

NIH Pediatric MRI Release 5.1 Notes

INTRODUCTION	2
GENERAL INFORMATION	3
Powered by NDAR	3
Data Structures	3
Data Dictionary.....	4
USING THE NDAR INFRASTRUCTURE	5
Overview	5
Query Tool.....	5
Create an NDAR Study.....	8
IMAGING DATA.....	9
Anatomic MRI.....	10
Spectroscopy Data.....	10
Phantom Data	11
Diffusion Tensor Imaging	12
TIMEPOINT	12
DEMOGRAPHICS.....	12
Child Race and Child Ethnicity	12
Adjusted Family Income Calculation	13
Biospecimen Dates	13
CLINICAL ASSESSMENTS	13
Missing Data; "NA" Designation	13
Bayley Scales of Infant Development II: Mental, Motor and Behavioral Rating Scale	13
Cambridge Neuropsychological Test Automated Battery (CANTAB)	13
Carey Temperament Scales.....	14
Child Behavior Checklist (CBCL).....	14
Computerized Diagnostic Inventory Schedule for Children (DISC) –Parent Version	14
Differential Ability Scales (DAS).....	14
Family Interview for Genetic Studies (FIGS).....	14
Handedness.....	15
NEPSY Semantic Verbal Fluency (Verbal Fluency).....	15
Neurological Exams (i.e., Physical Neurological Exams).....	15
Parental Stress Index.....	15
Pre-School Language Scale	15
Purdue Pegboard.....	16
Forms and Data not available in the Database	16

Introduction

As of July 9, 2012 the NIH Pediatric MRI Data Repository – Release 5.1 is available to qualified researchers.

New in Release 5.1 are high resolution (2.5mm) expanded diffusion tensor imaging (eDTI) data and raw MRSI data. Stereotaxically aligned anatomic MRI files that were previously excluded in Release 5 have been made available again to researchers in Release 5.1.

Release 5, made available in March 2012, included longitudinally-registered anatomic MRI images and derived volumetric measures, processed single-voxel MRS data which has been corrected for CSF voxel content, raw single-voxel MRSI files, low resolution diffusion tensor imaging (DTI), additional demographic measures and revised biospecimen collection dates. The additional demographics measures added to this release were child race, child ethnicity, and a continuous adjusted family income variable. The biospecimen dates are now included as the month of collection or month of lab processing. Release 5 also featured the database's migration to the National Database for Autism Research (NDAR) infrastructure. With this migration, variable names were shortened and revised; and the data dictionary, revised. The new infrastructure and procedures for querying data are described in these notes.

Release 4, made available in June 2010, included non-longitudinally registered anatomic MRI scans and derived volumes, cortical surface analysis data for Objective 1 (ages 4 ½ and older) and for a portion of Objective 2 (down to 2 years of age) for all timepoints, as well as age-specific brain atlases. Included as part of the cortical surface analysis data were surfaces in native space, transformation to stereotaxic space, gray and white surfaces, cortical thickness at each of 40,962 vertices per hemisphere, gyrification index, and surface mean curvature. MRI atlas templates included both symmetric and asymmetric atlases for the following overlapping age groups, generally corresponding to pubertal stages: 4.5 to 8.5 years (pre-pubertal), 7 to 11 years (pre- to early puberty), 5.5 to 13.5 years (pre- to mid-puberty), 10 to 14 years (early to advanced puberty), and 13 to 18.5 years (mid-puberty through post-puberty). These, as well as the clinical/behavioral data included in release 4, remain unchanged in Release 5.

Release 3, made available in September 2009, included anatomic MRI scans and data for all timepoints and clinical/behavioral data for all subjects, aged 10 days post delivery to young adult (both "Objective 1" and "Objective 2").

This release replaces (rather than supplements) earlier releases. Users should use this version for new analyses and, in all cases, reference the release version of the dataset used in any publications. The project homepage has also now moved to www.pediatricmri.nih.gov.

White papers describing data acquisitions and processing are available for each of the four imaging modalities as well as for the clinical/behavioral measures

General Information

Powered by NDAR

The Pediatric MRI Data Repository is now hosted within the National Database for Autism Research (NDAR) infrastructure. With this migration, former users will notice a significantly different database interface. These release notes will instruct former and new users how to query and download data through NDAR's Query Tool.

Data Structures

The data structures, listed below, are available for querying in this release using the new database interface powered by NDAR. Some of these data structures have been renamed from prior releases.

1. Demographics
2. Neuropsychological
3. Physical
4. Psychiatric & Personality
5. Diffusion Tensor Imaging (DTI)
6. Expanded Diffusion Tensor Imaging (eDTI)
7. Longitudinally Registered Anatomic MRI Variables
8. Non-longitudinally Registered Anatomic MRI Variables
9. Spectroscopy
10. Bayley Scales of Infant Development, Second Edition, Behavior Rating Scale
11. Bayley Scales of Infant Development, Second Edition, Mental Scale
12. Bayley Scales of Infant Development, Second Edition, Motor Scale
13. Brief Telephone Screening Interview
14. BRIEF: Behavior Rating Inventory of Executive Function - Adult Version (Informant Report)
15. BRIEF: Behavior Rating Inventory of Executive Function - Adult Version (Self Report)
16. BRIEF: Behavior Rating Inventory of Executive Function - Parent Form
17. CANTAB: Cambridge Neuropsychological Test Automated Battery
18. Carey Temperament Scales, Behavioral Style Questionnaire (3 to 7 Years)
19. Carey Temperament Scales, Early Infancy Temperament Questionnaire (1 to 4 Months)
20. Carey Temperament Scales, Revised Infant Temperament Questionnaire, (4 to 11 Months)
21. Carey Temperament Scales, Toddler Temperament Scale (1 and 2 Years)
22. C-DISC-4: Computerized Diagnostic Interview Schedule for Children – Parent and Youth Version
23. CVLT-C: California Verbal Learning Test for Children
24. CVLT-II: California Verbal Learning Test, Second Edition
25. DAS: Differential Ability Scales
26. DPS-4: DISC Predictive Scales
27. Family Biographical History Form (0:0 to 4:5 y:m)
28. FIGS MRI: Family History Interview for Genetic Studies - MRI - Version 1
29. FIGS MRI: Family History Interview for Genetic Studies - MRI - Version 2

NIH Pediatric MRI Release 5.1 Notes

30. Full Telephone Screening Interview - Version 1
31. Full Telephone Screening Interview - Version 2
32. Handedness (1:0 to 2:11)
33. Handedness (3:0 to 5:11) Part 1
34. Handedness (3:0 to 5:11) Part 2
35. Handedness (6:0+)
36. JTCI: Junior Temperament and Character Inventory - Parent Report
37. JTCI: Junior Temperament and Character Inventory - Self Report
38. MRI Child History Form (4:6+)
39. NEPSY Verbal Fluency (Semantic & Phonemic)
40. NEPSY Verbal Fluency (Semantic)
41. Physical and Neurological: (0:0 to 0:1 y:m) Examination
42. Physical and Neurological: (0:2 to 0:11 y:m) Examination
43. Physical and Neurological: (1:0 to 2:11 y:m) Examination
44. Physical and Neurological: (3:0 to 4:5 y:m) Examination
45. Physical/Neurological Examination
46. PLS-3: Preschool Language Scale-3
47. PSI: Parenting Stress Index
48. Pubertal Status Questionnaire
49. Purdue Pegboard - Full Board
50. Purdue Pegboard - Half Board
51. Screening and Exclusion Form (0:0 to 4:5 y:m)
52. Spectroscopy Scans
53. TCI: Temperament and Character Inventory - Parent Report
54. TCI: Temperament and Character Inventory - Self Report
55. Urine and Saliva
56. WAIS-R: Wechsler Adult Intelligence Scale - Revised - Digit Span and Digit Symbol
57. WASI: Wechsler Abbreviated Scale of Intelligence
58. WISC-III: Wechsler Intelligence Scale for Children, Third Edition - Digit Span and Coding
59. WJ3: Woodcock-Johnson III

Data Dictionary

An updated data dictionary accompanies Release 5. Updates to the data dictionary were made for a variety of reasons including new data structures and fields and enhanced field descriptions.

Additionally, many variable/element names (aka field names) were shortened due to character length restrictions in NDAR. Both the revised Release 5 and previous Release 4 variable/elements names have been included in the data dictionary spreadsheet. The significant changes to the data dictionary are included within these notes.

Using the NDAR Infrastructure

Overview

The National Database for Autism Research (NDAR) hosts data from hundreds of projects, of which this study, the NIH MRI Study of Normal Brain Development, is one. Therefore, some of the functionality of NDAR will not be applicable to Pediatric MRI users.

New database terminology, including “Collection” and “Study”, may be unfamiliar to previous users. An NDAR Collection contains all data related to a project. The collection containing all data from the NIH MRI Study of Normal Brain Development is titled “Pediatric MRI.” The NDAR Study is a tool with which researchers can save their queries and define any study specific to a publication on which they are working. This functionality is described in more detail later.

Additionally, the “GUID”, or NDAR’s Global Unique Identifier, labels all subjects within NDAR. The Pediatric MRI subjects were assigned these IDs in addition to their original project study IDs when the database migrated to NDAR. Within data structures, users will see the GUID labeled as “subjectkey” and the original Pediatric MRI study IDs labeled as “src_subject_id.”

Query Tool

From the Pediatric MRI Data Repository login, users will now be redirected to the NDAR Query page shown below in **Figure 1**. Users can access white papers and other documentation for the project directly from this page. To download study data, select the “Download Data” button.

The screenshot shows the NDAR Query page for the 'Pediatric MRI' collection. The page includes a navigation bar with links like Home, Query, Harmonization Tools, etc. Below the navigation, there are tabs for 'Query Data', 'Data from Labs', 'Data from Papers', and 'Query Instructions'. The main content area displays the collection title 'Pediatric MRI' and a list of investigators. To the right, there are three pie charts representing data categories: Genomics (1,544), Neuroimaging (676), and Phenotype (3,300). A 'Download Data' button is highlighted with a red box. Below this, there is a section for 'Supporting Documentation' with a table of files.

File Name	File Type	Description	Share File	Approve
Anatomic_MRI_White_Paper_R5.pdf	Methods	Anatomic MRI white paper and project overview.	General Public	Approved
Clinical_Behavioral_White_Paper_R5.pdf	Methods	Clinical/behavioral white paper.	General Public	Approved
DTI_White_Paper_RS.1.pdf	Methods	Diffusion tensor imaging (DTI) white paper.	General Public	Approved

Figure 1: New query site powered by NDAR

NIH Pediatric MRI Release 5.1 Notes

This will prompt the user to enter his or her login credentials as show in **Figure 2**.

login

Username:

Password:

[Reset Password](#)

NDAR provides a single access to de-identified autism research data. For permission to download data, you will need an NDAR account with approved access to NDAR or a connected repository (AGRE, IAN, or the ATP). For NDAR access, you need to be a research investigator sponsored by an NIH recognized institution with federal wide assurance. See [Request Access](#) for more information.

Warning Notice
This is a U.S. Government computer system, which may be accessed and used only for authorized Government business by authorized personnel. Unauthorized access or use of this computer system may subject violators to criminal, civil, and/or administrative action.

[Show More](#)

login

Figure 2: Login to NDAR

Users can then select categories of data structures for download. When all desired data structure categories are selected, the user should select the “Create Package” button as indicated in **Figure 3**.

NDAR National Database for Autism Research
Serving the autism research community

Home Query Harmonization Tools Contribute Request Access Policy Tutorials About FAQ cshuqars | logout

Based upon the filters applied, subjects contained in the Data Structures and Collections are available for download. Please note: filters identify the subjects included for download. Once included, all data for a subject for the selected data structure(s) and collection(s) will be included in the download package.

[Return](#) **Create Package**

DATA STRUCTURE BY CATEGORY [Collapse All](#) | [Deselect All](#)

- Neuroimaging
 - Peds - Diffusion Tensor Imaging (274 subjects)
 - Peds - Expanded Diffusion Tensor Imaging (eDTI) (152 subjects)
 - Peds - Longitudinally Registered aMRI Variables (553 subjects)
 - Peds - Non-longitudinally Registered aMRI Variables (553 subjects)
 - Peds - Spectroscopy (553 subjects)
 - Image (2 subjects)
- Behavior
 - Peds - BSID II Behavior Rating Scale (554 subjects)
 - Peds - BRIEF Adult Version (Self Report) (554 subjects)
 - Peds - Brief Telephone Screening Interview (554 subjects)
 - Peds - Carey Temperament Scales, TTS (1 to 2 years) (554 subjects)
 - Peds - Carey Temperament Scales, EITQ (1 to 4 months) (554 subjects)
 - Peds - Carey Temperament Scales, BSQ (3 to 7 years) (554 subjects)
 - Peds - Carey Temperament Scales, RITQ (4 to 11 months) (554 subjects)
 - Peds - Child Behavior Checklist (CBCL) (554 subjects)
 - Peds - BRIEF Parent Form (554 subjects)
 - Peds - BRIEF Adult Version (Informant Report) (554 subjects)
- Cognitive
 - Peds - BSID II Motor Scale (554 subjects)
 - Peds - BSID II Mental Scale (554 subjects)
 - Peds - California Verbal Learning Test, 2nd ed. (554 subjects)
 - Peds - NEPSY Verbal Fluency (Semantic and Phonemic) (554 subjects)
 - Peds - NEPSY Verbal Fluency (Semantic) (554 subjects)

COLLECTIONS BY PERMISSION GROUP [Collapse All](#) | [Deselect All](#)

- Pediatric MRI
 - Pediatric MRI (556 subjects)
 - Investigators:** PEDIATRIC STUDY CENTERS: Michael J. Rivkin, M.D., Boston Children's Hospital; William S. Ball, M.D., Children's Hospital Medical Center of Cincinnati; Dah-Jyuu Wang, Ph.D., Children's Hospital of Philadelphia; James T. McCracken, M.D., University of California at Los Angeles; Michael Brandt, Ph.D. and Jack Fletcher, Ph.D., University of Texas Health Science Center; Robert McKinstry, M.D., Washington University. DATA COORDINATING CENTER: Alan Evans, Ph.D., Montreal Neurological Institute Center. CLINICAL COORDINATING CENTER: Kelly Botteron, M.D., Washington University. DIFFUSION TENSOR PROCESSING CENTER: Carlo Pierpaoli, M.D., National Institute of Child Health and Human Development. SPECTROSCOPY PROCESSING CENTER: Joseph O'Neill, Ph.D., University of California at Los Angeles. (Owner: MRI, Pediatric)
 - Description:** Pediatric MRI Release 5.0. Visit our website at <http://pediatricmri.nih.gov> for more information about this project. [more](#)
 - Shared (556 subjects)

Figure 3: Create a query package

NIH Pediatric MRI Release 5.1 Notes

Users will then need to type a package name and check if they would like to include data files in the download, i.e. the actual imaging files. One may wish not to do this initially for faster download and preview of metadata. Selecting the “Create Package” button as shown in **Figure 4** will direct users to NDAR’s download manager.

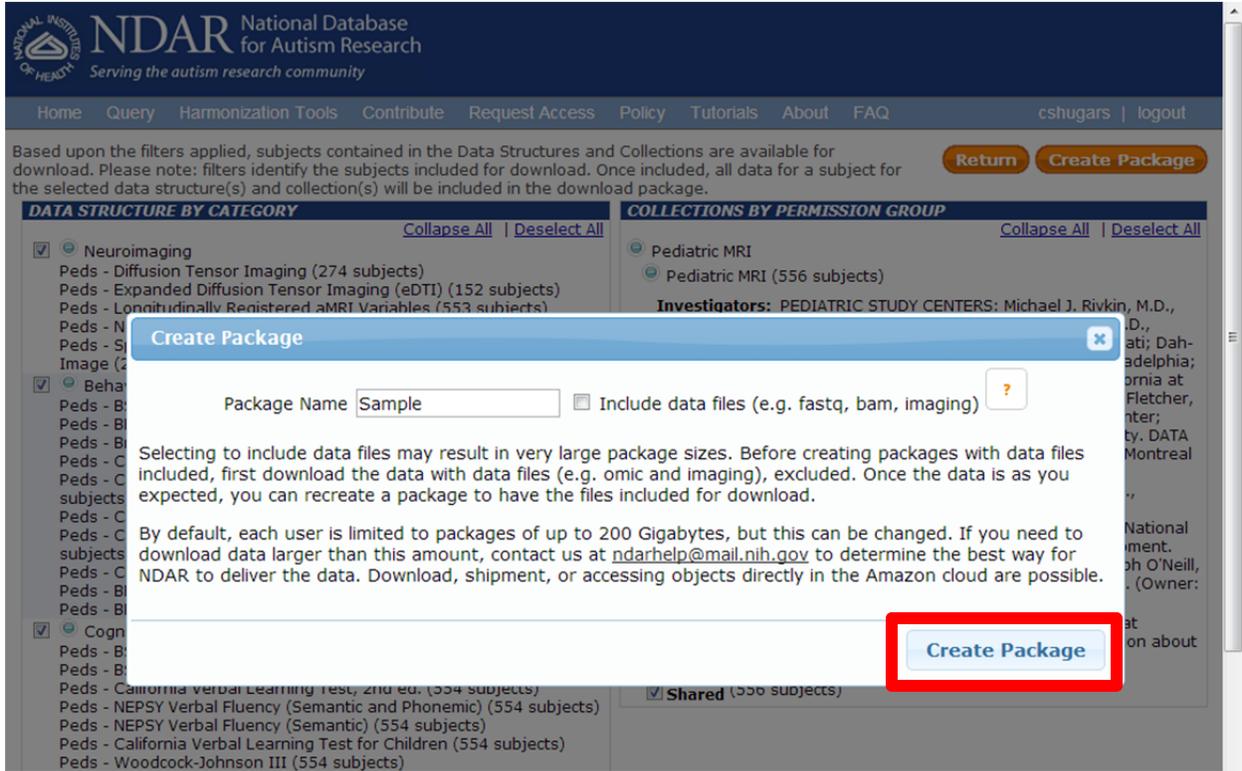


Figure 4: Name the query package

Users should then launch the Download Manager as shown in **Figure 5** below. *NOTE: The NDAR Download Manager requires Java 6 to be installed on your computer.*

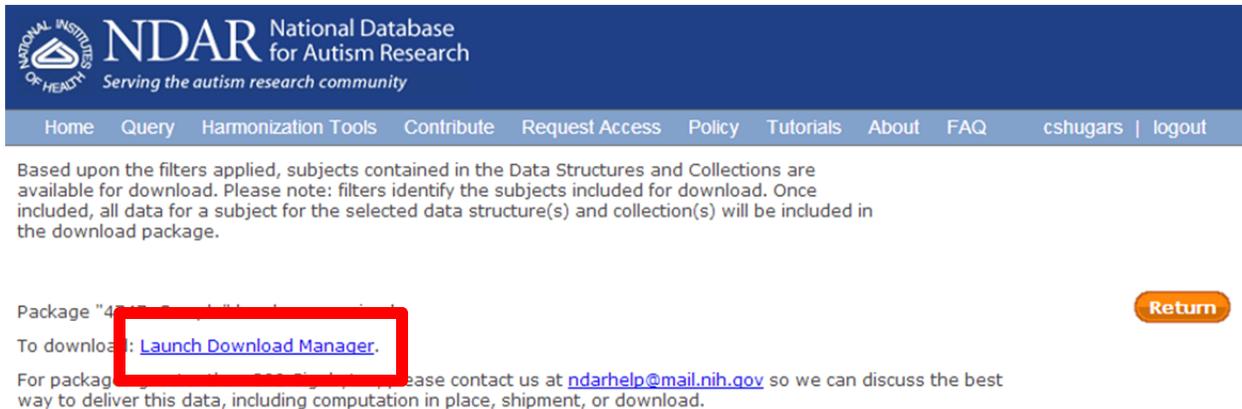


Figure 5: Launch Download Manager

NIH Pediatric MRI Release 5.1 Notes

When the Download Manager launches, users will need to reenter their login credentials and accept the user agreement. The NDAR Download Manager allows users to select a download directory, as shown below in **Figure 6**. The “Status” column will indicate when the package is ready for download. Note that the status may take several minutes to update for larger query packages. When the status is “Ready for Download”, put a check mark beside the package and then select the “Start Downloads” button. The download is complete once the status shows “Download Complete.”

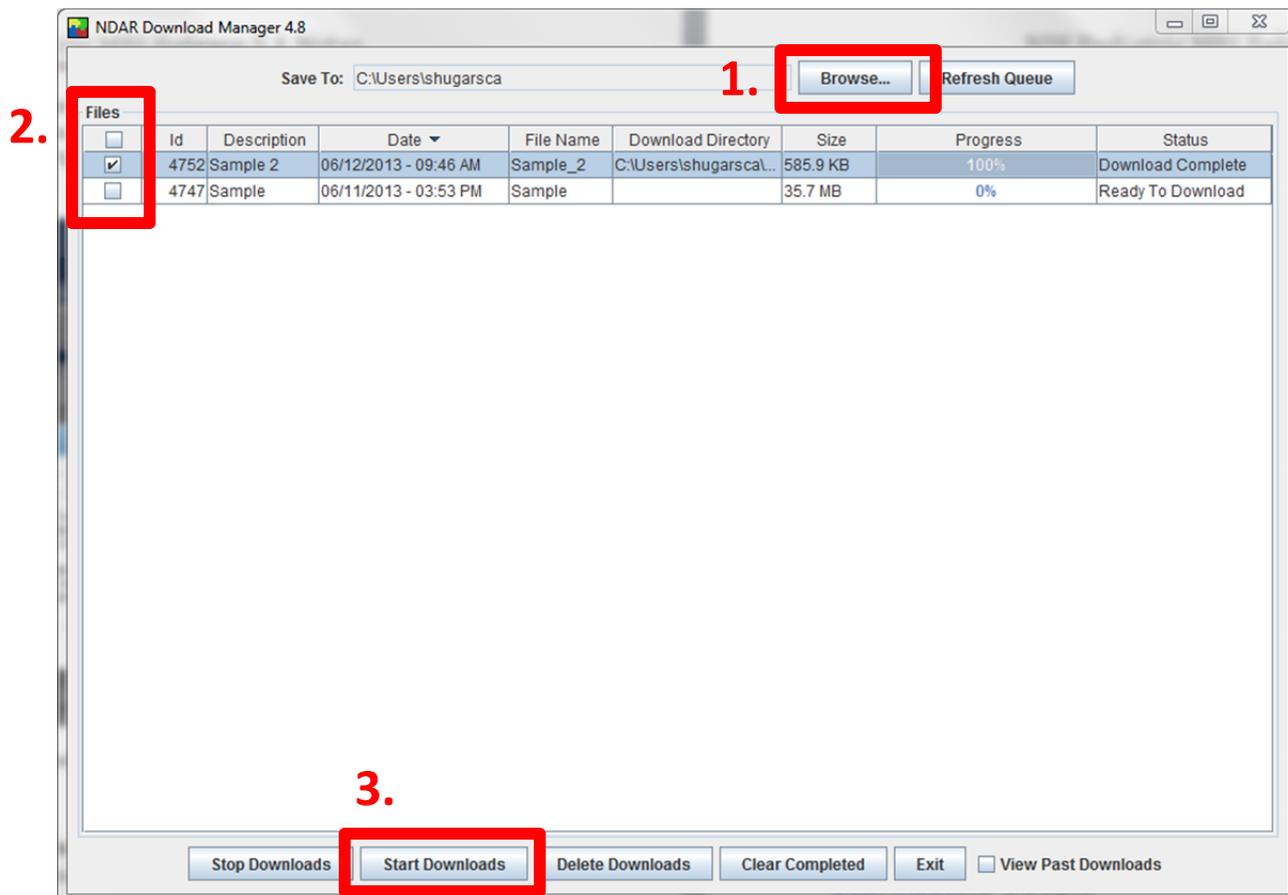


Figure 6: Download packages

To learn more about the NDAR Query Tool, view the online tutorial at http://ndar.nih.gov/ndarpublicweb/help/querytool/Query_Tool_Tutorial.html. The tutorial will demonstrate how to design more advanced queries, select specific datasets, and filter query results. *Note that there is a limit of 200GB per download. If more data is needed, contact pedsMRI@mail.nih.gov.*

Create an NDAR Study

The NDAR Study tool allows researchers to define the specific subjects associated with their project or study and later associate this data with a publication. This will also be a useful tool for users to save their queries in the Query Tool. Please view the online tutorial at <http://ndar.nih.gov/ndarpublicweb/training/html> for instructions on creating an NDAR Study. Whenever

a user's publication results become available in Pubmed, NDAR will also provide a link from Pubmed to the user's NDAR Study.

Imaging Data

The imaging modalities available are anatomic MRI (aMRI), proton MR spectroscopy (MRS) and MRS imaging (MRSI), diffusion tensor imaging (DTI), and expanded diffusion tensor imaging (eDTI). Approximate scan counts available with this release are as follows:

aMRI:

1381 datasets from 538 subjects, inclusive of 1316 longitudinally-registered datasets from 527 subjects

Single-voxel MRS:

273 datasets (4 voxels each) corrected for CSF voxel content from 145 subjects

MRS imaging (MRSI):

57 Single-voxel datasets from 34 subjects

DTI:

498 datasets from 274 subjects

eDTI:

193 datasets from 152 subjects

Left-right orientation:

All scans with the exception of MRSI and DTI follow the neurologic convention (left side of brain appear on left of image). A watermark on the most inferior slice of the anatomic MRI scans has been provided to indicate this orientation. MRSI and DTI used the radiologic convention (left brain on right side of image).

All imaging data is accessible through NDAR's Query Tool as described in the previous section. Release 5.1 images of subjects are divided into five data structures: Diffusion Tensor Imaging (version 3), Expanded Diffusion Tensor Imaging (eDTI), Non-longitudinally Registered aMRI Variables, Longitudinally Registered aMRI Variables, and Spectroscopy. Phantom imaging data is divided into two structures: Image, and Diffusion Tensor Imaging (version 4).

All imaging structures contain queryable data elements and file names. Once a query package is selected for download, the path names will appear in the CSV output file. Use this file path to find the imaging files in the downloaded query package as shown in **Figure 7**.

NIH Pediatric MRI Release 5.1 Notes

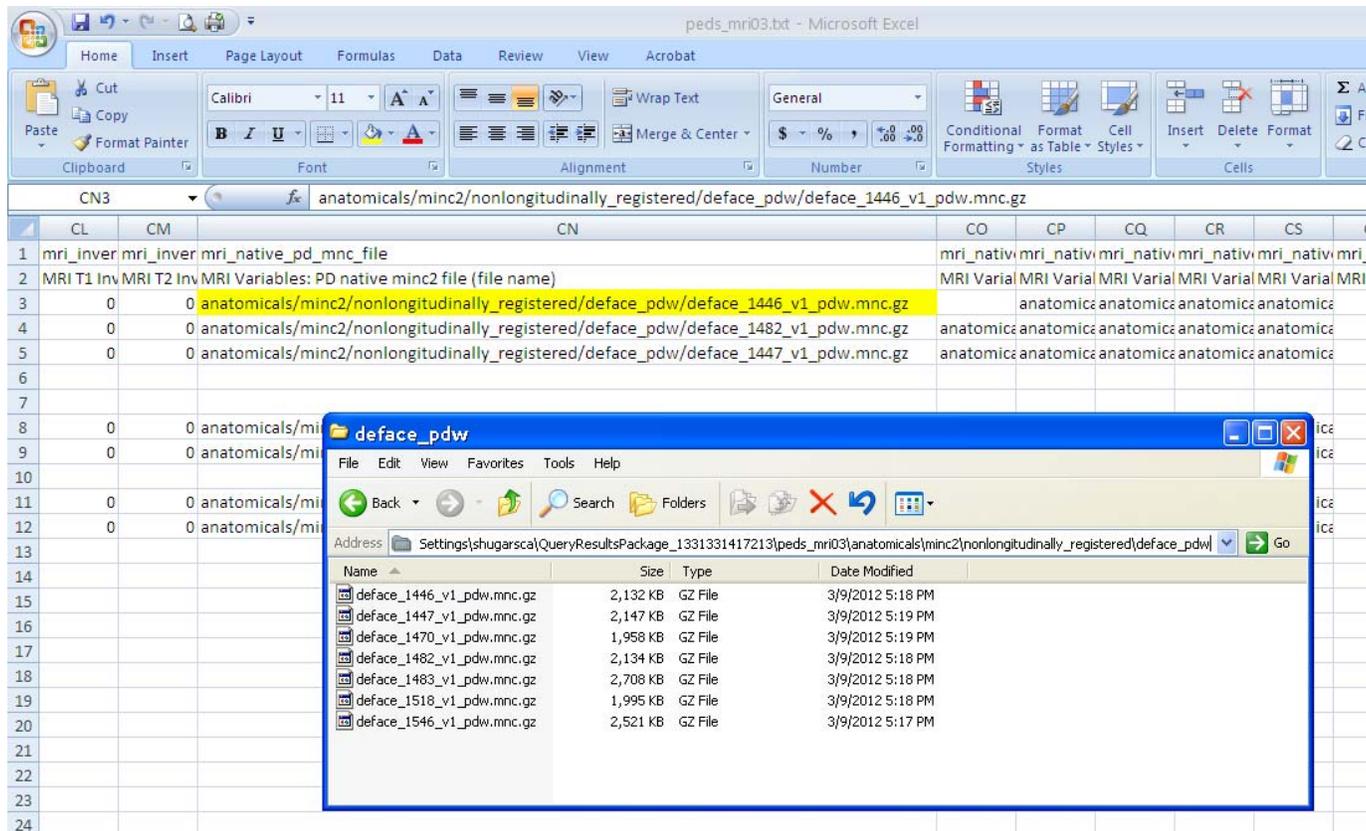


Figure 7: Use the file names in the query package .TXT file to find the image files

Anatomic MRI

For the Non-longitudinally Registered Anatomic MRI Variables and Longitudinally Registered Anatomic MRI Variables data structures, both MINC2 and NIFTI files are available for download through the NDAR Query Tool. Both sets of file names have a .gz suffix. The .gz indicates that these files have been compressed using the gzip standard.

Volumetric measures for whole brain white matter, gray matter and cerebrospinal fluid (CSF) and for several regional volumes, e.g., each cerebral lobe (e.g., right frontal white matter, left occipital gray matter), cerebellum, brainstem, are also provided. Also included are cortical thickness and surface analysis measures. See anatomic MRI white paper.

Spectroscopy Data

Both raw and processed proton spectroscopy files are listed in the Spectroscopy data structure. Processed images are in .png format and should be viewable using a standard web browser. Raw files are in .pfile, DICOM, and .rda formats. These may be viewed using LCMoDel (available at <http://s-provencher.com/pages/lcmoDel.shtml>) or MRUI (available at <http://sermn02.uab.es/mrui/>).

Release 5 reflects the enhanced, reprocessed single-voxel proton spectroscopy data which has now been corrected for voxel CSF content. Data contained in release 4 had not been corrected for voxel CSF

NIH Pediatric MRI Release 5.1 Notes

content and has been replaced. In addition, a small amount of new data has been added from one of the sites.

The spectroscopy data includes 273 single-voxel datasets from 145 Objective 1 subjects (ages 4 ½ to 18 at time 1) across their three timepoints. The four voxels sampled were from frontal white matter, thalamus, and parietal white matter (all on the left) and midline occipital gray matter. Spectra with their corresponding LC model printouts were provided, along with scalar values for the levels of three neurometabolites-- *N*-acetyl-aspartate+*N*-acetyl-aspartyl-glutamate (“NAA”), creatine+phosphocreatine (“Cr”), and choline-containing compounds (“Cho”), relative to the magnitude of the unsuppressed and suppressed water in the voxel and relative to creatine+phosphocreatine and quality-control (QC) parameters, principally the spectral linewidth (expressed as full-width at half-max; FWHM) and the spectral signal-to-noise ratio (S/N).

Release 5 also includes files that display and verify the proper anatomic placement of the MRS voxels on the anatomic MR images and tissue composition (gray matter, white matter, CSF) and correction of metabolite levels for voxel CSF-content. MRS data not meeting spectral and/or tissue QC criteria have now been excluded from the database.

Note the metabolite NGA+NAAG values are rounded in the Query Tool. That data is not rounded in the image files.

Phantom Data

As defined in procedure manuals, a series of images were acquired from a living human phantom who was scanned at all sites at several timepoints. These have been made available in both MINC2 and NIFTI formats and reside in the “Release 5.0 – Phantom Data” dataset. They can be downloaded through the “Image” and “Diffusion Tensor Imaging (NDAR version 4)” data structures. All phantom records are assigned a GUID number NDAR_INVCH863XA1. Each file name defines the phantom subject ID, the site at which the data was collected for that individual, and the time in days between scans. For example, the file name /LIVING_PHANTOM_1_SITE_2_8518 represents living phantom 1, Site 2.

The number 8518 represents the duration in days from an arbitrary date defined for all phantom scans. The actual number is meaningless except that it is relative to the other phantom scans. For example, living phantom scan PHANTOM_1_SITE_6_8872 occurred exactly 354 days after LIVING_PHANTOM_SCAN_1_SITE_2_8518. This can be determined by subtracting 8518 from 8872.

A standard American College of Radiology (ACR) physical phantom was also scanned in the project at each site across time. The user may compare the number of days from the arbitrary date for the subject against those for the ACR phantom using the `acr_age_days_arbitrary_of_t1` in the Non-longitudinally Registered MRI Variables section of the database. This variable corresponds to the last number in the living phantom IDs.

The Living phantom data was acquired at intervals of about 1 year on the scanner that was being used to acquire data at each site at that time. The living phantom data that has been released has passed the visual quality control inspection.

NIH Pediatric MRI Release 5.1 Notes

Diffusion Tensor Imaging

This Diffusion Tensor Imaging (DTI) data release contains data collected with higher resolution (2.5mm) acquisition protocol and data collected with lower resolution (3mm) or the “conventional” DTI protocol. DTI data was sometimes collected in sessions that were separated from the other measure in time. When this occurred, a visit number was assigned based on age and proximity to the other data acquisitions. In some instances, the corresponding clinical/behavioral data for that visit number may be unavailable.

For the low resolution DTI data, two sets of data are available for download; DTI release 1 (REL01) contains datasets that were virtually artifact-free and therefore did not need additional post-processing with artifact remediation strategies (85 scans from 62 unique subjects), while DTI release 2 (REL02) contains all DTI data that passed set quality criteria (498 scans from 274 unique subjects).

The higher resolution, expanded DTI (eDTI) data includes low ($b=0s/mm^2$) and high ($b=1100s/mm^2$) image volumes, but also includes a number of intermediate b-values as well ($b=100s/mm^2$, $b=300s/mm^2$, $b=500s/mm^2$, and $b=800s/mm^2$). Up to nine (9) files of tensor derived quantities are included in the database: Directionally encoded color maps (DEC); Eigenvalues (EV) as 4D image file with volumes in the order of $\lambda_1, \lambda_2, \lambda_3$; Fractional anisotropy (FA); Lattice index (LI); Relative anisotropy (RA); Trace of the diffusion tensor (TR) (equal to $3 \times$ Mean Diffusivity); Chi-Squared map of the fitting (CS) – a measure of the goodness of fit of the tensor model; Outlier map (OUT) – indicates the percentage of data points identified as outliers and removed from the tensor fitting on a voxel-by-voxel basis; and Brain mask (MS).

Timepoint

The Timepoint data structure when combined with other data structures enables a user to identify the objective (1 or 2), timepoint (1-3 for Objective 1; up to 10 for Objective 2), subject id and age in days.

Visits beginning with the letter “i” (or “i” followed by another letter, e.g., ia, ib, ic) indicate an incomplete visit and should correspond to age in days for an appropriate timepoint.

Demographics

Demographics contains information about each research participant’s demographic profile. The following changes/issues are identified with demographics in this release.

Child Race and Child Ethnicity

Child race and child ethnicity variables were derived from parent race and ethnicity variables and added to Release 5. If the child’s parents were of the same race and/or ethnicity, then the child is listed as also being of that race or ethnicity. If the parents were of different races, then both races are listed under

NIH Pediatric MRI Release 5.1 Notes

child_race_. If only one parent was Hispanic or Latino, then the child was listed as "Part Hispanic or Latino".

Adjusted Family Income Calculation

The categorical income variable used for sampling and included in earlier releases ("HUD_Adjusted_Family_Income") has been replaced with a more continuous income variable labeled adjusted_family_income in Release 5. The methods and formula for deriving this new income variable are included in the Clinical/Behavioral White Paper – Release 5.

Biospecimen Dates

All biospecimen collection and lab processing dates have been converted to months for Release 5. This affects all dates within the Urine and Saliva data structure. The months are listed numerically (1=January, 2=February, etc).

Clinical Assessments

The following notes are intended to give the user an overview of issues related to the clinical assessments and alert the user to changes made to the data dictionary for this data release.

Missing Data; "NA" Designation

Data may be "missing" for a variety reasons, e.g., (1) Instrument inappropriate for the age of the child, (2) instrument appropriate but not administered or not completed by the child, or (3) instrument administration did not pass quality control checks. Generally, and throughout the database, the use of "NA" does not specifically indicate anything other than the fact that no data has been inputted for that cell. Thus, NA or a blank cell each communicates only "no data."

Bayley Scales of Infant Development II: Mental, Motor and Behavioral Rating Scale

Users should refer to the Objective 2 procedure manual for clarification of how ages were calculated for the BSID-Mental Development Index (MDI), Psychomotor Development Index (PDI) and BRS.

Children coming in for a 36 month testing visit (e.g., between 1073 and 1122 days old) were given the BSID-PDI only.

Cambridge Neuropsychological Test Automated Battery (CANTAB)

CANTAB data is included only for individual ages 1643 days (4.67 years) and older. Data for children between 4:6 and 7:0 should be used with caution due to variations in data collection procedures in comparison to older children.

Following Visit 1 data collection with the CANTAB, the publishers upgraded the CANTAB to the CANTAB eclipse version which was used for Visits 2 and 3. Please see the following General CANTAB or subtest specific sections for definitions and explanations of specific CANTAB variables used during Visit-1. Also, see CANTAB website (<http://www.cantabeclipse.com>).

NIH Pediatric MRI Release 5.1 Notes

Carey Temperament Scales

The Carey was utilized only in Objective 2.

Child Behavior Checklist (CBCL)

Four versions of the CBCL were used, corresponding to ages 1.5 to 5, 4 to 18, 6 to 18, and the Young Adult Self Report.

Note that an erroneous value of 255 was found in a number of result variables at the time of this release. More information will be posted related to this data error as that information becomes known. .

Computerized Diagnostic Inventory Schedule for Children (DISC) –Parent Version

The following nonexclusionary DISC diagnoses are included with Release 5.

DISC_P_Encopresis_Past_Month	DISC_P_Encopresis_Past_Year
DISC_P_Enuresis_Diurnal_Past_Month	DISC_P_Enuresis_Diurnal_Past_Year
DISC_P_Enuresis_Nocturnal_Past_Month	DISC_P_Enuresis_Nocturnal_Past_Year
DISC_P_Nicotine_Dependence	DISC_P_Social_Phobia_Past_Month
DISC_P_Social_Phobia_Past_Year	DISC_P_Specific_Phobia_Past_Month
DISC_P_Specific_Phobia_Past_Year	DISC_Y_Encopresis_Past_Month
DISC_Y_Encopresis_Past_Year	DISC_Y_Enuresis_Diurnal_Past_Month
DISC_Y_Enuresis_Diurnal_Past_Year	DISC_Y_Enuresis_Nocturnal_Past_Month
DISC_Y_Enuresis_Nocturnal_Past_Year	DISC_Y_Nicotine_Dependence
DISC_Y_Oppositional_Defiant_Disorder_Past_Month	DISC_Y_Oppositional_Defiant_Disorder_Past_Year
DISC_Y_Separation_Anxiety_Disorder_Past_Month	DISC_Y_Separation_Anxiety_Disorder_Past_Year
DISC_Y_Social_Phobia_Past_Month	DISC_Y_Social_Phobia_Past_Year
DISC_Y_Specific_Phobia_Past_Month	DISC_Y_Specific_Phobia_Past_Year

Differential Ability Scales (DAS)

The Differential Ability Scales spans both Objective 1 and Objective 2 as it was used for children 3:0 to 5:11. Users should refer to the Objective 2 procedure manual for clarification of how ages were calculated for the DAS.

The Differential Ability Scales has two parts, the School Age DAS and the Preschool DAS. Only the Preschool Age DAS was used in the project. The Preschool DAS is further divided into the Lower Preschool and the Upper Preschool DAS. Both versions of the Preschool DAS were used for the project. Refer to the DAS manual for an explanation of the differences between the two versions.

The Differential Ability Scales does **not** use traditional raw scores (i.e., the subject is **not** given credit for items that appear below the starting point). In order to calculate a continuous raw score in the current version of the release (e.g., a raw score similar to one found in the WASI), users must manually calculate the raw scores to correct for the DAS variation.

Family Interview for Genetic Studies (FIGS)

Minor modifications were made to the FIGS following the first visit for Objective 1. Thus, two versions of the FIGS were utilized in the project. In Objective 1, all subjects' Visit 1 FIGS data was collected using Version 1 of the FIGS, following which a transition was made to Version 2. However, approximately half of the Objective 2 subjects utilized Version 1 FIGS for their first visit, and approximately half utilized Version 2.

NIH Pediatric MRI Release 5.1 Notes

These two versions are referenced elsewhere as Year 1 (meaning Version 1) and Year 3 (corresponding to Version 2).

Handedness

Handedness was measured via three age-appropriate assessments on the project, i.e., 1:0-2:11; 3:0-5:11, 6:0 and older. Handedness 1:0-2:11 was only used on Objective 2; Handedness 3:0-5:11/Handedness Form A covers both objectives, Handedness 6:0/Handedness Form B (ages 6:0 and older) was only used on Objective 1.

Handedness 3:0-5:11 is currently divided into two files in the database: Handedness 3:0-5:11 – Part 1 for Objective 2 and Handedness 3:0-5:11 – Part 2 for Objective 1. Thus, in order to download the complete dataset, it is necessary to download and concatenate both files to achieve Handedness 3:0-5:11.

It should be noted that in Handedness 3:0-5:11, the child actually performs the tasks, while for Handedness 6:0 and up, the child pantomimes the task. Handedness 3:0-5:11 and Handedness 6:0 and up share five common tasks (i.e., hand used to write; hand used to eat; hand used to cut with scissors, hand used to hammer, and hand used to throw a ball). This provides an age-range from 3:0 through 18:0 and older.

Handedness has very precise administration and measurement directions. These directions are critical to successful administration and scoring of all three of the handedness assessments.

NEPSY Semantic Verbal Fluency (Verbal Fluency)

The NEPSY is divided into two versions—the Semantic and the Semantic and Phonemic. Children between 3:0 and 6:11 were administered the Semantic portion of the NEPSY only. Data for the Semantic version span both Objective 1 and Objective 2.

To create a dataset of Semantic data that spans all ages above 3:0, users must download both the Semantic and Semantic and Phonemic versions of the NEPSY and concatenate them together.

The word list (e.g., word responses of the child) are included in the database for the 3:0 to 6:11 Semantic version, but are not included for the Semantic and Phonemic versions (7:0 and up).

Neurological Exams (i.e., Physical Neurological Exams)

Four age-specific neurological exams were used for Objective 2 (but only one for Objective 1).

Parental Stress Index

Available scores are limited to raw data for Summary Scores (e.g., Total Stress Raw Score).

Pre-School Language Scale

Users should refer to the Objective-2 procedure manual for clarification of how ages were calculated for the PLS-3.

NIH Pediatric MRI Release 5.1 Notes

Rather than beginning testing at one year below the child's chronological age as suggested in the PLS-3 manual, examiners began testing with items six month below the child's targeted testing age to help streamline the time required for the Objective 2 testing battery (see Objective 2 procedure manual for more information regarding targeted testing ages).

Purdue Pegboard

Two versions of this instrument were used—the Half Board (for younger children) and the Whole Board (for older children). The Half Board was administered to children 3:0 to 5:11 and spans both objective, i.e., includes data from young children in both objectives 1 and 2. If a child was unable to complete a trial, the trial was left blank and the administration was discontinued.

It is possible to make a dataset of Purdue data which spans all ages above 3:0 by downloading data for the Purdue Half Board and the Purdue Whole Board.

Forms and Data not available in the Database

Several forms available in the procedure manuals (most used for internal tracking of data collection) are not included in the database. For example, for objective 2, such forms include the Contact Sheet, Consent and Assent Forms, etc.