



The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry

Kate Brody Nooner^{1,2}, Stanley J. Colcombe¹, Russell H. Tobe¹, Maarten Mennes^{3,4}, Melissa M. Benedict¹, Alexis L. Moreno¹, Laura J. Panek¹, Shaquanna Brown¹, Stephen T. Zavitz¹, Qingyang Li⁵, Sharad Sikka⁵, David Gutman⁶, Saroja Bangaru³, Rochelle Tziona Schlachter¹, Stephanie M. Kamiel¹, Ayesha R. Anwar⁵, Caitlin M. Hinz⁵, Michelle S. Kaplan⁵, Anna B. Rachlin⁵, Samantha Adelsberg³, Brian Cheung⁵, Ranjit Khanuja⁵, Chaogan Yan⁵, Cameron C. Craddock⁷, Vincent Calhoun^{8,9}, William Courtney⁹, Margaret King⁹, Dylan Wood⁹, Christine L. Cox³, A. M. Clare Kelly³, Adriana Di Martino³, Eva Petkova^{1,3}, Philip T. Reiss^{1,3}, Nancy Duan⁵, Dawn Thomsen⁵, Bharat Biswal¹⁰, Barbara Coffey^{1,6}, Matthew J. Hoptman^{1,6}, Daniel C. Javitt^{1,11}, Nunzio Pomara^{1,6}, John J. Sidtis^{1,6}, Harold S. Koplewicz⁵, Francisco Xavier Castellanos^{1,3}, Bennett L. Leventhal¹ and Michael P. Milham^{1,5 *}

¹ Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA

² Psychology Department, University of North Carolina, Wilmington, NC, USA

³ Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, New York University Child Study Center, New York University Langone Medical Center, New York, NY, USA

⁴ Donders Institute for Brain, Cognition and Behavior, Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands

⁵ Center for the Developing Brain, Child Mind Institute, New York, NY, USA

⁶ Department of Psychiatry, New York University Langone Medical Center, New York, NY, USA

⁷ Virginia Tech Carilion Research Institute, Roanoke, VA, USA

⁸ Department of Psychiatry, University of New Mexico, Albuquerque, NM, USA

⁹ The Mind Research Network, Albuquerque, NM, USA

¹⁰ Department of Radiology, New Jersey Medical School, Newark, NJ, USA

¹¹ Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, USA

Edited by:

Satrajit S. Ghosh, Massachusetts Institute of Technology, USA

Reviewed by:

Yaroslav O. Halchenko, Dartmouth College, USA

Gabriele Lohmann, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

*Correspondence:

Michael P. Milham, Nathan S. Kline Institute, 140 Old Orangeburg Road, Orangeburg, NY 19062, USA; Child Mind Institute, 445 Park Avenue, New York, NY 10022, USA.
e-mail: mmilham@nki.rfmh.org

The National Institute of Mental Health strategic plan for advancing psychiatric neuroscience calls for an acceleration of discovery and the delineation of developmental trajectories for risk and resilience across the lifespan. To attain these objectives, sufficiently powered datasets with broad and deep phenotypic characterization, state-of-the-art neuroimaging, and genetic samples must be generated and made openly available to the scientific community. The enhanced Nathan Kline Institute-Rockland Sample (NKI-RS) is a response to this need. NKI-RS is an ongoing, institutionally centered endeavor aimed at creating a large-scale ($N > 1000$), deeply phenotyped, community-ascertained, lifespan sample (ages 6–85 years old) with advanced neuroimaging and genetics. These data will be publically shared, openly, and prospectively (i.e., on a weekly basis). Herein, we describe the conceptual basis of the NKI-RS, including study design, sampling considerations, and steps to synchronize phenotypic and neuroimaging assessment. Additionally, we describe our process for sharing the data with the scientific community while protecting participant confidentiality, maintaining an adequate database, and certifying data integrity. The pilot phase of the NKI-RS, including challenges in recruiting, characterizing, imaging, and sharing data, is discussed while also explaining how this experience informed the final design of the enhanced NKI-RS. It is our hope that familiarity with the conceptual underpinnings of the enhanced NKI-RS will facilitate harmonization with future data collection efforts aimed at advancing psychiatric neuroscience and nosology.

Keywords: fMRI, DTI, lifespan, brain, phenotype, psychiatry, discovery, open science

INTRODUCTION

Discovery science promises to transform our understanding of human brain function and the impact of neuropsychiatric illness. Traditionally focused on the generation and testing of specific hypotheses, the neuroimaging community is increasingly realizing the value of data exploration techniques capable of uncovering previously unappreciated links between behavior and brain function (Van Horn and Gazzaniga, 2002; Bilder et al., 2009; Cichon

et al., 2009; Biswal et al., 2010; Van Dijk et al., 2010). Beyond the identification of novel brain-behavior associations, discovery approaches have the potential to provide the bases for normative trajectories of brain structure and function across the lifespan (Giedd, 1999, 2004; Gogtay et al., 2004; Sowell et al., 2004; Dosenbach et al., 2010; Giedd and Rapoport, 2010; Thompson et al., 2011). Similar to the use of normative physical growth (weight, height) charts in pediatric medicine (Falkner, 1958; Nellhaus,

1968), such trajectories would facilitate the identification and characterization of pathophysiological processes contributing to the emergence of neuropsychiatric illness (Castellanos et al., 2002; Ge et al., 2002; Evans and Brain Development Cooperative Group, 2006; Shaw et al., 2006a,b, 2008, 2010; Gogtay et al., 2007; Dosenbach et al., 2010; Giedd and Rapoport, 2010; Giedd et al., 2010; Thompson et al., 2011). Potentially most exciting is the possibility of revealing markers in early life that have predictive value for the later emergence of illness (Riverol and López, 2011; Shim and Morris, 2011; Taber et al., 2011). Whether during childhood, early adulthood (e.g., Autism, Schizophrenia), or later in life (e.g., Alzheimer's Disease, Parkinson's Disease), identification of early biomarkers could transform the delivery of health care by helping to tailor resources and technology to the needs of an individual (i.e., personalized medicine), thus maximizing the likelihood of success for prevention and early intervention strategies.

Successful implementation of discovery science in the imaging community hinges on the accrual of large-scale imaging datasets from individuals who are phenotyped both deeply and broadly (Gogtay et al., 2004; Tracy, 2008; Lanktree et al., 2010). Unfortunately, such datasets are rare. An obvious initial hurdle is the cost of neuroimaging. Magnetic Resonance Imaging (MRI) scanning for research costs between \$400 and \$750 per hour throughout the world. Unlike genetics, where good quality samples can be obtained nearly anywhere with relative ease, the acquisition of high quality MRI data can be impacted by a variety of issues ranging from scanner-related variation to human factors. Data loss due to factors such as motion is substantial, particularly in child, aging, and clinical populations (Epstein et al., 2007; Power et al., 2012; Van Dijk et al., 2012). Scanner-related anxiety (e.g., claustrophobia) and difficulties following instructions (e.g., participants with severe autism, intellectual disability, psychosis, depression, or mania) represent additional obstacles to obtaining high quality data. Accordingly, funding for sufficiently large samples is beyond the scale of most grant mechanisms. Additionally, there are numerous impediments to deep and broad phenotypic characterizations of large groups of individuals, including recruitment and assessment costs, participant burden, staffing requirements (especially for handling developing and aging populations), and data management (e.g., collection, scoring, storage, retrieval).

The comprehensive characterization of human brain function and structure across the lifespan carries additional challenges with respect to experimental design (National Institute of Mental Health, 2008; Giedd and Rapoport, 2010). Imaging studies tend to focus on narrow comparisons (e.g., pediatrics vs. young adult, or adult vs. aging), precluding the presentation of the full spectrum of typical development, maturation, and aging within a single study (Sowell et al., 2004; Evans and Brain Development Cooperative Group, 2006; Fair et al., 2007; Dosenbach et al., 2010; Tamnes et al., 2010; Allen et al., 2011). This limitation is likely due to the logistical and financial challenges associated with the large sample sizes required to adequately encompass the lifespan with sufficient coverage to achieve statistically significant results. Moreover, the scarcity of imaging and phenotypic assessment tools validated for use across the entire lifespan hampers effective and meaningful acquisition of such data. Additionally, most researchers tend to gravitate toward the beginning or the end of the

lifespan, where age-related changes are greatest (Thompson et al., 2011). This produces gaps in our understanding of age-related trajectories of brain structure and function across the lifespan.

The Nathan S. Kline Institute for Psychiatric Research (NKI), funded and operated by the New York State Office of Mental Health, is attempting to address the challenges and capitalize on the opportunities of a lifespan study through the creation of the Nathan Kline Institute-Rockland Sample (NKI-RS). The NKI-RS is intended to be a large-scale, community-ascertained lifespan sample comprised of neuroimaging and genetic data coupled with neurocognitive, physiologic, behavioral, and psychiatric measurements. This initiative brings together researchers from a broad range of disciplines (e.g., basic and systems neuroscience, biostatistics, engineering, computer science, psychiatry, psychology, social work), with interests spanning a range of disorders. The NKI-RS group is focused on developing a unique neuroimaging and genetic sample, linked with descriptive metadata that incorporates solutions for the many challenges facing discovery. First, the age range for the design spans human development from childhood to late adulthood (6–85 years old). Second, the project is applying state-of-the-art resting-state structural and functional MRI (R-fMRI) and diffusion tensor imaging (DTI) techniques (Feinberg et al., 2010; Moeller et al., 2010; Feinberg and Yacoub, 2012; Smith et al., 2012), which minimize the obsolescence of these data along with participant burden. Third, imaging and genetic data are accompanied by a comprehensive phenotypic characterization (e.g., psychiatric, neurocognitive, psychological, and behavioral), to facilitate identification of developmental patterns (Evans and Brain Development Cooperative Group, 2006; Shaw et al., 2006a,b; Giedd and Rapoport, 2010; Giedd et al., 2010). Fourth, the project is grounded on the principles of open neuroscience, with the goal of prospective, pre-publication sharing of all collected data. Finally, while it is generally common practice in imaging studies to overlook concerns about the representativeness of datasets (Szklo, 1998; Evans and Brain Development Cooperative Group, 2006), the second phase of the NKI-RS (i.e., the enhanced NKI-RS) has been designed as a community-ascertained sample closely paralleling U.S. demographic distributions, thus, minimizing potential sampling biases and maximizing representativeness.

While no single effort can create the large-scale datasets necessary to deliver the entirety of normative assessments of the lifespan, it is our intent that this initial effort will highlight the challenges and provide a model through which such data can be acquired and shared. In the following sections, we elaborate on the challenges presented by discovery science and describe the strategies we have adopted in response. We discuss design considerations related to sampling, assessment, and other methodological choices that best characterize brain structure and function across the lifespan and which are most amenable to data sharing. Last, we discuss the specific steps we took to create the first, pilot phase of the NKI-RS, and then describe the conceptual underpinnings of phase two, the enhanced NKI-RS, now being executed.

MOTIVATION AND CHALLENGES FOR A LIFESPAN SAMPLE DESIGN

The developmental origins of most neuropsychiatric illnesses are increasingly being appreciated. Nearly 75% of mental illness in

adults originates prior to age 24 years (Kessler et al., 2005), and numerous links have been identified between the presence of childhood psychiatric problems and the later onset of adult illness (e.g., pediatric anxiety is associated with increased risk of adult depression; Drevets, 2003; Milham et al., 2005). Whether considering disorders affecting children, adolescents, adults, or the elderly, early detection of disease risk and/or onset is the critical first step in prevention and treatment, respectively (Kessler et al., 2005; Kessler and Wang, 2008). In this regard, the imaging community is increasingly hopeful that normative assessments of brain development, maturation, and aging can be obtained (Evans and Brain Development Cooperative Group, 2006; Fair et al., 2007, 2008, 2009; Church et al., 2009a,b; Kelly et al., 2009; Dosenbach et al., 2010; Giedd and Rapoport, 2010; Zuo et al., 2010; Allen et al., 2011). We anticipate that these normative trajectories will facilitate the identification of markers of pathologic development capable of 1 day informing multiple aspects of clinical assessment and decision-making – ranging from determinations of risk, diagnosis, and prognosis, to the selection and timing of interventions, as well as treatment response monitoring.

While conceptually attractive, building a large-scale imaging dataset that comprehensively samples the lifespan poses daunting challenges. The gold standard for studies of trajectories is the longitudinal design (Kraemer et al., 2000; Thompson et al., 2011). Unfortunately such designs are expensive and generally impractical due to time requirements and cumulative attrition (e.g., loss to follow-up from one time-point to the next, missing data).

A more tractable approach is a lifespan, cross-sectional design that assesses individuals spanning a broad age range to infer developmental, maturational, and aging trajectories. Although more practical, cross-sectional methods can be biased by differential recruitment along the lifespan (e.g., unintended differences in socioeconomic, intellectual, or behavioral characteristics among age-cohorts; Kraemer et al., 2000; Pediatric Imaging, Neurocognition, and Genetics (PING), 2011; Thompson et al., 2011). Hybrid, longitudinal, cross-lag designs that involve sampling individuals cross-sectionally across the lifespan, but following each of them longitudinally, albeit for briefer periods (e.g., 3 or 5 years), hold the greatest potential (Shaw et al., 2010; Thompson et al., 2011). However, these designs still engender significant costs and are hindered by increased potential for data loss.

FROM LABS TO COLLABORATIVES

It is fair to say that the majority of advances in clinical neuroscience over the past century have emerged through the accumulated contributions of individual labs, each collecting, analyzing, and interpreting its own data independently. However, as the scale and complexity of scientific inquiry increase, collaborative efforts are increasingly essential in order to attain samples of sufficient size and adequate statistical power.

A number of models have emerged to foster the necessary collaboration. For example, the multi-investigator, multi-center model, commonly employed by the pharmaceutical industry, has been effectively implemented by efforts such as the Biomedical Informatics Research Network (BIRN; Helmer et al., 2011) or the Alzheimer Disease Neuroimaging Initiative (ADNI; Mueller et al.,

2005). Although successful, such initiatives typically require considerable investment, limiting their growth. Recently, The Brain Genomics Superstruct Project (Buckner, 2012) demonstrated that efforts of similar or even grander scale can be undertaken while containing expenses. Specifically, the Superstruct effort added an optimized 15-min imaging acquisition protocol to ongoing studies at multiple sites and rapidly generated thousands of imaging datasets (Yeo et al., 2011). Web-based questionnaire and performance protocols were included to obtain comprehensive phenotyping while minimizing costs. The Mind Research Network (MRN; The Mind Research Network for Neurodiagnostic Discovery, 2012) provided another cost-effective model for large-scale data-generation by forming a collaborative (i.e., collaboration of laboratories) of independent investigators within and across multiple institutions united through the usage of a common informatics platform. Within the MRN, investigators can opt to share data with specific members, the larger collaborative or more broadly (e.g., see Allen et al., 2011). Finally, in recent years, uncoordinated, multi-center aggregation efforts such as the 1000 Functional Connectomes Project and its International Neuroimaging Data-sharing Initiative have emerged as open science solutions to the challenge of large-scale data aggregations (1000 Functional Connectomes Project (FCP), 2009; Biswal et al., 2010; Dolgin, 2010; Milham, 2012).

It is against this background that the NKI-RS emerged with the goal of building an institution-based open sharing model. The NKI-RS effort was designed to pool the global resources of an institution for the purpose of generating a large-scale, deeply phenotyped dataset reflective of the interests of its many investigators. Simultaneously, the open sharing of datasets was intended to facilitate investigations around the world and promote the generation and sharing of large-scale datasets at institution levels. The notion is that overlap of phenotypic protocols among openly shared datasets could rapidly accelerate the pace of discovery.

NKI-ROCKLAND SAMPLE

In conceptualizing the NKI-RS initiative, a major focus was the creation of an institution-wide resource, reflective of the diverse interests of the NKI faculty, spanning pediatric, adult, and geriatric psychiatric illnesses. One key purpose was to create a data repository that could be used to test existing hypotheses as well as for generating novel hypotheses to spark new endeavors. The initial (pilot) phase of the NKI-RS was designed to demonstrate the feasibility of an institutionally based, discovery science project. Institutional support and resources were central to the success of the pilot, as it was conducted without dedicated external funding. The success of the pilot phase was instrumental in the attainment of NIMH funding for phase two, which embodies more sophisticated recruitment and sampling strategies, phenotyping, imaging, and neuroinformatics. The following section details the strengths and limitations of the first phase of the NKI-RS, as well as the design choices for the second phase.

THE NKI-RS PILOT (PHASE I)

The goal of the pilot phase was to obtain diagnostic and behavioral assessments, tissue for genetic studies, and brain imaging (i.e., structural MRI, R-fMRI, diffusion imaging, and morphometry) on

250 individuals aged 4–89 years. More than 300 phenotypic variables were obtained across 26 psychiatric, behavioral, and cognitive domains. Additionally, participants consented in writing to unrestricted distribution of anonymous data through the International Neuroimaging Data-Sharing Initiative (INDI¹; 1000 Functional Connectomes Project (FCP), 2009). Prospective data sharing for the pilot phase of the project began on a regular basis in October of 2010. In 11 months, the NKI-RS collected and released data from 250 individuals, demonstrating that the pace (approximately five datasets released per week) and the process of open pre-publication data sharing were feasible.

Despite its successes, the pilot phase of the NKI-RS also highlighted areas with substantial room for enhancement and innovation. First, phase one relied on a convenience sample consisting of any individual who was willing to participate within the designated age range (ages 4–89). This approach is vulnerable to recruitment biases that can diminish the representativeness of the acquired sample (Szklo, 1998; Evans and Brain Development Cooperative Group, 2006). Such biases can compromise the generalizability and reproducibility of findings. Efforts such as the NIH Normal Brain Development Study have demonstrated the feasibility and value of increasing representativeness through tracking and balancing regional demographics for participants based on zip code (Waber et al., 2007). Second, the phenotypic battery consisted primarily of convenience assessments based on current practices and availability, as well as the interests of individual NKI-RS investigators. In retrospect, it would have been preferable to use commonly available, normed, and validated assessments to increase their utility and overlap with those employed by other efforts in the research community. Third, the pilot phase used paper and pencil assessments, which required scoring, entering, and checking data by hand, a time-consuming and error-prone task. Further, the data entered into the database was limited to summary scores. The following sections discuss steps taken in the construction of the enhanced NKI-RS to address these limitations.

THE ENHANCED NKI-RS (PHASE II): SAMPLING AND RECRUITMENT

Phase two began in March 2012, with an anticipated 4-year project period to recruit 1000 participants. It was designed to yield a community-ascertained, lifespan sample (0.32% of the population) in which age, ethnicity, and socioeconomic status are representative of Rockland County, New York. Rockland County is a suburban/rural county 20 miles northwest of New York City, with a population of 311,687 per the 2010 Census. Fortunately, ethnic and economic demographics of Rockland County resemble those of the United States (U.S. Census Bureau, 2010), increasing the generalizability of the NKI-RS to the broader U.S. population (Table 1).

Following the model of efforts such as the NIH Normal Brain Development Study, zip code based recruitment (e.g., advertisement flyer mailings, posting of materials in local shops and meeting places) and enrollment efforts are being used to avoid over-representation of any portion of the community, and to

Table 1 | 2010 United States census data: Rockland County versus United States.

People facts (Census, 2010)	Rockland county	USA
Population	311,687	308,745,538
Persons under 5 years old	7.6%	6.5%
Persons under 18 years old	28.1%	24.0%
Persons 65 years old and over	13.4%	13.0%
Female persons, percent, 2010	51.0%	50.8%
White	73.2%	72.4%
Black or African American	11.9%	12.6%
American Indian/Alaska Native	0.3%	0.9%
Asian	6.2%	4.8%
Native Hawaiian or other Pacific Islander	0.0%	0.2%
Two or more races reported	2.5%	2.9%
Hispanic or Latino	15.7%	16.3%
White, non-Hispanic	65.3%	63.7%
Foreign born, 2006–2010	22.1%	12.7%
Language other than English spoken at home	35.6%	20.1%
High school graduates	87.9%	85.0%
Bachelor's degree or higher	40.7%	27.9%
Persons per household, 2006–2010	3.02	2.59
Median household income, 2006–2010	\$82,534	\$51,914
Per capita money income, 2006–2010	\$34,304	\$27,334
Persons below poverty level	11.3%	13.8%

This figure contains 2010 census data for Rockland County in the State of New York as well as for the United States of America (USA). The purpose of this figure is to demonstrate that the census composition of Rockland County is similar to that of the USA as a whole. Therefore, data from this discovery science project based in Rockland County is likely to generalize to the USA.

ensure faithful representation of Rockland County. We monitor and adjust enrollment as necessary to ensure that the relative proportions of age, sex, and ethnicity accrued remain stable throughout the 4-years of this project, thereby minimizing potential cohort biasing effects (e.g., enrolling from one sub-population primarily in year one, and from another in year four).

Practical and logistical limitations dictated some constraints on the age range of our sample. Although children as young as 4 years of age were imaged successfully in the initial NKI-RS, we selected 6 years of age as the lower age limit to balance data losses with scientific yield. Similarly, we opted to truncate the upper limit at age 85 (versus 89) because of the dramatically increased rate of chronic illness above age 85. Additionally, the second phase of this project has intentional oversampling of the extremes of the lifespan (youngest and oldest) to increase statistical power for ages characterized by greatest changes (Table 2).

THE ENHANCED NKI-RS: ASSESSMENT

Phase two of the NKI-RS project will contain broader and deeper phenotypic characterization of participants, with a focus on key psychiatric and neurocognitive constructs. The battery for the second phase was constructed based on discussions with assessment developers, expert consultants, and a formal presentation

¹http://fcon_1000.projects.nitrc.org

Table 2 | Enrollment strategy for enhanced NKI-RS.

Age range	Target enrollment
6–10	150
11–20	150
21–30	100
31–40	75
41–50	100
51–60	125
61–70	150
71–85	150

This figure shows the age range in years with corresponding enrollment targets for the NKI-RS. The total enrollment is 1000 individuals ages 6–85 years in a 4-year period. There is intentional oversampling at the ends of the age range due to the rapidity of developmental changes associated with youth and older age.

to the Child Mind Institute's Scientific Research Council (SRC)². We prioritized inclusion of empirically validated measures in the public domain, as those are most amenable to widespread adoption in other studies. Additionally, we prioritized use of measures that could be administered and compared across the lifespan. In attempting to serve as a resource for future studies, the second phase compares commonly used assessments that measure the same construct, behavior, or disorder. We also compare proprietary and non-proprietary assessments (e.g., the Conners ADHD Scales (Conners et al., 1997; Conners, 1999) versus Strengths and Weaknesses of ADHD-Symptoms and Normal-Behavior Scale (SWAN; Hay et al., 2007, respectively), and can assess the construct validity of different assessments (e.g., the Computerized Neurocognitive Battery (Gur et al., 2001) versus the Delis–Kaplan Executive Functioning System (Delis et al., 2004; Figure 1).

Our approach, similar to that of the Brain Genomics Super-struct (Yeo et al., 2011), differs from the more common model of centering collaborative efforts on a particular disorder or set of disorders, which can limit the applicability of a comparison sample. Given the current focus on developing a dimensional framework for psychiatric illnesses and patterns of comorbidity (Chabernaude et al., 2012), we adopted broad phenotypic characterization for phase two of NKI-RS. By employing a common protocol that covers a wide array of domains of psychiatric, cognitive, and behavioral functions, we can make direct comparisons between psychiatric illnesses and increase the feasibility of determining overlap and distinctions among their neural correlates. During review of the finalized phenotyping protocol by the SRC, a key concern that emerged was that the comprehensiveness of the phenotyping protocol increased the burden to participants and experimenters – potentially endangering its effectiveness due to factors such as fatigue and increased data management needs. To address these concerns, the NKI-RS protocol was decompressed from a 1- to a 2-day format. As discussed below, state-of-the-art computer based data entry, scoring, and management capabilities were added, thereby minimizing burden on both participants

and experimenters³. Additionally, we carried out focus-group testing prior to initiation of the sample, and are obtaining customer satisfaction surveys and monitoring participant feedback as we progress so that small tweaks to the protocol can be made as necessary (preferably within the first 100 participants; Figures 2 and 3).

A key innovation of the second phase NKI-RS is the implementation of fast repetition time (0.645 and 1.4 s TR) and high-resolution (3 and 2 mm isotropic voxels) multiband R-fMRI (10 min per scan), and DTI (137-direction, 2 mm isotropic) measures provided by the Center for Magnetic Resonance Research at the University of Minnesota for the Human Connectomes Project (Feinberg et al., 2010; Smith et al., 2012; see⁴). Additionally, NKI-RS included a brief visual-checkerboard stimulation scan (duration = 2 min) for each of the multiband sequences to allow for assessment of the contrast to noise ratio. The NKI-RS and HCP efforts are independent projects, the former primarily focused on examination of brain-behavior relationships across the lifespan, and the latter focused on a twin and family based study of genetic-brain-behavior relationships in young adults (Van Essen et al., 2012). Despite these differences in focus, it is anticipated that inclusion of fast TR protocols will increase the ability of scientists to maximize the areas of overlap and compare or aggregate data obtained from the HCP and NKI-RS samples.

THE ENHANCED NKI-RS: DATABASE AND DATA SHARING

The broad and deep phenotyping of the enhanced NKI-RS raised several issues regarding data entry and administration. First, we could not continue paper and pencil approaches with research assistants entering summary scores into spreadsheets, as this inevitably leads to errors, which are expensive to find, correct, or prevent. Additionally, the practice of logging summary scores alone is inherently flawed, as potentially valuable item-level information is lost and typically too expensive to recover later. Second, integrating phenotypic and imaging data is non-trivial, and an undesirable potential source of error for investigators (Marcus et al., 2007). Fortunately, packages such as XNAT (Marcus et al., 2007), LORIS (Longitudinal Online Research and Imaging System (LORIS), 2011), HID (Helmer et al., 2011), and COINS (Scott et al., 2011) have emerged as viable options, though their usage is still relatively limited.

From these options, the NKI-RS team selected the Collaborative Informatics and Neuroimaging Suite (COINS) developed by the Mind Research Network (Scott et al., 2011; The Mind Research Network for Neurodiagnostic Discovery, 2012). COINS was created to facilitate communication and cultivate a data-sharing community by providing researchers with an open source information system that includes web-based tools to manage studies, subjects, imaging, and phenotypic data. This suite of tools has an intuitive ease of use and offers versatile data upload/import/entry

²<http://www.childmind.org/en/directory/src/>

³Measures taken to minimize subject and experimenter burden in response to CMI SRC recommendations (e.g., decompression from 1- to 2-day protocol, enhancement of informatics) were made possible by a grant provided by the Child Mind Institute, Inc (1FDN2012-1).

⁴http://fcon_1000.projects.nitrc.org/indi/pro/eNKI_RS_TRT/FrontPage.html for protocol specifications

General Information	Behavioral Measures
• Demographic Questionnaire (DEMOS)* • Edinburgh Handedness Inventory (EHI) • Hollingshead Four Factor Index of Socioeconomic Status (SES)* • Medical History Questionnaire (Med-Hist) • Medical Conditions Questionnaire* • Medications Questionnaire*	• 21-Item Peters et al. Delusions Inventory (PDI-21)(13+) • Autism Spectrum Screening Questionnaire (ASSQ)(6-17) • Behavioral Assessment System for Children (BASC-2)(6-17) • Cambridge-Hopkins Restless Leg Syndrome Questionnaire (CHRLS)(13+) • Child Behavioral Checklist (CBCL)(6-17); Achenbach Youth Self Report (YSR)(11-17); Adult Self Report (ASR)(18-59); Older Adult Self Report (OASR)(60+) • Child Eating Behavior Questionnaire (CEBQ)(6-11); Three Factor Eating Questionnaire (TFEQ)(12+) • Children's Behavior Questionnaire (CBQ)(6-8); Early Adolescent Temperament Questionnaire Parent Report (EATQ)(9-15); Adult Temperament Questionnaire (ATQ)(16+) • Children's Depression Inventory (CDI-II)(7-17); Beck Depression Inventory-II (BDI)(18-64); Geriatric Depression Scale-Long Form (GDS-LF)(65+) • Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS)(6-17)*; Yale-Brown Obsessive Compulsive Scale (Y-BOCS)(18+)* • Cognitive Failures Questionnaire (CFQ)(15+) • Comprehensive Adolescent Severity Inventory – Alcohol and Other Drugs (CASI-AOD)(11+) • Conners' Parent Rating Scale-Revised-Short (CPRS-R-S)(6-17); Conner-Wells' Adolescent Self-Report Scale- Short (CASS-S)(8-17); Conners Adult ADHD Rating Scales (CAARS)(18+) • DOSPERT Risk Taking Scale (DOSPERT)(18+) • Eating Disorder Examination Questionnaire (EDEQ)(13+) • Fagerstrom Tolerance Questionnaire for Adolescents (FTAQ)(13-17); Fagerstrom Test for Nicotine Dependence (FTND)(18+) • International Physical Activity Questionnaire (IPAQ)(15+) • Interpersonal Reactivity Index (IRI)(13+) • Inventory of Callous-Unemotional Traits Parent Report (ICU-P)(6-17); Inventory of Callous-Unemotional Traits Youth Version (ICU-Y)(13+) • Multidimensional Anxiety Scale for Children (MASC)(8-17); State Trait Anxiety Inventory (STAI)(18+) • NEO Five Factor Inventory (NEO-FFI)(12+) • Pittsburgh Sleep Quality Index (PSQI)(13+) • Repetitive Behavior Scale-Revised (RBS-R)(6-17) • Social Responsiveness Scale, Parent Report (SRS)(6-17) • Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior Scale Parent Version (SWAN)(6-17) • Trauma Symptom Checklist for Children (TSC-C)(8-17); Trauma Symptom Checklist (TSC-40)(18+) • UCLA PTSD Reaction Index – Parent version (UCLA-RI-P)(6-17); UCLA PTSD Reaction Index for Children and Adolescents (UCLA-RI)(8+) • UPPS Impulsive Behavior Scale (UPPS-P)(18+) • Vineland Adaptive Behavior Scales Parent Rating Form, Second Edition (Vineland-II)(6+)* • Yale Global Tic Severity Scale (YGTSS)(6+)* • Youth Risk Behavior Surveillance System (YRBSS)(11-21)
Physical Measures	
• Actigraphy • Bike Test* • Blood draw: chemistry profile, lipid profile, thyroid profile, CBC with differential, lead level, genetics, pregnancy test* • Urine sample (Drug Test)(11+)* • Height/Weight* • Hip/Waist Measurements* • Ishihara Color Vision Test (Color)* • MRI Mock Scan* • MRI Scan* • Tanner Staging (TANN)(6-17) • Vital Signs*	
Neurocognitive Tasks*	
• Attention Network Task (ANT) • Computerized Neurocognitive Battery (CNB) • Delis-Kaplan Executive Function System (D-KEFS)(8+) • Grip strength • Grooved Purdue Pegboard • Wechsler Abbreviated Scale of Intelligence-II (WASI-II) • Wechsler Individual Achievement Test-II-Abbreviated (WIAT-II-A)	
Diagnostic Assessments*	
• Adult ADHD Clinical Diagnostic Scale (ACDS)(18+) • Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS-PL)(6-17) • Structured Clinical Interview for DSM-IV – Non-Patient edition (SCID-NP)(18+)	

Assessments for all participants ages 6 years and older unless indicated
(*) terms: Clinician administered

FIGURE 1 | Assessment protocol for NKI-RS. This figure illustrates all of the assessments that are included in the 2-day enhanced Nathan Kline Institute-Rockland Sample (NKI-RS) protocol. There are five broad domains of

assessment: General, Physical, Neurocognitive, Diagnostic, and Behavioral. Within the table are the names, abbreviations, and age ranges in years for each of the assessments.

options, rapid and secure sharing of data among investigators, querying of data types and assessments, real-time reporting, and study-management tools. Among its many features, the web-based assessments, automated data scoring, and integrated management of phenotypic and imaging data are potentially the most attractive. Web-based assessment entry completed by participants and research staff increases efficiency and accuracy by eliminating the need for intermediate data entry (i.e., paper to computer). Equally important, individual item-level responses are coded in the database, providing researchers with a far richer phenotypic dataset for exploration. In addition, protected health information can be unlinked within COINS to facilitate data sharing while maximally protecting participant anonymity. Of note, COINS is in compliance with Health Insurance Portability and Accountability Act (HIPAA) standards and implementation rules.

Confidentiality was a paramount consideration in planning data-sharing requirements for phase two. Protecting participant privacy while also providing access to extensively revealing data was a goal. In the pilot phase, all imaging data were

fully anonymized in compliance with HIPAA by removing any potential protected health information identifiers, including identifying facial features from anatomical images, and randomizing the timing of release. It is important to note that data users must be aware of the possible negative impact of defacing on some analysis toolkits (e.g., FreeSurfer), and exercise additional care when producing such images and/or sharing pre-processed surfaces. Summary scores from all such measures were made publicly available, along with individuals' imaging data⁵. The same anonymization and distribution protocol was used in December of 2012, when the Enhanced NKI-Rockland team released a 24-participant multiband imaging test-retest pilot dataset, created to evaluate the cutting edge "fast TR" sequences provided by the Human Connectome Project for usage in the enhanced sample⁶.

Although successful for the pilot efforts, the phase two NKI-RS effort has two unique features that called for reconsideration of

⁵http://fcon_1000.projects.nitrc.org/indi/pro/nki.html

⁶http://fcon_1000.projects.nitrc.org/indi/pro/nKI_RS_TRT/FrontPage.html

Time		Time	
8:15am	Arrival, Orientation, Consent, MRI Screening, Contact Information	8:45am	Welcome Back
8:45am	Vitals, Blood Draw, Drug Test, Height/Weight, Hip/Waist, Grip Strength, Color, Medical Conditions Questionnaire, Medications Questionnaire, Actigraphy	9:00am	MRI
9:00am	Breakfast	10:00am	MRI Questionnaire
9:15am	Structured Clinical Interview for DSM-IV-Non-Patient Edition, ACDS, SES	10:15am	Y-BOCS, YGTSS, Grooved Purdue Pegboard, Vineland-II
10:15am	MRI Mock Scan	10:30am	ANT
11:00am	Computerized Neurocognitive Battery	10:45am	CASI-AOD, CHRLS, DOSPERT, EDEQ, IRI, PDI-21, PSQI, TFEQ, UPPS-P, YRBSS
12:00am	WIAT-II-A	12:00pm	Lunch
12:15pm	Lunch	12:30pm	Delis-Kaplan Executive Function System
12:45pm	WASI-II		
1:30pm	Bike	2:30pm	Satisfaction Questionnaire & Wrap-Up
1:45pm	ASR/OASR, ATQ, BDI/GDS-LF, CAARS, CFQ, Demographics, EHI, FTND, ICU-Y, IPAQ, Medical History Questionnaire, NEO-FFI, STAI, TSC-40, UCLA-RI		
2:45pm	Wrap-Up		

FIGURE 2 | Sample schedule for adult participants in NKI-RS. This figure illustrates the 2-day assessment schedule for adult participants (ages 18–85 years) in the Nathan Kline Institute-Rockland Sample (NKI-RS) protocol. Abbreviations for the assessments: ANT, Attention Network Task; ASR, Adult Self Report; ATQ, Adult Temperament Questionnaire; BDI, Beck Depression Inventory-II; CAARS, Conners' Adult ADHD Rating Scales; CASI-AOD, Comprehensive Adolescent Severity Inventory – Alcohol and Other Drugs; CHRLS, Cambridge-Hopkins Restless Leg Syndrome Questionnaire; CFQ, Cognitive Failures Questionnaire; DOSPERT, DOSPERT Risk Taking Scale; GDS-LF, Geriatric Depression Scale-Long Form; OASR, Older Adult Self Report; EDEQ, Eating Disorder Examination Questionnaire; EHI, Edinburgh Handedness Inventory; FTND, Fagerstrom Test for Nicotine Dependence;

ICU-Y, Inventory of Callous-Unemotional Traits Youth Version; IPAQ, International Physical Activity Questionnaire; IRI, Interpersonal Reactivity Index; NEO-FFI, NEO Five Factor Inventory; PDI-21, 21-Item Peters et al. Delusions Inventory; PSQI, Pittsburgh Sleep Quality Index; STAI, State Trait Anxiety Inventory; TFEQ, Three Factor Eating Questionnaire; TSC-40, Trauma Symptom Checklist; UCLA-RI, UCLA PTSD Reaction Index for Children and Adolescents; UPPS-P, UPPS Impulsive Behavior Scale; Vineland-II, Vineland Adaptive Behavior Scales Parent Rating Form, Second Edition; WASI-II, Wechsler Abbreviated Scale of Intelligence-II; WIAT-II-A, Wechsler Individual Achievement Test-II-Abbreviated; Y-BOCS, Yale-Brown Obsessive Compulsive Scale; YGTSS, Yale Global Tic Severity Scale; YRBSS, Youth Risk Behavior Surveillance System.

the data-sharing policy. First, the Enhanced NKI-RS differs from the pilot sample in that it is being obtained using a community-ascertained epidemiologic design, which requires residence in Rockland County. Accordingly, a given participant's residential location is identifiable to the level of a county, which is not in keeping with the definition of complete de-identification based on HIPAA's 18 protected health identifiers. Second, the concurrently supplied psychometric data in the Enhanced NKI-RS phenotypic protocol will include individual item-level data and an increased breadth of phenotypic sampling relative to the pilot NKI-RS efforts. The high-dimensionality of these data increases risk of identification far beyond that posed by revealing some of the 18 protected health identifiers specified by HIPAA. These concerns support the need for implementation of a data use agreement.

Given these considerations, the Enhanced NKI-RS Sample is requiring a data usage agreement for access to the data – a requirement similar to efforts such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the National Database for Autism Research (NDAR) (2009)⁷. The adoption of a data usage agreement is not intended to limit the specific analyses a researcher can perform; users will only need to specify the broad range of analyses they may pursue with the data (e.g., association studies between DTI, R-fMRI, and behavior), not a specific analysis or set of analyses. The intent of the agreement is to ensure that data users agree to protect participant confidentiality when handling data that contains potentially identifying information and that they

⁷<http://ndar.nih.gov/>

Time	Child	Parent
8:15am	Arrival, Orientation, Consent, MRI Screening, Contact Information	
8:45am	Vitals, Blood Draw, Drug Test, Height/Weight, Hip/Waist, TANN, Grip Strength, Color, Medical Conditions Questionnaire, Medications Questionnaire, Actigraphy	
9:00am	Breakfast	
9:15am	Computerized Neurocognitive Battery	K-SADS-PL
10:00am	K-SADS-PL	ASSQ, BASC-2, CBCL
10:30am	MRI Mock Scan	
11:15am	WASI-II	Vineland-II, SES
12:00pm	Lunch	
12:30pm	WIAT-II-A	CEBQ, CBQ, CPRS-R-S, Demographics, EATQ, ICU-P, Medical History Questionnaire, RBSR, SRS, SWAN, UCLA-RI-P
1:00pm	Bike	
1:15pm	ATQ, CDI-II, CASS-S, EHI, FTAQ, ICU-Y, IPAQ, MASC, TFEQ, TSC-C, UCLA-RI, YSR	
2:15pm	Wrap-Up	

Time	Child	Parent
8:45am	Welcome Back	
9:00am	MRI Scan	
10:00am	CASI-AOD, CFQ, EDEQ, IRI, MRI-Q, NEO-FFI, PDI-21, PSQI, YRBSS	
11:00am	CY-BOCS, YGTSS	
11:15am	Grooved Purdue Pegboard	
11:30am	ANT	
11:45am	Lunch	
12:15pm	Delis-Kaplan Executive Function System	
2:15pm	Satisfaction Questionnaire & Wrap-Up	

FIGURE 3 | Sample schedule for child and parent participants in NKI-RS.

This figure illustrates the 2-day assessment schedule for child and parent participants (children ages 6–17 years) in the Nathan Kline Institute-Rockland Sample (NKI-RS) protocol. Abbreviations for the assessments: ANT, Attention Network Task; ASSQ, Autism Spectrum Screening Questionnaire; ATQ, Adult Temperament Questionnaire; BASC-2, Behavioral Assessment System for Children; CASI-AOD, Comprehensive Adolescent Severity Inventory-Alcohol and Other Drugs; CASS-S, Conner-Wells' Adolescent Self-Report Scale-Short; CBCL, Child Behavioral Checklist; CBQ, Children's Behavior Questionnaire; CFQ, Cognitive Failures Questionnaire; CDI-II, Children's Depression Inventory-II; CEBQ, Child Eating Behavior Questionnaire; CPRS-R-S, Conners' Parent Rating Scale-Revised-Short; CY-BOCS, Children's Yale-Brown Obsessive Compulsive Scale; EATQ, Early Adolescent Temperament Questionnaire Parent Report; EDEQ, Eating Disorder Examination Questionnaire; EHI, Edinburgh Handedness Inventory; FTAQ, Fagerstrom Tolerance Questionnaire for Adolescents; ICU-P, Inventory of Callous-Unemotional Traits Parent Report; ICU-Y, Inventory of

Callous-Unemotional Traits Youth Version; IPAQ, International Physical Activity Questionnaire; IRI, Interpersonal Reactivity Index; K-SADS-PL, Kiddie Schedule for Affective Disorders and Schizophrenia; MASC, Multidimensional Anxiety Scale for Children; MRI-Q, Magnetic Resonance Imaging Questionnaire; NEO-FFI, NEO Five Factor Inventory; PSQI, Pittsburgh Sleep Quality Index; RBSR, Repetitive Behavior Scale-Revised; SES, Hollingshead Four Factor Index of Socioeconomic Status; SRS, Social Responsiveness Scale-Parent Report; SWAN, Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal-Behavior Scale-Parent Version; TANN, Tanner Staging; TFEQ, Three Factor Eating Questionnaire; TSC-C, Trauma Symptom Checklist for Children; UCLA-RI, UCLA PTSD Reaction Index for Children and Adolescents; UCLA-RI-P, UCLA PTSD Reaction Index-Parent version; Vineland-II, Vineland Adaptive Behavior Scales Parent Rating Form, Second Edition; WASI-II, Wechsler Abbreviated Scale of Intelligence-II; WIAT-II-A, Wechsler Individual Achievement Test-II-Abbreviated; YGTSS, Yale Global Tic Severity Scale; YRBSS, Youth Risk Behavior Surveillance System; YSR, Achenbach Youth Self Report.

will agree to take the necessary measures to prevent breaches of privacy. The specific agreement to be employed for the Enhanced NKI-RS are those previously defined by the New York State Office of Mental Health, which consist of two straightforward components: A Data Exchange Agreement and a Non-Disclosure of Confidential Information Agreement (forms can be found at^{8,9}).

⁸http://rocklandsample.rfmh.org/RocklandSample_dea.pdf

⁹http://rocklandsample.rfmh.org/RocklandSample_cnda.pdf

Unlike the NDAR agreement, institutional review board (IRB) approval is not required for transfer of the data; it will be up to the individual data user to satisfy any additional requirements specified by their local IRB or ethics committee, prior to using the NKI-RS. Given that local IRB approval is not required as part of an individual's application for access to the NKI-RS, there is no need for an individual's IRB to have a federal-wise assurance number – which can limit recipients of the NDAR datasets.

ETHICS STATEMENT

Institutional Review Board Approval was obtained for this project at the Nathan Kline Institute (Phase I #226781 and Phase II #239708) and at Montclair State University (Phase I #000983A and Phase II #000983B). Written informed consent was obtained for all study participants. Written consent and assent was also obtained from minor/child participants and their legal guardian.

A NOTE ON ERROR-HANDLING

Fears and attitudes regarding the reporting of data errors represent a major obstacle to open data sharing (Poldrack, 2011). Despite any group's best efforts to prevent errors from arising in the process of acquiring, handling, and distributing data, errors will undoubtedly arise. The pilot phase of the NKI-RS had to face such a challenge, when a slight deviation in the scoring of the DKEFS was noted. Although this error was relatively inconsequential, the team worked to rapidly report¹⁰ and correct it. Indeed, even relatively small errors in publicly shared data must be reported as soon as they are discovered, so that the community can be confident that the accuracy of the shared data is the best attainable. Without monitoring and ongoing open error reporting, errors will be perpetuated and their impact potentially magnified (Friedman and Glover, 2006).

CONCLUSION

Although the generation and release of data for 1000 broadly and deeply phenotyped participants with extensive neuroimaging data and archived genetic samples requires substantial effort, it is just the beginning of one project among the many needed to fully unravel the neurobiology of psychiatric disorders across the lifespan. The NKI-RS is intended to serve as a jumping off point for research that goes beyond individual institutions and has the power to obtain the truly large numbers needed to create normative trajectories in psychiatry. Attainment of normative lifespan trajectories will have a transformational effect on the way in which neuropsychiatric research is conducted (Gogtay et al., 2004; Sowell et al., 2004; Evans and Brain Development Cooperative

Group, 2006; Shaw et al., 2008, 2010; Biswal et al., 2010; Giedd and Rapoport, 2010).

Projects like the NKI-RS have the power to inform our understanding of the entire spectrum of psychiatric illness. It is our hope that the NKI-RS framework will inspire other institutions to join the era of discovery and revolutionize clinical practice for all of psychiatry, from children and adolescents to older adults. The ability to relate dimensional phenotypic measures to statistically normed brain relationships will support the identification of underlying pathophysiological mechanisms, which may ultimately transform psychiatric nosology, guide the diagnostic process, inform treatment selection, and permit tracking of therapeutic efficacy.

ACKNOWLEDGMENTS

We would like to thank Lawrence Maayan for his key role in the design and execution of the pilot NKI-RS, David Guilfoyle for his assistance with human imaging, and Raj Sangoi and Caxia Hu for the many hours contributed to scanning and providing invaluable feedback. Also, special thanks to Randy Buckner for providing us the specifications for the Brain Genomics Superstruct's Common Acquisition Protocol, which formed the basis of the imaging protocol in the pilot phase; to Kamil Urgurbil, Eddie Auerbach, Junquian Gordon Xu, and Steen Moeller from the HCP for assistance with the implementation of the multiplexed sequences and reconstruction algorithms; and to Satrajit Ghosh for his guidance in the development of our DICOM anonymization protocol.

Principal support for the enhanced NKI-RS project is provided by the NIMH BRAINS R01MH094639-01 (PI Milham). Funding for key personnel also provided in part by the New York State Office of Mental Health and Research Foundation for Mental Hygiene. Funding for the decompression and augmentation of administrative and phenotypic protocols provided by a grant from the Child Mind Institute (1FDN2012-1). Additional personnel support provided by the Center for the Developing Brain at the Child Mind Institute, as well as NIMH R01MH081218, R01MH083246, and R21MH084126. Project support also provided by the NKI Center for Advanced Brain Imaging (CABI), the Brain Research Foundation (Chicago, IL), and the Stavros Niarchos Foundation.

¹⁰http://www.nitrc.org/forum/forum.php?thread_id=2600&forum_id=1735

REFERENCES

- 1000 Functional Connectomes Project (FCP). (2009). Available at: http://fcon_1000.projects.nitrc.org/index.html [accessed February 2, 2012].
- Allen, E. A., Erhardt, E. B., Damaraju, E., Gruner, W., Segall, J. M., Silva, R. F., et al. (2011). A baseline for the multivariate comparison of resting-state networks. *Front. Syst. Neurosci.* 5:2. doi:10.3389/fnsys.2011.00002
- Bilder, R. M., Sabb, F. W., Cannon, T. D., London, E. D., Jentsch, J. D., Parker, D. S., et al. (2009). Phenomics: the systematic study of phenotypes on a genome-wide scale. *Neuroscience* 164, 30–42.
- Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739.
- Buckner, R. (2012). *The Brain Genomics Superstruct Project*. Available at: <http://sfari.org/funding/grants/abstracts/the-brain-genomics-superstruct-project> [accessed February 5, 2012].
- Castellanos, F. X., Lee, P. P., Sharp, W., Jeffries, N. O., Greenstein, D. K., Clasen, L. S., et al. (2002). Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *JAMA* 288, 1740–1748.
- Chabernaud, C., Mennes, M., Kelly, C., Nooner, K., Di Martino, A., Castellanos, F. X., et al. (2012). Dimensional brain-behavior relationships in children with attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 71, 434–442.
- Church, J. A., Fair, D. A., Dosenbach, N. U., Cohen, A. L., Miezin, F. M., Petersen, S. E., et al. (2009a). Control networks in paediatric Tourette syndrome show immature and anomalous patterns of functional connectivity. *Brain* 132, 225–238.
- Church, J. A., Wenger, K. K., Dosenbach, N. U., Miezin, F. M., Petersen, S. E., and Schlaggar, B. L. (2009b). Task control signals in pediatric tourette syndrome show evidence of immature and anomalous functional activity. *Front. Hum. Neurosci.* 3:38. doi:10.3389/fnhum.2009.0038.2009
- Cichon, S., Craddock, N., Daly, M., Faraone, S. V., Gejman, P. V., Kelsoe, J., et al. (2009). Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am. J. Psychiatry* 166, 540–556.
- Conners, C. K. (1999). Clinical use of rating scales in diagnosis and treatment of attention-deficit/hyperactivity disorder. *Pediatr. Clin. North Am.* 46, 857–870.

- Conners, C. K., Wells, K. C., Parker, J. D., Sitarenios, G., Diamond, J. M., and Powell, J. W. (1997). A new self-report scale for assessment of adolescent psychopathology: factor structure, reliability, validity, and diagnostic sensitivity. *J. Abnorm. Child Psychol.* 25, 487–497.
- Delis, D. C., Kramer, J. H., Kaplan, E., and Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan Executive Function System: an update. *J. Int. Neuropsychol. Soc.* 10, 301–303.
- Dolgin, E. (2010). This is your brain online: the Functional Connectomes Project. *Nat. Med.* 16, 351.
- Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., et al. (2010). Prediction of individual brain maturity using fMRI. *Science* 329, 1358–1361.
- Drevets, W. C. (2003). Neuroimaging abnormalities in the amygdala in mood disorders. *Ann. N. Y. Acad. Sci.* 985, 420–444.
- Epstein, J. N., Casey, B. J., Tonev, S. T., Davidson, M., Reiss, A. L., Garrett, A., et al. (2007). Assessment and prevention of head motion during imaging of patients with attention deficit hyperactivity disorder. *Psychiatry Res.* 155, 75–82.
- Evans, A. C., and Brain Development Cooperative Group. (2006). The NIH MRI study of normal brain development. *Neuroimage* 30, 184–202.
- Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., et al. (2008). The maturing architecture of the brain's default network. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4028–4032.
- Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U., Church, J. A., Miezin, F. M., et al. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Comput. Biol.* 5, e1000381. doi:10.1371/journal.pcbi.1000381
- Fair, D. A., Dosenbach, N. U., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507–13512.
- Falkner, F. (1958). Some physical measurements in the first three years of life. *Arch. Dis. Child.* 33, 1.
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Gunther, M., et al. (2010). Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS ONE* 5, e15710. doi:10.1371/journal.pone.0015710
- Feinberg, D. A., and Yacoub, E. (2012). The rapid development of high speed, resolution and precision in fMRI. *Neuroimage* 62, 720–725.
- Friedman, L., and Glover, G. H. (2006). Report on a multicenter fMRI quality assurance protocol. *J. Magn. Reson. Imaging* 23, 827–839.
- Ge, Y., Grossman, R. I., Babb, J. S., Rabin, M. L., Mannon, L. J., and Kolson, D. L. (2002). Age-related total gray matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis. *AJNR Am. J. Neuroradiol.* 23, 1327–1333.
- Giedd, J. N. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Neuroimage* 2, 861–863.
- Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Ann. N. Y. Acad. Sci.* 1021, 77–85.
- Giedd, J. N., and Rapoport, J. L. (2010). Structural MRI of pediatric brain development: what have we learned and where are we going? *Neuron* 67, 728–734.
- Giedd, J. N., Stockman, M., Weddle, C., Liverpool, M., Alexander-Bloch, A., Wallace, G. L., et al. (2010). Anatomic magnetic resonance imaging of the developing child and adolescent brain and effects of genetic variation. *Neuropsychol. Rev.* 20, 349–361.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179.
- Gogtay, N., Greenstein, D., Lenane, M., Clasen, L., Sharp, W., Gochman, P., et al. (2007). Cortical brain development in nonpsychotic siblings of patients with childhood-onset schizophrenia. *Arch. Gen. Psychiatry* 64, 772–780.
- Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., et al. (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology* 25, 766–776.
- Hay, D. A., Bennett, K. S., Levy, F., Sergeant, J., and Swanson, J. (2007). A twin study of attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses of ADHD-symptoms and normal-behavior (SWAN) scale. *Biol. Psychiatry* 61, 700–705.
- Helmer, K. G., Ambite, J. L., Ames, J., Ananthakrishnan, R., Burns, G., Chervenak, A. L., et al. (2011). Enabling collaborative research using the Biomedical Informatics Research Network (BIRN). *J. Am. Med. Inform. Assoc.* 18, 416–422.
- Kelly, A. M. C., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., et al. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cereb. Cortex* 19, 640–657.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* 62, 593–602.
- Kessler, R. C., and Wang, P. S. (2008). The descriptive epidemiology of commonly occurring mental disorders in the United States. *Annu. Rev. Public Health* 29, 115–129.
- Kraemer, H. C., Yesavage, J. A., Taylor, J. L., and Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *Am. J. Psychiatry* 157, 163–171.
- Lanktree, M. B., Hassell, R. G., Lahiry, P., and Hegele, R. A. (2010). Phenomics: expanding the role of clinical evaluation in genomic studies. *J. Invest. Med.* 58, 700–706.
- Longitudinal Online Research and Imaging System (LORIS). (2011). Available at: <http://cbrain.mcgill.ca/loris> [accessed February 1, 2012].
- Marcus, D. S., Olsen, T. R., Ramaratnam, M., and Buckner, R. L. (2007). The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics* 5, 11–34.
- Milham, M. P. (2012). Open neuroscience solutions for the connectome-wide association era. *Neuron* 73, 214–218.
- Milham, M. P., Nugent, A. C., Drevets, W. C., Dickstein, D. P., Leibenluft, E., Ernst, M., et al. (2005). Selective reduction in amygdala volume in pediatric anxiety disorders: a voxel-based morphometry investigation. *Biol. Psychiatry* 57, 961–966.
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., et al. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* 63, 1144–1153.
- Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., et al. (2005). Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement.* 1, 55–66.
- National Database for Autism Research (NDAR). (2009). Available at: <http://ndar.nih.gov/> [accessed February 14, 2012].
- National Institute of Mental Health. (2008). *National Institute of Mental Health Strategic Plan*. Available at: <http://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml> [accessed February 10, 2012].
- Nellhaus, G. (1968). Head circumference from birth to eighteen years. Practical composite international and interracial graphs. *Pediatrics* 41, 106–114.
- Pediatric Imaging, Neurocognition, and Genetics (PING). (2011). Available at: <http://ping.chd.ucsd.edu/> [accessed February 5, 2012].
- Poldrack, R. A. (2011). The future of fMRI in cognitive neuroscience. *Neuroimage* 62, 2016–2020.
- 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.
- Riverol, M., and López, O. L. (2011). Biomarkers in Alzheimer's disease. *Front. Neurol.* 2, 46.
- Scott, A., Courtney, W., Wood, D., de la Garza, R., Lane, S., King, M., Wang, R., et al. (2011). COINS: an innovative informatics and neuroimaging tool suite built for large heterogeneous datasets. *Front. Neuroinform.* 5:33. doi:10.3389/fninf.2011.00033
- Shaw, P., Gogtay, N., and Rapoport, J. (2010). Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. *Hum. Brain Mapp.* 31, 917–925.
- Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., et al. (2006a). Intellectual ability and cortical development in children and adolescents. *Nature* 440, 676–679.
- Shaw, P., Lerch, J., Greenstein, D., Sharp, W., Clasen, L., Evans, A., et al. (2006b). Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* 63, 540–549.
- Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., et al. (2008). Neurodevelopmental

- trajectories of the human cerebral cortex. *J. Neurosci.* 28, 3586–3594.
- Shim, Y. S., and Morris, J. C. (2011). Biomarkers predicting Alzheimer's disease in cognitively normal aging. *J. Clin. Neurol.* 7, 60–68.
- Smith, S. M., Miller, K. L., Moeller, S., Xu, J., Auerbach, E. J., Woolrich, M. W., et al. (2012). Temporally-independent functional modes of spontaneous brain activity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3131–3136.
- 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.
- Szklo, M. (1998). Population-based cohort studies. *Epidemiol. Rev.* 20, 81–90.
- Taber, K. H., Salpekar, J., Wong, A. H., and Hurley, R. A. (2011). Developmental origins for neuropsychiatric illness. *J. Neuropsychiatry Clin. Neurosci.* 23, 1–5.
- Tamnes, C. K., Ostby, Y., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., and Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb. Cortex* 20, 534–548.
- The Mind Research Network for Neurodiagnostic Discovery. (2012). Available at: <http://www.TheMindInstitute.org> [accessed February 5, 2012].
- Thompson, W., Hallmayer, J., and O'Hara, R. (2011). Design considerations for characterizing psychiatric trajectories across the lifespan: application to effects of APOE- $\epsilon 4$ on cerebral cortical thickness in Alzheimer's disease. *Am. J. Psychiatry* 168, 894–903.
- Tracy, R. P. (2008). "Deep phenotyping": characterizing populations in the era of genomics and systems biology. *Curr. Opin. Lipidol.* 19, 151–157.
- U.S. Census Bureau. (2010). *U.S. Census Bureau State & County Quickfacts: U.S.A.* Available at: <http://quickfacts.census.gov/qfd/states/00000.html> [accessed March 3, 2012].
- Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., and Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* 103, 297–321.
- Van Dijk, K. R., Savunvu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431–438.
- Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., et al. (2012). The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62, 2222–2231.
- 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.
- Waber, D. P., De Moor, C., Forbes, P. W., Almli, C. R., Botteron, K. N., Leonard, G., et al. (2007). The NIH MRI study of normal brain development: performance of a population based sample of healthy children aged 6 to 18 years on a neuropsychological battery. *J. Int. Neuropsychol. Soc.* 13, 729–746.
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165.
- Zuo, X. N., Kelly, C., Di Martino, A., Mennes, M., Margulies, D. S., Bangaru, S., 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.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 13 July 2012; accepted: 21 September 2012; published online: 16 October 2012.
- Citation:** Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, Panek LJ, Brown S, Zavitz ST, Li Q, Sikka S, Gutman D, Bangaru S, Schlachter RT, Kamiel SM, Anwar AR, Hinz CM, Kaplan MS, Rachlin AB, Adelsberg S, Cheung B, Khanuja R, Yan C, Craddock CC, Calhoun V, Courtney W, King M, Wood D, Cox CL, Kelly AMC, Di Martino A, Petkova E, Reiss PT, Duan N, Thomsen D, Biswal B, Coffey B, Hoptman MJ, Javitt DC, Pomara N, Sidtis JJ, Koplewicz HS, Castellanos FX, Leventhal BL and Milham MP (2012) The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* 6:152. doi: 10.3389/fnins.2012.00152
- This article was submitted to *Frontiers in Brain Imaging Methods*, a specialty of *Frontiers in Neuroscience*.
- Copyright © 2012 Nooner, Colcombe, Tobe, Mennes, Benedict, Moreno, Panek, Brown, Zavitz, Li, Sikka, Gutman, Bangaru, Schlachter, Kamiel, Anwar, Hinz, Kaplan, Rachlin, Adelsberg, Cheung, Khanuja, Yan, Craddock, Calhoun, Courtney, King, Wood, Cox, Kelly, Di Martino, Petkova, Reiss, Duan, Thomsen, Biswal, Coffey, Hoptman, Javitt, Pomara, Sidtis, Koplewicz, Castellanos, Leventhal and Milham. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.