CANDIShare: A Resource for Pediatric Neuroimaging Data

David N. Kennedy,
Division of Neuroinformatics and the Child and Adolescent NeuroDevelopment Initiative (CANDI), Department of Psychiatry, University of Massachusetts Medical School, 356 Plantation St, Biotech 1, Suite 100, Worcester, MA 01605, USA, David.Kennedy@umassmed.edu

Christian Haselgrove,
Division of Neuroinformatics and the Child and Adolescent NeuroDevelopment Initiative (CANDI), Department of Psychiatry, University of Massachusetts Medical School, 356 Plantation St, Biotech 1, Suite 100, Worcester, MA 01605, USA

Steven M. Hodge,
Division of Neuroinformatics and the Child and Adolescent NeuroDevelopment Initiative (CANDI), Department of Psychiatry, University of Massachusetts Medical School, 356 Plantation St, Biotech 1, Suite 100, Worcester, MA 01605, USA

Pallavi S. Rane,
Division of Neuroinformatics and the Child and Adolescent NeuroDevelopment Initiative (CANDI), Department of Psychiatry, University of Massachusetts Medical School, 356 Plantation St, Biotech 1, Suite 100, Worcester, MA 01605, USA

Nikos Makris, and
Center for Morphometric Analysis, Departments of Neurology & Psychiatry, Massachusetts General Hospital, Boston, MA, USA

Jean A. Frazier
Division of Neuroinformatics and the Child and Adolescent NeuroDevelopment Initiative (CANDI), Department of Psychiatry, University of Massachusetts Medical School, 356 Plantation St, Biotech 1, Suite 100, Worcester, MA 01605, USA

Keywords

There are numerous psychiatric disorders that can plague the development of children. Each of these disorders manifests as a distinct pattern of clinical, behavioral, etiological, neuroanatomic and neurofunctional characteristics that challenge the management of the individual patient, as well as the development of successful intervention and prevention strategies. In the area of neuroimaging, a substantial number of studies have been performed to date; and while much has been learned from this investment, this represents only the tip-of-the-iceberg of the information that can be gleaned from the data. Unfortunately, most of this additional, untapped information resource is lost due to ineffective sharing of the data that is possible by application of current technologies.

The Child and Adolescent NeuroDevelopment Initiative (CANDI) at University of Massachusetts Medical School is making available a series of structural brain images, as well as their anatomic segmentations and demographic data, as a coordinated set of related morphometric resources. The initial data set is a release of 103 subjects (T1-weighted MRI scans and anatomic segmentation) that comprised the neuroanatomic data published in 2008 in an article by Frazier, et al.[1] The subjects include 57 males and 46 females, aged 4–17, and come from four diagnostic groups: Healthy Controls (*N*=29), Schizophrenia Spectrum (*N*=20), Bipolar Disorder with Psychosis (*N*=19), and Bipolar Disorder without Psychosis (*N*=35).

Both the McLean Hospital and Cambridge Health Alliance Institutional Review Boards approved the original data acquisition research protocols. All subjects signed assent forms, and their parents/legal guardians signed informed consent forms. The University of Massachusetts Medical School Institutional Review Board approved the data sharing protocol. Images were acquired from 1996–2006 at the McLean Hospital Brain Imaging Center on a 1.5 Tesla General Electric Signa Scanner. Structural imaging was performed using a three-dimensional inversion recovery-prepared spoiled gradient recalled echo in the coronal plane with 124 1.5 mm thick slices, repetition time 10 ms, TE 3, flip angle 25°, field of view 24 cm, acquisition matrix 256×192 and 2 excitations.[2] All scans were reviewed by a clinical neuroradiologist to rule out gross pathology. The MR images released have undergone analysis at the Center for Morphometric Analysis (CMA) at the Massachusetts General Hospital which includes: preprocessing ('positional normalization' to put the image into the standard orientation of the Talairach coordinate space,[3] and bias field correction[4]) and 'general segmentation' following the CMA segmentation protocol.[5] Segmented regions include: Cerebral Cortex and White Matter, Cerebellum Cortex and White Matter, Lateral Ventricle, Thalamus, Ventral Diencephalon, Caudate, Putamen, Pallidum, Accumbens, Hippocampus and Amygdala, bilaterally; as well as Brain Stem, 3rd and 4th ventricles. Basic demographic details for each subject include diagnosis, age, gender, and handedness (Fig. 1).

The release is provided under the Creative Commons: Attribute license,[6] and is structured as four bundles (tar) of imaging data, one for each diagnostic group. Within each bundle, there are separate directories for each subject that includes the data. Release version V1.0 includes only the imaging data and basic demographics and is accessible with no limitations. Version V1.1 adds the segmentation data as an indexed label file in register with the MR image for each subject, and requires provision of a contact email address (via the NITRC registration process) and requires a 'clickthrough' acceptance of the licensing terms. We do not perceive these extra access 'hoops' as burdensome, but they are necessary. While we hope that the data provided are perfect and free of any errors, experience has taught that some unforeseen errors may be present. If found, we will correct any error, and re-release a new version of the data and are obligated to inform all prior users of the data of any errors and corrections that were necessitated in order to limit the impact of any errors.

In addition to the image releases, the CANDIShare portal begins to establish a more richly interconnected set of related informatics resources. The first example of this is the inclusion of linkages between each data release and their representation in the Internet Brain Volume Database (IBVD).[7] The IBVD is a web-based database of published brain neuroanatomic volumetric observations. As the images and segmentation in this data release supports the volumetric observations in a specific publication, this linkage completes the connection between the high-level, group volumetric observation and the detailed individualized data that supports it. As the IBVD is also interoperable with PubMed via the 'Link-Out' function, a PubMed user can get from the publication listing at PubMed, to the summary volumetric observations reported in IBVD, to the raw data that supports the observation in the

CANDIShare data release. Each of the diagnostic group bundles are linked to their respective group pages in the IBVD via the link exposed at the download page of the NITRC project.

This image and segmentation data release for a specific publication represents the first in a series of data representing the pediatric brain in health and disease from data collected in our lab over the past 15 years. Additional publications will be added, and more resource interconnectivity will be included to maximize the utility of the released data with other classes of available data. This release of information is designed to be dramatically greater than merely 'making the images available': each image is associated with substantial analytic results, many of which have been utilized in the preparation of various publications and comparisons. Moreover, these data will be most effectively shared with the research community when shared in a way that preserves the linkages between the images, the resultant analytic data and meta-data, and its relationships to other public sources of related information. The utility of this shared data includes potentially expanding the numbers of subjects available for other studies (when the imaging protocols are comparable), facilitating additional research findings in this specific data set through reuse by other laboratories, and additional community access to accurate, manually-labeled datasets to promote development and testing of data analysis software. Access to data of this sort has already been valuable in development of numerous analysis techniques.[8] In short, this represents a 'Knowledge Management' environment that will facilitate traversal of these data and linkages.

Acknowledgments

References

1. Frazier JA, Hodge SM, Breeze JL, Giuliano AJ, Terry JE, Moore CM, Kennedy DN, Lopez-Larson MP, Caviness VS, Seidman LJ, Zablotsky B, Makris N. Diagnostic and sex effects on limbic volumes in early-onset bipolar disorder and schizophrenia. Schizophr Bull. 2008 Jan; 34(1):37–46. [PubMed: 18003631]

2. Frazier JA, Chiu S, Breeze JL, et al. Structural Brain Magnetic Resonance Imaging of Limbic and Thalamic Volumes in Pediatric Bipolar Disorder. Am J Psychiatry. 2005 Jul.162:7.

3. Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers; 1988.

4. Worth AJ, Makris N, Patti MR, Goodman JM, Hoge EA, Caviness VS Jr, Kennedy DN. Precise segmentation of the lateral ventricles and caudate nucleus in MR brain images using anatomically driven histograms. IEEE Trans Med Imaging. 1998 Apr; 17(2):303–310. [PubMed: 9688163]

5. Filipek PA, Richelme C, Kennedy DN, Caviness VS Jr. The young adult human brain: an MRI-based morphometric analysis. Cereb Cortex. 1994 Jul-Aug;4(4):344–360. [PubMed: 7950308]

6. http://creativecommons.org/about/licenses/.

7. http://www.nitrc.org/projects/ibvd.

8. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage. 2011 Jun 1; 56(3):907–922. [PubMed: 21352927] Babalola KO, Patenaude B, Aljabar P, Schnabel J, Kennedy D, Crum W, Smith S, Cootes T, Jenkinson M, Rueckert D. An evaluation of four automatic methods of segmenting the subcortical structures in the brain. Neuroimage. 2009 Oct 1; 47(4):1435–1447. [PubMed: 19463960] Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002 Jan 31; 33(3):341–355. [PubMed: 11832223]
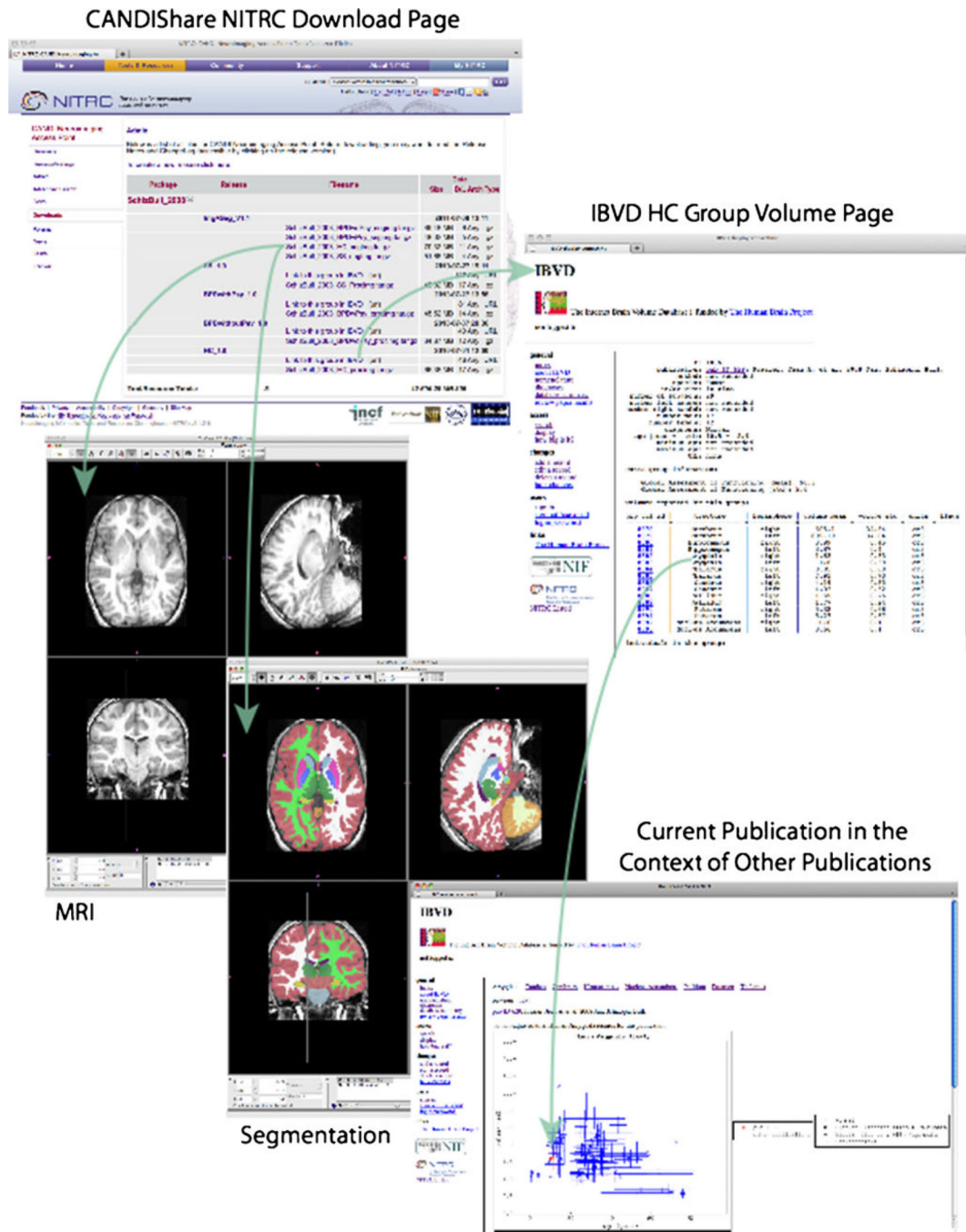
Fig. 1.
The figure shows the data and resource relationships accessible from the CANDIShare data release. From the CANDIShare NITRC download page, both version V1.0 and V1.1 can be accessed. Accessing V1.1 for the healthy control group (HC) gains the user access to the 29 individual subject MRI and segmentation results. Linkage is also provided to the group volumetric records pages in the IBVD, the example shown here is for the HC group. Finally, all publications in the IBVD can then be shown in the context of all other indexed volumetric publications for each of the anatomic structures; amygdala is demonstrated in this example