

Harmonised collection of data in youth mental health: Towards large datasets

Suzie Lavoie^{1,2}, Kelly Allott^{1,2}, Paul Amminger^{1,2},
Cali Bartholomeusz^{1,2}, Maximus Berger^{1,2}, Michael Breakspear³,
Anjali K Henders⁴, Rico Lee⁵, Ashleigh Lin⁶, Patrick McGorry^{1,2},
Simon Rice^{1,2}, Lianne Schmaal^{1,2} and Stephen J Wood^{1,2,7}

Australian & New Zealand Journal of Psychiatry
2020, Vol. 54(1) 46–56
DOI: 10.1177/0004867419844322

© The Royal Australian and
New Zealand College of Psychiatrists 2019
Article reuse guidelines:
sagepub.com/journals-permissions
journals.sagepub.com/home/anp



Editor's Choice

Abstract

Objective: The current international trend is to create large datasets with existing data and/or deposit newly collected data into repositories accessible to the scientific community. These practices lead to more efficient data sharing, better detection of small effects, modelling of confounders, establishment of sample generalizability and identification of differences between any given disorders. In Australia, to facilitate such data-sharing and collaborative opportunities, the Neurobiology in Youth Mental Health Partnership was created. This initiative brings together specialised researchers from around Australia to work towards a better understanding of the cross-diagnostic neurobiology of youth mental health and the translation of this knowledge into clinical practice. One of the mandates of the partnership was to develop a protocol for harmonised prospective collection of data across research centres in the field of youth mental health in order to create large datasets.

Methods: Four key research modalities were identified: clinical assessments, brain imaging, neurocognitive assessment and collection of blood samples. This paper presents the consensus set of assessments/data collection that has been selected by experts in each domain.

Conclusion: The use of this core set of data will facilitate the pooling of psychopathological and neurobiological data into large datasets allowing researchers to tackle important questions requiring very large numbers. The aspiration of this transdiagnostic approach is a better understanding of the mechanisms underlying mental illnesses.

Keywords

Harmonisation, data sharing, transdiagnostic, neurocognition, neuroimaging

Introduction

About 75% of mental health conditions emerge before age 25 (Kessler et al., 2007), and just under one in four young Australians meet criteria for a probable serious mental illness (Mission Australia and the Black Dog Institute, 2017). Despite a range of privately and publicly funded mental health services as well as evidence-based therapies being available for young people, many continue to suffer because they do not easily respond to social, psychological and biological interventions (Bhui, 2017).

It is widely recognised that psychiatry has not yet achieved detailed, satisfactory models of how neurobiological, psychological and social factors interact and drive the onset of psychiatric disorder. Our understanding of the neurobiological mechanisms underlying symptoms of mental

¹Orygen, The National Centre of Excellence in Youth Mental Health, Parkville, VIC, Australia

²Centre for Youth Mental Health, The University of Melbourne, Parkville, VIC, Australia

³University of Newcastle, Newcastle, NSW, Australia

⁴Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, Australia

⁵Brain and Mental Health Research Hub, Monash University, Clayton, VIC, Australia

⁶Telethon Kids Institute, The University of Western Australia, Perth, WA, Australia

⁷School of Psychology, University of Birmingham, Birmingham, UK

Corresponding author:

Suzie Lavoie, Orygen, The National Centre of Excellence in Youth Mental Health, 35 Poplar Road, Parkville, VIC 3052, Australia.
Email: suzie.lavoie@orygen.org.au

illnesses and associated neurocognitive impairments has largely been obtained from the study of adults with long-established and differentiated syndromes. Adult cases tend to present with long histories of illness and treatment that can confound interpretation. Furthermore, studies in adults, by definition, do not allow investigation of normative development and deviations in neurodevelopmental trajectories that contribute to mental disorders, despite this clearly being of great importance given that half of all mental illnesses begin by the age of 14 and three-quarters by mid-20s according to the World Health Organization.

In order to provide the most targeted and successful treatment to young people, a better understanding of the neurobiological changes occurring during the onset and early course of mental illnesses is needed. Key issues that have contributed to our limited understanding of the exact neurobiological underpinnings of mental health are underpowered studies and the lack of reproducible findings. Underpowered studies have led to inconsistent and poorly replicated neurobiological findings in psychiatry. To circumvent these issues, large datasets are required. Creating large datasets lead to (1) better detection of small effects, (2) modelling of confounders, (3) establishment of sample generalizability and (4) identification of differences between any given disorders.

Globally, researchers tend to work in silos, focusing on a theoretical construct of interest for a given categorical diagnosis or looking at a specific psychological, functional or biological outcome. As a consequence, the data they collect cannot always be shared because different tools or approaches have been used. Site-specific data often need imputation, interpolation, etc. in order to be pooled with data from another site, inevitably adding to the variance in the data that cannot be attributed to the studied disorder. Other difficulties arising from the pooling of existing data are the ethical and computational issues with regard to data sharing, as well as science and data-sharing policies that may vary from one research institute to another. This paper offers a solution to these issues by suggesting an approach to harmonise the collection of data across research centres in Australia.

Another matter that is addressed by the current proposal is the categorisation of patients based on their clinical presentation, e.g. *Diagnostic and Statistical Manual of Mental Disorders (DSM)* and *International Classification of Diseases (ICD)*. These categorical systems imply that a patient can be diagnosed with a disorder that is distinct and independent from other disorders, assumptions that are not justified based on clinical experience and empirical research. To better understand the neurobiological changes occurring as a young person's mental health deteriorates, a transdiagnostic approach is more appropriate. Specifically, in order to better understand the nature of mental health and illness, the National Institute of Mental Health (NIMH) suggests the use of a Research Domain Criteria (RDoC) approach, which

integrates many levels of information (from genomics and circuits to behaviour and self-reports).¹

This paper presents an initiative and methods to harmonise the collection of data across research centres in the field of youth mental health in Australia using an RDoC approach. The proposed assessments and measures were selected and recommended by experts in the field of youth mental health from across Australia in the following key domains: clinical assessment, neurocognitive assessment, neuroimaging data collection and biospecimens collection. Using the same protocol to collect data improves data sharing between collaborators and will enable participating researchers to answer questions that cannot be tackled currently due to the lack of statistical power.

Methodology

A consortium of researchers and clinicians working in the field of youth mental health in Australia, the Neurobiology in Youth Mental Health Partnership,² was created. Partners were selected based on their expertise as researchers or clinicians in the field of youth mental health. Twenty-six individuals including professors, psychiatrists, senior and junior researchers, as well as research support staff participated in the selection of measures and assessments for the harmonisation of prospective data collection. They represented 13 independent institutions (Brisbane Institute for Molecular Bioscience; Brisbane Metro North Mental Health Service; Deakin University; Melbourne Neuropsychiatry Centre [The University of Melbourne]; Monash University; Orygen, The National Centre of Excellence in Youth Mental Health [The University of Melbourne]; Orygen Youth Health; QIMR Berghofer; Sunshine Coast Mind and Neuroscience Thompson Institute [The University of Sunshine Coast]; Telethon Kids Institute; The University of Melbourne; The University of Sunshine Coast; Brain and Mind Centre [The University of Sydney], across Australia [New South Wales, Queensland, Victoria, Western Australia]). See Supplementary Table 1 for a full list of the partners involved.

Their task was to make recommendations about the type of data and corresponding measures that should be collected in young people with mental health issues participating in research, with the aim to produce large datasets and attempt to answer important questions in relation to the neurobiology of illnesses.

Panels of experts were formed to work on four specific areas: clinical presentation, neurocognition, neuroimaging and biospecimens. Panel members met face-to-face, via teleconference or through emails over a period of 2 years, in order to reach consensus within each domain. The technique used to reach consensus was a variant of the Nominal Group Technique, whereby participants were asked to answer the following question: 'If you had 30 minutes with a research participant, what assessments would you conduct?' Experts were invited to actively participate in idea-generating

activities, basing their propositions on the literature. Every suggested assessment was carefully considered by the panel members until everyone agreed to include it or not based on the relevance and importance of the assessment, and on the time restriction. Assessments also needed to be applicable across diagnoses and be relevant to young people aged 12 to 25 years. Wherever possible, cost-effectiveness had to be considered. The battery was to provide a comprehensive neurobiological, psychological and neurocognitive snapshot of a given individual at a given point in time regardless of their clinical presentation. The experts were asked to choose measurements accessible in the public domain whenever possible. However, it was expected that costs would be associated with the collection of neuroimaging and some neurocognitive data, as well as for the collection, storage and processing of biospecimens. Effort was made to choose assessments that have been validated in young people, while keeping in mind that, where possible, measures should ideally leverage recent technological advances in each of the relevant domains.

In order to keep the primary assessments battery short, recommendations were made to collect information about equally important domains in secondary sets of assessments or measures. This will enable harmonisation of data collection in these domains too.

Clinical assessment

The primary clinical battery addresses the domains most commonly assessed across diagnoses. The broad symptom domains used in the primary battery were based on the Psychiatric Domain of the Phenotypes and Exposures project (PhenX),³ funded by the National Institutes of Health's (NIH) National Human Genome Research Institute. The domains were depression, mania, suicidal ideation, anxiety, psychological distress, psychosis, personality disorders, quality of life, sleep and substance use. By assessing the mental state of young people in all those domains, researchers can obtain a high-quality snapshot of the clinical picture of the participants.

The recommended clinical assessments were conceptualised to complement interviewer-administered assessments for determining past and current clinical diagnoses assessed by structured clinical interviews (i.e. Structured Clinical Interview for *DSM-5* [SCID-5], Comprehensive Assessment of At-Risk Mental States [CAARMS], Social and Occupational Functioning Assessment Scale [SOFAS], Clinical Global Impressions [CGI], etc.). It is also recommended that clinical phenotyping be enhanced by multiple sources (e.g. relatives, clinical file audit, etc.). The Clinical Assessment Panel used the following criteria for recommending specific measures for each domain: valid measures that adequately cover the chosen domains, appropriate for youth populations and brief. The Clinical Assessment Panel also made the decision to recommend only assessments

which are available in the public domain (i.e. no financial/licencing costs involved in usage) in order to ensure ease of access.

Table 1 presents the assessments that constitute the primary battery of clinical assessments, which should be administered in full. The total administration time for the primary clinical assessment battery varies between 36 and 58 minutes, which is longer than the targeted 30 minutes. However, it was agreed by the experts that the self-report assessments can be taken home by the research participants and can be filled out in their own time. Self-report assessments are practical, cost-efficient and enable the collection of larger amounts of data. However, the downside is that the data collected may not be as accurate as data collected in face-to-face assessments due to the 'social desirability bias' and the protection of privacy. Reduced reliability can be minimised by reassuring the participant that their privacy is protected. Another issue with the use of self-reports is that participants may not understand a given question or the interpretation of a given statement may vary between participants. To minimise this issue, the panel members have chosen assessments that are well validated in the targeted population.

In order to keep the primary battery as targeted and efficient as possible, it was developed to include prevalent disorders or broader clinical syndromes. More specialised areas can be addressed in the secondary battery of assessments. The assessments included in this secondary battery are presented in Table 2.

The Clinical Assessment Panel has set a long-term vision of developing and validating a new, transdiagnostic self-report tool that would have the ability to map on to the clinical stages of psychiatric illnesses. The panel's vision is for this new measure to take a broad youth-specific assessment approach, furthering the model adopted by the Brief Psychiatric Rating Scale (BPRS). This would align with recent conceptualisations of the structure of psychopathology, based on analysis of longitudinal birth cohort data (Caspi et al., 2014). The Clinical Assessment Panel views the establishment of a nationally adopted minimum dataset as a significant step towards this.

Neurocognitive assessment

Neurocognitive impairments are common and associated with poorer functional outcomes for individuals with a mental illness (Lee et al., 2017). Comprehensive meta-analyses indicate that clinically and statistically significant neurocognitive impairments (e.g. in attention, learning, memory, processing speed and executive function) are evident early in the course of mental illnesses, including psychosis (Mesholam-Gately et al., 2009), bipolar disorder (Lee et al., 2014a), eating disorders (Zakzanis et al., 2010) and depression (Goodall et al., 2018). However, there is significant heterogeneity in neurocognitive functioning

Table 1. Primary clinical assessment battery.

| Domain | Instrument | No. of items | Time taken | Reasons for inclusion |
|-----------------------------|--|--------------|--------------|--|
| Depression | Quick Inventory of Depression Symptoms (QIDS-16-SR; Rush et al., 2003) | 16 | 5-7 minutes | The QIDS is a widely used measure of depression. It offers a combined assessment severity and frequency of symptoms, and enables comparison with exiting major studies of youth depression conducted throughout Australia. |
| Mania (including hypomania) | Altman Self-Rating Mania Scale (ASRM; Altman et al., 1997) | 5 | 3 minutes | The ASRM is compatible with the widely used YMRS and the DSM diagnostic criteria, and it can be used effectively as a screening and diagnostic instrument despite its brevity. The cut-off score demonstrates good sensitivity and specificity. |
| Suicidal ideation | Suicidal Ideation Screen (SIS; Hetrick et al., 2017) | 3 | 2 minutes | The 3-item SIS has been validated in young people and demonstrated good properties relative to the standard licenced Suicidal Ideation Questionnaire. |
| Anxiety (generalised) | (a) Generalised Anxiety Disorder Scale (GAD-7; Spitzer et al., 2006) | 7 | 2-5 minutes | The GAD-7 provides a self-report assessment of the frequency of the seven DSM-5 diagnostic criteria for anxiety. It is widely used and brief. |
| | (b) Overall Anxiety Severity and Impairment Scale (OASIS; Norman et al., 2006) | 5 | 2-3 minutes | The OASIS is also a brief measure of anxiety symptoms, but it has the advantage of providing an assessment of severity. |
| Psychological distress | Kessler-10 Psychological Distress Scale (K-10; Andrews and Slade, 2001) | 10 | 3 minutes | The K-10 is a widely used transdiagnostic measure, providing a global assessment of psychological distress. Population-level comparisons can be drawn. |
| Psychosis (prodrome) | Prodromal Questionnaire (PQ-16; Ising et al., 2012) | 16 | 3-5 minutes | The PQ-16 is one of the few available self-report measures for assessing at-risk mental state (i.e. ultra-high risk for psychosis; UHR). |
| Personality | Personality Inventory for DSM-5 (PID-5-Brief; Krueger et al., 2013) | 25 | 5 minutes | The PID-5 is recommended by the American Psychiatric Association for assessing emerging personality disorder. The short form includes subscales for negative affect, detachment, antagonism, disinhibition and psychoticism. |
| Quality of life | Assessment of Quality of Life-8 Dimensions (AQoL-8D; Richardson et al., 2014) | 35 | 5 minutes | The AQoL-8D provides significant information related to quality of life. The psychometric properties of the AQoL-8D have been shown to be superior to other measures of quality of life, and it enables detailed economic evaluation of interventions. |
| Sleep | Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989) | 19 | 5-10 minutes | The PSQI is considered the gold standard self-report assessment of sleep quality. It provides a global score in addition to measures of sleep efficiency. |
| Substance use | (a) Alcohol, Smoking and Substance Involvement Screening Test (ASSIST; WHO ASSIST Working Group, 2002) | Varies | 1-10 minutes | Both the ASSIST and the AUDIT measures are widely used and endorsed for the assessment of drug and alcohol use by the World Health Organization. NOTE: If assessing both drug and alcohol, use the ASSIST; if assessing only alcohol, use the AUDIT. |
| | (b) Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993) | 10 | | |

YMRS: Young Mania Rating Scale; DSM: Diagnostic and Statistical Manual of Mental Disorders.

Table 2. Secondary clinical assessment battery.

| Domain | Instrument | No. of items | Reasons for inclusion |
|---------------------------|--|--------------|---|
| OCD | Brief Obsessive Compulsive Scale (BOCS; Bejerot et al., 2014) | 15 | The BOCS provides assessment of current and past behaviours and cognitions related to obsessions and compulsions. It correlates well with the significantly longer gold standard Y-BOCS, and is able to discriminate OCD from non-OCD disorders. |
| PTSD | Impact of Events Scale–Revised (IES-R; Weiss and Marmar, 1997) | 22 | The IES-R includes three subscales for intrusion, avoidance and hyperarousal and has been used in adult and youth populations as a measure of subjective distress caused by traumatic events. It is recommended that a member of the research or clinical team helps the participant identifying and describing the trauma. |
| Social phobia | Social Phobia and Social Interaction Anxiety Scales (SPS/SIAS; Mattick and Clarke, 1998) | 20/19 | The SPS/SIAS are recommended as a measure of social anxiety. The SPS assesses fears of being scrutinised during routine activities (eating, drinking, writing, etc.), while the SIAS assesses fears of more general social interaction. Both measures are widely used within the field of youth mental health. |
| Social isolation | UCLA Loneliness Scale (Russell, 1996) | 20 | The UCLA Loneliness Scale assesses subjective feelings of loneliness and feelings of social isolation. It is widely used within the field and demonstrates excellent psychometric properties. |
| Autism | Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001) | 50 | The AQ is a widely used autism self-report measure, and has been effectively used in the screening of autism in youth populations. |
| ADHD | Adults ADHD Self-Report Scale (ASRS; Kessler et al., 2005) | 18 | The ASRS is used to evaluate adults or adolescents for ADHD. It contains all based DSM-5 inattentive and hyperactive-impulsive symptom criteria reworded to be more appropriate for adolescents and adults. |
| Perceived clinical change | Patient Global Impression Improvement (PGI-I) | 1 | Based on the Clinical Global Impression (CGI) scale, the PGI-I enables the respondent to indicate perceived global change. It was considered a salient outcome especially given its brevity, but can only be used at follow-up assessments. |
| Eating disorders | Eating Disorder Examination Questionnaire (EDEQ; Fairburn and Beglin, 1994) | 38 | The EDEQ is widely viewed as the ‘gold standard’ measure of eating disorder psychopathology and is routinely used at the Royal Children’s Hospital (Melbourne). It has the advantage of a corresponding parent version. |
| Personality disorders | Personality Inventory for DSM-5—full version (PID-5; Krueger et al., 2011) | 220 | The PID-5 is recommended by the American Psychiatric Association for assessing emerging personality disorder. The long version provides a comprehensive self-report assessment, including 25 subscales. |
| Physical activity | International Physical Activity Questionnaire—Short Form (IPAQ-SF; Craig et al., 2003) | 7 | The IPAQ provides brief assessment of physical activity and inactivity. It has undergone extensive cross-nation reliability and validity testing, enabling comparison with population norms. |

OCD: obsessive compulsive disorder; PTSD: post-traumatic stress disorder; UCLA: University of California, Los Angeles; ADHD: attention-deficit/hyperactivity disorder; DSM: *Diagnostic and Statistical Manual of Mental Disorders*.

between and within individuals (regardless of the diagnosis) (Lee et al., 2014b), hence the need to move towards coordinated transdiagnostic approaches to research into neurocognition in mental illness.

The aim of the neurocognition battery is to assess the domains of neurocognition that are most relevant transdiagnostically in youth aged 12 to 25 years, including processing

speed, attention, working memory, verbal learning and memory, executive functioning and social cognition (Cotter et al., 2018; Lee et al., 2015). A number of criteria were considered in designing the battery including (1) brevity, (2) ease of administration and scoring by non-specialists, (3) applicable to age 12 to 25 years, (4) applicable to various cultures, (5) preference for computerised over paper-and-pencil tests, (6)

Table 3. Primary neurocognitive assessment battery.

| Domain | Instrument | Time taken | Reasons for inclusion |
|--|--|------------|--|
| Processing speed | Detection task (CogState; Lees et al., 2015) | 3 minutes | The Detection test measures processing speed using a simple reaction time paradigm. It is a relatively pure measure of speed likely to tap into state-related impairment, and it is sensitive to change. For a more comprehensive assessment of processing speed, add the Symbol Digit Modalities Test from the secondary battery. |
| Attention | Identification task (CogState) | 3 minutes | The Identification test measures attention and processing speed using a choice reaction time paradigm. It is sensitive to change. |
| Working memory | 1-Back task 2-Back task (CogState) | 6 minutes | The 1-Back and 2-Back tests measure working memory using an n-back paradigm. They are brief and sensitive to change. |
| Verbal learning and memory | Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996) | 10 minutes | The RAVLT is arguably the most common test used in clinical and research psychiatry to assess verbal learning and memory. Memory impairments can be present early in the illness course, may further decline with progression of illness and are linked to functional outcome. |
| Executive functioning | Groton Maze Learning task (CogState) | 7 minutes | The Groton Maze Learning Test measures executive function using a maze-learning paradigm. It provides a range of scores including errors (planning) and perseverative (set-shifting) responding and it is linked to functional outcome. |
| Social cognition – emotion recognition | Penn Emotion Recognition Task (ER-40; Gur et al., 2002) | 4 minutes | The ER-40 is a computerised task that assesses facial affect recognition ability. Good validity, recommended by a social cognition panel in schizophrenia, and linked to functioning. For more in-depth assessment of social cognition, add the Hinting Task from the secondary battery. |

alternate forms for repeat testing and (7) sensitive to change. A number of alternatives were considered by the panel including Cogstate, CANTAB, IntegNeuro, WebNeuro, WebCNP, NIH Toolbox and individual paper-and-pencil tasks. Several of the tests in the final chosen battery are from Cogstate because these tasks met all of the above criteria. Furthermore, support is provided within Australia.

The chosen assessments (Table 3) have broad applicability and have been used in populations of diverse cultural backgrounds (including Aboriginal and Torres Strait Islander peoples) (Cairney et al., 2007; Pearce et al., 2014), ages, and clinical presentations (Maruff et al., 2009). Apart from the Rey Auditory Verbal Learning Test (RAVLT), they do not rely heavily on language and verbal ability. The tasks have been shown to be sensitive to change, which is important in clinical trials and longitudinal studies (Collie et al., 2007). They have also been shown to be sensitive to subtle neurocognitive impairment, which is of high importance in young people with emerging or subthreshold mental disorder (Patel et al., 2017).

The primary cognitive battery was designed to have utility for predicting or informing a range of clinical outcomes, including psychopathology, psychosocial functioning, severity of illness, state, trait and ‘scar’/progressive

markers of illness and treatment response. However, it should be noted that the Cogstate assessments have costs associated.

Table 3 presents the primary neurocognitive tasks to be administered in full. Alike the clinical assessments, to ensure the primary battery remains as targeted and efficient as possible, additional neurocognitive domains can be assessed using to the secondary battery.

The administration of the secondary battery is supplementary and optional and it can also be administered in part. Assessment included in the secondary battery is presented in Table 4.

Neuroimaging data collection

The Neuroimaging panel identified two important aspects for the decision-making process: (1) the extent of sequence harmonisation across scan sites, and (2) the selection of imaging modalities. Several models for scan sequence harmonisation were identified ranging from full harmonisation to a very permissive model. The ENIGMA consortium⁴ is a good example of the permissive model, e.g. same imaging modalities, no control around scan protocols, common pipeline for quality assurance (QA), data processing and

Table 4. Secondary neurocognitive assessment battery.

| Domain | Instrument | Time taken | Reasons for inclusion |
|-----------------------------------|--|------------|---|
| IQ | Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II) two subtest version (Wechsler, 2011) | 15 minutes | The WASI-II is a brief, reliable measure of general cognitive ability for use in research settings. Record forms need to be purchased. |
| Perceived neurocognitive function | The Perceived Deficits Questionnaire–Depression (PDQ-D; Fehnel et al., 2016) | 10 minutes | The PDQ-D can be used to measure participant subjective perspective of neurocognitive function. This is clinically important to measure as subjective and objective neurocognitive measures often have a low correlation. The PDQ has not been validated in mental illnesses other than depressive disorders, but none of the questions are specific to depression. |
| | Neuropsychological Symptoms Self-Report (NSSR; Fisher et al., 2016) | 3 minutes | The NSSR has been developed for and used in youth depression, and correlates with symptom severity. It is very quick to administer. |
| Processing speed | Symbol Digit Modalities Test (SDMT; Smith, 1982) | 5 minutes | The SDMT measures processing speed which is a lower-order cognitive function, commonly affected in mental health conditions and which may underlie impairments in other neurocognitive domains. The SDMT can be purchased from Psychological Assessments Australia. |
| Theory of mind | Hinting Task (Corcoran et al., 1995) | 15 minutes | The Hinting Task assesses theory of mind, considered an important domain of social cognition. Specifically, it examines the ability to infer the real intentions behind indirect speech utterances. |

statistical analysis. It involves data sharing at the level of summary statistics or derived imaging measures. The restrictive model, or ‘Simple up’, implies identical scan protocols in the same imaging domains and the same common pipeline for QA. The restrictive model mandates a simple common set of scan sequences and aims for complete data sharing with the purpose to build a central database. Finally, the aspirational model, or ‘Smart down’, uses the same imaging modalities, with a set of scan protocols optimised for high-end scanner platforms but designed to be broadly compatible with lower-end scanner platforms.

It was agreed by the neuroimaging panel that although the permissive model is a very useful for the purpose of retrospective pooling of already collected data, we should aim for more extensive harmonisation than the permissive model. However, full harmonisation and standardisation of scan sequences across all sites (i.e. restrictive model) is likely not realistic, since this would necessarily restrict participation to sites where a particular scanner vendor or platform was available. It was therefore agreed that the aspirational model would be used, i.e. a set of sequences have been defined with the protocols optimised for high-end platforms with a second set of similar sequences, which can be implemented should the chosen vendor/operating system not support the high-end sequences. The requirements for high-end platforms are (1) high-density head

coils (32 channel or 64 head and neck), (2) fast gradient switching and (3) slice acceleration (for further details, see below). Sites with access to scanners that cannot support these capabilities should aim to acquire data at a similar spatial resolution while compromising in other aspects, e.g. temporal resolution for functional magnetic resonance imaging (fMRI), number of encoding directions and/or maximum *b*-value for diffusion MRI.

We chose to include state-of-the-art (or ‘high-end’) sequences to optimise the balance between high image quality given and a limited time of scanning and to be consistent and further harmonise with other major international big data neuroimaging initiatives such as UK Biobank,⁵ Human Connectome Project (HCP)⁶ and the Adolescent Brain Cognitive Development Study.⁷ The proposed brain imaging data acquisition includes six modalities, covering structural, diffusion and resting-state functional imaging.

The advantages of harmonising diffusion-weighted imaging and resting-state fMRI are the consistency of structural and functional brain networks across different subject groups, analysis methods and types of scanning protocols, and their application includes minimal demands on patient compliance and short acquisition times. Sufficiently high spatial resolution of structural MRI enables the user to choose between traditional volume-based analyses and HCP-like surface-based analysis. Compulsory

Table 5. Primary set of neuroimaging modalities.

| Modality | Indicative scan time (min) | Description |
|--|----------------------------|--|
| Localizer | < 1 | A three-plane localizer or 'scout' scan is a scan to find the subject's head. It is also used for prescription for the subsequent scans. |
| T1-weighted structural image | 6 | A T1-weighted imaging is a structural technique with high-resolution depiction of brain anatomy, having strong contrast between grey and white matter. It can easily be harmonised across different scanners and platforms and is critical for processing other imaging modalities. This modality provides primarily measures of brain tissues and structures (e.g. volumes, cortical thickness, cortical surface area, shape measures, etc.). It is also critical for calculations of cross-subject and cross-modality alignments, needed in order to process all other brain modalities. |
| Diffusion-weighted images | 7 | Diffusion-weighted imaging is a structural technique that measures the ability of water molecules to move within their local tissue environment and that allows the study of neural fibre tract anatomy and structural brain connectivity. Water diffusion is measured along a range of orientations, providing various types of measures. For example, local (voxel-wise) estimates of diffusion properties reflect the integrity of microstructural tissue compartments (e.g. diffusion tensor estimates). Long-range estimates based on tract-tracing (tractography) reflect structural connectivity between brain regions. |
| Reverse encoding diffusion-weighted imaging (b0 directions only) | 1 | This is an essential sequence to perform distortion correction on the diffusion-weighted images. |
| Multiband resting-state functional MRI images (multiband acceleration 6) | 8 | Resting-state functional MRI (rsfMRI) measures changes in blood oxygenation associated with intrinsic brain activity (i.e. in the absence of an explicit task or sensory stimulus). Derived measures estimate the connectivity between brain regions, as reflected in the presence of spontaneous co-fluctuations in signal. |
| Gradient Echo Fieldmaps PA and AP (reverse phase encoding) BOLD (multiband acceleration 1) | 1 | These are two brief scans to correct the distortion that comes with the high multiband factor used in the resting-state fMRI sequence above. |
| Spin Echo Fieldmaps PA and AP (reverse phase encoding) BOLD imaging (multiband acceleration 1) | 1 | Same as above, used for distortion correction, but with spin echo instead of gradient echo. |

MRI: magnetic resonance imaging.

modalities are presented in Table 5. The current paper presents an overview of selected modalities, but a more detailed description of the acquisition parameters and scan protocols can be found in the supplemental material.

To keep the total scan time of the primary imaging assessments within 30 minutes, it was decided that task-based fMRI would only be included in the secondary set of modalities. The choice for a specific task included in the second set of modalities will be dependent on the personal interest of researchers and/or on the aims of individually funded projects. However, to promote data sharing of task-related fMRI data, the fMRI tasks that are currently being used by the different members of the neuroimaging panel can be found on the Neurobiology Partnership website. Secondary set of modalities also includes T2-weighted structural images, which is a structural technique with contrast

dominated by signal decay from interactions between water molecules (T2 relaxation times). T2 images depict alterations to tissue compartments typically associated with pathology (e.g. white matter lesions). The final modality is susceptibility-weighted imaging, which is a structural technique that is sensitive to magnetised tissue constituents (magnetic susceptibility). Data from one scan (including phase and magnitude images from two echo times) can be processed in multiple ways to reflect venous vasculature, microbleeds or aspects of microstructure (e.g. iron, calcium and myelin).

Collection of biospecimens

Peripheral biomarkers, i.e. biomarkers that can be quantified from peripheral tissue (such as plasma, blood cells or

Table 6. Primary biospecimens collection.

| Type of specimen | Biomarker category | Target biomarkers | Rationale |
|-------------------------------|---|--|---|
| Buffy coat/whole blood (EDTA) | Genetic biomarkers | DNA | The collection of this material enables: analysis of genetic variants association discovery (in combination with other datasets) polygenic risk scoring using genetic information generated in external datasets. generation of genetic risk score for psychiatric disorders and other traits of relevance analysis of epigenetic markers such as DNA methylation |
| Plasma (EDTA)/serum (SST II) | Immune system Oxidative stress Infectious disease | Inter-leukins Cytokines Antibodies | Alterations of the levels of several immune proteins and oxidative stress markers have been described in patients with psychotic and mood disorders, and such patterns may be indicative of trait or state characteristics. Serological studies suggest that some cases of schizophrenia and bipolar disorder may be associated with exposure to microbial infections. New 'omic technologies now allow cost-effective screens of lipids, metabolites, proteins. |
| Dried blood spot | Membrane fatty acids | PUFAs MUFAs | Membrane polyunsaturated fatty acids (PUFAs) represent a candidate marker for mental illnesses risk, and change in PUFA levels might be an early indicator of response to therapeutic interventions. Dried blood spot test enables the quantification of a variety of PUFAs and of monounsaturated fatty acid (MUFA). |

EDTA: ethylenediaminetetraacetic acid.

hair), are promising indicators of the pathophysiological processes that underlie mental disorders (McGorry et al., 2013). As such, peripheral biomarkers may serve multiple purposes, including indicating risk for onset or progression to a more advanced stage, delineating diagnostic entities or informing treatment choice.

Several meta-analyses of cross-sectional studies support alterations in peripheral biomarkers in schizophrenia/psychosis (Berger et al., 2016; Goldsmith et al., 2016), mood and anxiety disorders (Berk, 2015). However, it is less clear whether these group-level differences extend to earlier illness stages, how the significant within-group heterogeneity can be accounted for and longitudinal studies are needed to delineate the role of biomarkers over time. The harmonisation of biospecimen collection has the potential to address these questions, as it allows the detection of smaller effects, better control of confounders and cross-validation of findings.

The primary purpose of the harmonisation of the biospecimen collection is to agree on a set of specimens that are supported as candidate biomarkers by the literature (McGorry et al., 2013), can be collected in ambulatory settings and are of relevance to a broad spectrum of disorders. As such, the selected biomarkers consist of commonly reported and novel biomarkers, and those that have shown prospective associations with illness trajectories. The selection criteria used by the biospecimens panel were as follows: established

association with risk and clinical outcomes, relevance to the pathophysiology of mental disorders (in particular considering a transdiagnostic approach), minimising the burden to the participant, and limiting the invasiveness of the specimen collection.

The collection of the selected primary biomarker simply requires a blood draw (Table 6). The secondary biospecimens collection requires a supplementary blood tube and the collection of a hair sample. Together, these can be completed within approximately 5 minutes.

Secondary biospecimen collection includes the collection of blood using the PAXgene tube, which enables the quantification of transcription products of genes relevant to psychiatric disorders. RNA quantification can be performed to determine differences in gene expression between affected and unaffected participants, as a response to treatment and more. The panel also recommends the collection of hair for cortisol measurement. Cortisol is the end product of the hypothalamic–pituitary–adrenal axis (HPA axis) and a major mediator of the body's response to stress. Meta-analyses show elevated morning cortisol levels (Girshkin et al., 2016) and flattened awakening response in patients with schizophrenia and first-episode psychosis (Berger et al., 2016). Hair cortisol reflects the average cortisol concentrations in the past month and is robust to the sampling bias inherent to plasma cortisol (sample collection may evoke a stress response).

Discussion

In the present paper, a list of assessments or measures are presented with a view to invite our collaborators working in the field of youth mental health to collect a core set of data in an harmonised way. The measures have been thoughtfully selected by experts in the field following multiple consultations and using a transdiagnosis approach. In the study of the neurobiology of mental illness, large datasets are necessary due to the heterogeneity of the studied populations and, too often, pooling data from independent studies is not possible due to the data collected not being comparable. The harmonised collection of data will enable the production of a large dataset holding data on the clinical and neurobiological presentation of young people in the early stages of a mental illness. For example, large datasets will enable the identification of biomarkers in a more efficient way. Biomarkers can predict risk, delineate diagnostic entities (diagnostic or trait biomarkers), correlate with severity (state biomarkers), correlate with staging (stage biomarkers), allow prediction of treatment response (treatment biomarkers), and/or can predict the course and outcome of an illness (prognostic biomarkers) (McGorry et al., 2013). An important aim of studying biomarkers is to systematically evaluate their potential in regards to the above classifications and in particular with relevance to transdiagnostic conceptualisations and clinical staging.

Further to the assessments presented here, standardised demographics, history of medical conditions and history of treatment are also available on the Neurobiology Partnership website.

A limitation to the use of the suggested batteries of assessment in youth with mental ill health is that psychopathology is a ‘western’ construct, and within the Australian context, these measures have not been validated in Aboriginal and Torres Strait Island populations, which is an important consideration for data harmonisation.

According to data-sharing policies currently being prepared and pending ethics and governance clearance, contributors to the dataset will have access to the data and also to the expertise of other researchers and/or clinicians involved in the Neurobiology in Youth Mental Health Partnership. The Neurobiology Partnership will provide support in the implementation of one or more of the four suggested core sets of data into your research projects. Any interested collaborator is invited to visit the website <https://NeurobiologyPartnership.org.au>.

Acknowledgements

The authors would like to acknowledge the members of the Neurobiology in Youth Mental Health Partnership for their participation in the various panels of experts.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

K.A. and A.L. are supported by National Health and Medical Research Council (NHMRC) Career Development Fellowships nos. 1141207 and 1148793, respectively.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. www.nimh.nih.gov
2. www.neurobiologypartnership.org.au
3. www.phenx.org/
4. <http://enigma.ini.usc.edu/>
5. www.ukbiobank.ac.uk/
6. www.humanconnectomeproject.org/
7. <https://abcdstudy.org/>

References

- Altman E, Hedeker D, Perterson J, et al. (1997) The Altman Self-Rating Mania Scale. *Biological Psychiatry* 42: 948–955.
- Andrews G and Slade T (2001) Interpreting scores on the Kessler Psychological Distress Scale (K10). *Australian and New Zealand Journal of Public Health* 25: 494–497.
- Baron-Cohen S, Wheelwright S, Skinner R, et al. (2001) The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders* 31: 5–17.
- Bejerot S, Edman G, Anckarsäter H, et al. (2014) The Brief Obsessive-Compulsive Scale (BOCS): A self-report scale for OCD and obsessive-compulsive related disorders. *Nordic Journal of Psychiatry* 68: 549–559.
- Berger M, Kraeuter A, Romanik D, et al. (2016) Cortisol awakening response in patients with psychosis: Systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews* 68: 157–166.
- Berk M (2015) The classification of biomarkers. *JAMA Psychiatry* 72: 2658–2667.
- Bhui K (2017) Treatment resistant mental illnesses. *British Journal of Psychiatry* 210: 443–444.
- Buysse DJ, Reynolds CF, Monk TH, et al. (1989) The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research* 28: 193–213.
- Cairney S, Clough A, Jaragba M, et al. (2007) Cognitive impairment in aboriginal people with heavy episodic patterns of alcohol use. *Addiction* 102: 909–915.
- Caspi A, Houts R, Belsky D, et al. (2014) The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science* 2: 119–137.
- Collie A, Darekar A, Weissgerber G, et al. (2007) Cognitive testing in early-phase clinical trials: Development of a rapid computerized test battery and application in a simulated Phase I study. *Contemporary Clinical Trials* 28: 391–400.
- Corcoran R, Mercer G and Frith C (1995) Schizophrenia, symptomatology and social inference: Investigating ‘theory of mind’ in people with schizophrenia. *Schizophrenia Research* 17: 5–13.
- Cotter J, Granger K, Backx R, et al. (2018) Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience and Biobehavioral Reviews* 84: 92–99.
- Craig C, Marshall A, Sjöström M, et al. (2003) International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise* 35: 1381–1395.

- Fairburn C and Beglin S (1994) Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders* 16: 363–370.
- Fehnel S, Forsyth B, Di Benedetti D, et al. (2016) Patient-centered assessment of cognitive symptoms of depression. *CNS Spectrums* 21: 43–45.
- Fisher C, Goodall J, Simmons M, et al. (2016) Subjective ratings of neurocognitive functioning in depressed young people undergoing treatment: Utility of a brief screening tool. *Early Intervention in Psychiatry* 10: 252–257.
- Girshkin L, O'Reilly N, Quidé Y, et al. (2016) Diurnal cortisol variation and cortisol response to an MRI stressor in schizophrenia and bipolar disorder. *Psychoneuroendocrinology* 67: 61–69.
- Goldsmith D, Rapaport M and Müller B (2016) A meta-analysis of blood cytokine network alterations in psychiatric patients: Comparisons between schizophrenia, bipolar disorder and depression. *Molecular Psychiatry* 21: 1696–1709.
- Goodall J, Fisher C, Hetrick S, et al. (2018) Neurocognitive functioning in depressed young people: A systematic review and meta-analysis. *Neuropsychology Review* 28: 216–231.
- Gur R, Sara R, Hagendoorn M, et al. (2002) A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods* 115: 137–143.
- Hetrick S, Goodall J, Yuen H, et al. (2017) Comprehensive online self-monitoring to support clinicians manage risk of suicide in youth depression. A pilot study. *Crisis* 38: 147–157.
- Ising H, Veling W, Loewy R, et al. (2012) The validity of the 16-Item Version of the Prodromal Questionnaire (PQ-16) to screen for ultra high risk of developing psychosis in the general help-seeking population. *Schizophrenia Bulletin* 38: 1288–1296.
- Kessler RC, Adler L, Ames M, et al. (2005) The World Health Organization Adult ADHD Self-Report Scale (ASRS): A short screening scale for use in the general population. *Psychological Medicine* 35: 245–256.
- Kessler RC, Amminger GP, Aguilar-Gaxiola S, et al. (2007) Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry* 20: 359–364.
- Krueger R, Derringer J, Markon K, et al. (2011) Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine* 42: 1879–1890.
- Krueger R, Derringer J, Markon K, et al. (2013) *The Personality Inventory for DSM-5 Brief Form (PID-5-BF)*. Washington, DC: American Psychiatric Association.
- Lee RS, Hermens D, Naismith S, et al. (2015) Neuropsychological and functional outcomes in recent-onset major depression, bipolar disorder and schizophrenia-spectrum disorders: A longitudinal cohort study. *Translational Psychiatry* 5: e555.
- Lee RS, Hermens D, Scott J, et al. (2014a) A meta-analysis of neuropsychological functioning in first-episode bipolar disorders. *Journal of Psychiatric Research* 57: 1–11.
- Lee RSC, Hickie IB and Hermens DF (2014b) Letter to the Editor: Neuropsychological subgroups are evident in both mood and psychosis spectrum disorders. *Psychological Medicine* 44: 2015.
- Lee RSC, Hermens DF, Scott J, et al. (2017) A transdiagnostic study of education, employment, and training outcomes in young people with mental illness. *Psychological Medicine* 47: 2061–2070.
- Lees J, Applegate E, Emsley R, et al. (2015) Calibration and cross-validation of MCCB and CogState in schizophrenia. *Psychopharmacology* 232: 3873–3882.
- McGorry P, Keshavan M, Goldstone S, et al. (2013) Biomarkers and clinical staging in psychiatry. *World Psychiatry* 13: 211–223.
- Maruff P, Thomas E, Cysique L, et al. (2009) Validity of the CogState brief battery: Relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Archives of Clinical Neuropsychology* 24: 165–178.
- Mattick R and Clarke J (1998) Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy* 36: 455–470.
- Meshulam-Gately R, Giuliano A, Faraone S, et al. (2009) Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology* 23: 315–336.
- Mission Australia and the Black Dog Institute (2017) *Youth Mental Health Report – Youth Survey 2012-16*. Sydney, NSW, Australia: Mission Australia.
- Norman S, Cissell S, Means-Christensen A, et al. (2006) Development and validation of an Overall Anxiety Severity And Impairment Scale (OASIS). *Depression and Anxiety* 23: 245–249.
- Patel S, Meier A, Fernandez N, et al. (2017) Convergent and criterion validity of the CogState computerized brief battery cognitive assessment in women with and without breast cancer. *Clinical Neuropsychologist* 31: 1375–1386.
- Pearce M, Mann K, Singh G, et al. (2014) Birth weight and cognitive function in early adulthood: The Australian Aboriginal birth cohort study. *Journal of Developmental Origins of Health and Disease* 5: 240–247.
- Richardson J, Jezzi A, Khan M, et al. (2014) Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. *The Patient: Patient-Centered Outcomes Research* 7: 85–96.
- Rush A, Trivedi M, Ibrahim H, et al. (2003) The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry* 54: 573–583.
- Russell D (1996) UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment* 66: 20–40.
- Saunders J, Aasland O, Babor T, et al. (1993) Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption – II. *Addiction* 88: 791–804.
- Schmidt M (1996) *Rey Auditory Verbal Learning Test. A Handbook*. Los Angeles, CA: Western Psychological Services.
- Smith A (1982) *Symbol Digit Modalities Test (SDMT). Manual (Revised)*. Los Angeles, CA: Western Psychological Services.
- Spitzer R, Kroenke K, Williams J, et al. (2006) A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine* 166: 1092–1097.
- Wechsler D (2011) *Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson.
- Weiss D and Marmar C (1997) *Assessing Psychological Trauma and Post Traumatic Stress Disorder: A Handbook for Practitioners*. New York: Guilford Press.
- WHO ASSIST Working Group (2002) The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, reliability and feasibility. *Addiction* 97: 1183–1194.
- Zakzanis K, Campbell Z and Polsinelli A (2010) Quantitative evidence for distinct cognitive impairment in anorexia nervosa and bulimia nervosa. *Journal of Neuropsychology* 4: 89–106.