barriers associated with cumbersome DUAs, RBC accelerates scientific discovery. However, it is important to note that while RBC imposes no DUA for data access, individual researchers may still be subject to local or institutional policies regarding the use of publicly shared human data. Users are therefore encouraged to consult with their own institutions to ensure compliance with any applicable regulations. RBC data *interoperability* is ensured by standard data structures such as BIDS and tidy tabular derivative data along with well-documented, open-source imaging pipelines. Moreover, RBC data can be redistributed without restriction, ensuring that it is *reusable*. For example, researchers can use RBC data to develop tools or integrate RBC data with other

datasets. Crucially, such efforts can be shared alongside RBC on INDI. These contributions may include additional derivatives generated using other BIDS-compliant processing pipelines (e.g., FastSurfer,¹⁶⁹ fMRIPrep,¹⁷⁰ XCP-D¹⁷¹, HippUnfold¹⁷², and Micapipe¹⁷³), helping to extend RBC's utility across a wider array of research approaches. RBC is also accompanied by a version-controlled website to help facilitate data access and maintenance (<https://reprobrainchart.github.io/>). We provide a “quick-start guide” on how to access and download RBC data on the RBC website (https://reprobrainchart.github.io/docs/get_data). We also provide detailed analysis workflows that allow users to replicate the results reported here (see “data and code availability”). For support, researchers can post questions to INCF NeuroStars using the RBC tag (<https://neurostars.org/tag/rbc>).

Beyond providing a large new open data resource for the community, RBC also offers a transparent and reproducible workflow for large-scale data integration and sharing, which may serve as a model for future multi-study efforts that others can adopt or adapt. While not all phenotypic measures could be included due to data-sharing restrictions, RBC serves as an important resource that future studies can build upon through expanded consent and tiered access models. Taken together, RBC accelerates large-scale, robust, and reproducible research in developmental and psychiatric neuroscience.

RESOURCE AVAILABILITY

Lead contact

Requests for further information should be directed to the lead contact, Theodore D Satterthwaite (sattertt@penmedicine.upenn.edu).

Materials availability

All harmonized phenotypes as well as raw and processed neuroimaging data are openly shared via DataLad and the International Neuroimaging Data-sharing Initiative (INDI). The released RBC data are accessible without any data use agreement (DUA) requirements and can be downloaded using DataLad via <https://github.com/ReproBrainChart>.

Data and code availability

All processing pipelines are shared using Docker containers for frictionless portability across platforms (<https://github.com/ReproBrainChart>) along with the analysis workflows and data used in this study (<https://github.com/ReproBrainChart/rbc-analysis-template>). RBC data release is also accompanied by a website to help facilitate data access and maintenance (<https://reprobrainchart.github.io/>). RBC website provides additional information and guidelines on how to access the data. Follow-up queries on RBC are closely monitored on INCF NeuroStars (<https://neurostars.org/tag/rbc>).

ACKNOWLEDGMENTS

RBC was funded by the National Institute of Mental Health (NIMH) via grant R01MH120482. Additional support was provided by R37MH125829, R01EB022573, R01MH112847, R01MH113550, RF1MH121867, R01MH123550, U24NS130411, P50MH109429, R01MH123440, and K08MH079364. G.S. was supported by a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR). We would like to thank the AWS Open Data Sponsorship Program for storage support. CCNP receives funding support from the STI 2030-the major projects of the Brain Science and Brain-Inspired Intelligence Technology (2021ZD0200500), the National Basic Science Data Center “Interdisciplinary Brain Database for In-vivo Population Imaging” (ID-BRAIN), the Key-Area Research and Development Program of Guangdong Province (2019B030335001), the Start-up Funds for Leading Talents at Beijing Normal

University, the Beijing Municipal Science and Technology Commission (Z161100002616023 and Z181100001518003), the Major Project of National Social Science Foundation of China (20&ZD296), the CAS-NWO Programme (153111KYSB20160020), the Guangxi BaGui Scholarship (201621), the National Basic Research (973) Program (2015CB351702), the Major Fund for International Collaboration of National Natural Science Foundation of China (81220108014), the Chinese Academy of Sciences Key Research Program (KSZD-EW-TZ-002), and the National Basic Research Program (2015CB351702). The Philadelphia neurodevelopmental Cohort was supported by NIMH RC2 MH089983 and RC2 MH089924. BHRC was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq grant numbers 573974/2008-0 and 465550/2014-2), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP grant numbers: 2008/57896-8, 2013/08531-5, 2014/50917-0, 2020/06172-1, 2021/05332-8, and 2021/12901-9) to the National Institute of Development Psychiatric for Children and Adolescent (INPD), the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no 337673.

AUTHOR CONTRIBUTIONS

Conceptualization: M.C., A.R.F., G.K., G.A.S., M.P.M., and T.D.S.; data curation: J.C., S.C., K.M., T.M.M., T.M.T., M.C., and G.K.; formal analysis: G.S., N.B.E., M.S.H., S.C., K.M., T.M.T., and M.C.; funding acquisition: M.S.H., T.M.M., M.E.C., S.C., C.D., R.E.G., R.C.G., P.M.P., A.P.J., A.R., L.A.R., R.T.S., N.T., X.N.Z., A.R.F., G.K., G.A.S., M.P.M., and T.D.S.; investigation: G.S., N.B.E., M.S.H., S.C., K.M., T.M.M., T.S., T.M.T., M.C., G.K., G.A.S., M.P.M., and T.D.S.; methodology: G.S., N.B.E., M.S.H., L.A., A.A.C., J.C., S.C., S.G., C.L., K.M., T.M.M., T.S., T.M.T., M.E.C., C.D., R.E.G., R.C.G., A.R., R.T.S., M.C., A.R.F., G.K., G.A.S., M.P.M., and T.D.S.; project administration: G.S., N.B.E., M.S.H., L.A., S.C., K.M., T.M.T., M.C., A.R.F., G.K., G.A.S., M.P.M., and T.D.S.; resources: N.B.E., M.E.C., S.C., R.E.G., R.C.G., P.M.P., A.P.J., L.A.R., N.T., X.N.Z., A.R.F., G.K., G.A.S., M.P.M., and T.D.S.; software: G.S., N.B.E., L.A., A.A.C., J.C., S.C., S.G., C.L., K.M., T.S., T.M.T., A.R., R.T.S., M.C., G.K., M.P.M., and T.D.S.; supervision: M.P.M. and T.D.S.; validation: G.S., N.B.E., M.S.H., and M.C.; visualization: G.S., N.B.E., and M.S.H.; writing – original draft: G.S., N.B.E., M.S.H., J.C., M.P.M., and T.D.S.; writing – review & editing: G.S., N.B.E., M.S.H., L.A., A.A.C., J.C., S.C., S.G., C.L., K.M., T.M.M., T.S., T.M.T., M.E.C., S.C., C.D., R.E.G., R.C.G., P.M.P., A.P.J., A.R., L.A.R., R.T.S., N.T., X.N.Z., M.C., A.R.F., G.K., G.A.S., M.P.M., and T.D.S.

DECLARATION OF INTERESTS

R.T.S. has received consulting income from Octave Bioscience and compensation for scientific reviewing from the American Medical Association. L.A.R. has received grant or research support from, served as a consultant to, and served on the speakers' bureau of Abdi Ibrahim, Abbott, Aché, Adium, Apsen, Bial, Cellera, EMS, Hypera Pharma, Knight Therapeutics, Libbs, Medice, Novartis/Sandoz, Pfizer/Upjohn/Viatris, Shire/Takeda, and Torrent in the last three years. The ADHD and Juvenile Bipolar Disorder Outpatient Programs chaired by L.A.R. have received unrestricted educational and research support from the following pharmaceutical companies in the last three years: Novartis/Sandoz and Shire/Takeda. Dr Rohde has received authorship royalties from Oxford Press and ArtMed.

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**
 - Study description
- **METHOD DETAILS**
 - Phenotypic data harmonization
 - Reproducible neuroimaging data curation and workflows
 - Structural MRI data processing and quality control

- Functional MRI data processing and quality control
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- Example analysis workflow with RBC data

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2025.08.026>.

Received: February 21, 2025

Revised: July 10, 2025

Accepted: August 27, 2025

REFERENCES

- Kessler, R.C., Petukhova, M., Sampson, N.A., Zaslavsky, A.M., and Wittchen, H.U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int. J. Methods Psychiatr. Res.* 21, 169–184. <https://doi.org/10.1002/mpr.1359>.
- Casey, B.J., Oliveri, M.E., and Insel, T. (2014). A neurodevelopmental perspective on the research domain criteria (rdoc) framework. *Biol. Psychiatry* 76, 350–353. <https://doi.org/10.1016/j.biopsych.2014.01.006>.
- Bethlehem, R.A.I., Seidlitz, J., White, S.R., Vogel, J.W., Anderson, K.M., Adamson, C., Adler, S., Alexopoulos, G.S., Anagnostou, E., Areces-Gonzalez, A., et al. (2022). Brain charts for the human lifespan. *Nature* 604, 525–533. <https://doi.org/10.1038/s41586-022-04554-y>.
- Paus, T.S., Zijdenbos, A., Worsley, K., Collins, D.L., Blumenthal, J., Giedd, J.N., Rapoport, J.L., and Evans, A.C. (1999). Structural maturation of neural pathways in children and adolescents: In vivo study. *Science* 283, 1908–1911. <https://doi.org/10.1126/science.283.5409.1908>.
- Paus, T. (2010). Population neuroscience: Why and how. *Hum. Brain Mapp.* 31, 891–903. <https://doi.org/10.1002/hbm.21069>.
- Van Horn, J.D., and Gazzaniga, M.S. (2002). Opinion: Databasing fMRI studies towards a 'discovery science' of brain function. *Nat. Rev. Neurosci.* 3, 314–318. <https://doi.org/10.1038/nrn788>.
- Milham, M.P. (2012). Open neuroscience solutions for the connectome-wide association era. *Neuron* 73, 214–218. <https://doi.org/10.1016/j.neuron.2011.11.004>.
- Nooner, K.B., Colcombe, S.J., Tobe, R.H., Mennes, M., Benedict, M.M., Moreno, A.L., Panek, L.J., Brown, S., Zavitz, S.T., Li, Q., et al. (2012). The nki-rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* 6, 152. <https://doi.org/10.3389/fnins.2012.00152>.
- Marek, S., Tervo-Clemmens, B., Calabro, F.J., Montez, D.F., Kay, B.P., Hatoun, A.S., Donohue, M.R., Foran, W., Miller, R.L., Hendrickson, T. J., et al. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature* 603, 654–660. <https://doi.org/10.1038/s41586-022-04492-9>.
- Zhou, Z.-X., Chen, L.-Z., Milham, M.P., and Zuo, X.-N.; Lifespan Brain Chart Consortium (LBCC) (2023). Six cornerstones for translational brain charts. *Sci. Bull.* 68, 795–799. <https://doi.org/10.1016/j.scib.2023.03.047>.
- Uddin, L.Q., Castellanos, F.X., and Menon, V. (2024). Resting state functional brain connectivity in child and adolescent psychiatry: Where are we now? *Neuropsychopharmacology* 50, 196–200. <https://doi.org/10.1038/s41386-024-01888-1>.
- Brown, S.A., Garavan, H., Jernigan, T.L., Tapert, S.F., Huber, R.S., Lopez, D., Murray, T., Dowling, G., Hoffman, E.A., and Uddin, L.Q. (2025). Responsible use of population neuroscience data: Towards standards of accountability and integrity. *Dev. Cogn. Neurosci.* 72, 101466. <https://doi.org/10.1016/j.dcn.2024.101466>.
- Biswal, B.B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. USA* 107, 4734–4739. <https://doi.org/10.1073/pnas.0911855107>.
- Howell, B.R., Styner, M.A., Gao, W., Yap, P.-T., Wang, L., Baluyot, K., Yacoub, E., Chen, G., Potts, T., Salzwedel, A., et al. (2019). The UNC/UMN Baby Connectome Project (BCP): An overview of the study design and protocol development. *NeuroImage* 185, 891–905. <https://doi.org/10.1016/j.neuroimage.2018.03.049>.
- Somerville, L.H., Bookheimer, S.Y., Buckner, R.L., Burgess, G.C., Curtiss, S.W., Dapretto, M., Elam, J.S., Gaffrey, M.S., Harms, M.P., Hodge, C., et al. (2018). The Lifespan Human Connectome Project in Development: A large-scale study of brain connectivity development in 5–21 year olds. *NeuroImage* 183, 456–468. <https://doi.org/10.1016/j.neuroimage.2018.08.050>.
- Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Loughhead, J., Prabhakaran, K., Calkins, M.E., Hopson, R., Jackson, C., Keefe, J., Riley, M., et al. (2014). Neuroimaging of the Philadelphia Neurodevelopmental Cohort. *NeuroImage* 86, 544–553. <https://doi.org/10.1016/j.neuroimage.2013.07.064>.
- Alexander, L.M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., et al. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* 4, 170181. <https://doi.org/10.1038/sdata.2017.181>.
- Tobe, R.H., MacKay-Brandt, A., Lim, R., Kramer, M., Breland, M.M., Tu, L., Tian, Y., Trautman, K.D., Hu, C., Sangoi, R., et al. (2022). A longitudinal resource for studying connectome development and its psychiatric associations during childhood. *Sci. Data* 9, 300. <https://doi.org/10.1038/s41597-022-01329-y>.
- Volkow, N.D., Koob, G.F., Croyle, R.T., Bianchi, D.W., Gordon, J.A., Koroshetz, W.J., Pérez-Stable, E.J., Riley, W.T., Bloch, M.H., Conway, K., et al. (2018). The conception of the ABCD study: From substance use to a broad NIH collaboration. *Dev. Cogn. Neurosci.* 32, 4–7. <https://doi.org/10.1016/j.dcn.2017.10.002>.
- Fair, D.A., Nigg, J.T., Iyer, S., Bathula, D., Mills, K.L., Dosenbach, N.U.F., Schlaggar, B.L., Mennes, M., Gutman, D., Bangaru, S., et al. (2012). Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. *Front. Syst. Neurosci.* 6, 80. <https://doi.org/10.3389/fnsys.2012.00080>.
- Di Martino, A., O'Connor, D., Chen, B., Alaerts, K., Anderson, J.S., Assaf, M., Balsters, J.H., Baxter, L., Beggiani, A., Benaerts, S., et al. (2017). Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci. Data* 4, 170010. <https://doi.org/10.1038/sdata.2017.10>.
- Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667. <https://doi.org/10.1038/mp.2013.78>.
- Yan, C.-G., Craddock, R.C., Zuo, X.-N., Zang, Y.-F., and Milham, M.P. (2013). Standardizing the intrinsic brain: Towards robust measurement of inter-individual variation in 1000 functional connectomes. *NeuroImage* 80, 246–262. <https://doi.org/10.1016/j.neuroimage.2013.04.081>.
- Mennes, M., Biswal, B.B., Castellanos, F.X., and Milham, M.P. (2013). Making data sharing work: The FCP/INDI experience. *NeuroImage* 82, 683–691. <https://doi.org/10.1016/j.neuroimage.2012.10.064>.
- Poldrack, R.A., and Poline, J.-B. (2015). The publication and reproducibility challenges of shared data. *Trends Cogn. Sci.* 19, 59–61. <https://doi.org/10.1016/j.tics.2014.11.008>.
- Gilmore, R.O., Lorenzo Kennedy, J.L., and Adolph, K.E. (2018). Practical solutions for sharing data and materials from psychological research.

- Adv. Methods Pract. Psychol. Sci. 1, 121–130. <https://doi.org/10.1177/2515245917746500>.
27. Zuo, X.-N., Anderson, J.S., Bellec, P., Birn, R.M., Biswal, B.B., Blautzik, J., Breitner, J.C.S., Buckner, R.L., Calhoun, V.D., Castellanos, F.X., et al. (2014). An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci. Data* 1, 140049. <https://doi.org/10.1038/sdata.2014.49>.
28. Milham, M.P., Vogelstein, J., and Xu, T. (2021). Removing the reliability bottleneck in functional magnetic resonance imaging research to achieve clinical utility. *JAMA Psychiatry* 78, 587–588. <https://doi.org/10.1001/jamapsychiatry.2020.4272>.
29. Xu, T., Kiar, G., Cho, J.W., Bridgeford, E.W., Nikolaidis, A., Vogelstein, J. T., and Milham, M.P. (2023). ReX: An integrative tool for quantifying and optimizing measurement reliability for the study of individual differences. *Nat. Methods* 20, 1025–1028. <https://doi.org/10.1038/s41592-023-01901-3>.
30. Laird, A.R. (2021). Large, open datasets for human connectomics research: Considerations for reproducible and responsible data use. *NeuroImage* 244, 118579. <https://doi.org/10.1016/j.neuroimage.2021.118579>.
31. Kang, K., Seidlitz, J., Bethlehem, R.A.I., Xiong, J., Jones, M.T., Mehta, K., Keller, A.S., Tao, R., Randolph, A., Larsen, B., et al. (2024). Study design features increase replicability in brain-wide association studies. *Nature* 636, 719–727. <https://doi.org/10.1038/s41586-024-08260-9>.
32. McElroy, E., Villadsen, A., Patalay, P., Goodman, A., Richards, M., Northstone, K., Fearon, P., Tibber, M., Gondek, D., and Ploubidis, G.B. (2020) Harmonisation and Measurement Properties of Mental Health Measures in Six British Cohorts.
33. Polanczyk, G.V., Willcutt, E.G., Salum, G.A., Kieling, C., and Rohde, L.A. (2014). ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *Int. J. Epidemiol.* 43, 434–442. <https://doi.org/10.1093/ije/dyt261>.
34. Luningham, J.M., Hendriks, A.M., Krapohl, E., Fung Ip, H., van Beijsterveldt, C.E.M., Lundström, S., Vuoksimaa, E., Korhonen, T., Lichtenstein, P., Plomin, R., et al. (2020). Harmonizing behavioral outcomes across studies, raters, and countries: Application to the genetic analysis of aggression in the ACTION Consortium. *J. Child Psychol. Psychiatry* 61, 807–817. <https://doi.org/10.1111/jcpp.13188>.
35. Bauer, D.J., and Hussong, A.M. (2009). Psychometric approaches for developing commensurate measures across independent studies: Traditional and new models. *Psychol. Methods* 14, 101–125. <https://doi.org/10.1037/a0015583>.
36. Curran, P.J., and Hussong, A.M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychol. Methods* 14, 81–100. <https://doi.org/10.1037/a0015914>.
37. Reise, S.P. (2012). Invited Paper: The rediscovery of bifactor measurement models. *Multivariate Behav. Res.* 47, 667–696. <https://doi.org/10.1080/00273171.2012.715555>.
38. Lahey, B.B., Applegate, B., Hakes, J.K., Zald, D.H., Hariri, A.R., and Rathouz, P.J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *J. Abnorm. Psychol.* 121, 971–977. <https://doi.org/10.1037/a0028355>.
39. Caspi, A., Houts, R.M., Belsky, D.W., Goldman-Mellor, S.J., Harrington, H., Israel, S., Meier, M.H., Ramrakha, S., Shalev, I., Poulton, R., et al. (2014). The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clin. Psychol. Sci.* 2, 119–137. <https://doi.org/10.1177/2167702613497473>.
40. Kotov, R., Krueger, R., Watson, D., Achenbach, T., Althoff, R., Bagby, M., Brown, T., Docherty, A., Miller, J., Simms, L., et al. (2017). The hierarchical taxonomy of psychopathology (hitop): A dimensional alternative to traditional nosologies (Center for Open Science). <https://doi.org/10.31234/osf.io/zaadn>.
41. Kotov, R., Krueger, R.F., Watson, D., Cicero, D.C., Conway, C.C., DeYoung, C.G., Eaton, N.R., Forbes, M.K., Hallquist, M.N., Latzman, R. D., et al. (2021). The hierarchical taxonomy of psychopathology (hitop): A quantitative nosology based on consensus of evidence. *Annu. Rev. Clin. Psychol.* 17, 83–108. <https://doi.org/10.1146/annurev-clinpsy-081219-093304>.
42. Hoffmann, M.S., Moore, T.M., Axelrud, L.K., Tottenham, N., Pan, P.M., Miguel, E.C., Rohde, L.A., Milham, M.P., Satterthwaite, T.D., and Salum, G.A. (2024). An evaluation of item harmonization strategies between assessment tools of psychopathology in children and adolescents. *Assessment* 31, 502–517. <https://doi.org/10.1177/10731911231163136>.
43. Hoffmann, M.S., Moore, T.M., Axelrud, L.K., Tottenham, N., Rohde, L.A., Milham, M.P., Satterthwaite, T.D., and Salum, G.A. (2023). Harmonizing bifactor models of psychopathology between distinct assessment instruments: Reliability, measurement invariance, and authenticity. *Int. J. Methods Psychiatr. Res.* 32, e1959. <https://doi.org/10.1002/mpr.1959>.
44. Scopel Hoffmann, M.S., Moore, T.M., Kvitko Axelrud, L., Tottenham, N., Zuo, X.-N., Rohde, L.A., Milham, M.P., Satterthwaite, T.D., and Salum, G.A. (2022). Reliability and validity of bifactor models of dimensional psychopathology in youth. *J. Psychopathol. Clin. Sci.* 131, 407–421. <https://doi.org/10.1037/abn0000749>.
45. Cuthbert, B.N., and Insel, T.R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Med.* 11, 126. <https://doi.org/10.1186/1741-7015-11-126>.
46. Carp, J. (2012). On the Plurality of (Methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments. *Front. Neurosci.* 6, 149. <https://doi.org/10.3389/fnins.2012.00149>.
47. Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature* 533, 452–454. <https://doi.org/10.1038/533452a>.
48. Nichols, T.E., Das, S., Eickhoff, S.B., Evans, A.C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M.P., Poldrack, R.A., Poline, J.-B., et al. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nat. Neurosci.* 20, 299–303. <https://doi.org/10.1038/nn.4500>.
49. Botvinik-Nezer, R., Holzmeister, F., Camerer, C.F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J.A., Adcock, R.A., et al. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* 582, 84–88. <https://doi.org/10.1038/s41586-020-2314-9>.
50. Bhagwat, N., Barry, A., Dickie, E.W., Brown, S.T., Devenyi, G.A., Hatano, K., DuPre, E., Dagher, A., Chakravarty, M., Greenwood, C.M.T., et al. (2021). Understanding the impact of preprocessing pipelines on neuroimaging cortical surface analyses. *GigaScience* 10, g1aa155. <https://doi.org/10.1093/gigascience/g1aa155>.
51. Li, X., Bianchini Esper, N., Ai, L., Giavasis, S., Jin, H., Feczko, E., Xu, T., Clucas, J., Franco, A., Sólón Heinsfeld, A., et al. (2024). Moving beyond processing- and analysis-related variation in resting-state functional brain imaging. *Nat. Hum. Behav.* 8, 2003–2017. <https://doi.org/10.1038/s41562-024-01942-4>.
52. Fischl, B. (2012). FreeSurfer. *NeuroImage* 62, 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>.
53. Cameron, C., Sharad, S., Brian, C., Ranjeet, K., Satrajit, G., Chaogan, Y., Qingyang, L., Daniel, L., Joshua, V., Randal, B., et al. (2013). Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (C-PAC). *Front. Neuroinform.* 7. <https://doi.org/10.3389/conf.fninf.2013.09.00042>.
54. Wagner, A.S., Waite, L.K., Wierzb, M., Hoffstaedter, F., Waite, A.Q., Poldrack, B., Eickhoff, S.B., and Hanke, M. (2022). FAIRly big: A framework for computationally reproducible processing of large-scale data. *Sci. Data* 9, 80. <https://doi.org/10.1038/s41597-022-01163-2>.
55. Halchenko, Y.O., Meyer, K., Poldrack, B., Solanky, D.S., Wagner, A.S., Gors, J., MacFarlane, D., Pustina, D., Sochat, V., Ghosh, S.S., et al. (2021). DataLad: Distributed system for joint management of code,

- data, and their relationship. *J. Open Source Softw.* 6, 3262. <https://doi.org/10.21105/joss.03262>.
56. Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D.S., and Craddock, R.C. (2017). The Neuro Bureau ADHD-200 Preprocessed repository. *NeuroImage* 144, 275–286. <https://doi.org/10.1016/j.neuroimage.2016.06.034>.
57. Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118–127. <https://doi.org/10.1093/biostatistics/kxj037>.
58. Fortin, J.-P., Cullen, N., Sheline, Y.I., Taylor, W.D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P.J., et al. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* 167, 104–120. <https://doi.org/10.1016/j.neuroimage.2017.11.024>.
59. Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M.A., Ruparel, K., Roalf, D.R., Satterthwaite, T.D., Gur, R.C., Gur, R.E., et al. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage* 161, 149–170. <https://doi.org/10.1016/j.neuroimage.2017.08.047>.
60. Pomponio, R., Erus, G., Habes, M., Doshi, J., Srinivasan, D., Mamourian, E., Bashyam, V., Nasrallah, I.M., Satterthwaite, T.D., Fan, Y., et al. (2020). Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *NeuroImage* 208, 116450. <https://doi.org/10.1016/j.neuroimage.2019.116450>.
61. Chen, A.A., Beer, J.C., Tustison, N.J., Cook, P.A., Shinohara, R.T., and Shou, H. (2021). Mitigating site effects in covariance for machine learning in neuroimaging data. *Hum. Brain Mapp.* 43, 1179–1195. <https://doi.org/10.1002/hbm.25688>.
62. Hu, F., Chen, A.A., Horng, H., Bashyam, V., Davatzikos, C., Alexander-Bloch, A., Li, M., Shou, H., Satterthwaite, T.D., Yu, M., et al. (2023). Image harmonization: A review of statistical and deep learning methods for removing batch effects and evaluation metrics for effective harmonization. *NeuroImage* 274, 120125. <https://doi.org/10.1016/j.neuroimage.2023.120125>.
63. Bridgeford, E.W., Powell, M., Kiar, G., Noble, S., Chung, J., Panda, S., Lawrence, R., Xu, T., Milham, M., Caffo, B., et al. (2025). When no answer is better than a wrong answer: A causal perspective on batch effects. *Imaging Neurosci.* (Camb) 3, imag_a_00458. https://doi.org/10.1162/imag_a_00458.
64. 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>.
65. 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>.
66. Van Dijk, K.R.A., Sabuncu, M.R., and Buckner, R.L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage* 59, 431–438. <https://doi.org/10.1016/j.neuroimage.2011.07.044>.
67. Murphy, K., Birn, R.M., and Bandettini, P.A. (2013). Resting-state fMRI confounds and cleanup. *NeuroImage* 80, 349–359. <https://doi.org/10.1016/j.neuroimage.2013.04.001>.
68. Gilmore, A.D., Buser, N.J., and Hanson, J.L. (2021). Variations in structural MRI quality significantly impact commonly used measures of brain anatomy. *Brain Inform.* 8, 7. <https://doi.org/10.1186/s40708-021-00128-2>.
69. Milham, M.P., Craddock, R.C., Son, J.J., Fleischmann, M., Clucas, J., Xu, H., Koo, B., Krishnakumar, A., Biswal, B.B., Castellanos, F.X., et al. (2018). Assessment of the impact of shared brain imaging data on the scientific literature. *Nat. Commun.* 9, 2818. <https://doi.org/10.1038/s41467-018-04976-1>.
70. White, T., Blok, E., and Calhoun, V.D. (2022). Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed. *Hum. Brain Mapp.* 43, 278–291. <https://doi.org/10.1002/hbm.25120>.
71. Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä., Pedaste, M., Raju, M., Astapova, A., Lukner, H., et al. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Sci. Data* 8, 192. <https://doi.org/10.1038/s41597-021-00981-0>.
72. Jwa, A.S., and Poldrack, R.A. (2022). The spectrum of data sharing policies in neuroimaging data repositories. *Hum. Brain Mapp.* 43, 2707–2721. <https://doi.org/10.1002/hbm.25803>.
73. Salum, G.A., Gadelha, A., Pan, P.M., Moriyama, T.S., Graeff-Martins, A. S., Tamanaha, A.C., Alvarenga, P., Valle Krieger, F.V., Fleitlich-Bilyk, B., Jackowski, A., et al. (2015). High risk cohort study for psychiatric disorders in childhood: Rationale, design, methods and preliminary results. *Int. J. Methods Psychiatr. Res.* 24, 58–73. <https://doi.org/10.1002/mpr.1459>.
74. Liu, S., Wang, Y.-S., Zhang, Q., Zhou, Q., Cao, L.-Z., Jiang, C., Zhang, Z., Yang, N., Dong, Q., Zuo, X.-N., et al. (2021). Chinese Color Nest Project: An accelerated longitudinal brain-mind cohort. *Dev. Cogn. Neurosci.* 52, 101020. <https://doi.org/10.1016/j.dcn.2021.101020>.
75. Fan, X.-R., Wang, Y.-S., Chang, D., Yang, N., Rong, M.-J., Zhang, Z., He, Y., Hou, X., Zhou, Q., Gong, Z.-Q., et al. (2023). A longitudinal resource for population neuroscience of school-age children and adolescents in China. *Sci. Data* 10, 545. <https://doi.org/10.1038/s41597-023-02377-8>.
76. Satterthwaite, T.D., Connolly, J.J., Ruparel, K., Calkins, M.E., Jackson, C., Elliott, M.A., Roalf, D.R., Hopson, R., Prabhakaran, K., Behr, M., et al. (2016). The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage* 124, 1115–1119. <https://doi.org/10.1016/j.neuroimage.2015.03.056>.
77. Achenbach, T.M., and Rescorla, L.A. (2000) Manual for the ASEBA preschool forms and profiles.
78. Calkins, M.E., Merikangas, K.R., Moore, T.M., Burstein, M., Behr, M.A., Satterthwaite, T.D., Ruparel, K., Wolf, D.H., Roalf, D.R., Mentch, F.D., et al. (2015). The Philadelphia Neurodevelopmental Cohort: Constructing a deep phenotyping collaborative. *J. Child Psychol. Psychiatry* 56, 1356–1369. <https://doi.org/10.1111/jcpp.12416>.
79. 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>.
80. Covitz, S., Tapera, T.M., Adebimpe, A., Alexander-Bloch, A.F., Bertolero, M.A., Feczko, E., Franco, A.R., Gur, R.E., Gur, R.C., Hendrickson, T., et al. (2022). Curation of BIDS (CuBIDS): A workflow and software package for streamlining reproducible curation of large BIDS datasets. *NeuroImage* 263, 119609. <https://doi.org/10.1016/j.neuroimage.2022.119609>.
81. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., and Ugurbil, K.; WU-Minn; HCP Consortium (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage* 80, 62–79. <https://doi.org/10.1016/j.neuroimage.2013.05.041>.
82. Glasser, M.F., Smith, S.M., Marcus, D.S., Andersson, J.L.R., Auerbach, E.J., Behrens, T.E.J., Coalson, T.S., Harms, M.P., Jenkinson, M., Moeller, S., et al. (2016). The Human Connectome Project's neuroimaging approach. *Nat. Neurosci.* 19, 1175–1187. <https://doi.org/10.1038/nn.4361>.
83. Esteban, O., Markiewicz, C.J., Blair, R., Poldrack, R.A., and Gorgolewski, K.J. (2024) sMRIPrep: Structural MRI PREProcessing workflows.

84. Zang, Y., Jiang, T., Lu, Y., He, Y., and Tian, L. (2004). Regional homogeneity approach to fMRI data analysis. *NeuroImage* 22, 394–400. <https://doi.org/10.1016/j.neuroimage.2003.12.030>.
85. Zang, Y.-F., He, Y., Zhu, C.-Z., Cao, Q.-J., Sui, M.-Q., Liang, M., Tian, L.-X., Jiang, T.-Z., and Wang, Y.-F. (2007). Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev.* 29, 83–91. <https://doi.org/10.1016/j.braindev.2006.07.002>.
86. Zou, Q.-H., Zhu, C.-Z., Yang, Y., Zuo, X.-N., Long, X.-Y., Cao, Q.-J., Wang, Y.-F., and Zang, Y.-F. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *J. Neurosci. Methods* 172, 137–141. <https://doi.org/10.1016/j.jneumeth.2008.04.012>.
87. Reuter, M., Tisdall, M.D., Qureshi, A., Buckner, R.L., van der Kouwe, A.J.W., and Fischl, B. (2015). Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *NeuroImage* 107, 107–115. <https://doi.org/10.1016/j.neuroimage.2014.12.006>.
88. Rosen, A.F.G., Roalf, D.R., Ruparel, K., Blake, J., Seelaus, K., Villa, L.P., Ciric, R., Cook, P.A., Davatzikos, C., Elliott, M.A., et al. (2018). Quantitative assessment of structural image quality. *NeuroImage* 169, 407–418. <https://doi.org/10.1016/j.neuroimage.2017.12.059>.
89. Kong, X.-Z., Zhen, Z., Li, X., Lu, H.-H., Wang, R., Liu, L., He, Y., Zang, Y., and Liu, J. (2014). Individual differences in impulsivity predict head motion during magnetic resonance imaging. *PLoS One* 9, e104989. <https://doi.org/10.1371/journal.pone.0104989>.
90. Keshavan, A., Yeatman, J.D., and Rokem, A. (2019). Combining citizen science and deep learning to amplify expertise in neuroimaging. *Front. Neuroinform.* 13, 29. <https://doi.org/10.3389/fninf.2019.00029>.
91. Klapwijk, E.T., van de Kamp, F., van der Meulen, M., Peters, S., and Wierenga, L.M. (2019). Qoala-T: A supervised-learning tool for quality control of FreeSurfer segmented MRI data. *NeuroImage* 189, 116–129. <https://doi.org/10.1016/j.neuroimage.2019.01.014>.
92. Elyounssi, S., Kunitoki, K., Clauss, J.A., Laurent, E., Kane, K., Hughes, D.E., Hopkinson, C.E., Bazer, O., Sussman, R.F., Doyle, A.E., et al. (2023). Uncovering and mitigating bias in large, automated MRI analyses of brain development (Cold Spring Harbor Laboratory). <https://doi.org/10.1101/2023.02.28.530498>.
93. Elyounssi, S., Kunitoki, K., Hughes, D., Bazer, O., Clauss, J., and Roffman, J. (2022). P438. Growing brains in motion: In-Depth quality control of structural MRI scans from 5,764 9 to 10-year-olds in the adolescent brain cognitive development (ABCD) study. *Biol. Psychiatry* 91, S265. <https://doi.org/10.1016/j.biopsych.2022.02.674>.
94. Sydnor, V.J., Larsen, B., Seidlitz, J., Adeimpe, A., Alexander-Bloch, A.F., Bassett, D.S., Bertolero, M.A., Cieslak, M., Covitz, S., Fan, Y., et al. (2023). Intrinsic activity development unfolds along a sensorimotor-association cortical axis in youth. *Nat. Neurosci.* 26, 638–649. <https://doi.org/10.1038/s41593-023-01282-y>.
95. Luo, A.C., Sydnor, V.J., Pines, A., Larsen, B., Alexander-Bloch, A.F., Cieslak, M., Covitz, S., Chen, A.A., Esper, N.B., Feczko, E., et al. (2024). Functional connectivity development along the sensorimotor-association axis enhances the cortical hierarchy. *Nat. Commun.* 15, 3511. <https://doi.org/10.1038/s41467-024-47748-w>.
96. Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., and Yeo, B.T.T. (2018). Local-Global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* 28, 3095–3114. <https://doi.org/10.1093/cercor/bhx179>.
97. Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. <https://doi.org/10.1152/jn.00338.2011>.
98. Gennatas, E.D., Avants, B.B., Wolf, D.H., Satterthwaite, T.D., Ruparel, K., Ciric, R., Hakonarson, H., Gur, R.E., and Gur, R.C. (2017). Age-Related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. *J. Neurosci.* 37, 5065–5073. <https://doi.org/10.1523/JNEUROSCI.3550-16.2017>.
99. Marshall, A.T., Adise, S., Kan, E.C., and Sowell, E.R. (2025). Longitudinal sex-at-birth and age analyses of cortical structure in the ABCD study. *J. Neurosci.* 45, e1091242025. <https://doi.org/10.1523/JNEUROSCI.1091-24.2025>.
100. Satterthwaite, T.D., Wolf, D.H., Roalf, D.R., Ruparel, K., Erus, G., Vandekar, S., Gennatas, E.D., Elliott, M.A., Smith, A., Hakonarson, H., et al. (2015). Linked sex differences in cognition and functional connectivity in youth. *Cereb. Cortex* 25, 2383–2394. <https://doi.org/10.1093/cercor/bhu036>.
101. Shanmugan, S., Seidlitz, J., Cui, Z., Adeimpe, A., Bassett, D.S., Bertolero, M.A., Davatzikos, C., Fair, D.A., Gur, R.E., Gur, R.C., et al. (2022). Sex differences in the functional topography of association networks in youth. *Proc. Natl. Acad. Sci. USA* 119, e2110416119. <https://doi.org/10.1073/pnas.2110416119>.
102. Duffy, K.A., Wiglesworth, A., Roediger, D.J., Island, E., Mueller, B.A., Luciana, M., Klimes-Dougan, B., Cullen, K.R., and Fiecas, M.B. (2025). Characterizing the effects of age, puberty, and sex on variability in resting-state functional connectivity in late childhood and early adolescence. *NeuroImage* 313, 121238. <https://doi.org/10.1016/j.neuroimage.2025.121238>.
103. Milham, P.M., D.F., Maarten, M., and Stewart, H.M. (2012). The ADHD-200 Consortium: A Model to Advance the Translational Potential of Neuroimaging in Clinical Neuroscience. *Front. Syst. Neurosci.* 6, 62. <https://doi.org/10.3389/fnsys.2012.00062>.
104. Cameron, C., Yassine, B., Carlton, C., Francois, C., Alan, E., András, J., Budhachandra, K., John, L., Qingyang, L., Michael, M., et al. (2013). The Neuro Bureau Preprocessing Initiative: Open sharing of preprocessed neuroimaging data and derivatives. *Front. Neuroinform.* 7. <https://doi.org/10.3389/conf.fninf.2013.09.00041>.
105. Watts, A.L., Greene, A.L., Bonifay, W., and Fried, E.I. (2024). A critical evaluation of the p-factor literature. *Nat. Rev. Psychol.* 3, 108–122. <https://doi.org/10.1038/s44159-023-00260-2>.
106. Caspi, A., and Moffitt, T.E. (2018). All for one and one for all: Mental disorders in one dimension. *Am. J. Psychiatry* 175, 831–844. <https://doi.org/10.1176/appi.ajp.2018.17121383>.
107. Sallis, H., Szekely, E., Neumann, A., Jolicoeur-Martineau, A., van Ijzendoorn, M., Hillegers, M., Greenwood, C.M.T., Meaney, M.J., Steiner, M., Tiemeier, H., et al. (2019). General psychopathology, internalising and externalising in children and functional outcomes in late adolescence. *J. Child Psychol. Psychiatry* 60, 1183–1190. <https://doi.org/10.1111/jcpp.13067>.
108. Allegrini, A.G., Cheesman, R., Rimpfeld, K., Selzam, S., Pingault, J.B., Eley, T.C., and Plomin, R. (2020). The p factor: Genetic analyses support a general dimension of psychopathology in childhood and adolescence. *J. Child Psychol. Psychiatry* 61, 30–39. <https://doi.org/10.1111/jcpp.13113>.
109. Waszczuk, M.A., Miao, J., Docherty, A.R., Shabalin, A.A., Jonas, K.G., Michelini, G., and Kotov, R. (2023). General v. specific vulnerabilities: Polygenic risk scores and higher-order psychopathology dimensions in the Adolescent Brain Cognitive Development (ABCD) Study. *Psychol. Med.* 53, 1937–1946. <https://doi.org/10.1017/S0033291721003639>.
110. Caspi, A., Houts, R.M., Fisher, H.L., Danese, A., and Moffitt, T.E. (2024). The General Factor of Psychopathology (p): Choosing Among Competing Models and Interpreting p. *Clin. Psychol. Sci.* 12, 53–82. <https://doi.org/10.1177/21677026221147872>.
111. 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.
112. 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>.
113. Macey, P.M., Macey, K.E., Kumar, R., and Harper, R.M. (2004). A method for removal of global effects from fMRI time series. *NeuroImage* 22, 360–366. <https://doi.org/10.1016/j.neuroimage.2003.12.042>.
114. Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *NeuroImage* 44, 893–905. <https://doi.org/10.1016/j.neuroimage.2008.09.036>.
115. Murphy, K., and Fox, M.D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *NeuroImage* 154, 169–173. <https://doi.org/10.1016/j.neuroimage.2016.11.052>.
116. Saad, Z.S., Gotts, S.J., Murphy, K., Chen, G., Jo, H.J., Martin, A., and Cox, R.W. (2012). Trouble at rest: How correlation patterns and group differences become distorted after global signal regression. *Brain Connect.* 2, 25–32. <https://doi.org/10.1089/brain.2012.0080>.
117. Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughhead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage* 64, 240–256. <https://doi.org/10.1016/j.neuroimage.2012.08.052>.
118. Power, J.D., Schlaggar, B.L., and Petersen, S.E. (2015). Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage* 105, 536–551. <https://doi.org/10.1016/j.neuroimage.2014.10.044>.
119. Uddin, L.Q. (2017). Mixed signals: On separating brain signal from noise. *Trends Cogn. Sci.* 21, 405–406. <https://doi.org/10.1016/j.tics.2017.04.002>.
120. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>.
121. Zhao, C., Chen, Y., Jarecka, D., Ghosh, S., Cieslak, M., and Satterthwaite, T.D. (2024). BABS in action: Merging reproducibility and scalability for large-scale neuroimaging analysis with BIDS apps. *Biol. Psychiatry* 95, S21–S22. <https://doi.org/10.1016/j.biopsych.2024.02.056>.
122. Sowell, E.R., Peterson, B.S., Thompson, P.M., Welcome, S.E., Henkenius, A.L., and Toga, A.W. (2003). Mapping cortical change across the human life span. *Nat. Neurosci.* 6, 309–315. <https://doi.org/10.1038/nn1008>.
123. Sowell, E.R., Thompson, P.M., Leonard, C.M., Welcome, S.E., Kan, E., and Toga, A.W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *J. Neurosci.* 24, 8223–8231. <https://doi.org/10.1523/JNEUROSCI.1798-04.2004>.
124. Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G.L., Greenstein, D., Clasen, L., Gogtay, N., and Giedd, J.N. (2011). How does your cortex grow? *J. Neurosci.* 31, 7174–7177. <https://doi.org/10.1523/JNEUROSCI.0054-11.2011>.
125. Shaw, P., Kabani, N.J., Lerch, J.P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J.L., et al. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *J. Neurosci.* 28, 3586–3594. <https://doi.org/10.1523/JNEUROSCI.5309-07.2008>.
126. Wierenga, L.M., Langen, M., Oranje, B., and Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *NeuroImage* 87, 120–126. <https://doi.org/10.1016/j.neuroimage.2013.11.010>.
127. Xu, T., Yang, Z., Jiang, L., Xing, X.-X., and Zuo, X.-N. (2015). A Connectome Computation System for discovery science of brain. *Sci. Bull.* 60, 86–95. <https://doi.org/10.1007/s11434-014-0698-3>.
128. Kelly, A.M.C., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., and Milham, M.P. (2008). Competition between functional brain networks mediates behavioral variability. *NeuroImage* 39, 527–537. <https://doi.org/10.1016/j.neuroimage.2007.08.008>.
129. Supekar, K., Musen, M., and Menon, V. (2009). Development of large-scale functional brain networks in children. *PLoS Biol.* 7, e1000157. <https://doi.org/10.1371/journal.pbio.1000157>.
130. Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U.F., Church, J.A., Miezin, F.M., Schlaggar, B.L., and Petersen, S.E. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Comput. Biol.* 5, e1000381. <https://doi.org/10.1371/journal.pcbi.1000381>.
131. Zuo, X.-N., Kelly, C., Di Martino, A., Mennes, M., Margulies, D.S., Bangaru, S., Grzadzinski, R., Evans, A.C., Zang, Y.-F., Castellanos, F. X., et al. (2010). Growing together and growing apart: Regional and sex differences in the lifespan developmental trajectories of functional homotopy. *J. Neurosci.* 30, 15034–15043. <https://doi.org/10.1523/JNEUROSCI.2612-10.2010>.
132. Vogel, A.C., Power, J.D., Petersen, S.E., and Schlaggar, B.L. (2010). Development of the brain’s functional network architecture. *Neuropsychol. Rev.* 20, 362–375. <https://doi.org/10.1007/s11065-010-9145-7>.
133. Hagmann, P., Grant, P.E., and Fair, D.A. (2012). MR connectomics: A conceptual framework for studying the developing brain. *Front. Syst. Neurosci.* 6, 43. <https://doi.org/10.3389/fnsys.2012.00043>.
134. Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V.J., Meuli, R., Thiran, J.P., and Grant, P.E. (2010). White matter maturation reshapes structural connectivity in the late developing human brain. *Proc. Natl. Acad. Sci. USA* 107, 19067–19072. <https://doi.org/10.1073/pnas.1009073107>.
135. Collin, G., and van den Heuvel, M.P. (2013). The ontogeny of the human connectome: Development and dynamic changes of brain connectivity across the life span. *Neuroscientist* 19, 616–628. <https://doi.org/10.1177/1073858413503712>.
136. Menon, V. (2013). Developmental pathways to functional brain networks: Emerging principles. *Trends Cogn. Sci.* 17, 627–640. <https://doi.org/10.1016/j.tics.2013.09.015>.
137. Betzel, R.F., Byrge, L., He, Y., Goñi, J., Zuo, X.-N., and Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage* 102, 345–357. <https://doi.org/10.1016/j.neuroimage.2014.07.067>.
138. Cao, M., Wang, J.-H., Dai, Z.-J., Cao, X.-Y., Jiang, L.-L., Fan, F.-M., Song, X.-W., Xia, M.-R., Shu, N., Dong, Q., et al. (2014). Topological organization of the human brain functional connectome across the lifespan. *Dev. Cogn. Neurosci.* 7, 76–93. <https://doi.org/10.1016/j.dcn.2013.11.004>.
139. Di Martino, A., Fair, D.A., Kelly, C., Satterthwaite, T.D., Castellanos, F.X., Thomason, M.E., Craddock, R.C., Luna, B., Leventhal, B.L., Zuo, X.-N., et al. (2014). Unraveling the Miswired Connectome: A developmental perspective. *Neuron* 83, 1335–1353. <https://doi.org/10.1016/j.neuron.2014.08.050>.
140. Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B., and Chahal, R. (2015). An integrative model of the maturation of cognitive control. *Annu. Rev. Neurosci.* 38, 151–170. <https://doi.org/10.1146/annurev-neuro-071714-034054>.
141. Fair, D.A., Dosenbach, N.U.F., Church, J.A., Cohen, A.L., Brahmbhatt, S., Miezin, F.M., Barch, D.M., Raichle, M.E., Petersen, S.E., and Schlaggar, B.L. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. USA* 104, 13507–13512. <https://doi.org/10.1073/pnas.0705843104>.

142. Gu, S., Satterthwaite, T.D., Medaglia, J.D., Yang, M., Gur, R.E., Gur, R.C., and Bassett, D.S. (2015). Emergence of system roles in normative neurodevelopment. *Proc. Natl. Acad. Sci. USA* 112, 13681–13686. <https://doi.org/10.1073/pnas.1502829112>.
143. Keller, A.S., Sydnor, V.J., Pines, A., Fair, D.A., Bassett, D.S., and Satterthwaite, T.D. (2023). Hierarchical functional system development supports executive function. *Trends Cogn. Sci.* 27, 160–174. <https://doi.org/10.1016/j.tics.2022.11.005>.
144. Goodkind, M., Eickhoff, S.B., Oathes, D.J., Jiang, Y., Chang, A., Jones-Hagata, L.B., Ortega, B.N., Zaiko, Y.V., Roach, E.L., Korgaonkar, M.S., et al. (2015). Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* 72, 305–315. <https://doi.org/10.1001/jamapsychiatry.2014.2206>.
145. Snyder, H.R., Hankin, B.L., Sandman, C.A., Head, K., and Davis, E.P. (2017). Distinct patterns of reduced prefrontal and limbic grey matter volume in childhood general and internalizing psychopathology. *Clin. Psychol. Sci.* 5, 1001–1013. <https://doi.org/10.1177/2167702617714563>.
146. Barch, D.M. (2017). The neural correlates of transdiagnostic dimensions of psychopathology. *Am. J. Psychiatry* 174, 613–615. <https://doi.org/10.1176/appi.ajp.2017.17030289>.
147. Kaczkurkin, A.N., Moore, T.M., Sotiras, A., Xia, C.H., Shinohara, R.T., and Satterthwaite, T.D. (2020). Approaches to defining common and dissociable neurobiological deficits associated with psychopathology in youth. *Biol. Psychiatry* 88, 51–62. <https://doi.org/10.1016/j.biopsych.2019.12.015>.
148. Mewton, L., Lees, B., Squeglia, L.M., Forbes, M.K., Sunderland, M., Krueger, R., Koch, F.C., Baillie, A., Slade, T., Hoy, N., et al. (2022). The relationship between brain structure and general psychopathology in preadolescents. *J. Child Psychol. Psychiatry* 63, 734–744. <https://doi.org/10.1111/jcpp.13513>.
149. Mattoni, M., Wilson, S., and Olino, T.M. (2021). Identifying profiles of brain structure and associations with current and future psychopathology in youth. *Dev. Cogn. Neurosci.* 51, 101013. <https://doi.org/10.1016/j.dcn.2021.101013>.
150. Romer, A.L., Elliott, M.L., Knodt, A.R., Sison, M.L., Ireland, D., Houts, R., Ramrakha, S., Poulton, R., Keenan, R., Melzer, T.R., et al. (2021). Pervasively thinner neocortex as a transdiagnostic feature of general psychopathology. *Am. J. Psychiatry* 178, 174–182. <https://doi.org/10.1176/appi.ajp.2020.19090934>.
151. Parkes, L., Moore, T.M., Calkins, M.E., Cook, P.A., Cieslak, M., Roalf, D. R., Wolf, D.H., Gur, R.C., Gur, R.E., Satterthwaite, T.D., et al. (2021). Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Transl. Psychiatry* 11, 232. <https://doi.org/10.1038/s41398-021-01342-6>.
152. Royer, J., Kebets, V., Piguet, C., Chen, J., Ooi, L.Q.R., Kirschner, M., Siffredi, V., Mistic, B., Yeo, B.T.T., and Bernhardt, B.C. (2024). Multimodal neural correlates of childhood psychopathology. *eLife* 13, e87992. <https://doi.org/10.7554/eLife.87992>.
153. Yarkoni, T. (2009). Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power—Commentary on Vul et al. (2009). *Perspect. Psychol. Sci.* 4, 294–298. <https://doi.org/10.1111/j.1745-6924.2009.01127.x>.
154. Poldrack, R.A., Baker, C.I., Durnez, J., Gorgolewski, K.J., Matthews, P.M., Munafò, M.R., Nichols, T.E., Poline, J.-B., Vul, E., and Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* 18, 115–126. <https://doi.org/10.1038/nrn.2016.167>.
155. Zuo, X.-N., Xu, T., and Milham, M.P. (2019). Harnessing reliability for neuroscience research. *Nat. Hum. Behav.* 3, 768–771. <https://doi.org/10.1038/s41562-019-0655-x>.
156. Gell, M., Eickhoff, S.B., Omidvarnia, A., Küppers, V., Patil, K.R., Satterthwaite, T.D., Müller, V.I., and Langner, R. (2024). How measurement noise limits the accuracy of brain-behaviour predictions. *Nat. Commun.* 15, 10678. <https://doi.org/10.1038/s41467-024-54022-6>.
157. Romer, A.L., Knodt, A.R., Houts, R., Brigidi, B.D., Moffitt, T.E., Caspi, A., and Hariri, A.R. (2018). Structural alterations within cerebellar circuitry are associated with general liability for common mental disorders. *Mol. Psychiatry* 23, 1084–1090. <https://doi.org/10.1038/mp.2017.57>.
158. Satterthwaite, T.D., Vandekar, S.N., Wolf, D.H., Bassett, D.S., Ruparel, K., Shehzad, Z., Craddock, R.C., Shinohara, R.T., Moore, T.M., Gennatas, E.D., et al. (2015). Connectome-wide network analysis of youth with Psychosis-Spectrum symptoms. *Mol. Psychiatry* 20, 1508–1515. <https://doi.org/10.1038/mp.2015.66>.
159. Xia, C.H., Ma, Z., Ciric, R., Gu, S., Betzel, R.F., Kaczkurkin, A.N., Calkins, M.E., Cook, P.A., la Garza, A.G.d., Vandekar, S., et al. (2018). Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat. Commun.* 9, 3003. <https://doi.org/10.1038/s41467-018-05317-y>.
160. Bassett, D.S., Xia, C.H., and Satterthwaite, T.D. (2018). Understanding the Emergence of Neuropsychiatric Disorders With Network Neuroscience. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 3, 742–753. <https://doi.org/10.1016/j.bpsc.2018.03.015>.
161. Baker, J.T., Dillon, D.G., Patrick, L.M., Roffman, J.L., Brady, R.O., Jr., Pizzagalli, D.A., Öngür, D., and Holmes, A.J. (2019). Functional connectomics of affective and psychotic pathology. *Proc. Natl. Acad. Sci. USA* 116, 9050–9059. <https://doi.org/10.1073/pnas.1820780116>.
162. Kebets, V., Holmes, A.J., Orban, C., Tang, S., Li, J., Sun, N., Kong, R., Poldrack, R.A., and Yeo, B.T.T. (2019). Somatosensory-Motor dysconnectivity spans multiple transdiagnostic dimensions of psychopathology. *Biol. Psychiatry* 86, 779–791. <https://doi.org/10.1016/j.biopsych.2019.06.013>.
163. Woo, C.-W., Chang, L.J., Lindquist, M.A., and Wager, T.D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nat. Neurosci.* 20, 365–377. <https://doi.org/10.1038/nn.4478>.
164. Feczko, E., Miranda-Dominguez, O., Marr, M., Graham, A.M., Nigg, J.T., and Fair, D.A. (2019). The heterogeneity problem: Approaches to identify psychiatric subtypes. *Trends Cogn. Sci.* 23, 584–601. <https://doi.org/10.1016/j.tics.2019.03.009>.
165. Sui, J., Jiang, R., Bustillo, J., and Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: Methods and promises. *Biol. Psychiatry* 88, 818–828. <https://doi.org/10.1016/j.biopsych.2020.02.016>.
166. Tian, Y., and Zalesky, A. (2021). Machine learning prediction of cognition from functional connectivity: Are feature weights reliable? *NeuroImage* 245, 118648. <https://doi.org/10.1016/j.neuroimage.2021.118648>.
167. Xavier, R.M., Calkins, M.E., Bassett, D.S., Moore, T.M., George, W.T., Taylor, J.H., and Gur, R.E. (2022). Characterizing youth-caregiver concordance and discrepancies in psychopathology symptoms in a US community sample. *Issues Ment. Health Nurs.* 43, 1004–1013. <https://doi.org/10.1080/01612840.2022.2099494>.
168. Jones, J.D., Boyd, R.C., Sandro, A.D., Calkins, M.E., Los Reyes, A.D., Barzilai, R., Young, J.F., Benton, T.D., Gur, R.C., Moore, T.M., et al. (2024). The general psychopathology 'p' factor in adolescence: Multi-Informant assessment and computerized adaptive testing. *Res Child Adolesc. Psychopathol.* 52, 1753–1764. <https://doi.org/10.1007/s10802-024-01223-8>.
169. Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., and Reuter, M. (2020). FastSurfer - A fast and accurate deep learning based neuroimaging pipeline. *NeuroImage* 219, 117012. <https://doi.org/10.1016/j.neuroimage.2020.117012>.
170. Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al. (2019). fMRIprep: A robust preprocessing pipeline for functional MRI. *Nat. Methods* 16, 111–116. <https://doi.org/10.1038/s41592-018-0235-4>.

171. Mehta, K., Salo, T., Madison, T.J., Adebimpe, A., Bassett, D.S., Bertolero, M., Cieslak, M., Covitz, S., Houghton, A., Keller, A.S., et al. (2024). XCP-D: A robust pipeline for the post-processing of fMRI data. *Imaging Neurosci. (Camb)* 2, imag-2-00257. https://doi.org/10.1162/imag_a_00257.
172. DeKraker, J., Haast, R.A.M., Yousif, M.D., Karat, B., Lau, J.C., Köhler, S., and Khan, A.R. (2022). Automated hippocampal unfolding for morphometry and subfield segmentation with HippUnfold. *eLife* 11, e77945. <https://doi.org/10.7554/eLife.77945>.
173. Cruces, R.R., Royer, J., Herholz, P., Larivière, S., Vos de Wael, R., Paquola, C., Benkarim, O., Park, B.-Y., Degré-Pelletier, J., Nelson, M. C., et al. (2022). Micapipe: A pipeline for multimodal neuroimaging and connectome analysis. *NeuroImage* 263, 119612. <https://doi.org/10.1016/j.neuroimage.2022.119612>.
174. Cox, R.W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. <https://doi.org/10.1006/cbmr.1996.0014>.
175. Muthén, L.K., and M.B.O. (2017) *Mplus User's Guide: Statistical Analysis with Latent Variables: User's Guide*, [Eighth Edition].
176. Hallquist, M.N., and Wiley, J.F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in mplus. *Struct. Equ. Modeling* 25, 621–638. <https://doi.org/10.1080/10705511.2017.1402334>.
177. Hu, L.T., and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. <https://doi.org/10.1080/10705519909540118>.
178. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31, 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>.
179. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178. <https://doi.org/10.1038/nature18933>.
180. Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., and Petersen, S.E. (2016). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex* 26, 288–303. <https://doi.org/10.1093/cercor/bhu239>.
181. Developers, C.-P. (2022) CPAC Pipeline Configuration YAML File: RBCv0.
182. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289. <https://doi.org/10.1006/nimg.2001.0978>.
183. Developers, T.N. (2021) *init_brain_extraction_wf*. NeuroImaging workflows (NiWorkflows): robust processing tools for MRI data.
184. Esteban, O., Markiewicz, C.J., Goncalves, M., Blair, R., Berleant, S.L., Poldrack, R.A., and Gorgolewski, K.J. (2020) *NIWorkflows*: NeuroImaging workflows.
185. 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., Gillen, D.L., et al. (2021). The ANTsX ecosystem for quantitative biological and medical imaging. *Sci. Rep.* 11, 9068. <https://doi.org/10.1038/s41598-021-87564-6>.
186. Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. <https://doi.org/10.1109/42.906424>.
187. Neuroscience, C.f.R. (2023). *EstimateReferenceImage*. NeuroImaging workflows (NiWorkflows): robust processing tools for MRI data.
188. Esteban, O., Markiewicz, C.J., Goncalves, M., Blair, R., Berleant, S.L., Poldrack, R.A., and Gorgolewski, K.J. (2021) *NIWorkflows*: NeuroImaging workflows.
189. Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. [https://doi.org/10.1016/s1361-8415\(01\)00036-6](https://doi.org/10.1016/s1361-8415(01)00036-6).
190. Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17, 825–841. [https://doi.org/10.1016/s1053-8119\(02\)91132-8](https://doi.org/10.1016/s1053-8119(02)91132-8).
191. developers, T.f. (2020). *init_bold_std_trans_wf*. fMRIPrep: A Robust Preprocessing Pipeline for fMRI Data.
192. 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>.
193. Ciric, R., Thompson, W.H., Lorenz, R., Goncalves, M., MacNicol, E.E., Markiewicz, C.J., Halchenko, Y.O., Ghosh, S.S., Gorgolewski, K.J., Poldrack, R.A., et al. (2022). TemplateFlow: FAIR-sharing of multi-scale, multi-species brain models. *Nat. Methods* 19, 1568–1571. <https://doi.org/10.1038/s41592-022-01681-2>.
194. Neuroscience, C.f.R. (2023). *init_enhance_and_skullstrip_bold_wf*. NeuroImaging Workflows (NiWorkflows): Robust Processing Tools for MRI Data.
195. Neuroscience, C.f.R. (2023). *TPM2ROI*. NeuroImaging Workflows (NiWorkflows): Robust Processing Tools for MRI Data.
196. Developers, C.-P. (2023). *Timeseries extraction*. C-PAC v1.8.5 documentation.
197. Developers, T.N. (2018) *Computing Functional Connectivity Matrices*. Nilearn.
198. Nilearn, c., Chamma, A., Frau-Pascual, A., Rothberg, A., Abadie, A., Abraham, A., Gramfort, A., Savio, A., Cionca, A., Sayal, A., et al. (2025). *nilearn*.
199. Developers, C.-P. (2023). *create_alff*. C-PAC v1.8.5 documentation.
200. Developers, C.-P. (2023). *Amplitude of Low Frequency Fluctuations (ALFF) and fractional ALFF (fALFF)*. C-PAC v1.8.5 documentation.
201. Brett, M., Hanke, M., Markiewicz, C., Côté, M.-A., McCarthy, P., Cheng, C., Halchenko, Y., Ghosh, S., Wassermann, D., Gerhard, S., et al. (2019). *nipy/nibabel*: 2.3.3.
202. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al. (2020). Array programming with NumPy. *Nature* 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>.
203. Developers, C.-P. (2023). *Regional Homogeneity*. C-PAC v1.8.5 documentation.
204. Developers, C.-P. (2023). *eXTensible Connectivity Pipeline-style quality control files*. C-PAC v1.8.5 documentation.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Reproducible Brain Charts (RBC) data resource	This paper; accompanying website: https://reprobrainchart.github.io	N/A
Brazil High Risk Cohort (BHRC)	Salum et al. ⁷³	N/A
Developmental Chinese Color Nest Project (CCNP)	Liu et al. ⁷⁴	N/A
Healthy Brain Network (HBN)	Alexander et al. ¹⁷	N/A
Nathan Kline Institute–Rockland Sample (NKI)	Tobe et al. ¹⁸	N/A
Philadelphia Neurodevelopmental Cohort (PNC)	Satterthwaite et al. ¹⁶ ; Satterthwaite et al. ⁷⁶	N/A
Software and algorithms		
CuBIDS	Covitz et al. ⁸⁰	N/A
FreeSurfer	Fischl ⁵²	RRID: SCR_001847
sMRIPrep	Esteban et al. ⁸³	N/A
C-PAC	Cameron et al. ⁵³	RRID: SCR_000862
fMRIPrep	Esteban et al. ¹⁷⁰	RRID: SCR_016216
Swipes for Science	Keshavan et al. ⁹⁰	N/A
DataLad	Halchenko et al. ⁵⁵	RRID: SCR_003931
AFNI tools	Cox ¹⁷⁴	RRID: SCR_005927
Python		RRID: SCR_008394
RStudio		RRID: SCR_000432

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study description

The RBC project contains data from 5 large studies of brain development (age range: 5–85 years old) from $N=6,346$ participants ($N=2,869$ Female). All studies included structural and functional Magnetic Resonance Imaging (MRI) data as well as phenotypic data. Studies were conducted in Brazil, China, and the United States of America. Specific studies include: Brazilian High Risk Cohort (BHRC⁷³; $n=610$), Developmental Chinese Color Nest Project (CCNP^{74,75}; $n=195$), Healthy Brain Network (HBN¹⁷; $n=2,611$), Nathan Kline Institute–Rockland Sample (NKI¹⁸; $n=1,329$), and Philadelphia Neurodevelopmental Cohort (PNC^{16,76}; $n=1,601$).

BHRC⁷³ is a sample of children and adolescents attending school in Brazil (Porto Alegre and São Paulo cities) that is aimed to be a random sample of the state-funded school-based community in addition to children with increased family risk of mental disorders. The study was approved by the ethics committee of the University of São Paulo.

CCNP^{74,75} is a sample of children and adolescents that is aimed to represent the population residing in multiple cities of China with varying economies and from different regions of the country (<https://ccnp.scidb.cn/en>). The ethical approval for this study was obtained from the Institutional Review Board of the Chinese Academy of Sciences (CAS) Institute of Psychology and Beijing Normal University.

HBN¹⁷ is a sample of children and adolescents residing in the New York City area (United States of America) that is aimed to represent a diverse sample of healthy and help-seeking individuals with heterogeneous metrics of developmental psychopathology. The HBN study included data from four different acquisition sites: Staten Island (SI - Mobile Scanner), Rutgers University (RU), The City University of New York (CUNY), and Citigroup Biomedical Imaging Center (CBIC). The study was approved by the Chesapeake Institutional Review Board.

NKI–Rockland Sample¹⁸ is aimed to represent a lifespan sample of individuals with varying demographic distributions residing in the United States. The Institutional Review Board approved this project at the Nathan Kline Institute.

PNC^{16,76} is a community sample of children and adolescents residing in the greater Philadelphia area (United States) that is aimed to represent a diverse developmental sample. The study was approved by the Institutional Review Boards of the University of Pennsylvania and the Children’s Hospital of Philadelphia.

Written informed consents were obtained from all participants (or their parents or legal guardians) by each study separately as part of the study-specific data collection procedure.

T1-weighted structural MRI and resting-state, task, and movie-watching functional MRI data were included in the RBC dataset. All scans were defaced and de-identified to ensure ethical compliance and protect participants' privacy. Data acquisition parameters varied by study and data acquisition sites. Detailed information about structural and functional MRI data acquisitions are summarized in [Tables S4–S5](#).

METHOD DETAILS

Phenotypic data harmonization

RBC includes psychiatric phenotyping data for each study, which were assessed using one of two different phenotypic questionnaires. The Child Behavior Checklist (CBCL⁷⁷) was used in the BHRC, CCNP, HBN, and NKI studies, while the GOASSESS interview⁷⁸ was used in the PNC study. To ensure that measures of psychopathology were consistent across study, we used a bifactor modeling strategy to harmonize differences between samples that used the same instrument (i.e., the CBCL) as well as differences between disparate instruments (i.e., GOASSESS vs. CBCL). The CBCL is a 120-item parent-report assessment of emotional and behavioral symptoms over the past 6 months, answered on a 3-point scale (0=not true, 1=somewhat/sometimes true, and 2=very true/often). It encompasses eight syndromes: anxious-depressed, withdrawn-depressed, somatic complaints, rule-breaking behavior, aggressive behavior, social problems, thought problems, and attention problems.⁷⁷ To harmonize CBCL with GOASSESS, CBCL scores of 1 and 2 were collapsed to generate a binary-scaled variable compatible with GOASSESS (i.e., 0 or 1). The GOASSESS is a structured screening interview administered to collateral informants (usually a caregiver) by trained assessors. It contains 112 unconditioned screening items based on DSM-IV constructs, including symptoms of mood disorders (Major Depressive Episode, Manic Episode), anxiety disorders (Generalized Anxiety Disorder, Separation Anxiety Disorder, Specific Phobia, Social Phobia, Panic Disorder, Agoraphobia, Obsessive-Compulsive Disorder, Post-traumatic Stress Disorder), Attention Deficit/Hyperactivity Disorder (ADHD), behavioral (Oppositional Defiant Disorder, Conduct Disorders) and eating disorders (Anorexia, Bulimia), and suicidal thinking and behavior. Items are scored as 0 (absent) or 1 (ever present). The instrument is abbreviated and modified from the epidemiologic version of the NIMH Genetic Epidemiology Research Branch Kiddie-SADS, and its development is described and tested elsewhere.⁷⁸

We tested a sequence of harmonization procedures to minimize between-cohort and between-questionnaire differences across studies, as well as to disentangle general and specific aspects of psychopathology. In a previous work, we thoroughly tested the impact of different bifactor model configurations on the resulting factor scores⁴⁴ in addition to six item-matching strategies that best harmonize different questionnaires.⁴² We then tested the best bifactor model configuration using several parameters to identify the model that best harmonized CBCL and GOASSESS questionnaires used in RBC in another previous work.⁴³ In brief, 12 CBCL bifactor models were first identified from previous literature, which varied in item and factor configurations.⁴⁴ The impact of these modeling choices was small for the general psychopathology factor (i.e., p -factor), but quite marked for the specific factors.⁴⁴ We then tested 6 item-matching strategies to harmonize items between questionnaires (i.e., CBCL and GOASSESS), where we found that the expert-based 1-to-1 semantic item-matching performed best for item harmonization.⁴² Finally, the extent to which the CBCL–GOASSESS harmonized models were similar to the original models was assessed across different models.⁴³ We selected the McElroy model as it: (1) demonstrated measurement invariance between the questionnaires; (2) retained the majority of original items during the harmonization process; (3) included harmonized items that were endorsed in all samples; and (4) was among the best four harmonized bifactor models in terms of factor reliability and authenticity (i.e., generated factor scores fairly correlated and close to the factor scores from full item set models).^{42–44}

In our previous RBC harmonization studies and for the publicly released RBC dataset, all factor scores were estimated with confirmatory factor analysis (CFA) using delta parameterization and Weighted Least Squares with diagonal weight matrix with standard errors and mean- and variance-adjusted Chi-square test statistics (WLSMV) estimators. RBC studies were used as clusters in the CFA in the bifactor model that included all RBC data. Analyses were carried out in Mplus 8.6¹⁷⁵ and implemented in R version 4.0.3 using the MplusAutomation package,¹⁷⁶ which was also used to extract factor scores generated in Mplus using maximum a posteriori method. The global model fit was evaluated using root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker–Lewis index (TLI), and standardized root mean square residual (SRMR). RMSEA lower than 0.060 and CFI or TLI values higher than 0.950 indicate a good-to-excellent model. SRMR lower than or equal to 0.080 indicates acceptable fit, and lower than 0.060 in combination with previous indices indicates good fit.¹⁷⁷ Only individuals that filled the CBCL and GOASSESS on the day of neuroimaging were included in the publicly released RBC dataset.

Reproducible neuroimaging data curation and workflows

We curated all raw neuroimaging data and metadata in a fully-reproducible fashion that conforms with the Brain Imaging Data Structure (BIDS⁷⁹) using the Curation of BIDS (CuBIDS⁸⁰) software package. CuBIDS provides a workflow for identifying unique combinations of imaging data acquisition parameters based on metadata, summarizing the heterogeneity in an MRI BIDS dataset, and reproducibly modifying scan filenames to reflect information about their imaging parameters. These curation steps are critical given that preprocessing pipelines designed for BIDS datasets (e.g., “BIDS-apps”⁷⁹) automatically configure workflows based on a

dataset's accompanying metadata. However, potential inaccuracies or missing information in metadata may lead to incorrectly configured pipelines that will run without any errors (i.e., pipelines that are reproducible but wrong). CuBIDS facilitates the identification of such inaccuracies or missing information in metadata and categorizes imaging data in subgroups (i.e., unique parameter groups) based on the heterogeneity of their metadata.

We applied CuBIDS to RBC data and generated a CuBIDS summary data file for each study included in the RBC dataset. The CuBIDS summary tables are provided along with the RBC BIDS data, available at <https://github.com/ReproBrainChart>. CuBIDS was then used to rename each BIDS data file based on the scanning parameters of the identified parameter groups, such that the new file names indicated the source of the variance between different parameter groups. Notably, all data curation steps were fully tracked via CuBIDS's wrapped use of DataLad⁵⁵ throughout, ensuring the tracking of version history and yielding a complete audit trail. After categorizing data based on heterogeneity of their acquisition parameters, we used an automated procedure implemented in CuBIDS to create an example dataset with one subject from each acquisition group for each study. The example dataset was then used to closely examine the heterogeneity of data and make necessary adjustments to the data curation step (e.g., if there were errors or missing information in metadata). Eventually, the example dataset was used to test image processing pipelines to ensure that they performed well on each combination of acquisition parameters present in the dataset. Finally, to run the RBC neuroimaging data through modality specific preprocessing pipelines, we employed the FAIRly-big workflow.⁵⁴ FAIRly-big is a DataLad-based open-source framework that is suitable for reproducible processing of large-scale datasets. We adapted the FAIRly-big workflow for RBC image processing to ensure that all data preparation and processing were accompanied by a full audit trail in DataLad.⁵⁵

Structural MRI data processing and quality control

Structural MRI (sMRI) data were processed using FreeSurfer v6.0.1⁵² and sMRIPrep v0.7.1,⁸³ yielding commonly used measures of brain structure. Specifically, structural images underwent correction for intensity non-uniformity and skull-stripping with ANT's brain extraction workflow. Brain surfaces were then reconstructed using FreeSurfer. RBC provides full FreeSurfer outputs as well as tabulated data parcellated using 35 anatomical, functional, and multimodal atlases that were included to align with the functional image processing. Atlases include the Desikan Killiany¹⁷⁸, Glasser¹⁷⁹, Gordon,¹⁸⁰ and multiple resolutions of the Schaefer⁹⁶ parcellation among others. Specific features include commonly used measures of brain structure such as regional surface area, cortical thickness, gray matter volume, and folding and curvature indices. Moreover, summary brain measures such as total intracranial volume, ventricle size, and mean and standard deviation of various measures (e.g., cortical thickness, surface area) are provided for the whole brain and per hemisphere. Tabulated data are also accompanied by .json files describing each structural feature in detail.

An important feature of RBC is the emphasis on harmonized measures of neuroimaging data quality control (QC). To achieve consistent quality ratings across all studies in RBC, every structural image was manually evaluated by 2-5 expert raters using Swipes for Science, a web application for binary image classification.⁹⁰ The expert rating workflow consisted of seven phases (Figure 3). Overall, 4 two-dimensional slices per participant (two axial and two sagittal slices extracted per structural scan) were created to rate each scan. To ensure consistent slice selection across participants, the anatomical images were registered (linear rigid body) to the MNI152 template space prior to slice selection. Expert raters (experience in brain imaging: range = 0.17-16 years; mean = 5.63 years; SD = 5.76 years) evaluated a total of 28,780 slices and assigned "Pass" or "Fail" to those slices based on visual inspection only. A "Pass" rating would be given if an image was deemed of sufficient quality for skullstripping and segmentation of cerebrospinal fluid (CSF), white, and gray matter. Slices were presented to the raters in random order.

More specifically, Phase 1 involved generating the ground truth for 200 images to either "Pass" (rating of 1) or "Fail" (rating of 0). Two senior raters evaluated these images and agreed on 96% of the slices. The images with disagreement were further evaluated after skullstripping and brain segmentation, using Configurable Pipeline for the Analysis of Connectomes (C-PAC⁵³) pipeline. The expert raters then reached 100% concordance through consensus of either "Pass" or "Fail" for each slice. In Phase 2, these ground-truth annotated images (i.e., outcome of Phase 1) were used to train four raters until they reached at least 85% concordance with the ground-truth. In Phase 3, all raters manually rated 10% of images and the reliability of ratings was assessed across raters. In Phase 4, the raters were divided into two groups with balanced experience and evaluated 20% of the remaining images independently, such that no overlapping images were assessed by the two groups. In Phase 5, the reliability of ratings was re-assessed (similar to Phase 3), where all raters manually rated another 10% of the remaining images. In Phase 6, the raters were divided into two groups again to rate 20% of the remaining images (similar to Phase 4). Finally, in Phase 7, a new set of images that were not included in the first 6 phases were evaluated by raters. To avoid sequence type or scanner parameters biases, each Swipes for Science instance for each phase contained data from all data acquisition sites and studies.

Following manual expert ratings, an overall QC determination score of "Pass", "Artifact", or "Fail" was assigned to each structural image as the final scan quality based on the average rating across raters. "Pass" or "Fail" labels were assigned to images if all raters agreed about their quality. For example, if all raters agreed that a given image was of adequate quality (i.e., all raters assigned "Pass" or rating of 1 to the image), the average rating across raters was equal to 1 and the image was labeled as "Pass". Similarly, if all raters assigned "Fail" or rating of 0 to an image, the image was labeled as "Fail" (i.e., average rating of 0 across raters). If there was any disagreement between raters about a given image's quality (i.e., a combination of 1 and 0 ratings for the image), the "Artifact" label was assigned to that image (i.e., an average value between 0 and 1). Summary information on the number of participants with structural scans before and after applying RBC's recommended QC is available in Figure S1.

In addition to the manual ratings, we calculated the Euler number from FreeSurfer.⁸⁸ The Euler number is calculated as part of the FreeSurfer processing pipeline and reflects the topological complexity of the initial reconstructed cortical surface.⁸⁸ Euler number is calculated as the sum of the vertices and faces subtracted by the number of faces. Lower values below an ideal value of 2 indicate more defects in surface reconstruction. Euler number can be used as a continuous measure of structural MRI QC (e.g., can be included as a covariate in downstream analysis) whereas manual ratings provide categorical QC labels. Previous reports have demonstrated that the Euler number is an accurate, fully-automated measure of data quality, capturing variation in data quality within the coarse categories provided by manual ratings.^{18,88,91–93}

Functional MRI data processing and quality control

As noted above, preprocessing steps were carried out in DataLad⁵⁵ using the FAIRly-big workflow⁵⁴ to track provenance and ensure the re-executability of processing workflows for all RBC studies. Following guidelines from extensive benchmarking and harmonization studies,⁵¹ functional MRI (fMRI) data were preprocessed using Configurable Pipeline for the Analysis of Connectomes (C-PAC v1.8.5.dev1⁵³), with a pipeline configuration file specifically developed for RBC¹⁸¹ including a fixed random state to maximize reproducibility. This configuration produces commonly used measures of functional MRI data such as fully processed fMRI time-series and functional connectivity matrices (e.g., Pearson and partial correlations between processed regional time-series). In addition, the functional data derivatives include regional measures based on processed time-series such as Regional Homogeneity (ReHo⁸⁴), Amplitude of Low Frequency Fluctuation (ALFF⁸⁵), and fractional ALFF (fALFF⁸⁶). Processed time-series and functional connectivity are available in parcellated format using 17 different atlases, including AAL¹⁸², Glasser¹⁷⁹ and Schaefer⁹⁶ parcellations. ReHo, ALFF, and fALFF are available in volumetric MNI space (MNI152NLin6ASym).

For anatomical preprocessing, the configuration specified C-PAC's reimplementation of the NiWorkflows ANTs brain extraction workflow^{183–185} and segmentation using FSL-FAST¹⁸⁶ with a 0.95 threshold for each tissue type. Anatomical registration was configured to use antsRegistration¹⁸⁵ on skull-stripped images at 1mm isotropic resolution with parameters specified in the configuration file.¹⁸¹

Functional registration was configured to first use C-PAC's reimplementation of the NiWorkflows reference image estimation method^{187,188} with boundary-based registration performed on skull-stripped images to MNI152NLin6ASym-template space using FSL-FLIRT^{189,190} with binarized partial volume white matter masks. It then used C-PAC's reimplementation of the fMRIprep single-step resampling workflow^{170,191} at 2mm isotropic resolution. Motion statistics were calculated before slice-timing correction using the previously established reference image and FSL-MCFLIRT¹⁹⁰ for motion estimation. Where field maps were present, distortion correction was performed using FSL-FUGUE¹⁹² or FSL-TOPUP¹⁹² depending on the type of field maps. Functional masking was configured to use a reference image from TemplateFlow¹⁹³ and C-PAC's reimplementation of the NiWorkflows BOLD masking method.^{188,194} A mean functional image was also generated using AFNI 3dTstat.¹⁷⁴ The first 2 timepoints were excluded from analysis.

AFNI 3dDespike¹⁷⁴ was run on template-space images. Prior to nuisance regression, the brain mask, CSF mask, and white matter mask were eroded using C-PAC's reimplementation of the NiWorkflows utility interface for converting tissue probability masks into regions of interest.^{188,195} Two nuisance regression strategies (named “36-parameter” and “aCompCor”, with global signal regression (GSR) or aCompCor as part of the strategy, respectively and exclusively) were run in parallel in template space (the full specific nuisance regression configurations can be found in the configuration file¹⁸¹).

Time-series extraction¹⁹⁶ was performed, and correlation matrices were generated using Nilearn's “correlation” and “partial correlation” methods^{197,198} separately for each of the specified region of interest atlases. ALFF⁸⁵ and fALFF⁸⁶ were calculated in template space using a sequence of AFNI tools.^{174,199,200} Regional Homogeneity⁸⁴ was calculated in template space using NiBabel²⁰¹ and NumPy²⁰² in Python.²⁰³

C-PAC outputs also include extensive measures of quality control, such as various in-scanner motion parameters (e.g., framewise displacement), functional image to structural image (e.g., to T1-weighted scan or MNI template) registration and normalization quality parameters, generated for each preprocessed BOLD image using C-PAC's reimplementation²⁰⁴ of the xcpEngine quality control dictionary.¹⁹³ As part of the RBC data release, we performed an initial quality assurance procedure using measures of in-scanner motion quantified as Framewise Displacement (FD) and normalization quality quantified as cross correlation from normalization of functional image to template image. Functional MRI runs with a median FD less than or equal to 0.2 and normalized cross correlation greater than or equal to 0.8 were considered of adequate quality. These thresholds were selected after extensive manual review of images. Similar to structural data, the majority of functional scans passed RBC's QC thresholds. Summary information on the number of participants with functional scans before and after applying RBC's recommended QC is available in Figure S1. Overall, about 90% of RBC data passed both structural and functional QC guidelines.

Following the initial structural and functional data QC procedure, three different versions of the RBC dataset were publicly released and are accessible depending on the user's choice of QC threshold: (1) structural images with “Pass” label only and functional images that passed the motion and image normalization QC threshold; (2) structural images with “Pass” and “Artifact” labels and functional images that passed the motion and image normalization QC threshold; (3) all structural and functional images including QC failures. Versions (1) and (2) are recommended by RBC; however, users may use Version (3) for research on QC or to apply their desired QC procedures using the extensive structural and functional QC information accompanying RBC data.

QUANTIFICATION AND STATISTICAL ANALYSIS

Example analysis workflow with RBC data

In addition to providing a publicly available data resource, we illustrated the utility of RBC data and the increased statistical power of the aggregated sample. Specifically, we examined the relationship between derived structural and functional neuroimaging features and both participant age and general psychopathology. Furthermore, we evaluated whether those associations are influenced by RBC's QC protocols and imaging data harmonization.

Neuroimaging data harmonization

RBC provides both raw and processed neuroimaging data for each dataset. However, due to differences in scanners and sequences used in data collection between studies, there is significant technical variance between studies and sites that requires harmonization. Therefore, users may want to statistically harmonize neuroimaging data across studies and data acquisition sites. However, neuroimaging data harmonization remains an ongoing research topic in the field with a growing number of statistical methods that aim to effectively harmonize data between multiple sites and studies. Given the lack of consensus on a single neuroimaging data harmonization technique, we released unharmonized neuroimaging data in RBC. This allows researchers to implement their method of choosing to harmonize the imaging data based on their specific research questions. In addition, given that statistical harmonization models require specification of covariates that are hypothesis-dependent, researchers can define the covariates they want to include in the harmonization analysis. While RBC's released neuroimaging data is unharmonized, below we provide an example workflow to harmonize the imaging data and illustrate its utility in our analysis. This workflow can be tailored by RBC users for their specific hypotheses.

We used CovBat-GAM^{57–61} to harmonize structural and functional MRI data across data acquisition sites. Specifically, we used the “covfam” function with Generalized Additive Model (GAM) from ComBatFamily in R 4.2.2. The “covfam” function uses CovBat⁶¹ (Correcting Covariance Batch Effects) to harmonize the mean and covariance of data across multiple batches (e.g., sites, datasets). In our developmental analysis, data acquisition sites were treated as batches. Covariates included in harmonization were age as a smooth term (to account for linear and nonlinear age effects), and sex and data quality (i.e., Euler number for structural data and median FD for functional data) as linear terms. We used a separate harmonization for the psychopathology analyses, where general psychopathology from the bifactor model (p -factor) was added as a linear term to the model. CovBat-GAM yields harmonized structural and functional data while protecting effects of model covariates. To evaluate the effect of harmonization, we repeated analysis before and after data harmonization.

Generalized Additive Models (GAM)

We used Generalized Additive Models (GAMs) to delineate linear and nonlinear developmental effects (i.e., age effects) and assess the relationship between neuroimaging data derivatives and general psychopathology (i.e., p -factor).^{60,94,95} The analysis was performed using the “mgcv” package in R 4.2.2. To ensure all studies contributed to the analysis, a subset of aggregated RBC data within the age range of 6–22 years old was used for this analysis. Structural features included cortical thickness (CT), surface area (SA), and gray matter volume (GMV), and functional features included between- and within-network resting-state functional connectivity. Structural and functional data parcellated into 400 brain regions using the Schaefer-400 atlas.⁹⁶ Functional networks were defined using the 7 Yeo-Krienen networks for the parcellated functional data.⁹⁷ Within-network connectivity was estimated as the average functional connectivity between regions from a given network and all the other regions from the same network. Between-network connectivity was estimated as the average connectivity between regions from a given network and regions from all the other networks. Note that we used the baseline functional scans of individuals who were scanned at multiple time points (e.g., BHRC and NKI). For individuals with multiple functional data at the baseline, we used the average functional connectivity across scans at the baseline.

Separate GAMs were fit for functional and structural features as dependent variables. Each GAM included age as a smooth term and sex and data quality as linear covariates. The linear covariate for data quality was the Euler number for structural data and median FD for functional data. For models evaluating associations with overall psychopathology, the p -factor was also included in GAMs as a linear term. In all GAM analyses, the maximum basis complexity was set to 3 for the smooth term (i.e., age) to avoid overfitting. The code invocation used to fit the model was: ‘mgcv::gam(feature ~ s(age, k=3, fx=F) + factor(sex) + data_quality)’. The GAM formula also included p -factor as a linear term for psychopathology analysis. The Restricted Maximum Likelihood (REML) approach was used to estimate the smoothing parameters in GAMs. We performed the analyses at two resolutions: (1) whole-brain models, where we examined average structural and functional features across the cortex for each participant; (2) regional and network level GAMs, where we evaluated regional structural features and network-level functional features (e.g., mean within- or between-network connectivity) for each participant.

The effect size of associations between imaging features and age were estimated as partial R^2 . We quantified partial R^2 as the difference in Residual Sum of Squares (SSE) between the full model that included the smooth term for age and a reduced model with no age term (i.e., only including model covariates) normalized by SSE of the reduced model. Normalization by SSE of the reduced model highlights the relative contribution of the predictor (e.g., age) to the reduced model. The linear effect size for the relationship between p -factor and neuroimaging features was also estimated in a similar manner, where partial R^2 was defined as the difference in SSE between the full model and a reduced model without p -factor. We used a signed version of partial R^2 , such that the sign reflected the directionality of observed effects (increase or decrease in neuroimaging features with increasing age or p -factor). To obtain

the directionality of partial R^2 for age analysis, we calculated the mean derivative of the model fit for the smooth term (i.e., age). A positive sign was assigned to partial R^2 if the mean derivative was positive, reflecting an overall increasing trend between the neuroimaging feature of interest and age, whereas a negative sign was assigned to partial R^2 where the mean derivative was negative. To obtain the directionality of partial R^2 for p -factor analysis, we used the sign of the t -value associated with the linear p -factor term. The statistical significance of the associations between neuroimaging features and age or p -factor was assessed using analysis of variance (ANOVA) between the full model and reduced model that excluded the term of interest. Results were corrected for multiple comparisons by controlling for the false discovery rate (FDR correction; $Q < 0.05$).

To assess how neuroimaging data QC and harmonization impacted our findings, we repeated all structural and functional GAM analyses with three different versions of data: First, we used all neuroimaging data without excluding individuals based on QC thresholds and without applying neuroimaging data harmonization. Second, we used structural MRI data from individuals with “Pass” or “Artifact” structural QC (i.e., excluding scans with “Failed” structural QC) and functional data that passed functional QC thresholds (median FD < 0.2 and normalization cross correlation > 0.8). However, we did not apply neuroimaging data harmonization in this version. Third and finally, we used neuroimaging data that passed structural and functional QC (as above) and performed GAM analyses after applying neuroimaging data harmonization.

Finally, we tested how participant sex and its interaction with age relate to harmonized structural and functional features. Similar to the original analysis, separate GAMs were fit for functional and structural features as dependent variables using within- or between-network connectivity for functional and CT, SA, and GMV for structural data. We first assessed the effects of participant sex on neuroimaging features by including biological sex as a linear two-level factor in GAMs. The effect size for the sex term was estimated in a similar manner to the original analysis, where partial R^2 was defined as the difference in SSE between the full model and a reduced model without sex for each cortical region for structural data and for resting-state networks for functional data. The directionality of the observed effects was assessed using the sign of the t -value associated with the sex term, such that a positive value corresponded to higher values in females than males. Multiple comparisons were accounted for using FDR. We also repeated the analysis after controlling for global cortical effects by including mean CT (for CT analysis) and total intracranial volume (TIV, for SA and GMV analyses) as model covariates.

We next examined how the interaction between biological sex and age influences the developmental patterns. Specifically, we added a factor-smooth interaction term to GAMs to estimate how age-related changes vary by sex. For structural data, we assessed the sex-by-age interaction on the whole-brain level using the average CT, SA, and GMV maps, as well as on the regional level. We also repeated the regional analysis after controlling for global cortical effects by including mean CT (for CT analysis) and TIV (for SA and GMV analyses) as model covariates. For functional data, we assessed the sex-by-age interaction on within- and between-network connectivity. The interaction effect size was quantified using the F -statistic from GAMs. All reported statistics were corrected for multiple comparisons using FDR.