FAQ

PUBLISHED
March 14, 2025

Data access/use

How do I access the data?

Our data are publicly shared with eligible researchers with a valid research use of the data at a research institution. All sharing of the data, as well as access to it, is handled by the NIMH Data Archive (NDA). The process for accessing the data starts by creating an account at NDA (if you do not already have one), and then requesting access to the ABCD Study data through the dashboard there (https://nda.nih.gov/user/dashboard/data_permissions.html).

This will generate a Data Use Agreement/Certification (DUC) that you can fill out with your research aims. The DUC also outlines all of the rules you must follow when using ABCD data. You will then sign this document electronically, and submit it to an Institutional Signing Officer at your institution for their signature as well. The exact office that handles these agreements differs by institution; for example, at UCSD this is handled by the Office of Contracts and Grants Administration.

You will then upload this document signed by you and your institution to the same dashboard you started the process with, and NDA will review your application. You will then be able to download the data from NDA, and I would encourage you to learn more about the process at our Wiki: https://wiki.abcdstudy.org/. You can find a bit more about the data sharing timelines on our website here: https://abcdstudy.org/scientists/data-sharing/. If you have any questions along the way, the NDA help desk (ndahelp@mail.nih.gov) can help answer any further questions.

I'm from a non-US institution - can I access the data?

Our data are publicly shared with eligible researchers with a valid research use of the data, and who are at an institution with an active Federal Wide Assurance, which many international institutions have (https://ohrp.cit.nih.gov/search/fwasearch.aspx?styp=bsc). Users from many other countries have successfully accessed and published with the ABCD Study data.

Is there a cost to access the data?

Does NDA still require an NDA study for public disclosure of ABCD study results?

Please note that NDA no longer requires an NDA Study to be created for public disclosure of results pertaining to ABCD analysis. This will be reinstated in future data releases.

Can I have a student or trainee on my DUC? I'm a student, can I have my own DUC?

In general, each person is able to get their own data access, but you may also list collaborators on a "Parent" DUC that then grants access to the data to all collaborators listed on the DUC. That said, it is up to your institution to determine whether they will "co-sign" an independent DUC for a student, though generally graduate students, postdocs, etc., have been able to have their own DUC. Some groups have the PI submit a DUC with trainees and other lab members listed as collaborators. If you go that route, then you just need to make sure the primary DUC holder doesn't forget to renew their DUC annually, otherwise all others on the DUC will lose access along with them when it expires.

If you go this route of having one DUC with collaborators listed, be sure that all collaborators also have NDA accounts, and list their email addresses associated with their NDA account, because that way when the DUC is approved by NDA, those people will be able to access the downloads themselves.

Do I need IRB approval to use these data?

Institutions vary on whether they consider use of the de-identified dataset to be human subjects research, with some requiring expedited-style or even exempt IRB reviews by their institution's IRB. Other institutions do not consider it to be human subjects research given the de-identified nature of the data. That is a question to ask your IRB.

I am writing a proposal to use ABCD data, can you write a letter of support?

Can I use ChatGPT or another generative AI tool to create figures and graphs representing ABCD data?

No, this is not permitted. Inputting ABCD data to generative AI tools, such as ChatGPT, is a violation of the terms of use outlined in the data use agreement.

Data download

I am having issues with downloading the .zip files on the NDA website?

The size of the ABCD release 5.0 raw genotyping data seems to differ from what is stated.

How can I download the ABCD 4.0 data?

Where can I download the ABCD raw data?

Protocol

Where can I find the protocol for the ABCD study?

You can find our protocol on the scientist section of our website here (https://abcdstudy.org/scientists/protocols/), but in general we are not able to share the actual measures outside of the study, as many are proprietary. You can read a lot more about the protocol development in a special issue of DCN on the ABCD Study here: https://www.sciencedirect.com/journal/developmental-cognitive-neuroscience/vol/32. Finally, you can also generally find the questions asked and answer options as part of the protocol on our new interactive Data Dictionary (https:// data-dict.abcdstudy.org/) and find more information on the ABCD Wiki: https://wiki.abcdstudy.org/

I am interested in proposing a new assay for collected bio-specimens?

Autism in ABCD Study sample

Imaging

Why is the cerebellum sometimes cutoff in fMRI and dMRI?

Where are the dMRI gradient tables?

When should I use field maps?

The “minimally processed” (mproc) images have been corrected for B0 distortion using the field maps. As a result, field maps are not shared in the mproc data releases.

If you are obtaining fast track data (i.e., unprocessed), you would want to correct using the field maps provided in Fast Track.

To know which field map goes with a particular scan, here are some pointers:

- The study date and series time are included in the fast track file names.
- The protocol calls for two scans per task-fMRI session (i.e., two for MID, two for SST, and two for nBack).
- Each pair scans is typically proceeded by a field map scan (or pair of field map scans).
- On GE scanners, the field map is an integrated sequence of forward and reverse phase-encode polarity scans (i.e., one field map scan for GE). On Philips and Siemens scanners, the forward and reverse scans are separate series (i.e., two field map scans for Philips and Siemens).

Here are some extra things to know:

- Sometimes field map scans have image quality issues and may fail raw QC. Those will be excluded from Fast Track downloads if you access the "recommended" Fast Track query.
- If there is no usable field map scan directly preceding a scan or pair of scans, we use the next closest field map in the imaging session.
- Regardless of whether the field map is collected immediately before the pair of tasks scans, or earlier, or later, it is advisable to correct for motion between the field map and the main scan.
- If you use FSL's TOPUP, it corrects for motion between the forward and reverse phase-encode polarity volumes, which occurs in many subjects as well.

Are FreeSurfer processing outputs included in the data sharing?

Sorry, FreeSurfer derivatives are not included in data shared via NDA, only "minimally processed" image volumes.

Where can I find more information about the imaging protocol?

The ABCD imaging protocol is described in Casey et al., 2018, The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev Cogn Neurosci. 32:43-54. doi: 10.1016/j.dcn.2018.03.001. Epub 2018 Mar 14. Review. PMID: 29567376.

Why do some imaging sessions, particularly from Philips scanners, have two or more dMRI series?

For imaging data from Philips scanners, the dMRI acquisition is split into two series because of a limitation of the Philips platform. Both scans have the same phase-encode polarity. They are meant to be concatenated together. In other rare cases, multiple dMRI scans may have been acquired, due to

acquisition problems in early scans. For the minimally processed data, one scan is selected for each session based on QC ratings, except for Philips scanners, in which case two are selected for packaging and sharing. All scans are available as raw DICOM files via fasttrack data sharing.

What is the "minimally processed" imaging data?

Minimally processed neuroimaging data includes:

- High-resolution structural data (3D T1w and T2w scans)
- Advanced diffusion MRI (multiple b-values and directions)
- Resting-State fMRI
- Task fMRI (Monetary Incentive Delay, Stop-Signal, and Emotional N-Back), with event files for each fMRI run.

These series have been run through standard modality-specific pre-processing stages including conversion from raw to compressed files, distortion correction, movement correction, alignment to standard space, and initial quality control (refer to *MRI Quality Control (QC)* Release Notes). This is to enable researchers to use the ABCD neuroimaging data in their own processing pipelines more quickly and efficiently than starting with raw data. Note that minimal processing is identical for rs-fMRI and task-fMRI and does not include analysis-specific pre-processing steps (e.g. removal of initial TRs, normalization by mean, etc.). Researchers intending to use minimally processed data should take note of the appropriate acknowledgment language to include in any public disclosure of results (refer to https://data-archive.nimh.nih.gov/abcd/results). The available minimally processed files are detailed in the *Other Imaging Instruments* Release Notes.

Preprocessed imaging data are packaged in archive files (tgz) for each image series containing BIDS formatted directory trees and NIfTI format data files (software to share preprocessed data: https://scicrunch.org/resolver/SCR_016016; consistent with BIDS specifications version 1.1.1: http://bids.neuroimaging.io/bids_spec.pdf). Imaging metadata derived from the original DICOM files are packaged along with each preprocessed data series as JSON files. The minimally processed T2w data are resampled into voxel-wise alignment with the T1w, which is rigid-body resampled into alignment with an atlas.

dMRI-specific information included diffusion gradients adjusted for head rotation (bvecs.txt), diffusion gradient strengths (bvals.txt), and a rigid-body transformation matrix specifying the registration between the dMRI image and the corresponding processed sMRI T1w image (stored in the JSON file). The dMRI minimally processed data are also kept in their original resolution, but reoriented into a standard alignment, based on registration to T1w, but not voxel-wise aligned with the T1w. A registration matrix supplied with the minimally processed dMRI data.

fMRI-specific information includes estimated motion time courses and a rigid-body transformation matrix specifying the registration between the fMRI image and the T_1w image (stored in the JSON file). The fMRI minimally processed data are kept in their original space and resolution, but a registration matrix is supplied with the minimally processed fMRI data. For task-fMRI series, event timing information is included as tab-separated value (tsv) files. The results of additional processing and ROI analysis are shared in tabulated form to the NDA database (https://scicrunch.org/resolver/SCR_016010), from which users can export spreadsheet files (tsv).

Information about this is included in the release notes and in our recent publication, Hagler et al., 2019, NeuroImage. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study (doi:10.1016/j.neuroimage.2019.116091). They also describe what processing steps are included in the "minimally processed" data shared on NDA.

There is currently no script available to run the ABCD minimal processing. There is a Docker that runs the complete processing and analysis pipeline available at https://www.nitrc.org/projects/mmps_docker/. Other useful software packages include https://github.com/ABCD-STUDY/abcd-dicom2bids and https://github.com/ABCD-STUDY/abcd-hcp-pipeline.

How can I download raw imaging data?

Unprocessed raw imaging DICOM files are made publicly available via the Fast Track mechanism. Only basic subject characteristics are provided with this service (i.e. sex, age, scan date). This is what is written about the Fast Track data in the ABCD 5.0 NDA Release Notes Imaging Instruments document:

"If unprocessed raw imaging data is needed, DICOM files are also made publicly available via a Fast Track mechanism. These DICOM files are released on NDA (https://data-archive.nimh.nih.gov/abcd) on a continual basis within approximately one month of data collection using the ABCD fast-track image sharing scripts (RRID: SCR_016021). DICOM files are arranged in BIDS-compliant directory trees and packaged in individual archive files (tgz) for each series. Metadata are included in the form of json-format text files, and for task-fMRI series, also included are the files containing stimulus and behavioral response timing information exported from the stimulus program (E-prime). A copy of the metadata is uploaded to NDA's image (version 03 – image03) database to link information to non-imaging-based assessments for the same participants. Raw DICOM Fast Track data have not undergone quality control or curation. The up-to-date index of Fast Track shared DICOM files can be downloaded in a separate image03 instrument."

In addition, the fastrackqc instrument has variables indicating whether data are usable or not. There is an option to download all "active" data or only the "recommended" data, based on the variables in fasttrackqc (see below). In

some cases, data are replaced on Fast Track; for example, if we are able to recover the missing files for a particular incomplete series. In those cases, the original series will still be available forever but we use the fasttrackqc instrument to mark it as “recalled” and not usable.

Here is a link to the Featured Dataset page.

It has two options. Option 2 allows you to select different pre-packaged datasets. Under “Raw Imaging Data”, select the item, “FastTrack Images. Recommended Active Series” (see attached). If you only want particular modalities, you can expand the menu item and pick only the ones you want.

Then you click “Add to Workspace”, etc., etc., and finally download a package from NDA. For Fast Track, this will not provide the actual imaging data, but a table with links to tgz files in AWS.

By the way, the DCAN Labs have developed tools to selectively download ABCD Fast Track DICOM data and convert it to actual nifti-BIDS format data: https://github.com/DCAN-Labs/abcd-dicom2bids

Which imaging data are recommended for use in analyses?

Please refer to the Imaging release notes. There are modality-specific image inclusion flags, like imgincl_t1w_include or imgincl_rsfmri_include, that are either 0 or 1. The fsqc_qc variable is used in the inclusion criteria to set the inclusion flags (criteria are listed in the release notes).

Why are post-processing QC scores available for only a subset of imaging data?

As described in the release notes, not all visits had manual post-processing QC; instead a sampling approach was used. For Release 3.0, post-processing QC (including FreeSurfer QC) was done for ~5-10% of visits for each modality. All visits that had failed QC in Release 3.0 were re-reviewed for Release 4.0. Because FreeSurfer version was unchanged between Release 4.0 and 5.0, reviews from Release 4.0 carried forward to Release 5.0. Additionally for Release 4.0 and 5.0, as described in the release notes, we selected additional visits for manual review based on automated QC metrics and multivariate outlier detection.

What methods were used for processing the imaging data to produce the minimally processed and tabulated imaging data?

Please refer to the ABCD Release Notes: Imaging Instruments. Image processing and analysis methods corresponding to ABCD Release 2.0.1 are described in Hagler et al., 2019, Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. Neuroimage, 202:116091. Changes to image processing and analysis methods are documented in the relevant release notes.

How can I download the abcd_fastqc01.csv file?

taken from https://github.com/ABCD-STUDY/abcd-dicom2bids?tab=readme-ov-file#how-to-download-abcd_fastqc01csv

1. Login to the NIMH Data Archive.
2. From the homepage, click the button labeled `GET DATA` to go to `Featured Datasets`.
3. Under the `Data Dictionary` heading in the sidebar, click `Data Structures`.
4. Add `abcd_fastqc01.csv` to the Filter Cart.
    1. Enter the spreadsheet file name into the `Text Search` box to find `ABCD Fasttrack QC Instrument`, then click its checkbox to select it.
    2. At the bottom of the page, click the `Add to Workspace` button.
5. At the top-right corner of the page under `logout` is a small icon. Click on it to open the `Selected Filters Workspace`.
6. Click `Submit to Filter Cart` at the bottom of the workspace.
7. Wait until the `Filter Cart` window at the top-right no longer says `Updating Filter Cart`.
8. Once the Filter Cart is updated, click `Package/Add to Study` in the `Filter Cart` window.
9. Click one of the buttons that says `Create Package`
    - Name each package something like **abcdQC**.
    - Select Only **Include documentation**.
    - Click **Create Package**.
10. From your NDA dashboard, click `Packages`.
11. Click the `Download Manager` button to download an executable `.jnlp` file. Once it downloads, run it to install the NDA Download Manager on your local machine.
12. Enter your NDA username and password in the window that pops up.
13. Accept the terms and conditions.
14. Click the checkbox at the left side of the window next to the file that you downloaded (e.g. `abcdQC`).
    - If you want to save the file to a different location than your home directory, click the `Browse` button at the top of the window.
15. Once the `Status` column of your package says `Ready to Download`, click `Start Downloads` at the bottom of the page to begin the download.
    - To track your download's progress at any given point, click the `Refresh Queue` button at the top of the window.
16. Once the `Status` column of your file says `Download Complete`, your file is ready.