

REVIEW

Clinically useful brain imaging for neuropsychiatry: How can we get there?

Michael P. Milham^{1,2} | R. Cameron Craddock^{1,2} | Arno Klein¹

¹Center for the Developing Brain, Child Mind Institute, New York, New York

²Center for Biomedical Imaging and Neuromodulation, Nathan S. Kline Institute for Psychiatric Research, New York, New York

Correspondence

Michael P. Milham, Phyllis Green and Randolph Cowen Scholar, Child Mind Institute, 445 Park Avenue, New York, NY 10022.
Email: michael.milham@childmind.org

Despite decades of research, visions of transforming neuropsychiatry through the development of brain imaging-based “growth charts” or “lab tests” have remained out of reach. In recent years, there is renewed enthusiasm about the prospect of achieving clinically useful tools capable of aiding the diagnosis and management of neuropsychiatric disorders. The present work explores the basis for this enthusiasm. We assert that there is no single advance that currently has the potential to drive the field of clinical brain imaging forward. Instead, there has been a constellation of advances that, if combined, could lead to the identification of objective brain imaging-based markers of illness. In particular, we focus on advances that are helping to (1) elucidate the research agenda for biological psychiatry (e.g., neuroscience focus, precision medicine), (2) shift research models for clinical brain imaging (e.g., big data exploration, standardization), (3) break down research silos (e.g., open science, calls for reproducibility and transparency), and (4) improve imaging technologies and methods. Although an arduous road remains ahead, these advances are repositioning the brain imaging community for long-term success.

KEYWORDS

assessment/diagnosis, biological markers, brain imaging/neuroimaging, dissemination/implementation, functional MRI, neuroimaging

As progress is made toward the prevention and treatment of cardiovascular disease and cancers, neuropsychiatric disorders are accounting for an increasing proportion of the unmet disease burden throughout the lifespan (Walker, McGee, & Druss, 2015; Whiteford, Ferrari, Degenhardt, Feigin, & Vos, 2015). Yet our approaches for diagnosing and treating these disorders remain grounded in advances dating back to the late 19th to mid-20th centuries. Over the past two decades, clinical brain imaging has sought to address this gap and has succeeded in revealing biological correlates for a broad array of illnesses (Castellanos & Proal, 2012; Etkin & Wager, 2007; Lagopoulos, Das, Saether, & Malhi, 2009; Rane et al., 2015; Savitz, Rauch, & Drevets, 2013; Strakowski, Delbello, & Adler, 2005). However, the inability to find clear relationships between abnormal brain function and traditional diagnostic classifications of neuropsychiatric illness has hindered the pace of progress. In particular, it has resulted in an inadequate biological understanding of these illnesses and difficulty for imaging to (1) identify meaningful biological indices of clinical diagnosis, prognosis, or risk (i.e., biomarkers), and (2) facilitate the development of novel targets for therapeutic interventions (Kapur, Phillips, & Insel, 2012). In short, despite decades of effort, visions of transforming clinical practice for neuropsychiatric illness through developing objective imaging-based “growth charts” or “lab tests” for the brain have remained out of reach.

In recent years, efforts focused on mapping the human connectome (the wiring diagram of the human brain) (Craddock et al., 2013; Van Essen & Ugurbil, 2012) and its functional interactions (Kelly, Biswal, Cameron Craddock, Castellanos, & Milham, 2012) have sparked renewed enthusiasm about the possibility of realizing clinically useful biomarkers for neuropsychiatric illness. We consider a biomarker to have clinical utility if it not only correlates with the presence of an illness, but can help to inform diagnosis, treatment selection/evaluation, assessments of risk, and prognosis (Fornito & Bullmore, 2015; Fox & Greicius, 2010; Rubinov & Bullmore, 2013). This renewed enthusiasm is not without its detractors, who ask what is different from the past, what is the “game changer?” Here, we argue that there is no single game changer. Instead, we review a constellation of advances that together improve the chances of delivering biological markers of illness for neuropsychiatric disorders.

1 | ELUCIDATION OF THE RESEARCH AGENDA FOR BIOLOGICAL PSYCHIATRY

The first key question for any pursuit focused on developing clinically useful measures from brain imaging is: *what are we trying to accomplish?* The vast majority of clinical brain imaging research has focused

on extreme comparisons between groups of individuals with a clean Diagnostic and Statistical Manual of Mental Disorders (DSM-5) diagnosis (such as major depressive disorder with no comorbidities) and “pure” controls (Kapur et al., 2012). Such studies do not reflect the “real world” challenges for the assessment of neuropsychiatric disorders. Nearly half of all individuals affected by mental illness meet criteria for two or more mental health disorders (Bijl, Ravelli, & van Zessen, 1998), and DSM classifications for these disorders likely include a heterogeneous mix of biological profiles. The real power of brain imaging will be in differentiating seemingly overlapping clinical disorders, as well as subtypes of the same disorder, rather than simply differentiating typical individuals from those affected by one or more disorders. This ability to differentiate could also help to improve the assignment, targeting, and monitoring of treatment interventions, and provides a more pragmatic scenario for translating brain imaging research to the clinic.

1.1 | A shift away from disorder-focused definitions of psychiatry

The biological psychiatry community has come to accept the limitations of efforts to map current psychiatric nosologies onto the brain (Insel et al., 2010; Kapur et al., 2012). Although a logical starting point for the field, efforts to map classifications in current diagnostic manuals (e.g., DSM 5) to the brain, or to other biological modalities (e.g., genes) have had mixed success. The field has yet to find reproducible and generalizable one-to-one mappings between diagnoses and patterns of brain abnormality. This is not surprising, as these disorders were defined based on the detection of clusters of signs and symptoms of psychiatric illness, not based on biology.

Although frustrating, this lack of specificity has given rise to a new agenda for the biological psychiatry community—namely, the development of a neuroscience-based diagnostic classification system for mental health and learning disorders. The NIMH Research Domain Criteria project (RDoC) has taken a lead role in pushing a neuroscience-focused agenda for neuropsychiatry (Cuthbert & Insel, 2013). The calls for an increased neuroscience focus are not necessarily intended to be a complete rejection of classical nosologies, but rather, a challenge to rethink the existing diagnostic boundaries. Within a given domain of illness such as ADHD, researchers are being asked to find biologically based subtypes that can help explain variation among affected individuals. These subtypes would have the potential to be more informative than the rudimentary subdivisions defined by clinical observation, such as inattentive versus hyperactive-impulsive versus combined for ADHD (Gates, Molenaar, Iyer, Nigg, & Fair, 2014).

Researchers are also being challenged to use neuroscience to rethink existing boundaries for domains of illness that extend across multiple diagnostic categories, such as depression (e.g., bipolar, unipolar, atypical, and chronic; Insel & Cuthbert, 2015). For example, a recent review by McTeague, Goodkind, and Etkin (2016) highlighted the presence of broad perturbations in behavioral indices of cognitive control for a range of disorders (mood, anxiety, psychotic, and substance use disorders); they argued that only the severity of the perturbations observed appears to depend on the specific disorder being examined. Consistent with these behavioral findings, imaging

studies have reported abnormalities within the frontoparietal multiple demands system (Duncan, 2010) across the range of disorders, particularly in dorsal anterior cingulate and bilateral insular cortices. Such observations motivate further research that looks across diagnostic labels in clinically diverse samples for commonalities in impairments and in biology (see Van Dam et al., 2016) for another recent example); this research in turn has the potential to guide development of transdiagnostic treatments for impairments of interest.

Researchers are also being challenged to consider comorbid illnesses, which too often are overlooked in applications of the current diagnostic nosology in brain imaging research, despite undeniably high rates. For example, a recent U.S.-based study found the following comorbidities for adults with ADHD (Willcutt, 2012)—social phobia: 29.3%, specific phobia: 22.7%, bipolar disorder: 19.4%, major depressive disorder: 18.6%, intermittent explosive disorder: 19.6%, and substance use disorder: 15.2%. The integration of neuroscientific data will allow us to determine when seemingly disparate symptoms are actually linked to a common biological origin, and when seemingly similar phenomena have different biological origins.

1.2 | The emerging concept of precision medicine

Over the past 5 years, the concept of precision medicine has emerged as a guiding approach for the future of healthcare research and practice (Collins & Varmus, 2015; Mirnezami, Nicholson, & Darzi, 2012). The medical community has traditionally relied on the use of a one-size-fits-all approach to the prevention and treatment of an identified disease, where the same treatment is applied to all individuals with a given diagnosis. Unfortunately, such strategies fail to take into account interindividual variability, resulting in a trial and error approach to figuring out who will respond well to a given intervention and who will not, as well as who will have an adverse reaction and who will not. Precision medicine attempts to minimize the trial and error aspect of interventions by stratifying individuals based on one or more factors that are predictive of response. Candidate predictors that could potentially guide stratification include biological factors such as genes, immunological profiles, and imaging findings, as well as lifestyle and environmental exposures. It is worth noting that precision medicine differs from the concept of personalized medicine in that it is not calling for us to engineer treatments for the individual, but rather to select from existing treatments the one that is most likely to work for an individual (Sipka, 2016). Arguably, this is a more attainable goal. A demonstration of the potential role for brain imaging in precision medicine comes from a recent study that was able to: (a) identify four biological subtypes of depression in a sample of 711 individuals using resting-state functional magnetic resonance imaging (fMRI) measures of brain connectivity, and (b) use subtype information to predict responsiveness to transcranial magnetic stimulation therapy (Drysdale et al., 2017).

By emphasizing stratification, proponents of precision medicine have fueled ambitions to make brain imaging more clinically useful. Psychiatric researchers have long sought to develop clinical tools that could be helpful in the diagnostic process for complex cases, much the way a strep test can be used to disambiguate the causes of a sore throat in primary care settings and guide treatment. Stratification heralds a

broader research agenda that focuses on the assessment of prognosis (e.g., likelihood of disease progression or treatment response) and risk (e.g., likelihood of developing a disease or comorbid disease). The development of objective measures of risk and prognosis are essential for efforts focused on early intervention or prevention.

Although the prospect of bringing the precision medicine model to psychiatry is a source of excitement (Hahn, Nierenberg, & Whitfield-Gabrieli, 2016; Insel & Cuthbert, 2015; Kapur et al., 2012), a number of logistical and infrastructural challenges exist for its implementation (Joyner & Paneth, 2017). The biggest (and most costly) challenge is obtaining longitudinal data from individuals to identify predictors of clinically relevant outcomes, especially in dedicated research studies. An attractive solution is to leverage the longitudinal clinical data obtained in clinical settings (past and future) for the purposes of research; however, the implementation of such strategies is hindered by privacy concerns, poor data quality, and limitations in interoperability among differing electronic health record systems. Efforts such as the NIH-funded Informatics for Integrating Biology and the Bedside (i2b2; <https://www.i2b2.org/>; Weber et al., 2009) are working to bridge these gaps and facilitate discovery efforts (see Lingren et al., 2017 and Doshi-Velez, Ge, & Kohane, 2017 for examples).

2 | SHIFTING RESEARCH MODELS FOR CLINICAL BRAIN IMAGING

2.1 | Bringing the individual into focus

There is a key confusion in the literature between brain-based differences that are informative to scientists and those that are useful to clinicians. Scientific tradition has long established the pursuit of statistical significance as the goal of identifying differences among individuals, with the threshold set at a minimum standard of $P < .05$ (following appropriate statistical correction for multiple comparisons). Consistent with this tradition, myriad clinical brain imaging studies have worked to demonstrate statistically significant brain differences between affected individuals and comparisons. These differences have proven to be highly useful in the generation and testing of scientific models of psychiatric illness, and are well justified for such pursuits. However, as highlighted by several recent works (Castellanos, Di Martino, Cameron Craddock, Mehta, & Milham, 2013; Kapur et al., 2012), population-level differences in brain or behavior indices have little direct clinical applicability, as they are not sufficiently informative about the individual standing in front of a clinician. A far more challenging set of standards is required to derive tools that can be used for clinical purposes. As highlighted in a recent review (Castellanos et al., 2013), effect sizes required for a clinically useful biomarker can range between a Cohen's d of 1.5 and 3.0 depending on the nature of the application (e.g., screening, diagnosis). Such effect sizes are rare in the clinical brain imaging literature (Müller et al., 2017; Schmaal et al., 2016; Thompson et al., 2017), though they represent the true goal for which studies focused on clinical applications should strive—not the more traditional pursuit of $P < .05$ significance. Consistent with the requirements of any laboratory test, sufficient accuracy/validity,

reliability/precision, sensitivity, and specificity must also be established for brain-based biomarkers before any conversation of clinical application should begin.

In recent years, the increasing popularity of machine learning in the imaging community (Davatzikos et al., 2005; Lemm, Blankertz, Dickhaus, & Müller, 2011; Pereira, Mitchell, & Botvinick, 2009) has brought with it shifts in the design of experiments from testing for clinical differences among populations to individual-level prediction of clinical variables, such as diagnostic status. For those focused on the development of clinical tools, this shift is necessary, as it brings the target closer to direct clinical applications. However, as highlighted in greater detail elsewhere (Castellanos et al., 2013), care needs to be taken in the handling of findings from prediction-based studies, as a number of factors can limit their readiness for clinical application. Such factors include the level of rigor involved in cross-validation approaches (e.g., leave-one-out, k -fold, split-half, and independent-sample; Varoquaux & Craddock, 2013) and how well a study sample reflects the real-world characteristics of a diseased population.

2.2 | The big data research model

Following decades of innovations in technology and methods, the neuroscience of the 21st century is increasingly being defined by large-scale efforts to understand complex brain systems using massive datasets (Craddock, Tungaraza, & Milham, 2015; Kandel, Markram, Matthews, Yuste, & Koch, 2013; Poldrack & Gorgolewski, 2014; Sejnowski, Churchland, & Movshon, 2014). This approach to scientific investigation, referred to as the "big data" research model, has already yielded unprecedented results in diverse domains such as genetics, physics, astronomy, and medicine. The brain imaging community—which historically has been criticized for its reliance on underpowered samples, approximate replications, and a research silo culture—has emerged as an entry point for the big data model in neuroscience.

The implications of the big data model are multifold for the future of clinical brain imaging. First, in contrast to the more traditional scientific model ingrained in the imaging community, questions, not hypotheses, are the guiding light for investigations. The exploratory focus of the big data model is not intended to minimize the value of hypothesis-driven research. Instead, it is intended to be a complementary approach, capable of leading to the formation of novel hypotheses through the detection of previously unknown associations. The big data model also enables researchers to look for multivariate relationships that may not be readily accessible with more traditional models (Varoquaux & Craddock, 2013). The application of unsupervised learning methods (e.g., cluster analysis) to neuroscientific data can be used to identify previously unappreciated groupings of individuals within a given domain of illness (e.g., novel subtypes for ADHD or depression), or to build links across behaviorally defined boundaries that do not reflect known underlying neurobiology (Gates et al., 2014; Miranda-Dominguez et al., 2014). Supervised learning techniques can be used to develop predictive models for precision medicine (Castellanos et al., 2013; Craddock, Holtzheimer, Hu, & Mayberg, 2009).

It is important to note that from a clinical application perspective, the value of large-scale datasets is not their ability to increase

statistical power. Our ability to detect small to moderate effects will undoubtedly increase with sample size; while these can be valuable for appreciating scientifically interesting population-level associations, as previously discussed, clinical utility hinges instead on the detection of large effects that can be used to meaningfully generate predictions about an individual. As such, the value of big data samples for clinical imaging lies in their ability to represent the broader range of clinical and biological heterogeneity in populations of interest, which is essential for creating more generalizable and informative research tools. Additionally, large sample sizes enable more rigorous cross-validation approaches, such as split-half cross-validation (Strother et al., 2002).

2.3 | Rethinking phenotyping

Arguably, the most valuable aspect of brain imaging data is actually the associated phenotypic data; a given anatomical, diffusion, or resting-state fMRI dataset has little value if we do not know anything about the individual from which the images were acquired or the context within which they were acquired. Unfortunately, many researchers are only beginning to realize the importance of careful phenotyping. Although the community has focused on the reliability of imaging measures for years, many cognitive and behavioral measures are brought directly into imaging studies without sufficiently establishing their reliability. Consistency in demographic, cognitive, and behavioral measures used across studies is also a challenging issue resulting from the lack of standards; for example, a multitude of questionnaires exist, each providing slightly different information. In response to these challenges, there is a growing focus on the development of standardized data elements that can be included in all studies, known as “common data elements” (“General Standards—NINDS Common Data Elements”; Simmons & Quinn, 2014; Thurmond et al., 2010), as well as standardized (semistructured and structured) clinical interviews with defined mechanisms for establishing inter-rater reliability (e.g., ADOS; Lord, Rutter, Dilavore, & Risi, 2008 for autism).

A common alternative to the use of DSM labels for the characterization of phenotypic variation is the use of dimensional psychiatric assessments (e.g., Achenbach, 2016). Typically reliant on questionnaires, the tendency of such approaches is to probe for the presence or absence of symptoms. However, this practice frequently results in the creation of truncated distributions, with little variance among those individuals who are without symptoms. Recent efforts are attempting to draw attention to the value of designing instruments that can index both strengths and weaknesses related to particular domains of function. For example, the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Rating Scales (SWAN) is a rating scale that was explicitly designed to index the broader range of attentional function across individuals, yielding near-normal distributions (Arnett et al., 2013; Hay, Bennett, Levy, Sergeant, & Swanson, 2007; Lakes, Swanson, & Riggs, 2012); such distributions are optimal for application in population studies. The Extended Strengths and Weaknesses Assessment of Normal Behavior (E-SWAN; <http://eswan.org>) is currently under development to extend the design concept from the SWAN to a broader range of psychiatric domains.

Researchers are also beginning to turn their attention to mobile technologies such as smartphone applications (Bot et al., 2016) and wearable devices to collect cognitive, behavioral, mood, and physiological data using active and passive data collection methods. Active methods include explicit self-reports, from occasional and detailed survey instruments to more frequent and brief interactions, sometimes referred to as “ecological momentary assessments,” as well as the performance of tasks probing various domains of function (e.g., memory). Passive methods usually involve sensors worn on some part of the body that can unobtrusively collect high temporal frequency data (e.g., accelerometry actigraphs, photoplethysmography, electrodermal activity, electromyography, etc.). Passive methods can collect objective, real-world data about participants to track cognitive, behavioral, mood, and physiological states, detect medication response, etc. (Chaibub Neto et al., 2017), in ways that could never be achieved before (Adams et al., 2016; Glenn & Monteith, 2014; Maetzler, Domingos, Srulijes, Ferreira, & Bloem, 2013). Collecting and analyzing these different sources of data help to foster a high-dimensional view of health.

3 | FROM RESEARCH SILOS TO A GLOBAL SCIENTIFIC COMMUNITY

As biological psychiatry research adopts an agenda with a greater focus on neuroscience and precision medicine, and an emphasis on research models that collect and analyze larger and more varied datasets, new challenges arise. In particular, increased demands for sophisticated informatics and analysis resources far surpass the capabilities of any single investigative team, consortium, or scientific discipline (Neuro Cloud Consortium, 2016). The field can overcome such seemingly insurmountable obstacles in the realization of clinically useful biomarkers by shifting away from the status quo, where discovery is hampered by the limited resources and capabilities of competing research entities, to a collaborative model that leverages the combined strengths of the research field as a whole. Widescale, effective collaboration demands significant sharing, which is in line with the tenets of open science—a commitment to providing unrestricted access to all ideas, data, tools, and other artifacts generated in the scientific process (Carvalho, 2015). In recognition of the benefits of this shift, funding institutions and research organizations are turning to, and in some cases mandating, open science as a means of addressing the massive requirements of this enterprise.

3.1 | Scaling up data resources

The natural prerequisite to successful implementation of the big data model is the accrual of massive-scale, heterogeneous, and deeply phenotyped datasets. Few labs can hope to generate more than a hundred high-quality, well-phenotyped brain image datasets in a year, which sets an unreasonably lengthy timeline for the big data agenda. However, the open science model can make very large datasets available much more quickly, and in many cases more cost effectively, by pooling data collected from many different laboratories. Data

sharing efforts such as the 1000 Functional Connectomes Project and its International Neuroimaging Data-sharing Initiative (Mennes, Biswal, Castellanos, & Milham, 2013) (FCP/INDI), ADNI (Mueller et al., 2005/7), and OpenfMRI (Poldrack et al., 2013) are quickly increasing the scale of analyses that can be performed by the brain imaging community. FCP/INDI projects focused on specific disorders, such as the Autism Brain Imaging Data Exchange (ABIDE) and ADHD-200 are testaments to the type of resources that can be aggregated when researchers in the same topic area collaborate rather than compete.

It is worth noting that the large datasets generated by successful data sharing initiatives require considerable image processing, making them potentially unusable by researchers with limited access to computational expertise and resources. One solution to this challenge is the creation of collaborative efforts such as the Preprocessed Connectomes Project (<http://preprocessed-connectomes-project.org/>), in which researchers who have these computational resources prospectively collaborate with the entire research field by preprocessing the data and sharing the results (Bellec et al., 2016). Data processing steps that require manual interventions can also be handled much more quickly and efficiently if performed collaboratively by a large number of researchers using crowdsourcing platforms such as Brain Box (<http://www.brainbox.co.uk/>) or through online games such as Eyewire (<http://eyewire.org/>).

3.2 | Directing the challenge to the broader scientific community

Beyond bolstering the efforts of existing community members, open science initiatives remove barriers to entry for members of the broader scientific community. Large-scale brain imaging data are rife with complexities that push the boundaries of what is possible with conventional inferential statistics. Given the need for advanced analysis methods and infrastructures that are outside of the areas of expertise for most clinician scientists, it is essential to recruit computer scientists, engineers, mathematicians, and statisticians to mental health research. The ADHD-200 (Milham et al., 2012) and ABIDE (Di Martino et al., 2014) initiatives demonstrated that making data available from a population with a given condition draws focus to that condition and engages individuals from a range of disciplines. Contests such as the ADHD-200 Global Competition or the Alzheimer's Disease Big Data DREAM Challenge (<https://www.synapse.org/Synapse:syn2290704>) can help to accelerate the recruitment of investigators from other disciplines (Allen et al., 2016; Milham, 2012). The integration of researchers from other domains requires education and training resources, however. Those from nonimaging, computational disciplines will need to be educated on the fundamentals of brain imaging data, while clinical researchers will need to improve their computational and analytical skills to promote the cohesiveness and quality of their research. This is being addressed by educational workshops such as Brainhack (Craddock et al., 2016) and Neurohackweek ("Neurohackweek," 2016), where interdisciplinary researchers come together to learn from one another as they work to address open problems in neuroscience.

3.3 | Increasing the focus on reproducible research and transparency

Current estimates suggest that as many as 50% of scientific findings cannot be reproduced (Ioannidis, 2005). Excepting rare cases of blatant fraud, the lack of reproducibility in research can be attributed to a variety of factors, ranging from underpowered sample sizes, differences in experimental methods, and heterogeneous samples, to liberal statistical thresholding, insufficient descriptions of study procedures, and errors in in-house software (Eklund, Nichols, & Knutsson, 2016; Müller et al., 2017; Woo, Krishnan, & Wager, 2014). Open science is increasingly heralded as a solution to these challenges. Concepts such as open data sharing, open source software, and open lab notebooks are promising to revolutionize research practices and dramatically augment reproducibility. From a practical view, such levels of transparency are likely to increase the standard of data quality and documentation; researchers will know that others will scrutinize their work, and have the opportunity to improve upon the original work through feedback and experimentation. Even the publication process is beginning to be transformed by open science, as preprint archives such as arXiv, bioRxiv, and Authorea are encouraging authors to upload their manuscripts for viewing soon after submission, rather than at the end of a review process—a practice that is dramatically speeding up the communication of findings. For example, the manuscript by Eklund et al. (2016) that raised concerns about the accuracy of commonly used statistical corrections for fMRI was openly available as a preprint 1 year prior to publication. Venues such as Academic Karma, Frontiers, Gigascience, and F1000 are pushing the idea of open review, where reviewers no longer blind themselves, removing a commonly cited source of bias in the publication process.

Although promising, the implementation of open science is not without its challenges. In particular, the field is still working through issues related to protection of privacy for high-dimensional datasets to ensure the safety of research participants. Mechanisms are still being worked out to promote academic accreditation in an open science community (e.g., data papers; Gorgolewski, Margulies, & Milham, 2013), as current tenure promotion practices tend to engender feelings of competition rather than promote sharing. The recent increase in attention to open science by funding agencies such as the NIH and NSF will likely prove to be the greatest catalyst to motivate change at all levels.

4 | BREAKING DOWN BARRIERS BY ADVANCING METHODS AND TECHNOLOGIES

4.1 | The maturation of pediatric brain imaging

The majority of neuropsychiatric illnesses have onsets in the first two decades of life (Kessler et al., 2005), when brain structure and function are actively developing and maturing. As such, the broader range of neuropsychiatric disorders is increasingly being conceptualized

as neurodevelopmental disorders (Di Martino et al., 2014), with origins possibly tracing back as early as the fetal and perinatal periods (O'Donnell & Meaney, 2016; Schlotz & Phillips, 2009). It follows that early detection and interventions, as well as prevention-focused efforts, could necessitate early life and childhood imaging (Koyama et al., 2016). As discussed in greater detail elsewhere (Di Martino et al., 2014; van den Heuvel & Thomason, 2016), advances in task-independent brain imaging methods (e.g., resting-state fMRI, morphometry MRI, diffusion MRI) are rapidly expanding the window of investigations focused on human brain function to early portions of the lifespan, including toddler, infant, neonatal, and fetal imaging. This expansion is invigorating the developmental imaging community and increasing enthusiasm about the potential to develop tools capable of guiding early intervention efforts for disorders such as autism and schizophrenia.

4.2 | Increasing harmonization of methods

Previously described efforts such as the FCP/INDI have provided examples of the potential to quickly amass large-scale datasets through the aggregation of data independently collected at different research facilities. However, the lack of coordination in the acquisition of imaging (and phenotypic) data can introduce site-related variation in the data that can limit our ability to detect underlying biological signals. To address this challenge, recent efforts have demonstrated the potential value of accounting for unintended site-related variation through statistical corrections (Keshavan et al., 2016), or explicitly optimizing image analysis pipelines and techniques to maximize the reproducibility of findings across different imaging sites (Abraham et al., n.d., in press; Eloyan, Crainiceanu, & Caffo, 2013; Fortin, Sweeney, Muschelli, Crainiceanu, & Shinohara, 2016; Yan et al., 2013).

Looking forward, prospective large-scale data collection consortia such as ADNI (Mueller et al., 2005/7), the Brain Genomics Superstruct (Holmes et al., 2015), the upcoming NIH Human Connectome Lifespan Project (Van Essen, 2011), PING (Jernigan et al., 2016), and the NIH ABCD Study ("ABCD Study," 2016) have worked to address the challenges of imaging site-related variation through harmonizing their acquisition protocols, as well as the use of phantoms for calibrating data between sites. Among these efforts, ADNI, PING, and the NIH ABCD Study have taken on the particular challenge of harmonizing protocols across different MRI scanner platforms. The success of these efforts is leading to calls in the brain imaging community for increased standardization of data acquisition protocols across sites in general. Although the provision of basic guidelines for acquisition can improve the overall quality and comparability of data, the benefits need to be balanced against the need for continued efforts to innovate data acquisition techniques; likely such decisions should be made on a case-by-case basis. At a minimum, encouraging and supporting the routine acquisition of data from standardized phantoms can facilitate efforts to calibrate data across scanner sites (DeDora et al., 2016; Gouttard, Styner, Prastawa, Piven, & Gerig, 2008; Gunter et al., 2017; Jovicich et al., 2017; Keenan et al., 2016).

4.3 | Improving MRI scanner capabilities

Modern advances in scanning technology are increasing the quality, resolution, and efficiency of data collection. Advances in head coil technologies have improved SNR and have enabled higher factors of parallel imaging, which combine to increase the resolution and speed for structural imaging acquisition. Newly developed simultaneous multi-slice (SMS) imaging sequences have increased the temporal and spatial resolutions achievable for fMRI (Feinberg & Setsompop, 2013). SMS combined with stronger magnetic field gradients are vastly improving the quality of diffusion MRI data that can be acquired during routine clinical evaluations. These recent significant gains in MRI technology serve as a valuable reminder that MRI is still a young technology and there are still opportunities for improvement.

In particular, continued innovation in MRI scanning equipment and sequences will be essential to fully realize the potential of brain imaging biomarkers and revolutionize clinical practice. Head motion is the nemesis of all brain imaging applications, but is particularly problematic when relying on quantitative measurements from the data (Friston, Williams, Howard, Frackowiak, & Turner, 1996; Reuter et al., 2015; Van Dijk, Sabuncu, & Buckner, 2011). Methods for prospectively correcting head motion in structural MRI are showing promise (Tisdall et al., 2012), but their efficacy in clinical settings has yet to be proven. For fMRI, while a motion-robust scanning sequence has yet to emerge, a variety of post hoc correction techniques have been proposed and evaluated (Friston et al., 1996; Power, Barnes, Snyder, Schlaggar, & Petersen, 2012; Pruim, Mennes, Buitelaar, & Beckmann, 2015; Yan, Craddock, He, & Milham, 2013). The challenges of head motion for diffusion MRI are similar to fMRI, and as such are becoming an increasing area of focus for future work (Yendiki, Koldewyn, Kakunoori, Kanwisher, & Fischl, 2014).

4.4 | Improved methods for large-scale image processing and analysis

With every advance in scanning technologies and increase of scale in data analysis comes the need for new and improved data processing and analysis tools. The growing availability of large-scale datasets is bringing about a change in the design of brain imaging tools, which up to now have focused primarily on the needs of smaller scale datasets (Craddock et al., 2015). Shortcomings of parametric models of brain imaging data have also highlighted the need to switch to more computationally intensive Monte Carlo methods for assessing statistical significance (Eklund, Andersson, Josephson, Johansson, & Knutsson, 2012; Eklund et al., 2016). To make this computation possible, a new emphasis is being placed on optimizing image processing to minimize its memory footprint and computational demands. Tools that implement and integrate with graphical processing units, cloud computing, and other computational infrastructure are emerging as potential solutions to the challenge of providing relatively low-cost computational resource options to users who do not necessarily have such expertise (Eklund et al., 2012). Making this work at a large scale will require a shift from relying exclusively on local collaborations to designing software that takes care of these details for the user.

5 | CONCLUDING REMARKS

We assert that there is no single advance that currently has the potential to drive the field of clinical neuroscience forward. Instead, there has been a constellation of advances that, if combined, could bring about the insights, momentum, and direction to discover biomarkers of neuropsychiatric illness. Although our focus was on advances that promise to transform multimodal imaging studies, there is a broader landscape of change and innovation across the range of disciplines focused on biomarker discovery (e.g., genetics, immunology, electron microscopy). Over time, the integration of these varied fronts may prove to bring about the most pivotal changes yet for clinical neuroscience.

ACKNOWLEDGMENTS

This work was supported by gifts to the Child Mind Institute from Phyllis Green, Randolph Cowen, and Joseph P. Healey.

REFERENCES

- "ABCD Study". (2016). Retrieved from <http://abcdstudy.org>
- Abraham, A., Milham, M., Di Martino, A., Cameron Craddock, R., Samaras, D., Thirion, B., & Varoquaux, G. (n.d.). Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. *NeuroImage*. doi:10.1016/j.neuroimage.2016.10.045
- Achenbach, T. (2016). 11.3 Multicultural findings on dimensional assessment of strengths and weaknesses. *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(10), S274–S275.
- Adams, Z., McClure, E. A., Gray, K. M., Danielson, C. K., Treiber, F. A., & Ruggiero, K. J. (2016). Mobile devices for the remote acquisition of physiological and behavioral biomarkers in psychiatric clinical research. *Journal of Psychiatric Research*, October. doi:10.1016/j.jpsychires.2016.10.019
- Allen, G. I., Amoroso, N., Anghel, C., Balagurusamy, V., Bare, C. J., Beaton, D., ... Alzheimer's Disease Neuroimaging Initiative. (2016). Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease. *Alzheimer's & Dementia*, 12(6), 645–653.
- Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelson, S., & Olson, R. K. (2013). The SWAN captures variance at the negative and positive ends of the ADHD symptom dimension. *Journal of Attention Disorders*, 17(2), 152–162.
- Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D. S., & Cameron Craddock, R. (2016). The Neuro Bureau ADHD-200 preprocessed repository. *NeuroImage*, July. doi:10.1016/j.neuroimage.2016.06.034
- Bijl, R. V., Ravelli, A., & van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: Results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*, 33(12), 587–595.
- Bot, B. M., Suver, C., Neto, E. C., Kellen, M., Klein, A., Bare, C., ... Trister, A. D. (2016). The mPower Study, Parkinson disease mobile data collected using ResearchKit. *Scientific Data*, 3(March), 160011.
- Cameron Craddock, R., Margulies, D. S., Bellec, P., Nichols, B., Alcauter, S., Barrios, F. A., ... Xu, T. (2016). Brainhack: A collaborative workshop for the open neuroscience community. *GigaScience*, 5(March), 16.
- Carvalho, J. (2015). FOSTER: Facilitate Open Science Training for European Research. In *COAR-SPARC Conference 2015*. Retrieved from <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>
- Castellanos, F., Di Martino, A., Cameron Craddock, R., Mehta, A. D., & Milham, M. P. (2013). Clinical applications of the functional connectome. *NeuroImage*, 80, 527–540.
- Castellanos, F., & Proal, E. (2012). Large-scale brain systems in ADHD: Beyond the prefrontal-striatal model. *Trends in Cognitive Sciences*, 16(1), 17–26.
- Chaibub Neto, E., Bot, B. M., Perumal, T., Omberg, L., Guinney, J., Kellen, M., ... Trister, A. D. (2017). Personalized hypothesis tests for detecting medication response in Parkinson disease patients using iPhone sensor data. PubMed–NCBI. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/26776193>
- Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine*, 372(9), 793–795.
- Craddock, R., Holtzheimer, P. E., Hu, X. P., & Mayberg, H. S. (2009). Disease state prediction from resting state functional connectivity. *Magnetic Resonance in Medicine*, 62(6), 1619–1628.
- Craddock, R., Jbabdi, S., Yan, C.-G., Vogelstein, J. T., Castellanos, F., Di Martino, A., ... Milham, M. P. (2013). Imaging human connectomes at the macroscale. *Nature Methods*, 10(6), 524–539.
- Craddock, R., Tungaraza, R. L., & Milham, M. P. (2015). Connectomics and new approaches for analyzing human brain functional connectivity. *GigaScience*, 4(1), 13.
- Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, 11(May), 126.
- Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughhead, J. W., ... Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, 28(3), 663–668.
- DeDora, D. J., Nedic, S., Katti, P., Arnab, S., Wald, L. L., Takahashi, A., ... Mujica-Parodi, L. R. (2016). Signal fluctuation sensitivity: An improved metric for optimizing detection of resting-state fMRI networks. *Frontiers in Neuroscience*, 10. doi:10.3389/fnins.2016.00180
- Di Martino, A., Fair, D. A., Kelly, C., Satterthwaite, T. D., Castellanos, F., Thomason, M. E., ... Zuo, X.-N. (2014). Unraveling the miswired connectome: A developmental perspective. *Neuron*, 83(6), 1335–1353.
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., ... Milham, M. P. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, 19(6), 659–667.
- Doshi-Velez, F., Ge, Y., & Kohane, I. (2017). Comorbidity clusters in autism spectrum disorders: An electronic health record time-series analysis. PubMed–NCBI. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/24323995>
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., ... Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, 23(1), 28–38.
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172–179.
- Eklund, A., Andersson, M., Josephson, C., Johannesson, M., & Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? An empirical study of 1484 rest datasets. *NeuroImage*, 61(3), 565–578.
- Eklund, A., Andersson, M., & Knutsson, H. (2012). fMRI analysis on the GPU-possibilities and challenges. *Computer Methods and Programs in Biomedicine*, 105(2), 145–161.
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, 113(28), 7900–7905.

- Eloyan, A., Crainiceanu, C. M., & Caffo, B. S. (2013). Likelihood-based population independent component analysis. *Biostatistics*, 14(3), 514–527.
- Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *American Journal of Psychiatry*, 164(10), 1476–1488.
- Feinberg, D. A., & Setsompop, K. (2013). Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of Magnetic Resonance*, 229, 90–100.
- Fornito, A., & Bullmore, E. T. (2015). Connectomics: A new paradigm for understanding brain disease. *European Neuropsychopharmacology*, 25(5), 733–748.
- Fortin, J.-P., Sweeney, E. M., Muschelli, J., Crainiceanu, C. M., Shinohara, R. T., & Alzheimer's Disease Neuroimaging Initiative. (2016). Removing inter-subject technical variability in magnetic resonance imaging studies. *NeuroImage*, 132(May), 198–212.
- Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in Systems Neuroscience*, 4(June), 19.
- Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, 35(3), 346–355.
- Gates, K. M., Molenaar, P. C. M., Iyer, S. P., Nigg, J. T., & Fair, D. A. (2014). Organizing heterogeneous samples using community detection of GIMME-derived resting state functional networks. *PLoS One*, 9(3), e91322.
- Glenn, T., & Monteith, S. (2014). New measures of mental state and behavior based on data collected from sensors, smartphones, and the internet. *Current Psychiatry Reports*, 16(12), 523.
- Gorgolewski, K. J., Margulies, D. S., & Milham, M. P. (2013). Making data sharing count: A publication-based solution. *Frontiers in Neuroscience*, 7(February), 9.
- Gouttard, S., Styner, M., Prastawa, M., Piven, J., & Gerig, G. (2008). Assessment of reliability of multi-site neuroimaging via traveling phantom study. *Medical Image Computing and Computer-Assisted Intervention*, 11(Pt 2), 263.
- Gunter, J. L., Bernstein, M. A., Borowski, B. J., Ward, C. P., Britson, P. J., Felmlee, J. P., ... Jack, C. R. (2017). Measurement of MRI scanner performance with the ADNI phantom. PubMed—NCBI. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/19610308>
- Hahn, T., Nierenberg, A. A., & Whitfield-Gabrieli, S. (2016). Predictive analytics in mental health: Applications, guidelines, challenges and perspectives. *Molecular Psychiatry*, November. doi:10.1038/mp.2016.201
- Hay, D. A., Bennett, K. S., Levy, F., Sergeant, J., & 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. *Biological Psychiatry*, 61(5), 700–705.
- Holmes, A. J., Hollinshead, M. O., O'Keefe, T. M., Petrov, V. I., Fariello, G. R., Wald, L. L., ... Buckner, R. L. (2015). Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific Data*, 2(July), 150031.
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167(7), 748–751.
- Insel, T. R., & Cuthbert, B. N. (2015). Medicine. Brain disorders? Precisely. *Science*, 348(6234), 499–500.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124.
- Jernigan, T. L., Brown, T. T., Hagler, D. J., Jr., Akshoomoff, N., Bartsch, H., & Newman, E. ... Pediatric Imaging, Neurocognition and Genetics Study. (2016). The Pediatric Imaging, Neurocognition, and Genetics (PING) data repository. *NeuroImage*, 124(Pt B), 1149–1154.
- Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., ... Dale, A. (2017). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. PubMed—NCBI. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/16300968>
- Joyner, M. J., & Paneth, N. (2017). Seven questions for personalized medicine. PubMed—NCBI. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/26098474>
- Kandel, E. R., Markram, H., Matthews, P. M., Yuste, R., & Koch, C. (2013). Neuroscience thinks big (and collaboratively). *Nature Reviews. Neuroscience*, 14(9), 659–664.
- Kapur, S., Phillips, A. G., & Insel, T. R. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17(12), 1174–1179.
- Keenan, K. E., Stupic, K. F., Boss, M. A., Russek, S. E., Chenevert, T. L., Prasad, P. V., ... Jackson, E. F. (2016). Comparison of T1 measurement using ISMRM/NIST system phantom. NIST. May 9. Retrieved from <https://www.nist.gov/publications/comparison-t1-measurement-using-ismrmnist-system-phantom>
- Kelly, C., Biswal, B. B., Cameron Craddock, R., Castellanos, F., & Milham, M. P. (2012). Characterizing variation in the functional connectome: Promise and pitfalls. *Trends in Cognitive Sciences*, 16(3), 181–188.
- Keshavan, A., Paul, F., Beyer, M. K., Zhu, A. H., Papinutto, N., Shinohara, R. T., ... Henry, R. G. (2016). Power estimation for non-standardized multisite studies. *NeuroImage*, 134(July), 281–294.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, 62(6), 593–602.
- Koyama, M. S., Di Martino, A., Castellanos, F. X., Ho, E. J., Marcelle, E., Leventhal, B., & Milham, M. P. (2016). Imaging the 'at-risk' brain: Future directions. *Journal of the International Neuropsychological Society*, 22(2), 164–179.
- Lagopoulos, J., Das, P., Saether, O., & Malhi, G. S. (2009). A review of imaging in psychiatry. *309(15)*, 15–20.
- Lakes, K. D., Swanson, J. M., & Riggs, M. (2012). The reliability and validity of the english and spanish strengths and weaknesses of ADHD and normal behavior rating scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders*, 16(6), 510–516.
- Lemm, S., Blankertz, B., Dickhaus, T., & Müller, K.-R. (2011). Introduction to machine learning for brain imaging. *NeuroImage*, 56(2), 387–399.
- Lingren, T., Chen, P., Bochenek, J., Doshi-Velez, F., Manning-Courtney, P., Bickel, J., ... Savova, G. (2017). Electronic health record based algorithm to identify patients with autism spectrum disorder. PubMed—NCBI. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/27472449>
- Lord, C., Rutter, M., Dilavore, P. C., & Risi, S. (2008). ADOS: Autism diagnostic observation schedule. Hogrefe.
- Maetzler, W., Domingos, J., Srulijes, K., Ferreira, J. J., & Bloem, B. R. (2013). Quantitative wearable sensors for objective assessment of Parkinson's disease. *Movement Disorders*, 28(12), 1628–1637.
- McTeague, L. M., Goodkind, M. S., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of Psychiatric Research*, 83(December), 37–46.
- Mennes, M., Biswal, B. B., Castellanos, F., & Milham, M. P. (2013). Making data sharing work: The FCP/INDI experience. *NeuroImage*, 82, 683–691.
- Milham, M. P. (2012). Open neuroscience solutions for the connectome-wide association era. *Neuron*, 73(2), 214–218.

Q16

Q17

- Milham, M. P., Fair, D., Mennes, M., Mostofsky, S. H. & Others. (2012). The ADHD-200 Consortium: A model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience*, 6, 62.
- Miranda-Dominguez, O., Mills, B. D., Carpenter, S. D., Grant, K. A., Kroenke, C. D., Nigg, J. T., & Fair, D. A. (2014). Connectotyping: Model based fingerprinting of the functional connectome. *PLoS One*, 9(11), e111048.
- Mirnezami, R., Nicholson, J., & Darzi, A. (2012). Preparing for precision medicine. *New England Journal of Medicine*, 366(6), 489–491.
- Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., ... Beckett, L. (2005/2007). Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's disease neuroimaging initiative (ADNI). *Alzheimer's & Dementia*, 1(1), 55–66.
- Müller, V. I., Cieslik, E. C., Serbanescu, I., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2017). Altered brain activity in unipolar depression revisited: Meta-analyses of neuroimaging studies. *JAMA Psychiatry*, 74(1), 47–55.
- Neuro Cloud Consortium. Electronic address: Jovo@jhu.edu, and Neuro Cloud Consortium. (2016). To the Cloud! A grassroots proposal to accelerate brain science discovery. *Neuron*, 92(3), 622–627.
- Neurohackweek. (2016). Retrieved from <http://neurohackweek.github.io/>
- O'Donnell, K. J., & Meaney, M. J. (2016). Fetal origins of mental health: The developmental origins of health and disease hypothesis. *American Journal of Psychiatry*, November, appiajp201616020138.
- Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*, 45(1 Suppl.), S199–S209.
- Poldrack, R. A., Barch, D. M., Mitchell, J. P., Wager, T. D., Wagner, A. D., Devlin, J. T., ... Milham, M. P. (2013). Toward open sharing of task-based fMRI data: The OpenfMRI project. *Frontiers in Neuroinformatics*, 7(July), 12.
- Poldrack, R. A., & Gorgolewski, K. J. (2014). Making big data open: Data sharing in neuroimaging. *Nature Neuroscience*, 17(11), 1510–1517.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154.
- Pruim, R. H. R., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2015). Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *NeuroImage*, 112(May), 278–287.
- Rane, P., Cochran, D., Hodge, S. M., Haselgrove, C., Kennedy, D. N., & Frazier, J. A. (2015). Connectivity in autism: A review of MRI connectivity studies. *Harvard Review of Psychiatry*, 23(4), 223–244.
- Reuter, M., Tisdall, Man., Qureshi, A., Buckner, R. L., van der Kouwe, A. J. W., & Fischl, B. (2015). Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *NeuroImage*, 107, 107–115.
- Rubinov, M., & Bullmore, E. (2013). Fledgling pathoconnectomics of psychiatric disorders. *Trends in Cognitive Sciences*, 17(12), 641–647.
- Sándor, S. (2016). [Precision and personalized medicine]. *Orvosi hetilap*, 157(44), 1739–1741.
- Savitz, J. B., Rauch, S. L., & Drevets, W. C. (2013). Clinical application of brain imaging for the diagnosis of mood disorders: The current state of play. *Molecular Psychiatry*, 18(5), 528–539.
- Schlotz, W., & Phillips, D. I. W. (2009). Fetal origins of mental health: Evidence and mechanisms. *Brain, Behavior, and Immunity*, 23(7), 905–916.
- Schmaal, L., Hibar, D. P., Sämann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., ... Veltman, D. J. (2016). Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Molecular Psychiatry*, May. doi:10.1038/mp.2016.60
- Sejnowski, T. J., Churchland, P. S., & Movshon, J. (2014). Putting big data to good use in neuroscience. *Nature Neuroscience*, 17(11), 1440–1441.
- Simmons, J. M., & Quinn, K. J. (2014). The NIMH Research Domain Criteria (RDoC) Project: Implications for genetics research. *Mammalian Genome*, 25(1–2), 23–31.
- Strakowski, S. M., Delbello, M. P., & Adler, C. M. (2005). The functional neuroanatomy of bipolar disorder: A review of neuroimaging findings. *Molecular Psychiatry*, 10(1), 105–116.
- Strother, S. C., Anderson, J., Hansen, L. K., Kjems, U., Kustra, R., Sidtis, J., ... Rottenberg, D. (2002). The quantitative evaluation of functional neuroimaging experiments: The NPAIRS data analysis framework. *NeuroImage*, 15(4), 747–771.
- Thompson, P. M., Andreassen, O. A., Arias-Vasquez, A., Bearden, C. E., Boedhoe, P. S., & Brouwer, R. M. ... ENIGMA Consortium. (2017). ENIGMA and the individual: Predicting factors that affect the brain in 35 countries worldwide. *NeuroImage*, 145(Pt B), 389–408.
- Thurmond, V. A., Hicks, R., Gleason, T., Miller, Ate., Szufliata, N., Orman, J., & Schwab, K. (2010). Advancing integrated research in psychological health and traumatic brain injury: Common data elements. *Archives of Physical Medicine and Rehabilitation*, 91(11), 1633–1636.
- Tisdall, Man, Hess, A. T., Reuter, M., Meintjes, E. M., Fischl, B., & van der Kouwe, A. J. W. (2012). Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magnetic Resonance in Medicine*, 68(2), 389–399.
- Van Dam, N. T., O'Connor, D., Marcelle, E. T., Ho, E. J., Craddock, R., Tobe, R. H., ... Milham, M. P. (2016). Data-driven phenotypic categorization for neurobiological analyses: Beyond DSM-5 labels. *Biological Psychiatry*, July. doi:10.1016/j.biopsych.2016.06.027
- van den Heuvel, M. I., & Thomason, M. E. (2016). Functional connectivity of the human brain in utero. *Trends in Cognitive Sciences*, November. doi:10.1016/j.tics.2016.10.001
- Van Dijk, K. R. A., Sabuncu, M. R., & Buckner, R. L. (2011). The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, 59(1), 431–438.
- Van Essen, D. C. (2011). The Human Connectome Project. *Journal of Vision*, 11(15), 8–8.
- Van Essen, D. C., & Ugurbil, K. (2012). The Future of the human connectome. *NeuroImage*, 62(2), 1299–1310.
- Varoquaux, G., & Craddock, R. (2013). Learning and comparing functional connectomes across subjects. *NeuroImage*, April. doi:10.1016/j.neuroimage.2013.04.007
- Walker, E. R., McGee, R. E., & Druss, B. G. (2015). Mortality in mental disorders and global disease burden implications: A systematic review and meta-analysis. *JAMA Psychiatry*, 72(4), 334–341.
- Weber, G. M., Murphy, S. N., McMurphy, A. J., Macfadden, D., Nigrin, D. J., Churchill, S., & Kohane, I. S. (2009). The Shared Health Research Information Network (SHRINE): A prototype federated query tool for clinical data repositories. *Journal of the American Medical Informatics Association*, 16(5), 624–630.
- Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2015). The global burden of mental, neurological and substance use disorders: An analysis from the Global Burden of Disease Study 2010. *PLoS One*, 10(2), e0116820.
- Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics*, 9(3), 490–499.
- Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, 91(May), 412–419.
- Yan, C.-G., Craddock, R., He, Y., & Milham, M. P. (2013). Addressing head motion dependencies for small-world topologies in functional connectomics. *Frontiers in Human Neuroscience*, 7(December), 910.

- Yan, C.-G., Craddock, R., Zuo, X.-N., Zang, Y.-F., & Milham, M. P. (2013). Standardizing the intrinsic brain: Towards robust measurement of inter-individual variation in 1000 functional connectomes. *NeuroImage*, 80(October), 246–262.
- Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., & Fischl, B. (2014). Spurious group differences due to head motion in a diffusion MRI study. *NeuroImage*, 88(March), 79–90.

How to cite this article: Milham MP, Craddock RC, Klein A. Clinically useful brain imaging for Neuropsychiatry: How can we get there? *Depress Anxiety*. 2017;00:1–10. <https://doi.org/10.1002/da.22627>

Q20