World Psychiatry

OFFICIAL JOURNAL OF THE WORLD PSYCHIATRIC ASSOCIATION (WPA)

Volume 17, Number 2 *June 2018*

http://www.arabpsynet.com/Journals/WJ/WP-June2018.Eng.pdf

WILEY **IMPACT FACTOR: 26.561** ISSN 1723-8617

The World Psychiatric Association (WPA)

The WPA is an association of national psychiatric societies aimed to increase knowledge and skills necessary for work in the field of mental health and the care for the mentally ill. Its member societies are presently 140, spanning 120 different countries and representing more than 200,000 psychiatrists.

The WPA organizes the World Congress of Psychiatry every three years. It also organizes international and regional congresses and meetings, and thematic conferences. It has 72 scientific sections, aimed to disseminate information and promote collaborative work in specific domains of psychiatry. It has produced several educational programmes and series of books. It has developed ethical guidelines for psychiatric practice, including the Madrid Declaration (1996).

Further information on the WPA can be found on the website www.wpanet.org.

WPA Secretariat
Geneva University Psychiatric Hospital, 2 Chemin du Petit Bel-Air, 1225 Chêne-Bourg, Geneva, Switzerland. Phone: +41223055737; Fax: +41223055735; E-mail: wpasecretariat@wpanet.org.

World Psychiatry

World Psychiatry is the official journal of the World Psychiatric Association. It is published in three issues per year and is sent free of charge to psychiatrists whose names and addresses are provided by WPA member societies and sections.

Research Reports containing unpublished data are welcome for submission to the journal. They should be subdivided into four sections (Introduction, Methods, Results, Discussion). References should be numbered consecutively in the text and listed at the end according to the following style:

1. Cuijpers P, Sijbrandij M, Koole SL et al. Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis. World Psychiatry 2014;13: 56-67.
2. McRae TW. The impact of computers on accounting. London: Wiley, 1964.
3. Fraeijs de Veubeke B. Displacement and equilibrium models in the finite element method. In: Zienkiewicz OC, Hollister GS (eds). Stress analysis. London: Wiley, 1965:145-97.

All submissions should be sent to the office of the Editor.

Office of the Editor – Department of Psychiatry, University of Naples SUN, Largo Madonna delle Grazie, 80138 Naples, Italy. Phone: +390815666502; Fax: +390815666523; E-mail: majmario@tin.it.

World Psychiatry **is indexed in PubMed, Current Contents/Clinical Medicine, Current Contents/Social and Behavioral Sciences, Science Citation Index, and EMBASE.**
All back issues of *World Psychiatry* can be downloaded free of charge from the PubMed system (http://www.pubmedcentral.nih.gov/tocrender.fcgi?journal=297&action=archive).

Why the clinical utility of diagnostic categories in psychiatry is intrinsically limited and how we can use new approaches to complement them

It is becoming increasingly evident that the usefulness of diagnostic categories in psychiatry has been overemphasized. These categories have been initially charged with implications in terms of pointing to a specific treatment and prospectively a specific etiology and/or pathogenesis, in complete analogy with the other branches of medicine. More recently, they have been more modestly charged with relative, not absolute, pragmatic implications in terms of guiding the formulation of a management plan and the prediction of outcomes (the two main elements of "clinical utility")[1]. The underlying concept has been that we are dealing with "patterns" of intercorrelated reported experiences (in medical jargon, symptoms) and observed behaviours (in medical jargon, signs) which allow significant inferences about further course and management, whereas there is no assumption that these patterns are all "natural kinds" (i.e., discrete disease entities marking a real division in nature)[2]. Indeed, improving the clinical utility of psychiatric diagnoses has been the declared main objective of both the DSM-5 and, even more explicitly, the ICD-11[3].

Unfortunately, even these more modest implications of diagnostic categories in psychiatry have turned out to be overestimated. This is not to say that our current diagnoses do not have clear implications in terms of treatment choice and prediction of outcomes. The fact is, however, that these implications are less significant than originally believed and still assumed by most treatment guidelines. A clear reflection of this state of affairs can be found in the survey by First et al[4] that appears in this issue of the journal, in which a large sample of users of either the ICD-10 or some edition of the DSM rated those diagnostic systems as having the *lowest* utility in "selecting a treatment" and "assessing probable prognosis", whereas they were perceived to be much more useful for meeting administrative requirements, communicating with other health professionals, and teaching trainees or students. Indeed, both research evidence and clinical experience tell us that patients sharing the same psychiatric diagnosis often respond differently to a given treatment, and patients with different psychiatric diagnoses may respond similarly to a given treatment (not to mention the wide variability of outcomes in people receiving the same diagnosis).

Alternative approaches to the ICD/DSM are currently being developed. They usually assume either: a) that the realm of psychopathology can be more efficiently described in terms of dimensions, or b) that the neurobiological underpinnings of psychopathology should be the major drivers of psychiatric classifications. These alternative approaches are being put forward both at the level of the entire realm of psychopathology (respectively, by projects such as the Hierarchical Taxonomy of Psychopathology, HiTOP[5] and the Research Domain Criteria, RDoC[6]) and at the level of specific areas of psychopathology (respectively, through models such as the "transdiagnostic psychosis spectrum"[7] and the "neurodevelopmental gradient"[8]).

These approaches, in order to really emerge in the future as a practical alternative to ICD/DSM-based diagnosis, will have to prove: a) to be reasonably applicable in ordinary clinical practice (also in various clinical settings and in the hands of different categories of professionals), and b) to be actually more clinically useful than current diagnostic practices, i.e., more efficient in guiding the choice of treatment and the prediction of outcomes. This evidence is *not* available at the moment.

But, are these approaches really "alternative" to the DSM/ICD systems, as they are usually proposed to be? I think it needs to be clarified that, in psychiatric practice, "diagnosis" (i.e., the application to an individual case of a given category or "type" from a classification) is (or should be) only one step in the process that leads to the formulation of the management plan and of prognosis. The other step is (or should be) the further characterization of the individual case with respect to a series of additional variables. This second step is at least as important as "diagnosis" in the management choices and the prediction of outcomes. Since the vast majority of our current diagnostic categories are unlikely to represent "natural kinds" (and the minority which may approximate that model are likely to be heterogeneous from the etiopathogenetic viewpoint), the information conveyed by "diagnosis" (i.e., the "type" to which the patient can to a variable extent be reconducted) is in itself insufficient for therapeutic and prognostic purposes. Hence the need for a more detailed psychopathological characterization of the individual case, as well as for an exploration of what is behind the "pattern" we have applied, in that specific case, with respect to vulnerability and protective factors.

The fact is, however, that up to now the first step (diagnosis) has received a lot of attention, with the production of several generations of tools providing systematic guidance to the clinician, whereas the second step (further characterization of the individual case) has been largely ignored, thus generating an inter-clinician variability in its implementation which is not less significant and deleterious than that described for the first step in the 1970s. The focus on diagnostic categories in most research and in virtually all clinical guidelines, as well as the emphasis on pharmacological interventions, for which a simplistic and stereotyped relationship between "diagnosis" and "treatment" can be more easily proposed, has certainly contributed to this situation.

Well, one could argue that the above "alternative" approaches may not have a significant chance in the future to "replace" our current diagnostic practices (i.e., to take their place in the first step of the above-mentioned process), while

they are much more likely to improve significantly the second step (the further characterization of the individual case), thus complementing current diagnoses.

What are, or may be, in fact the main elements of that second step? They include the characterization of the individual case with respect to the relevant psychopathological dimensions and possibly to the current stage of development of the diagnosed disorder (see McGorry et al[9] in this issue of the journal); an assessment of the severity of the clinical picture which is less generic and more evidence-based than that currently provided by the ICD and the DSM; the exploration of antecedent variables such as family history of mental illness, other parental factors, perinatal factors, early environmental exposures, psychomotor development, premorbid social adjustment, psychopathological antecedents, and possibly in the future polygenic risk scores; and the assessment of concomitant variables such as personality traits, cognitive functioning, social functioning (including personal resources such as resilience and coping strategies), soft neurological signs, substance abuse, recent environmental exposures, and possibly in the future some biological markers. It is with respect to the assessment of these latter elements that clinicians need today a systematic guidance, which current diagnostic systems and related tools do not provide, or do not provide satisfactorily (thus contributing to a therapeutic practice which, being guided just by a diagnostic label, is oversimplified and stereotyped).

I would therefore envisage that the approaches which are currently regarded as alternative to the ICD and DSM may not turn out to be, in the future, a basis for a clinically useful reclassification of psychopathology, but that elements of those approaches may be increasingly incorporated in the further characterization of the individual case, which is at least as important as the application of a diagnostic label in the management choices and the formulation of prognosis.

The message may be, therefore, that we do need current diagnostic categories (which can certainly be much improved, but without which we would either be lost in a *mare magnum* of variables, or presented with synthetic formulations which are less efficient, in addition to being potentially controversial and not rooted in clinical tradition), but that those categories are intrinsically insufficient in pursuing the "clinical utility" objectives of the DSM-5 and the ICD-11, because the act of diagnosis is only one step in the process leading to the key aims of the optimal formulation of the management plan (especially if this does not include just the selection of a medication) and the prediction of outcomes (especially if this is meant to cover not only clinical variables, but also elements concerning social functioning and personal recovery).

We should start to promote the construction and validation – in addition to structured interviews leading to a given diagnosis – of tools guiding the clinician systematically in the characterization of the individual case, with a special focus on the assessment of psychopathological dimensions, the reliable evaluation of the severity of the clinical picture, and the exploration of a series of antecedent and concomitant variables. We should try to incorporate in this effort – already now and increasingly in the future – elements of the approaches that are currently presented as alternative to the ICD/DSM. The entire mental health field should ideally contribute to this endeavour, declaring a moratorium on self-defeatism and parochial struggles.

Mario Maj
Department of Psychiatry, University of Campania "L. Vanvitelli", Naples, Italy

1. Jablensky A. World Psychiatry 2016;15:26-31.
2. Maj M. World Psychiatry 2016;15:1-2.
3. Reed GM. Prof Psychol Res Pr 2010;41:457-63.
4. First MB, Rebello TJ, Keeley JW. World Psychiatry 2018;17:187-95.
5. Kotov R, Krueger RF, Watson D et al. World Psychiatry 2018;17:24-5.
6. Cuthbert BN. World Psychiatry 2014;8:28-35.
7. van Os J, Reininghaus U. World Psychiatry 2016;15:118-24.
8. Owen MJ, O'Donovan MC. World Psychiatry 2017;16:227-35.
9. McGorry PD, Hartmann JA, Spooner R et al. World Psychiatry 2018;17:133-42.

DOI:10.1002/wps.20512

Experience sampling methodology in mental health research: new insights and technical developments

Inez Myin-Germeys[1], Zuzana Kasanova[1], Thomas Vaessen[1], Hugo Vachon[1], Olivia Kirtley[1], Wolfgang Viechtbauer[1,2], Ulrich Reininghaus[2,3]

[1]Center for Contextual Psychiatry, Department of Neurosciences, KU Leuven, Leuven, Belgium; [2]School for Mental Health & Neuroscience, Maastricht University, Maastricht, The Netherlands; [3]Centre for Epidemiology and Public Health, Health Service and Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

In the mental health field, there is a growing awareness that the study of psychiatric symptoms in the context of everyday life, using experience sampling methodology (ESM), may provide a powerful and necessary addition to more conventional research approaches. ESM, a structured self-report diary technique, allows the investigation of experiences within, and in interaction with, the real-world context. This paper provides an overview of how zooming in on the micro-level of experience and behaviour using ESM adds new insights and additional perspectives to standard approaches. More specifically, it discusses how ESM: a) contributes to a deeper understanding of psychopathological phenomena, b) allows to capture variability over time, c) aids in identifying internal and situational determinants of variability in symptomatology, and d) enables a thorough investigation of the interaction between the person and his/her environment and of real-life social interactions. Next to improving assessment of psychopathology and its underlying mechanisms, ESM contributes to advancing and changing clinical practice by allowing a more fine-grained evaluation of treatment effects as well as by providing the opportunity for extending treatment beyond the clinical setting into real life with the development of ecological momentary interventions. Furthermore, this paper provides an overview of the technical details of setting up an ESM study in terms of design, questionnaire development and statistical approaches. Overall, although a number of considerations and challenges remain, ESM offers one of the best opportunities for personalized medicine in psychiatry, from both a research and a clinical perspective.

Key words: Experience sampling methodology, ecological momentary interventions, mental health, context, psychopathology, person-environment interaction, variability, treatment evaluation, mixed-effect models

(World Psychiatry 2018;17:123–132)

The essence of psychiatric symptoms is that they are natural experiences emerging in the realm of ordinary daily life, often in interaction with contextual factors.

In the mental health field, there is a growing awareness that the study of psychiatric symptoms in the context of everyday life may provide a powerful and necessary addition to more conventional research approaches[1]. Increasingly, studies are using techniques such as experience sampling methodology (ESM)[2,3] or ecological momentary assessment (EMA)[4,5] to study psychopathology in daily life.

This paper provides a comprehensive review of the principles and applications of ESM, and an update on its design and techniques in the mental health field.

PRINCIPLES OF ESM

ESM is a structured self-report diary technique assessing mood, symptoms, context and appraisals thereof as they occur in daily life[1,3]. It typically requires participants to complete a momentary questionnaire several times a day over a number of days.

ESM is rooted in ecological psychology, which states that behaviour is radically situated, i.e., it can only be understood in relation to the context in which it occurs[6]. In order to fully understand experience and behaviour, they need to be investigated in the real-world context, outside the laboratory.

The use of ESM to investigate experiences within, and in interaction with, the real-world context is also consistent with a more recent emphasis on embodiment and embeddedness in the cognitive sciences[7]. This approach claims that "the way humans (and other organisms) perceive, think and act is determined by the kinds of bodies they have and the kind of actions they perform or are capable of"[8]. Therefore, in order to understand or explain experiences, one must examine them in interaction with the context. ESM directly addresses this, in that it allows us to study the actual experience as it occurs in everyday environments, rather than assessing people's self-reflections on who they are or how they usually behave.

As experience and behaviour are at the core of psychopathology, ESM was quickly adopted by mental health researchers[3]. In addition to its benefits in terms of examining experience and behaviour in interaction with the real-life context, ESM also allows us to study these in the moment, overcoming the problem of retrospective recall bias, and prospectively, allowing us to investigate temporal variability and associations. ESM may therefore fundamentally strengthen the behavioural study of mental health problems and contribute to a contextual approach to personalized medicine in psychiatry.

APPLICATIONS OF ESM IN THE MENTAL HEALTH RESEARCH FIELD

In this section, we review how zooming in on the micro-level of experience and behaviour using ESM can help in improving our understanding of the phenomenology and aetiol-

ogy of psychopathology and in strengthening and changing clinical practice.

Improving understanding of symptoms

Although psychiatry has developed a common terminology to describe different aspects of psychopathology (e.g., as reflected in features described in the ICD or DSM), our understanding of the nature of these phenomena is still fairly limited, partly due to the biases introduced by the retrospective recall of symptoms. ESM addresses this issue by capturing symptoms as they occur. Indeed, studies that compared retrospective and ESM assessment of symptoms reported that the former assessment tends to under- or over-estimate depressive symptoms in patients with depression[9]. Furthermore, only moderate correlations were found between retrospective and momentary assessment of affect in patients with psychosis[10,11].

Besides providing a more accurate assessment, ESM may be instrumental for obtaining a deeper understanding of how symptoms unfold in daily life over time. Good examples are anhedonia and avolition, both of which have long been reported to form part of the phenomenology of psychotic disorders, depression and bipolar disorder[12].

Anhedonia is generally described as a diminished capacity to experience pleasure. However, what does this mean for our experience and behaviour in daily life? A decreased level of positive affect in daily life – which has been found in some studies in patients with psychosis[13,14] – may reflect a diminished capacity to experience pleasure. Yet, decreased levels of positive affect may also result from patients' lives being less enjoyable. Indeed, patients with psychosis do report, on average, a lower number of pleasant events in their daily life than healthy controls[13]. In order to disentangle this, positive affect has been examined in moments when people do report pleasant events. ESM research in individuals with psychosis revealed an intact ability to generate positive affect upon experiencing pleasant events in daily life[15-17], which does not support the widely held assumption that anhedonia reflects a general incapacity to experience pleasure.

So, what does anhedonia then relate to? Gard et al[18] distinguished experiencing positive affect in the moment (consummatory pleasure) from pleasure related to future activities (anticipatory pleasure), and found the latter to be particularly reduced in patients with psychosis. This distinction may partly explain why patients with psychosis and students with social anhedonia (assessed with observer-rated measures) reported higher levels of positive affect when in the company of others compared to when alone (i.e., suggesting higher consummatory pleasure in social situations), but still spent more time on their own[13,19].

This brings us to avolition, commonly defined as a lack of motivation or, put differently, an incapacity to translate positive emotional experience into productive goal-directed behaviour. The longitudinal design as well as the assessment of both mood and activities in ESM allows us to directly relate emotional experience to subsequent activities and behaviour. For example, it has been shown that positive affect experienced in social contexts or during physical activity increases the odds of engaging in these behaviours at the next measurement moment in healthy women[20]. Interestingly, this reward-oriented behaviour of positive affect in one moment driving future behaviour was absent in the everyday behaviour of individuals with anhedonia[16], supporting the idea that a lack of anticipatory pleasure (i.e., anhedonia) may result in a reduced drive for seeking out these activities (i.e., avolition).

An emerging trend in affective neuroscience has been to increase functional relevance of experimental findings by investigating hedonic and goal-directed behaviour using laboratory paradigms in conjunction with ESM. Moran et al[21] linked blunted daily-life experience of pleasure and motivation to poorer performance on effort and reward learning tasks in individuals with schizophrenia. Our group combined ESM with neuromolecular imaging in healthy individuals with increased familial risk for psychosis. We found that intact striatal dopaminergic modulation of reward learning predicted daily-life reward-oriented behaviour in both groups, which may point to a neurochemical and behavioural mechanism of resilience in those predisposed to psychosis[22,23].

In sum, the accumulating ESM accounts of hedonic and volition function in clinical populations have contributed to a more fine-grained understanding of the affective and behavioural dynamics compromising productive goal-directed behaviour, thus opening promising avenues for equally fine-grained prevention and treatment strategies.

Capturing emotional variability

Many mental disorders are characterized by dynamic fluctuations in emotions. The resolution of traditional self-report measures for capturing these fluctuations is limited. Multiple measurements within one person in ESM may help to assess affective variability in more detail, as well as to identify the context in which (mal)adaptive emotion regulation strategies are used[24,25].

A meta-analysis investigating dynamic fluctuations in emotions showed that lower levels of psychological well-being were associated with greater intensity of emotions, larger moment-to-moment fluctuations and a slower recovery to a normative state (i.e., inertia), and this was particularly true for negative affect[26]. Indeed, studies in large samples of patients revealed intense and variable negative and positive affect in patients with psychotic disorder[13,14]. The slower recovery to a normative state possibly reflects inadequate emotion regulation strategies.

A recent ESM study suggested that patients with schizophrenia demonstrated adequate effort to control their emotions in their daily life, but these efforts were unsuccessful[27]. Another study in bipolar disorder examined coping mechanisms or response styles to both positive and negative mood.

Low mood predicted elevated rumination, which subsequently dampened mood further. High positive affect, on the other hand, predicted high-risk taking, which resulted in even higher positive affect, demonstrating a vicious cycle of escalating mood levels[28].

Capturing variability may not just be helpful in identifying a worse psychological state or an inadequate emotion regulation strategy. It may also predict future course of symptoms. Affective variability at baseline predicted the recurrence of depressive symptoms over a three-year period in remitted depressed patients[29]. Similarly, elevated temporal auto-correlation and variance, as well as increased correlation between emotions, increased the probability for an upcoming shift between a normal and a depressed state[30].

Identifying internal and situational determinants of psychopathology

Variability does not only pertain to affect. Most symptoms observed in patients with severe psychiatric disorder show meaningful and widespread variation over time. For example, intensity of visual and auditory hallucinations or delusions is highly variable over time[31,32]. Identifying what drives this variation, either internally or contextually, may be very helpful in detecting targets for treatment. At the same time, it may help patients to become aware of their own patterns of behaviour.

The longitudinal nature of ESM data makes it excellently suited to examine temporal associations between context, experience and behaviour. It has, for example, been shown that increases in paranoia are preceded by increases in anxiety, reductions in self-esteem and engagement in experiential avoidance[33-35]. Similarly, an ESM study into non-suicidal self-injury (NSSI) – the non-accidental damaging of one's own bodily tissue without suicidal intent[36] – found increased feelings of numbness and rejection to predict NSSI behaviour. These studies thus identify potential targets for treatment; improving self-esteem or diminishing feelings of rejection may help reduce levels of paranoia or NSSI behaviour.

ESM is not restricted to inner mental states as a possible predictor of symptoms. Contextual variables can also be taken into account. Collip et al[37] found paranoid thinking to be context-dependent in individuals with low or medium levels of trait paranoia. Paranoid thoughts increased when people were in the company of strangers. Yet, for those with high levels of trait paranoia, momentary paranoia became autonomous and independent of social reality. With respect to NSSI in adolescents, it was found that they were more often alone when they started thinking about NSSI, and being alone was also a significant predictor of engaging in NSSI[38]. Making people aware of these behavioural patterns may, again, be very instrumental for treatment.

Examining the consequences of events may also improve our understanding of real-life dynamics. In NSSI research, it has been shown that negative affect rises substantially prior to an episode of NSSI and then decreases directly after, underlining the mood-regulating function of the NSSI behaviour[39]. Similarly, an ESM study examining the association between cannabis use and psychosis found higher increases in both positive affect and psychotic symptoms following cannabis use in patients compared to controls, possibly explaining the vicious cycle of deleterious use in these patients[40].

Importantly, these patterns of association may substantially differ within persons. Our group examined the individual data of 64 persons with psychotic disorder and found clear inter-individual differences in the temporal order of mood and paranoia, with findings for each case deviating from the overall group findings[41]. ESM is highly attuned to individual patterns of associations, which may lead to person-tailored psychoeducation and identification of individual targets for treatment, thus providing opportunities for personalized medicine.

Zooming in on person-environment interaction: sensitivity to stress

An important putative psychological mechanism for many psychiatric disorders is sensitivity to minor stressors in daily life. While most epidemiological studies have focused only on the association between stressors and presence of psychopathology, ESM allows us to investigate the way an individual reacts to a stressor and how this is associated with psychopathology.

Several ESM studies have shown altered reactivity to daily-life stressors in patients with psychopathology. In major depressive disorder, patients showed increased negative affective reactivity to daily stressors compared to non-depressed individuals[42] and remitted patients[43]. Increased affective reactivity to stressors has also been reported across the psychosis continuum[44-50]. Interestingly, the most pronounced increases in stress sensitivity have been found in individuals with a clinical high risk state for psychosis[48,49,51], in part since psychotic symptoms may particularly add to the experience of stress in this group[51]. Moreover, stress is also associated with momentary increases in psychotic experiences in both patients and their family members[52].

As affective disturbances mediate the effect of daily hassles on psychotic experiences[53], an affective pathway to psychosis has been proposed[54]. In line with this, ESM studies have shown that exposure to social adversity, such as childhood trauma[50,55-59] and major life events[60,61], is associated with increased reactivity even to minor stressors in daily life, suggesting a process of behavioural sensitization[62]. In addition, these traumatic experiences were also linked to stronger psychotic reactivity to threat anticipation in daily life in individuals at the more severe end of the psychosis continuum[50]. Interestingly, mediation analyses revealed that the effect of daily stressors on psychotic experiences was not only mediated by affective reactivity, but also by threat anticipation in first-episode psychosis patients[53].

Increased reactivity to daily-life stressors also has prognostic value, as it has been found to predict persistence of psychotic

symptoms[63], and onset of depressive disorder[64] one year later in adults. One study even found an increased likelihood of reporting an affective disorder ten years later[65]. Reactivity to the smallest of daily hassles has been shown to be predictive of persisting psychopathology in adolescents and young adults[66].

As ESM provides multiple assessments over time within one individual, more complex analysis techniques have been adopted to study the temporal dynamics of stress, affect and symptoms. Using a network analysis, Klippel et al[67] showed that risk for psychosis is associated with changes in contextual networks where daily stressors have a central position and predict psychotic experiences while reducing physical activity. In depression, a network approach indicated an important role for the experience of social pressure[68]. Automatically generated personalized models that require intensive sampling have revealed different patterns of affect dynamics, including stress reactivity, between subclinically depressed individuals with and without anhedonic complaints[69].

Finally, an evolving field in ESM stress research incorporates biological measures. Previous studies have looked at associations between such measures and daily-life stress. For instance, structural changes in the pituitary and hippocampus have been directly associated with increased daily-life stress reactivity in psychosis[70,71]. Similarly, several studies were able to predict daily-life stress reactivity using functional neuroimaging[72-74]. More recent studies have embedded physiological monitoring in the ambulatory setting, measuring heart rate, blood pressure, cortisol and α-amylase in real life. Initial results indicated increased cardiovascular stress reactivity in posttraumatic stress disorder[75,76], blunted cortisol responsivity in depression[77] and psychosis[78], and increased cortisol reactivity in people at familial risk for psychosis[44]. In this evolving field, technological developments allow for novel passive monitoring approaches for continuous measurement of physiology, which can provide unique insight into the role of stress in the aetiology of mental disorders.

Examining real-world social interactions

ESM assessments also provide an opportunity for gaining more insight in activities and social interactions of people in daily life. One study found that patients with psychosis spend more time alone and at home, and are more often doing nothing compared to a healthy control sample[79]. This was also the case for patients meeting criteria for symptomatic recovery[80]: despite the reduction of symptoms, they were still more isolated and less engaged in goal-directed activities compared to healthy controls. Another study detected that individuals with psychosis set more pleasure-based and fewer effort-based goals, and, similarly, engage in more pleasurable and less effortful activities throughout their daily lives[18]. In a general population sample of women, it was found that avoiding social contact after appraising company more negatively increased the risk for the development of major depressive disorder in the following 20 months[81].

When comparing a standard social functioning questionnaire, the Social Functioning Scale (SFS), with ESM measures of social functioning, the SFS did show some degree of ecological validity for assessing the broad aspects of social functioning, but ESM measures offered a much more detailed and rich alternative[79]. This may be helpful for clinical practice, but can also further our theoretical knowledge. It has, for example, been shown that different social cognition tasks such as emotion recognition and theory of mind are not related to fine-grained measures of interaction in daily life[82,83]. Similarly, subjective quality of life measured in the moment was more consistently associated with affect, social interaction and activity compared to self-reported quality of life as assessed with a retrospective questionnaire[84].

Evaluation of treatment

As ESM provides a fine-grained picture of mental state and functioning, it may be much more sensitive to capturing change and, thereby, significantly improve assessment of outcomes in studies investigating therapeutic effects of biological, psychological and social interventions in psychiatry[85]. Moore et al[86], in a study evaluating the effect of a mindfulness intervention in depressed individuals, reported that ESM measures were much more sensitive to change, particularly for depressive symptoms and mindfulness, for which the number-needed-to-treat was 25 to 50% lower than for traditional questionnaire.

In addition, ESM expands outcome and process measures beyond the reach of conventional assessments. One study in patients with major depressive disorder found clear dose-response effects in increases of positive affect and enhanced responsiveness to pleasant daily-life activities over 18 weeks of antidepressant treatment[87]. This effect was also observed in patients participating in mindfulness-based cognitive therapy[88] and in cognitive behavioural therapy[89]. The latter group also endorsed increased resilience to stress in daily life. Another study in psychosis found different dimensions of delusions changing at different rates over time in response to antipsychotic treatment[90].

ESM's sensitivity to change may also allow earlier detection of side effects. One study investigating the association between the dosage of antipsychotic medication and affect in daily life found significant decreases in positive affect at a much lower medication dose than was needed for the occurrence of more pronounced extrapyramidal symptoms[91].

A more detailed baseline assessment, that can be achieved using ESM, may also improve prediction of treatment outcome. Forbes et al[92] found lower negative affect and higher positive affect at baseline to be predictive of better treatment response in children and adolescents with affective problems. There is also evidence of diminished emotional reactivity to be associated with a lower likelihood of recovery over a period of 18 months in adults suffering from major depressive disorder[93].

In addition, early processes of change can be identified that predict later outcome. Early changes in positive rather than

negative affect during the first week of an antidepressant treatment predicted treatment response at six weeks[94]. Furthermore, responding to treatment was associated with an increase in reward experience, suggesting that response to treatment in depression may be conditional on a recovery of hedonic capacity[95]. With an even more sophisticated analysis, the same authors found that stronger reductions in negative affect following a peak of positive affect during the day was associated with a more favourable course of depression[96].

Opportunity for new interventions

Over the last years, ESM has also been used to deliver treatment in real life. Ecological momentary interventions (EMIs) use mobile devices to deliver treatment in the daily life of patients, thus extending the therapy beyond the clinical setting and into daily life[97].

The content of these interventions is highly variable. Some are developed to augment face-to-face contacts with EMI components, such as the recently developed EMI Acceptance and Commitment in Daily Life, where therapeutic sessions are followed by three days of real-life exercises using an ESM app[98,99]. An example of a fully automated EMI is FOCUS, which has been specifically developed to provide automated real-time and real-world illness management support for psychosis[100,101].

Some EMIs integrate assessment of symptoms using ESM with a real-world delivery of treatment. PRISM (Personalized Real-time Intervention for Stabilizing Mood), for example, prompts individuals with bipolar disorder to fill out a survey on current context and mood state twice a day, which then triggers predefined and personalized action steps[102]. Another study in depression provided personalized feedback based on aggregated ESM information to increase awareness and induce behavioural change[103,104], thus reducing depressive symptoms over time.

Although the field of EMI is still in its early stages, recent systematic reviews suggest a high acceptance and feasibility in individuals with severe mental illness[97,105,106]. With regard to efficacy, there is only a limited amount of research available to date. Overall, the limited evidence supports the efficacy of EMIs in mental health[101,102,104,107-111]. Evidence seems to point towards greater efficacy when EMI is integrated with real-life assessment using ESM, preferentially tailoring the intervention to the specific needs of the individual as well as to those moments when intervention is most needed[97,112]. Evidence from exploratory and definitive randomized controlled trials is now required to further elucidate the efficacy and effectiveness of EMIs.

Summary

In summary, ESM has many advantages. It increases patient empowerment by identifying the individual as the expert of his/her experiences. Focusing on the micro-level dynamics of symptoms[113], it improves our understanding of their nature, their variability over time and their patterns of associations, both at group level and at the level of the individual. Furthermore, ESM can improve evaluation of treatment due to its sensitivity to capture change, as well as provide opportunities for the development of new interventions.

DESIGNS, METHODOLOGY AND STATISTICS

In this section, we provide updated information on the technical details of setting up an ESM study.

Designs

ESM uses event-contingent, time-contingent or hybrid designs.

In event-contingent designs, the sampling units are predefined events, i.e., ESM assessments are triggered by certain events such as panic attacks, social interactions or cannabis use. These designs offer the advantage of a comprehensive sample of the events under investigation. The disadvantage, however, is that it is sometimes difficult to define discrete events in a way that is readily accessible to, and recalled by, participants. It is also difficult to establish compliance, i.e., whether participants completed assessments for all predefined events. Finally, while a comprehensive sample of predefined events may be obtained (assuming optimal compliance), other relevant aspects of experience and behaviour may be missed. Therefore, most studies use time-contingent designs as an alternative approach.

In time-contingent designs, time is used as the sampling unit, i.e., participants are asked to complete a questionnaire contingent on time instead of an event. Sampling schedules of time-contingent designs can either be fixed or random.

A fixed sampling schedule requires participants to complete questionnaires at equal time intervals (e.g., every two hours). Although fixed sampling schedules ensure that target constructs are comprehensively assessed and allow for longitudinal statistical analyses, they have two major limitations. First, fixed sampling schedules may increase reactivity to the method, as participants know when they have to complete an assessment, which may lead them to adapt their daily routines. Second, they do not allow calculation of time budgets, as assessments at equal intervals are not necessarily representative of other moments that are not sampled.

A random sampling schedule signals the participant at random intervals during the day, making use of random blocks to ensure that time is evenly sampled across the day. A significant advantage of random sampling schedules is that they provide a representative sample of the target construct. They significantly decrease reactivity to the method, as participants do not know when the next signal will occur. Also, as random sampling schedules allow for calculating time budgets, they

offer the advantage of providing estimates of the average time that people spend in certain contexts.

Hybrid designs combine event-contingent and time-contingent designs. For example, measuring mood using a time-contingent design with a blocked random time schedule, and substance use as a discrete event in an event-contingent design, or time-contingent designs with different (fixed or random) time schedules.

Sampling frequency and period determines the resolution required for assessing the target construct. This primarily depends on the known or expected variability of the target construct over time. Many experience sampling studies in psychiatry conducted to date have used a sampling frequency and period of ten assessments per day over a period of six consecutive days, given the resolution required for assessing highly variable constructs (e.g., mood). However, this needs to be considered in detail for each individual study and balanced with assessment burden for participants.

ESM designs were, for a long time, implemented using pen and paper diaries to assess target constructs and wristwatches to implement time-contingent designs. After a brief period of using personal digital assistants or other handheld devices, most, if not all, studies now use smartphone applications for implementing ESM designs[114].

Questionnaire development

The development of questionnaires is an important aspect of ESM research, as targeting momentary experiences is very different from the global and retrospective approach in cross-sectional questionnaires. ESM questionnaires, therefore, follow their own rules and logic. Overall, completing a questionnaire should not take longer than 2 minutes. Questionnaires contain on average 30-60 items, depending on the item format. Including more items may minimize reactivity to the method by diverting attention away from specific items of interest, but increases burden. On the other hand, when the explicit goal is to make people aware of their patterns of behaviour, such as in clinical therapy, it may be helpful to reduce the number of items.

ESM questionnaires in psychiatry often include questions on current mental states (e.g., thoughts, mood and symptoms), behaviour, context, and appraisals of these contexts. Items are preferentially presented in this order, moving from more to less transient items. Sometimes "between moment" questions are included at the end of the questionnaire, inquiring about the time between the previous and the current report. Although it may be of interest to include ratings of these in-between periods (e.g., drug use or daily hassles), these questions should be limited, as they may again be subject to recall bias, even if minimal, and do not directly reflect interactions in context.

When developing a questionnaire, the first imperative is that the questions should inquire about momentary states. This may seem obvious, yet is not always easy to achieve. For example, including "right now" before a global statement is not transferring this statement into an item of a momentary state. "Right now, I have a number of good qualities" remains a global statement and, therefore, will result in little variation over time.

Another imperative is that the language should reflect how people think about and describe their own behaviour and experience. Lexicon commonly used by professionals, such as attribution, coping or dissociation, is best avoided. ESM questionnaires could still substantially improve in this respect. We are currently conducting focus groups with people with experience of psychosis to more accurately grasp their actual experiences, and thus improve the assessment of psychosis using ESM, but also the way we assess their social interactions.

Third, it is conceptually important to focus on processes that pertain to common situations in daily life. "Did you initiate the contact?" may be informative in very specific situations but, when you are in the kitchen having breakfast with your husband, it is difficult to answer.

Finally, ESM optimally aims to capture patterns of behaviour that people are not necessarily aware of. Therefore, it is important to avoid using reflective questions, such as "How do you feel in this company?", but rather to inquire purely about the momentary states, e.g. "How do you feel right now?" and "With whom are you right now?", which can later be correlated, either cross-sectionally or over time, to establish behavioural patterns.

Statistical approaches

ESM studies typically yield a substantial amount of data. A study with 100 subjects, each assessed ten times per day over the course of six days using a random time sampling schedule, yields a dataset with 6,000 rows of data, where each row corresponds to a particular assessment moment for a given subject. Data of this type can be used at a more descriptive level, e.g. to estimate the mean level of a particular variable and its variability within and between study participants, or to study group differences, e.g. whether patients report on average elevated levels of negative affect compared to controls. However, the full strength of ESM comes into play once we start to examine the within-person relationship between some outcome of interest (e.g., negative affect) and some time-varying predictor (e.g., stress) and how the strength of such a relationship may differ across groups (e.g., whether the relationship between negative affect and stress differs for patients versus controls).

Although an ESM study is essentially a repeated measures design, classical analysis procedures such as repeated measures analysis of variance (ANOVA) or multivariate analysis of variance (MANOVA) are not directly applicable, as they cannot easily handle the complexities involved (e.g., missing data, unequally spaced time points, time-varying covariates, auto-correlated observations). Instead, given that ESM data adhere to a multilevel structure (with repeated assessments nested within days, which in turn are nested within subjects), multilevel/mixed-effects models are typically the method of choice for their analysis[115].

Mixed-effects models extend the standard regression model by allowing for the inclusion of additional "random effects", which can be used to account for person- and day-level differences in the model coefficients (i.e., intercepts and slopes). Such models also allow us to disaggregate between- and within-person relationships[116]. For example, the degree to which those who experience higher levels of stress overall also tend to report higher levels of negative affect (between-person) may be quite different from the degree to which negative affect varies within a subject in relation to the perceived stressfulness of particular situations (within-person).

Standard mixed-effects models used for the analysis of ESM data[117] can be extended in various ways: for example, by allowing mixture distributions for the random effects[118], by adding predictors for the amount of within- and between-subject variability[119], and by allowing model coefficients to change smoothly over time using splines[120,121]. Vector autoregressive models using ESM data of single subjects[122] or groups of subjects[123] can provide further insights into the dynamics of psychopathology from a network perspective[124]. Recently, the use of mixture latent Markov models has also been suggested as an alternative approach to analysing ESM data[125]. Given the increasing use of ESM in research, we expect to see further developments in the analysis approaches in the coming years.

CONSIDERATIONS AND FUTURE PROSPECTS

While there has been a rapidly increasing number of studies using ESM to investigate highly important areas of research in psychiatry, it is important to take stock and critically appraise what has been achieved in order to move on and build upon what can be gleaned from studies using this method.

One important consideration is the replicability and consistency of findings that have emerged from previous ESM studies. Most experience sampling studies to date have focused on novel questions using different designs and questionnaires, resulting in heterogeneity in terms of conceptual definitions as well as findings across different populations. Therefore, as in many other areas of psychiatric research, there is a need for direct replications of findings for the same study population. In order to achieve this, there is a pressing need for greater consistency in the definition and operationalization of target constructs, as well as robust psychometric research on the structural validity of ESM measures of key constructs as a basis for deriving composite scores and reducing heterogeneity in findings. Similar to other areas of research where small samples are common, such as neuroimaging research, careful sampling of participants is required to minimize selection bias.

One way of moving the field forward is to develop questionnaires with good psychometric properties (e.g., the Maastricht Mood Questionnaire[126]) that can be used in ESM studies. Additionally, more methodological research is needed providing robust evidence on design issues for key constructs to achieve standardization and replicability. This will enable ESM researchers to work in larger networks and consortia, as has been the case in other areas of psychiatric research[127], to generate consistent and generalizable findings across countries.

A further consideration is that ESM data collection is time-intensive and may be associated with assessment burden for participants, which raises the question of whether this method can be used in all populations and, in particular, in vulnerable populations. However, there is strong evidence in support of the feasibility of using ESM in vulnerable populations, including individuals with (severe) mental health problems[3], which may be due to the nature of ESM as a structured inquiry about current mental states with clear ecological appeal. However, further developments are needed for the use of ESM in children and in older populations, e.g. people in the early stages of dementia. An interesting prospect here is the development of dyadic approaches, combining self-report within one person with observational and context data provided by an informant.

Furthermore, numerous researchers have raised the question of whether being repeatedly asked about particular thoughts and behaviours may, in fact, induce those thoughts and behaviours or may cause participants to alter their behaviour[3,5,128]. Whilst participants may feel positively about repeated questioning during ESM, this may nevertheless result in them consciously or unconsciously altering their behaviour. Measurement reactivity is a key challenge for ESM research, yet remains an under-researched phenomenon[129]. However, as outlined above, there are ways of minimizing reactivity by selecting appropriate ESM designs and measures.

Finally, most findings to date relate to evidence of associations based on cross-sectional modelling of ESM data. This now needs to be probed further to investigate the strength of the evidence in support of other important criteria for establishing causality, such as temporal order or experimental evidence using ecological interventionist causal models[130]. Also, the further development and implementation of new statistical techniques is crucial here. As ESM is collecting a large amount of data from each individual person, this may be linked to other sources, such as big data, to monitor and recognize an individual's state[108] and context to eventually provide person-tailored contextualized interventions (although this may require an even larger number of observations than are usually collected).

There is further immense emerging potential for combining ESM with physical remote monitoring technologies. Combining ESM with wearables assessing, for example, physical activity, heart rate variability or sleep may provide even richer and more detailed insights at various levels of causality (biological, psychological, social)[131,132]. Another step is to include data from context-aware systems using sensor data that automatically provide input on relevant context variables[108]. A large ongoing study in Europe, RADAR-CNS (https://www.radar-cns.org), set up to examine the relevance of both active and passive remote monitoring approaches in predicting and understanding the clin-

ical course of central nervous system disorders, including depression, may be an important step in this respect.

In addition, one of the most important, but also most challenging, next steps is to bridge the gap between research and clinical care that would allow the implementation of ESM in routine monitoring and outcome measurement in mental health services. ESM has enormous potential to contribute to, and improve upon, clinical care. Yet, to date, it has hardly been implemented, due to issues related to data safety, data ownership, privacy and consent, access to technology, as well as integration of data management systems across mental health services. More implementation initiatives are needed to bridge this gap.

CONCLUSIONS

In summary, we have shown that ESM is an indispensable methodology in psychiatry research. It adds new insights and additional perspectives to standard approaches, enriches our understanding of psychopathological phenomena and their associated mechanisms, and offers clear opportunities for improving and changing clinical practice.

A number of considerations and challenges remain and, with the growing body of research in this field, there is a pressing need for methodological advances. However, as using ESM creates the possibility to study and analyze temporal associations in everyday social contexts, as well as tailor treatment to individual needs, it offers one of the best opportunities for personalized medicine in psychiatry, from both a research and a clinical perspective.

ACKNOWLEDGEMENTS

The second and third author contributed equally to this paper. This work was supported by a Research Foundation Flanders (FWO) Odysseus grant (G0F8416N) and a European Research Council consolidator grant (ERC-2012-StG, project 309767 - INTERACT) to I. Myin-Germeys, and a Veni grant from the Netherlands Organization for Scientific Research (grant no. 451-13-022) to U. Reininghaus.

REFERENCES

1. Mehl MR, Conner TS (eds). Handbook of research methods for studying daily life. New York: Guilford, 2012.
2. Hektner JM, Schmidt JA, Csikszentmihalyi M. Experience sampling method: measuring the quality of everyday life. Thousand Oaks: Sage, 2007.
3. Myin-Germeys I, Oorschot M, Collip D et al. Experience sampling research in psychopathology: opening the black box of daily life. Psychol Med 2009;39:1533-47.
4. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. Ann Behav Med 1994;16:199-202.
5. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32.
6. Barker RG. Ecological psychology; concepts and methods for studying the environment of human behavior. Stanford: Stanford University Press, 1968.
7. Shapiro L (ed). The Routledge handbook of embodied cognition. New York: Routledge, 2014.
8. Myin E, Van Eemeren J. Embodied and embedded cognition. In: Stam H, Hooren H, de Jong H (eds). The Sage handbook of theoretical psychology. Thousand Oaks: Sage (in press).
9. Ben-Zeev D, Young MA. Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports: an experience sampling study. J Nerv Ment Dis 2010;198:280-5.
10. Ben-Zeev D, McHugo GJ, Xie H et al. Comparing retrospective reports to real-time/real-place mobile assessments in individuals with schizophrenia and a nonclinical comparison group. Schizophr Bull 2012;38:396-404.
11. Blum LH, Vakhrusheva J, Saperstein A et al. Depressed mood in individuals with schizophrenia: a comparison of retrospective and real-time measures. Psychiatry Res 2015;227:318-23.
12. Strauss GP, Cohen AS. A transdiagnostic review of negative symptom phenomenology and etiology. Schizophr Bull 2017;43:712-9.
13. Oorschot M, Lataster T, Thewissen V et al. Emotional experience in negative symptoms of schizophrenia – no evidence for a generalized hedonic deficit. Schizophr Bull 2013;39:217-25.
14. Myin-Germeys I, Delespaul PA, deVries MW. Schizophrenia patients are more emotionally active than is assumed based on their behavior. Schizophr Bull 2000;26:847-54.
15. Heininga VE, van Roekel E, Wichers M et al. Reward and punishment learning in daily life: a replication study. PLoS One 2017;12:e0180753.
16. Heininga VE, Van Roekel E, Ahles JJ et al. Positive affective functioning in anhedonic individuals' daily life. J Affect Disord 2017;218:437-45.
17. van Roekel E, Bennik EC, Bastiaansen JA et al. Depressive symptoms and the experience of pleasure in daily life: an exploration of associations in early and late adolescence. J Abnorm Child Psychol 2016;44:999-1009.
18. Gard DE, Kring AM, Gard MG et al. Anhedonia in schizophrenia: distinctions between anticipatory and consummatory pleasure. Schizophr Res 2007;93:253-60.
19. Brown LH, Silvia PJ, Myin-Germeys I et al. When the need to belong goes wrong: the expression of social anhedonia and social anxiety in daily life. Psychol Sci 2007;18:778-82.
20. Wichers M, Kasanova Z, Bakker J et al. From affective experience to motivated action: tracking reward-seeking and punishment-avoidant behaviour in real-life. PLoS One 2015;10:e0129722.
21. Moran EK, Culbreth AJ, Barch DM. Ecological momentary assessment of negative symptoms in schizophrenia: relationships to effort-based decision making and reinforcement learning. J Abnorm Psychol 2017;126:96-105.
22. Kasanova Z, Ceccarini J, Frank M et al. Intact striatal dopaminergic modulation of reward learning and daily-life reward-oriented behaviour in relatives of individuals with psychotic disorder. Psychol Med 2017;13:1-6.
23. Kasanova Z, Ceccarini J, Frank MJ et al. Striatal dopaminergic modulation of reinforcement learning predicts reward-oriented behavior in daily life. Biol Psychol 2017;127:1-9.
24. Aldao A. The future of emotion regulation research: capturing context. Perspect Psychol Sci 2013;8:155-72.
25. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in Emotion Regulation Scale. J Psychopathol Behav Assess 2004;26:41-54.
26. Houben M, van den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: a meta-analysis. Psychol Bull 2015;41:901-30.
27. Visser KF, Esfahlani FZ, Sayama H et al. An ecological momentary assessment evaluation of emotion regulation abnormalities in schizophrenia. Psychol Med (in press).
28. Pavlickova H, Varese F, Smith A et al. The dynamics of mood and coping in bipolar disorder: longitudinal investigations of the inter-relationship between affect, self-esteem and response styles. PLoS One 2013;8:e62514.
29. Timm C, Ubl B, Zamoscik V et al. Cognitive and affective trait and state factors influencing the long-term symptom course in remitted depressed patients. PLoS One 2017;12:e0178759.
30. van de Leemput IA, Wichers M, Cramer AOJ et al. Critical slowing down as early warning for the onset and termination of depression. Proc Natl Acad Sci USA 2014;111:87-92.
31. Oorschot M, Lataster T, Thewissen V et al. Temporal dynamics of visual and auditory hallucinations in psychosis. Schizophr Res 2012;140:77-82.
32. Thewissen V, Bentall RP, Lecomte T et al. Fluctuations in self-esteem and paranoia in the context of daily life. J Abnorm Psychol 2008;117:143-53.
33. Thewissen V, Bentall RP, Oorschot M et al. Emotions, self-esteem, and paranoid episodes: an experience sampling study. Br J Clin Psychol 2011; 50:178-95.

34. Udachina A, Thewissen V, Myin-Germeys I et al. Understanding the relationships between self-esteem, experiential avoidance, and paranoia: structural equation modelling and experience sampling studies. J Nerv Ment Dis 2009;197:661-8.
35. Udachina A, Varese F, Myin-Germeys I et al. The role of experiential avoidance in paranoid delusions: an experience sampling study. Br J Clin Psychol 2014;53:422-32.
36. Nock M, Favazza A. Non-suicidal self-injury: definition and classification. In: Nock MK (ed). Understanding non-suicidal self-injury: origins, assessment, and treatment. Washington: American Psychological Association, 2009:9-18.
37. Collip D, Oorschot M, Thewissen V et al. Social world interactions: how company connects to paranoia. Psychol Med 2011;41:911-21.
38. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: a real-time ecological assessment study among adolescents and young adults. J Abnorm Psychol 2009;118: 816-27.
39. Armey MF, Crowther JH, Miller IW. Changes in ecological momentary assessment reported affect associated with episodes of nonsuicidal self-injury. Behav Ther 2011;42:579-88.
40. Henquet C, van Os J, Kuepper R et al. Psychosis reactivity to cannabis use in daily life: an experience sampling study. Br J Psychiatry 2010;196:447-53.
41. Oorschot M, Lataster T, Thewissen V et al. Mobile assessment in schizophrenia: a data-driven momentary approach. Schizophr Bull 2012;38: 405-13.
42. Myin-Germeys I, Peeters F, Havermans R et al. Emotional reactivity to daily life stress in psychosis and affective disorder: an experience sampling study. Acta Psychiatr Scand 2003;107:124-31.
43. van Winkel M, Nicolson NA, Wichers M et al. Daily life stress reactivity in remitted versus non-remitted depressed individuals. Eur Psychiatry 2015; 30:441-7.
44. Collip D, Nicolson NA, Lardinois M et al. Daily cortisol, stress reactivity and psychotic experiences in individuals at above average genetic risk for psychosis. Psychol Med 2011;41:2305-15.
45. Lataster T, Collip D, Lardinois M et al. Evidence for a familial correlation between increased reactivity to stress and positive psychotic symptoms: stress-reactivity and psychotic symptoms. Acta Psychiatr Scand 2010;122: 395-404.
46. Lataster T, Wichers M, Jacobs N et al. Does reactivity to stress cosegregate with subclinical psychosis? A general population twin study. Acta Psychiatr Scand 2009;119:45-53.
47. Myin-Germeys I, Van Os J, Schwartz J et al. Emotional reactivity to daily life stress in psychosis. Arch Gen Psychiatry 2001;58:1137-44.
48. Palmier-Claus JE, Dunn G, Lewis SW. Emotional and symptomatic reactivity to stress in individuals at ultra-high risk of developing psychosis. Psychol Med 2012;42:1003-12.
49. Reininghaus U, Kempton MJ, Valmaggia L et al. Stress sensitivity, aberrant salience, and threat anticipation in early psychosis: an experience sampling study. Schizophr Bull 2016;42:712-22.
50. Reininghaus U, Gayer-Anderson C, Valmaggia L et al. Psychological processes underlying the association between childhood trauma and psychosis in daily life: an experience sampling study. Psychol Med 2016;46: 2799-813.
51. van der Steen Y, Gimpel-Drees J, Lataster T et al. Clinical high risk for psychosis: the association between momentary stress, affective and psychotic symptoms. Acta Psychiatr Scand 2017;136:63-73.
52. Myin-Germeys I, Delespaul P, Van Os J. Behavioural sensitization to daily life stress in psychosis. Psychol Med 2005;35:733-41.
53. Klippel A, Myin-Germeys I, Chavez-Baldini U et al. Modeling the interplay between psychological processes and adverse, stressful contexts and experiences in pathways to psychosis: an experience sampling study. Schizophr Bull 2017;43:302-15.
54. Myin-Germeys I, van Os J. Stress-reactivity in psychosis: evidence for an affective pathway to psychosis. Clin Psychol Rev 2007;27:409-24.
55. Cristóbal-Narváez P, Sheinbaum T, Ballespí S et al. Impact of adverse childhood experiences on psychotic-like symptoms and stress reactivity in daily life in nonclinical young adults. PLoS One 2016;11:e0153557.
56. Glaser JP, van Os J, Portegijs PJM et al. Childhood trauma and emotional reactivity to daily life stress in adult frequent attenders of general practitioners. J Psychosom Res 2006;61:229-36.
57. Lardinois M, Lataster T, Mengelers R et al. Childhood trauma and increased stress sensitivity in psychosis. Acta Psychiatr Scand 2011;123:28-35.
58. Rauschenberg C, van Os J, Cremers D et al. Stress sensitivity as a putative mechanism linking childhood trauma and psychopathology in youth's daily life. Acta Psychiatr Scand 2017;136:373-88.
59. van Nierop M, Lecei A, Myin-Germeys I et al. Stress reactivity links childhood trauma exposure to an admixture of depressive, anxiety, and psychosis symptoms. Psychiatry Res 2018;260:451-7.
60. Myin-Germeys I, Krabbendam L, Delespaul PA et al. Do life events have their effect on psychosis by influencing the emotional reactivity to daily life stress? Psychol Med 2003;33:327-33.
61. Wichers M, Schrijvers D, Geschwind N et al. Mechanisms of gene-environment interactions in depression: evidence that genes potentiate multiple sources of adversity. Psychol Med 2009;39:1077-86.
62. Collip D, Myin-Germeys I, Van Os J. Does the concept of "sensitization" provide a plausible mechanism for the putative link between the environment and schizophrenia? Schizophr Bull 2008;34:220-5.
63. Collip D, Wigman JTW, Myin-Germeys I et al. From epidemiology to daily life: linking daily life stress reactivity to persistence of psychotic experiences in a longitudinal general population study. PLoS One 2013;8:e62688.
64. Wichers M, Geschwind N, Jacobs N et al. Transition from stress sensitivity to a depressive state: longitudinal twin study. Br J Psychiatry 2009;195: 498-503.
65. Charles ST, Piazza JR, Mogle J et al. The wear and tear of daily stressors on mental health. Psychol Sci 2013;24:733-41.
66. Vaessen T, van Nierop M, Decoster J et al. Is sensitivity to daily stress predictive of onset or persistence of psychopathology? Eur Psychiatry 2017; 45:167-73.
67. Klippel A, Viechtbauer W, Reininghaus U et al. The cascade of stress: a network approach to explore differential dynamics in populations varying in risk for psychosis. Schizophr Bull 2018;44:328-37.
68. Dejonckheere E, Bastian B, Fried EI et al. Perceiving social pressure not to feel negative predicts depressive symptoms in daily life. Depress Anxiety 2017;34:836-44.
69. Bos FM, Blaauw FJ, Snippe E et al. Exploring the emotional dynamics of subclinically depressed individuals with and without anhedonia: an experience sampling study. J Affect Disord 2018;228:186-93.
70. Collip D, Habets P, Marcelis M et al. Hippocampal volume as marker of daily life stress sensitivity in psychosis. Psychol Med 2013;43:1377-87.
71. Habets P, Collip D, Myin-Germeys I et al. Pituitary volume, stress reactivity and genetic risk for psychotic disorder. Psychol Med 2012;42:1523-33.
72. Fischer S, Breithaupt L, Wonderlich J et al. Impact of the neural correlates of stress and cue reactivity on stress related binge eating in the natural environment. J Psychiatr Res 2017;92:15-23.
73. Hernaus D, Collip D, Lataster J et al. Psychotic reactivity to daily life stress and the dopamine system: a study combining experience sampling and [18F]fallypride positron emission tomography. J Abnorm Psychol 2015; 124:27-37.
74. Tully LM, Lincoln SH, Hooker CI. Lateral prefrontal cortex activity during cognitive control of emotion predicts response to social stress in schizophrenia. Neuroimage Clin 2014;6:43-53.
75. Buckley TC, Holohan D, Greif JL et al. Twenty-four-hour ambulatory assessment of heart rate and blood pressure in chronic PTSD and non-PTSD veterans. J Trauma Stress 2004;17:163-71.
76. Edmondson D, Sumner JA, Kronish IM et al. The association of posttraumatic stress disorder with clinic and ambulatory blood pressure in healthy adults. Psychosom Med 2018;80:55-61.
77. Peeters F, Nicholson NA, Berkhof J. Cortisol responses to daily events in major depressive disorder. Psychosom Med 2003;65:836-41.
78. Vaessen T, Kasanova Z, Hernaus D et al. Cortisol reactivity to daily-life stressors in psychosis. Submitted for publication.
79. Schneider M, Reininghaus U, van Nierop M et al. Does the Social Functioning Scale reflect real-life social functioning? An experience sampling study in patients with a non-affective psychotic disorder and healthy control individuals. Psychol Med 2017;47:2777-86.
80. Oorschot M, Lataster T, Thewissen V et al. Symptomatic remission in psychosis and real-life functioning. Br J Psychiatry 2012;201:215-20.
81. van Winkel M, Wichers M, Collip D et al. Unraveling the role of loneliness in depression: the relationship between daily life experience and behavior. Psychiatry 2017;80:104-17.
82. Janssens M, Lataster T, Simons CJP et al. Emotion recognition in psychosis: no evidence for an association with real world social functioning. Schizophr Res 2012;142:116-21.
83. Schneider M, Myin E, Myin-Germeys I. Is social cognition a prerequisite for social interaction? A study in psychotic disorder. Submitted for publication.

84. Leendertse P, Myin-Germeys I, Lataster T et al. Subjective quality of life in psychosis: evidence for an association with real world functioning? Psychiatry Res 2018;261:116-23.
85. Myin-Germeys I, Birchwood M, Kwapil T. From environment to therapy in psychosis: a real-world momentary assessment approach. Schizophr Bull 2011;37:244-7.
86. Moore RC, Depp CA, Wetherell JL et al. Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. J Psychiatr Res 2016;75:116-23.
87. van Os J, Delespaul P, Barge D et al. Testing an mhealth momentary assessment routine outcome monitoring application: a focus on restoration of daily life positive mood states. PLoS One 2014;9:e115254.
88. Geschwind N, Peeters F, Drukker M et al. Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: a randomized controlled trial. J Consult Clin Psychol 2011;79:618-28.
89. Eddington KM, Burgin CJ, Silvia PJ et al. The effects of psychotherapy for major depressive disorder on daily mood and functioning: a longitudinal experience sampling study. Cogn Ther Res 2017;41:266-77.
90. So SH, Peters ER, Swendsen J et al. Changes in delusions in the early phase of antipsychotic treatment – an experience sampling study. Psychiatry Res 2014;215:568-73.
91. Lataster J, Thewissen V, Bak M et al. Emotional experience and estimates of D2 receptor occupancy in psychotic patients treated with haloperidol, risperidone, or olanzapine: an experience sampling study. J Clin Psychiatry 2011;72:1397-404.
92. Forbes EE, Stepp SD, Dahl RE et al. Real-world affect and social context as predictors of treatment response in child and adolescent depression and anxiety: an ecological momentary assessment study. J Child Adolesc Psychopharmacol 2012;22:37-47.
93. Peeters F, Berkhof J, Rottenberg J et al. Ambulatory emotional reactivity to negative daily life events predicts remission from major depressive disorder. Behav Res Ther 2010;48:754-60.
94. Geschwind N, Nicolson NA, Peeters F et al. Early improvement in positive rather than negative emotion predicts remission from depression after pharmacotherapy. Eur Neuropsychopharmacol 2011;21:241-7.
95. Wichers MC, Barge-Schaapveld D, Nicolson NA et al. Reduced stress-sensitivity or increased reward experience: the psychological mechanism of response to antidepressant medication. Neuropsychopharmacology 2009;34:923-31.
96. Wichers M, Lothmann C, Simons CJP et al. The dynamic interplay between negative and positive emotions in daily life predicts response to treatment in depression: a momentary assessment study: emotional dynamics and future treatment response. Br J Clin Psychol 2012;51:206-22.
97. Myin-Germeys I, Klippel A, Steinhart H et al. Ecological momentary interventions in psychiatry. Curr Opin Psychiatry 2016;29:258-63.
98. Batink T, Bakker J, Vaessen T et al. Acceptance and commitment therapy in daily life training: a feasibility study of an mhealth intervention. JMIR MHealth UHealth 2016;4:e103.
99. Steinhart H, Vaessen T, Batink T et al. ACT in daily life: a momentary intervention approach. Submitted for publication.
100. Ben-Zeev D, Kaiser SM, Brenner CJ et al. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. Psychiatr Rehabil J 2013;36:289.
101. Ben-Zeev D, Brenner CJ, Begale M et al. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014;40:1244-53.
102. Depp CA, Ceglowski J, Wang VC et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015;174:23-30.
103. Snippe E, Simons CJP, Hartmann JA et al. Change in daily life behaviors and depression: within-person and between-person associations. Health Psychol 2016;35:433-41.
104. Kramer I, Simons CJ, Hartmann JA et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. World Psychiatry 2014;13:68-77.
105. Menon V, Rajan TM, Sarkar S. Psychotherapeutic applications of mobile phone-based technologies: a systematic review of current research and trends. Indian J Psychol Med 2017;39:4-11.
106. Naslund JA, Marsch LA, McHugo GJ et al. Emerging mhealth and ehealth interventions for serious mental illness: a review of the literature. J Ment Health 2015;24:321-32.
107. Wenze SJ, Armey MF, Miller IW. Feasibility and acceptability of a mobile intervention to improve treatment adherence in bipolar disorder: a pilot study. Behav Modif 2014;38:497-515.
108. Burns MN, Begale M, Duffecy J et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13:e55.
109. Roepke AM, Jaffee SR, Riffle OM et al. Randomized controlled trial of SuperBetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015;4:235-46.
110. Watts S, Mackenzie A, Thomas C et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013;13:49.
111. Granholm E, Ben-Zeev D, Link PC et al. Mobile assessment and treatment for schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. Schizophr Bull 2012;38:414-25.
112. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010;15:1-39.
113. Wichers M, Hartmann JA, Kramer IMA et al. Translating assessments of the film of daily life into person-tailored feedback interventions in depression. Acta Psychiatr Scand 2011;123:402-3.
114. Conner TS. Experience sampling and ecological momentary assessment with mobile phones. http://www.otago.ac.nz/psychology/otago047475.
115. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003;54:579-616.
116. Wang LP, Maxwell SE. On disaggregating between-person and within-person effects with longitudinal data using multilevel models. Psychol Methods 2015;20:63-83.
117. Bolger N, Laurenceau J. Intensive longitudinal methods: an introduction to diary and experience sampling research. New York: Guilford, 2013.
118. Verbeke G, Lesaffre E. A linear mixed-effects model with heterogeneity in the random-effects population. J Am Stat Assoc 1996;91:217-21.
119. Hedeker D, Mermelstein RJ, Demirtas H. Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. Stat Med 2012;31:3328-36.
120. Tan X, Shiyko MP, Li R et al. A time-varying effect model for intensive longitudinal data. Psychol Methods 2012;17:61-77.
121. Dziak JJ, Li R, Tan X et al. Modeling intensive longitudinal data with mixtures of nonparametric trajectories and time-varying effects. Psychol Methods 2015;20:444-69.
122. Bak M, Drukker M, Hasmi L et al. An n=1 clinical network analysis of symptoms and treatment in psychosis. PLoS One 2016;11:e0162811.
123. Bringmann LF, Vissers N, Wichers M et al. A network approach to psychopathology: new insights into clinical longitudinal data. PLoS One 2013;8:e60188.
124. Borsboom D. A network theory of mental disorders. World Psychiatry 2017;16:5-13.
125. Crayen C, Eid M, Lischetzke T et al. Exploring dynamics in mood regulation – mixture latent Markov modeling of ambulatory assessment data. Psychosom Med 2012;74:366-76.
126. Viechtbauer W, Lataster T, Rintala A et al. Evidence for a two-factor positive and negative affect structure in daily life. Submitted for publication.
127. O'Donovan MC. What have we learned from the Psychiatric Genomics Consortium. World Psychiatry 2015;14:291-3.
128. Palmier-Claus JE, Myin-Germeys I, Barkus E et al. Experience sampling research in individuals with mental illness: reflections and guidance. Acta Psychiatr Scand 2011;123:12-20.
129. Wray TB, Merrill JE, Monti PM. Using ecological momentary assessment (EMA) to assess situation-level predictors of alcohol use and alcohol-related consequences. Alcohol Res Curr Rev 2014;36:19.
130. Reininghaus U, Depp CA, Myin-Germeys I. Ecological interventionist causal models in psychosis: targeting psychological mechanisms in daily life. Schizophr Bull 2016;42:264-9.
131. Krieger N. A glossary for social epidemiology. J Epidemiol Community Health 2001;55:693-700.
132. Kendler KS. Explanatory models for psychiatric illness. Am J Psychiatry 2008;165:695-702.

DOI:10.1002/wps.20513

Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry

Patrick D. McGorry, Jessica A. Hartmann, Rachael Spooner, Barnaby Nelson

Orygen, The National Centre of Excellence in Youth Mental Health, and Centre for Youth Mental Health, University of Melbourne, Parkville, Australia

The "at risk mental state" for psychosis approach has been a catalytic, highly productive research paradigm over the last 25 years. In this paper we review that paradigm and summarize its key lessons, which include the valence of this phenotype for future psychosis outcomes, but also for comorbid, persistent or incident non-psychotic disorders; and the evidence that onset of psychotic disorder can at least be delayed in ultra high risk (UHR) patients, and that some full-threshold psychotic disorder may emerge from risk states not captured by UHR criteria. The paradigm has also illuminated risk factors and mechanisms involved in psychosis onset. However, findings from this and related paradigms indicate the need to develop new identification and diagnostic strategies. These findings include the high prevalence and impact of mental disorders in young people, the limitations of current diagnostic systems and risk identification approaches, the diffuse and unstable symptom patterns in early stages, and their pluripotent, transdiagnostic trajectories. The approach we have recently adopted has been guided by the clinical staging model and adapts the original "at risk mental state" approach to encompass a broader range of inputs and output target syndromes. This approach is supported by a number of novel modelling and prediction strategies that acknowledge and reflect the dynamic nature of psychopathology, such as dynamical systems theory, network theory, and joint modelling. Importantly, a broader transdiagnostic approach and enhancing specific prediction (profiling or increasing precision) can be achieved concurrently. A holistic strategy can be developed that applies these new prediction approaches, as well as machine learning and iterative probabilistic multimodal models, to a blend of subjective psychological data, physical disturbances (e.g., EEG measures) and biomarkers (e.g., neuroinflammation, neural network abnormalities) acquired through fine-grained sequential or longitudinal assessments. This strategy could ultimately enhance our understanding and ability to predict the onset, early course and evolution of mental ill health, further opening pathways for preventive interventions.

Key words: At risk mental state, psychosis, ultra high risk, transition, transdiagnostic psychiatry, clinical staging, CHARMS, prediction strategies, network theory, dynamical systems theory, joint modelling

(World Psychiatry 2018;17:133–142)

Traditional approaches to psychiatric diagnosis have struggled to guide the care of patients and to illuminate the causes and mechanisms underlying mental ill health. Consequently and appropriately they are under constant critique. There has been very little innovation in over a century in how we conceptualize and classify mental illness, and what has passed for advances really only represent efforts to buttress a flawed paradigm.

How can we transcend a century of stagnation to pave the way for more effective mental health care which makes sense to clinicians, researchers and the public? A quarter of a century after the formulation of the concept of the "at risk mental state", we could be on the cusp of transforming how we approach the challenge of defining and treating mental illnesses. In this paper we discuss how this transformational concept can point the way to a radical rethink with greater clarity, utility and validity.

"AT RISK MENTAL STATES": ORIGINS AND BALANCED REVIEW

How did the at risk/clinical high risk/ultra high risk (UHR) concept originate, and what was the strategic intent behind it?

It had been well known for over a century that severe forms of mental disorder, notably schizophrenia, are typically preceded by a relatively non-specific period of symptoms, which are subthreshold in nature and of insufficient severity and clarity to justify a diagnosis. Seen through the deterministic lens of 19th century nosology, the term "prodrome", with its sense of inevitable progression, seemed to capture the concept well.

However, if a preventive approach to treatment of potentially serious mental disorders was going to be developed, this deterministic and fatalistic mindset had to change. The deadly nexus between diagnosis and prognosis within the concept of schizophrenia had to be severed or dramatically loosened. Prognosis had to be regarded as something that was malleable, and recovery as something possible. This objective was behind the decision to widen the focus to early psychosis and include the full spectrum of psychotic disorders, remaining agnostic about the future evolution of disorder[1,2].

This approach made further sense because so many clinical pictures were admixtures of mood and psychotic disorder and could only be arbitrarily assigned according to a binary schizophrenia/psychotic mood disorder system. Only around 60% of first episode psychosis patients met operational criteria for schizophrenia or schizophreniform disorder[3]. First episode psychosis was viewed as an early stage of psychotic illness which could have heterogeneous outcomes, from full remission to evolution in either direction along a spectrum from psychotic mood disorder to schizophrenia, with variable levels of associated functional impairment. The fact that overlapping and fluctuating outcomes did frequently occur supported the decision to select a wide boundary for entry.

This thinking extended to the re-conceptualization of the prodromal period as an "at risk mental state" rather than as a fixed entity, one that might resolve fully, persist, or progress in

several possible directions. This has been borne out by empirical data showing that approximately 36% of "at risk mental state" patients transition to psychosis within three years, approximately a third have persistent attenuated psychotic symptoms, and a third remit from symptoms[4,5]. Transition was felt to be a crucial concept to operationally define the progression from subthreshold or inconsistent positive psychotic symptoms to sustained full threshold symptoms.

While our goal in treatment is optimizing functional outcomes, transition is a significant event connoting a likely more serious illness and certainly mandating a change in treatment, namely the use of antipsychotic medications. This was transition to *psychosis,* not schizophrenia, and it was felt to be important to define it in this particular way, to link it to a critical treatment decision. It is a potentially different question whether there might be a qualitative change in underlying neurobiology at that same specific point or any other[6]. Once again, only 60% of those transitioning would attract a diagnosis of schizophrenia or schizophreniform disorder.

As it later developed, however, the early psychosis field remained somewhat split as to whether to broaden the focus to the full spectrum of psychosis, with many, especially in North America and parts of Europe, still adhering to faith in the validity of the schizophrenia concept. Hence, many first episode psychosis programs were essentially aiming to be first episode schizophrenia programs, which had a flow on effect when they later embraced the UHR paradigm.

The target of the UHR strategy is not schizophrenia, but psychosis. The tenacity of the schizophrenia focus has fuelled in part some recent critiques, including that by van Os and Guloksuz[7] in a previous issue of this journal. We support the main thrust of that critique, and most of its conclusions. However, in their intent to accelerate the demise of the increasingly fragile schizophrenia concept, those authors seem to have misinterpreted some aspects of the evidence in relation to the UHR field. A more balanced critique and synthesis is needed to highlight the real value of what has emerged from two decades of heuristic research and pave the way for genuine and exciting progress in pre-emptive care. We do not wish to mount a line by line defence of the UHR field here, but do need to clarify some issues.

First, transition has been robustly and operationally defined, based on the generally agreed (though arbitrary) timing of a key treatment change. This definition has received significant validation through studies showing that a range of neurobiological markers differ at baseline in those who make a transition vs. those who do not, and sometimes longitudinally change in those who make the transition compared to those who do not[6,8-11]. These studies, however, do not allow us to define the optimal transition point from a neurobiological point of view. While functional outcome is worse in those who make the transition, this is not the only predictor or correlate of this point in the evolution of disorder[12,13].

If the sample is enriched to at least the level of 20% "true positives" for subsequent first episode psychosis, it is statistically possible to predict who is at especially high risk for transition[14] and even assign individuals to different "risk classes"[15]. Indeed, the UHR research paradigm has been a very productive approach, illuminating risk factors, predictive markers and pointing towards aetiological mechanisms involved in onset of psychotic disorders, albeit with some limitations that might now be effectively addressed and a translation into clinical care that might be more effectively implemented (see below)[14,16,17].

Interventions during the UHR stage of disorder are effective in not only reducing the risk of transition for at least 1-2 years, but also in improving functional outcomes[18-20]. There is increasing recognition in the field that transition to psychosis should in fact not be the sole focus of intervention, and that the variety of unfavourable trajectories, including poor functional outcome, should be critical targets[21-23]. Recent work identified as many as seventeen clinical trajectories in a UHR sample, with 43% of patients having favourable (recovery or remission from UHR state) and 57% unfavourable (recurrence, relapse, no-remission, transition) outcomes over one year[24].

In addition, it has been increasingly recognized that the "at risk mental state" should be regarded as a syndrome in its own right, in addition to being seen as connoting risk for disorder progression. It is a symptomatic state (albeit with psychotic symptoms below threshold for traditional diagnostic categories) associated with distress, functional impairment and diminished quality of life, closer in level to other coded psychiatric disorders and first episode psychosis than to the state of healthy controls[25]. Indeed, this was one of the reasons that its formulation in DSM-5 was as "attenuated psychosis syndrome" rather than as a risk category[26].

Another key learning which has opened a pathway for wider utility and progress is that, in addition to transition to psychosis and longer term psychotic disorder or persistent subthreshold psychotic symptoms, progression to persistent mood, anxiety, personality and/or substance use disorders is also a very common outcome[27,28]. Hence, at this subthreshold or early stage of illness, extending the boundary beyond psychosis (both at the case identification point and as a preventive target) is likely to be essential. Cuijpers[29] anticipated this in proposing a widening of the target syndromes based largely on power considerations and efficiency of prediction.

Complementary to this notion is the recognition that it is not uncommon for onset of mental disorders to follow a heterotypic course (i.e., symptoms of one type/category evolving into another type/category). This is illustrated by the fact that onset of first episode psychosis can emerge out of non-psychotic precursor states. A review by Lee et al[30] demonstrated that people at risk of *non*-psychotic disorders (identified through the presence of subthreshold non-psychotic symptoms) were at elevated risk of *psychotic* disorder (3.87% three-year incidence rate) – not as high as people meeting UHR criteria (24.63% three-year incidence rate), but substantially higher (77.4-fold) than the general population.

On the one hand, the UHR criteria do have greater valence for psychosis outcomes[31], but also have some valence for per-

sistent or incident non-psychotic disorders[27,32,33]. On the other, full-threshold psychosis may emerge out of risk states not characterized by attenuated psychotic symptoms[30,34].

HOW DO MENTAL DISORDERS EMERGE AND EVOLVE?

When people are floridly psychotic, manic or deeply depressed, it is obvious that they are ill and in need of care. But how did they get there? How did the pathway to obvious and severe illness start? Some authors' choice of where to define the illness border has been driven by concerns of overdiagnosis, overtreatment and labelling. While these problems do exist in some pockets and jurisdictions, the serious treatment gap that exists in every single country in the world, with only a minority of those in need of quality care being able to access it, indicates that the problem of underdiagnosis and failure to deliver treatment is a dramatically more urgent issue.

Defining a boundary is nevertheless important, because it is linked to a categorical decision of whether treatment or at least some kind of help is indicated and should be offered. We argue that this should be a fuzzy boundary in which the patient has a major say, not only health professionals, funders and polemicists[35,36]. There should be a soft entry policy but safeguards linked to proportional treatment, balancing benefits vs. risks, guided by the maxim *primum non nocere*.

Defining a boundary or border zone must be complemented by an understanding of the dynamics of how people move from being "well" to "ill"[37]. Eaton et al[38] have described how this occurs in very clear terms. People develop symptoms either by intensification of existing traits or features within the normal range of experience, such as anxiety or sadness, or the acquisition of novel subjective experiences such as hallucinations or obsessional thoughts, or a combination of the two. Syndromes or constellations of symptoms develop through the concurrent or sequential accumulation of such experiences and behaviours, and when they manifest some coherence and stability.

The key characteristics for determining whether there is a disorder are severity and persistence[39], though some argue that distress and/or functional impairment must also be present. In real life, these phenomena emerge in sporadic or gradual ways, often ebbing and flowing, sometimes following familiar trajectories and sequences, other times in a more fluid and reversible manner. How they stabilize or fade, how they attract other features and comorbid patterns and behaviours has not been systematically studied as yet.

In the early stages of mental ill health, diffuse and unstable subthreshold states of anxiety and depression are common, but often commingle with other features, including psychotic-like disturbances of salience and perception, and emotional dysregulation, to produce a kaleidoscopic series of microphenotypes[39,40]. We have not yet defined which set of variables to include in systematic studies of this stage of illness development, but they could include traditional symptom concepts, momentary emotional and perceptual states, self or corporeal disturbances, and sleep and motor activity changes. In this sea of emerging psychopathology, we already know that early psychotic symptoms, particularly if persistent in nature[41], indicate enhanced risk, not only for traditional psychotic disorders, for which they do have a greater valence, but also for other syndromal and functional outcomes[42-44].

In addition to the emergence and evolution of symptoms and syndromes, patienthood, help-seeking and need for care are influenced and defined by sociological factors[37], notably prejudice, stigma and illness behaviour[45-47]. Financial constraints can have a strong influence on where the bar is set by governments, social welfare agencies and health insurers for access to financial coverage for care. Ideological forces also seek to deny the reality of need for care, by asserting against all available evidence that mental ill health is actually part of the human condition (the "worried well") and naturally heals through "resilience". The same could be said about limb fractures, which are common, subject to a natural health process, and yet require professional intervention for optimal healing. These factors are arguably more potent in the mental health field in distorting the definition of need for care and the boundary between health and illness. More subtle variants of this invalidation involve the unhelpful distinction between high and low prevalence disorders.

THE NEW DIAGNOSIS: WHY CATEGORIES STILL MATTER AND HOW TO DEFINE THEM TO GUIDE TREATMENT AND RESEARCH

Psychiatric diagnosis is once again experiencing a crisis of confidence, which has been created by a range of forces. Some derive from fundamental issues, including our failure to bridge the mind/body dichotomy of Descartes and the complications of what philosophers call the "explanatory gap" or the "hard problem of consciousness"[48]. Others involve the notion that psychiatry can be shoehorned into mainstream medical practice without thoughtful and serious redesign, and the related overreach of biological psychiatry[49]; the invalidity of reifying syndromal descriptions as disease entities, and the naïve and diluted phenomenological and psychological constructs partly associated with the "operational revolution" of DSM-III onwards[50]; the polemics of antipsychiatry; and, most tellingly, the fact that diagnosis has rather low utility for treatment decisions. These forces have combined to fuel this crisis, which reached a peak during the launch period for DSM-5. The question has been quite reasonably raised: why do we need diagnosis anyway?

The fact that in large transdiagnostic samples there is a general psychopathology factor (the "p" factor) which has good predictive validity[51], and that most domains of psychopathology appear to conform to dimensional rather than categorical models, seem to favour a unitary or at least a non-categorical approach. This thinking has helped to inspire the creation of

the Research Domain Criteria (RDoC) project, which has embraced a transdiagnostic approach in research, attempting to base psychiatric nosology on neuroscience and behavioural science rather than DSM-defined diagnostic categories[52].

In our view, this approach overly downplays the role of clinical phenotype-based classification and overamplifies the role of neuroscience and behavioural constructs, which, although no doubt contribute to the understanding of the aetiology of psychiatric disorder, should be regarded as complementary rather than central to the "object" of psychiatric research and clinical practice. As we have argued elsewhere[50], part of the frustration with phenotype-based classification, and the perceived roadblock that it has introduced to research progress, may be attributable not to that classification *per se* but rather to the oversimplified and broad nature of contemporary psychopathological descriptions present in DSM-III onwards and in many of the instruments used to measure psychopathology in research studies[48]. To borrow geological terminology, focusing on plate tectonics (underlying neurobiology) should not replace or compensate for poor characterization of topography (phenomenology). In addition, the RDoC approach as yet confers no diagnostic benefit to clinical care, and its feasibility in many clinical settings is questionable.

Another related approach has been the Hierarchical Taxonomy of Psychopathology (HiTOP), which attempts to provide a hierarchical dimensional approach to psychiatric classification[53]. Although these approaches may contribute to mapping and describing nature (although, as noted above and elsewhere[48,50], there are reservations on this front and the jury is still out), they are of no help when it comes to making key decisions in patient care, which will always depend on binary or categorical 0/1 approaches.

It is all too easy to look at such issues and data sets from a population health or epidemiological perspective and critique concepts like "transition"[7], but clinicians and patients who have to make decisions about treatment approaches and life goals need to be more pragmatic. How do we harness the reality of dimensional ebbs and flows of symptoms across a wide range to make decisions about which treatments and in which sequence and combination to offer to which patients[23]? This is where clinical staging provides a solution.

We have described clinical staging in several previous papers[9,54,55]. Its key goal is to provide a more accurate guide to treatment selection (and also to prognosis). It also serves to organize research into psychosocial risk factors, neurocognitive variables, and biomarkers (both of current stage and risk for stage development). The model attempts to determine the position of an individual along a continuum of illness, defined according to stages: Stage 0 = no current symptoms, Stage 1a = help-seeking with distress, Stage 1b = attenuated (i.e., subthreshold) syndrome, Stages 2-4 = full threshold disorder with varying degrees of recurrence and severity.

The best known application of clinical staging has been in oncology. There one could argue that the progression or resolution of cancer is also a dimensional issue, but we have imposed categories or stages in a successful effort to intervene proportionally and preventively to reduce the risk of extension of the disease and ultimately death. The risk/benefit ratio is a guide to how aggressively to intervene, with the balance in favour of slight overtreatment at each stage, rather than waiting for treatment failure and then stepping up the intensity, as with "stepped care" in mental health, which responds often very belatedly to treatment resistance.

It might still be an open question whether discrete traditional syndromes such as bipolar disorder, schizophrenia and severe depression have utility at any stage, given the ubiquitous comorbidity that manifests across all stages. Other key influences on the complexion of intervention strategies are developmental and personal goals, such as vocational pathways and individuation and identity formation, that people identify and struggle with, and which are equally transdiagnostic. These might also correlate more with stage of illness than with individual syndrome or classical diagnosis.

SOLVING THE PREVENTION PARADOX: UNLOCKING THE SECRET TO PRE-EMPTIVE CLINICAL CARE

The prevention paradox

The prevention paradox refers to the fact that, with low incidence events such as suicide, transition to psychosis, or onset of anorexia nervosa, numerically more of the ultimately true positive cases will develop from lower risk than higher risk groups. van Os and Guloksuz[7] applied this logic to transition to psychosis in quoting a recent study[56] which found that only a very small proportion (4.1%) of patients who developed a first episode psychotic disorder attending local mental health services had been in previous contact with the local UHR service.

Our own data suggest that this may be a particularly low-end case example reflecting local clinic service pathways. In the case of Orygen Youth Health Clinical Program in Melbourne, a public mental health specialist early intervention service, 12.5% of first episode psychosis patients over a three-year period were referred from our UHR service (the PACE clinic) and 7% from other Orygen clinics.

According to van Os and Guloksuz[7], the above low percentage indicates that "the impact of prodromal services in public health terms may be negligible in relation to their costs". While the authors fail to note that the UHR service may well have prevented onset for a number of first episode psychosis cases (i.e., the "false false positive" cases[57]), there are likely to be better clinical outcomes for first episode psychosis cases who have previously been seen at a UHR clinic compared to those who have not[58], and these services have been shown to be cost-effective[59,60]. The fact remains, however, that UHR services see only a minority proportion of those who develop first episode psychosis.

If we aim to address the falling transition rate in UHR samples[57], we should seek to increase efficiency of risk detection

by enhancing methods of predicting psychosis within the UHR group. There are a number of ways in which this might be achieved. One approach is to improve screening and enrichment strategies. Screening tools such as the Prodromal Questionnaire[61] have been found to identify UHR cases who transition with high sensitivity (87%) and specificity (87%) and have also been found to detect a more enriched sample for psychosis risk[62]. Another approach is to apply new analytic strategies to data collected at study entry. There are currently several consortia-based efforts underway (e.g., PSYSCAN[63] and PRONIA[64]) applying machine learning approaches to develop clinical translation tools for enhanced prediction of psychosis onset in the UHR population.

Another important advance has been to use iterative probabilistic multimodal models to combine assessment domains, such as patient history, clinical assessments and biomarkers. This approach incorporates data from different modalities to increase predictive strength. For example, a probabilistic multimodal model in a UHR cohort using a combination of patient history, clinical assessment and fatty-acid biomarkers was able to identify over 70% of UHR cases who transitioned within one year[65]. However, it is unlikely that this approach will widen the entry channel, such that a higher percentage of first episode psychosis cases will pass through the UHR service portal.

Another response is to accept that a UHR service with a focus on psychosis risk and early warning signs of psychosis may be too narrow a channel to attract many of the young people experiencing and manifesting this phenotype. Such clinics struggle to detect and engage more than a small percentage of those expected within a given population in this stage of illness. On the other hand, with broad spectrum youth mental health care primary care platforms, such as headspace[66,67], we now know that a much higher number of such young people can be engaged. In a recent study, we found that 38% of young people accessing these services reported attenuated psychotic symptoms likely to be in the UHR range[68].

Also, a recent retrospective study by Shah et al[34] reported that 32% of their first episode psychosis sample did *not* undergo a period of subthreshold psychotic experiences prior to the onset of frank psychosis, and that the most prevalent early symptomatology was depression, anxiety and low functioning. Together, these findings suggest that a broader identification approach could overcome the prevention paradox by also identifying lower risk cases with possibly different phenotypic pathways to first episode psychosis[34,69], and also at risk of other full threshold or Stage 2 disorders. This would pave the way to a truly transdiagnostic approach.

Transdiagnostic risk: the Clinical High At Risk Mental State (CHARMS) approach

The high prevalence and impact of mental disorders in young people, the limitations of current diagnostic systems and risk identification approaches, the diffuse symptom patterns in early stages, and their pluripotent, transdiagnostic trajectories all indicate the need to develop a new diagnostic and predictive strategy. The approach we have recently adopted, guided by the clinical staging model and consistent with broad spectrum youth clinical service structures such as headspace, is an adaptation of the original "at risk mental state" approach to encompass a broader range of inputs and output target syndromes.

This Clinical High At Risk Mental State (CHARMS), as it has been named, is a broad composite definition of a syndrome warranting treatment in its own right due to help-seeking and distress associated with presenting symptoms, albeit below DSM/ICD-defined threshold for diagnosis. Figures 1 and 2 show the shift in approach from the traditional UHR to the CHARMS paradigm in the context of clinical staging.

The subthreshold (Stage 1b) states covered in the criteria at present include attenuated psychotic symptoms, subthreshold bipolar states, mild-moderate depression, and borderline personality features of reduced range and shorter duration than full diagnostic threshold[70]. The trait vulnerability of the UHR criteria is expanded to include history of serious mental disorder in a first degree relative, in addition to functional decline or chronic low functioning in the young person. Early data indicate a ~30% transition rate to Stage 2 disorder over a 6-12 month period in young people meeting these criteria and receiving treatment in our headspace clinical services, as opposed to <5% transition rate in help-seeking young people below this threshold (Stage 1a).

The data also indicate that evolution of symptoms may not necessarily follow a homotypic course (e.g., subthreshold psychosis evolving into threshold psychosis), but may be heterotypic in nature (e.g., moderate depression without attenuated psychotic symptoms at entry evolving into first episode psychosis), consistent with the pluripotent model. While this heterotypic course has been regarded as a shortcoming in the UHR approach (i.e., indicating lack of specificity of the criteria), it is welcomed within the CHARMS approach, because the target is *any* Stage 2 "exit syndrome" rather than a specific disorder outcome.

Importantly, this broad input-output approach can still support research into "narrowing" down on predictors and mechanisms at play in specific disorders or symptom clusters: the UHR subgroup, for example, can be identified within the broad Stage 1b cohort, and specific predictors of outcome within this subgroup or specific Stage 2 outcomes, such as psychosis, can be studied, and predictors of this specific outcome within the broad Stage 1b at risk group can be researched.

This pluripotent risk paradigm tackles many of the shortcomings associated with the UHR approach. It addresses the low transition to psychosis rates observed in recent years, allowing for capturing a broad range of outcomes and therefore a higher "transition rate" to serious mental disorder generally. It also places attenuated psychotic symptoms within the context of a range of multidimensional psychopathology, deemphasising these symptoms as a form of "schizophrenia light"[71].

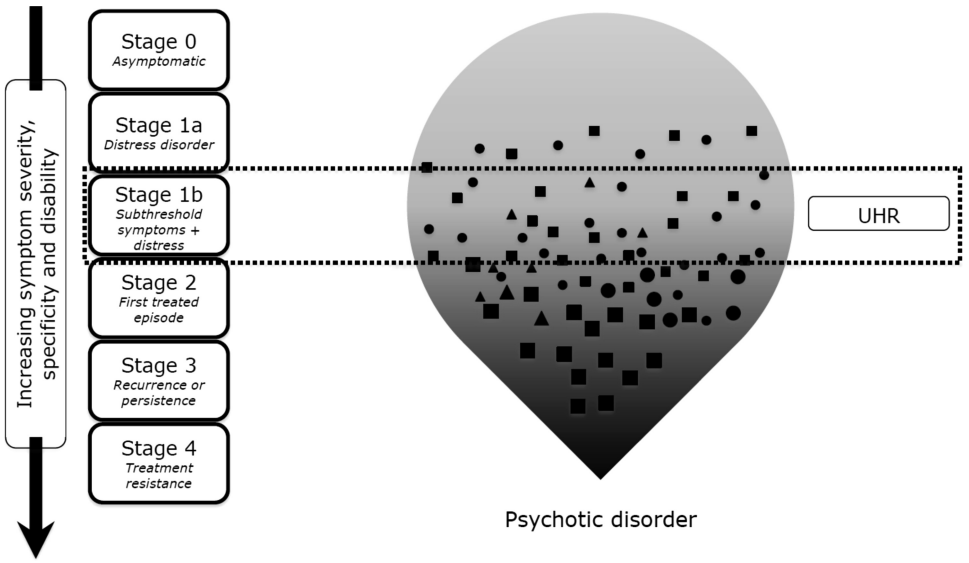

Figure 1 Traditional ultra high risk (UHR) paradigm in the context of clinical staging. The shapes represent different types of symptoms

It also provides a clinical identification approach for transdiagnostic preventive intervention trials. Such trials, which may consist of psychosocial or biological interventions or combinations and/or sequences of the two, would target the range of presenting symptomatology, rather than focus on a particular set of symptoms. In reality, this is what has occurred in UHR intervention trials anyway, particularly cognitive-behavioural therapy trials, where it is counterproductive to separate attenuated psychotic symptoms from the rest of the clinical picture (which is often more clinically distressing[72]) and focus treatment exclusively on those symptoms.

A suitable trial design for such studies are Sequential Multiple Assignment Randomized Trials ("SMART"), used in several recent large-scale studies in psychiatry to develop an evidence base to support adaptive clinical care[73]. This trial design methodology is a good fit with the clinical staging model, as it involves multiple intervention stages that correspond to the critical decisions involved in adaptive interventions. These are interventions in which the type or dosage is individualized on the basis of patient characteristics, such as psychological features, clinical presentation or mechanism linked biomarkers, and then is repeatedly adjusted over time in response to pa-

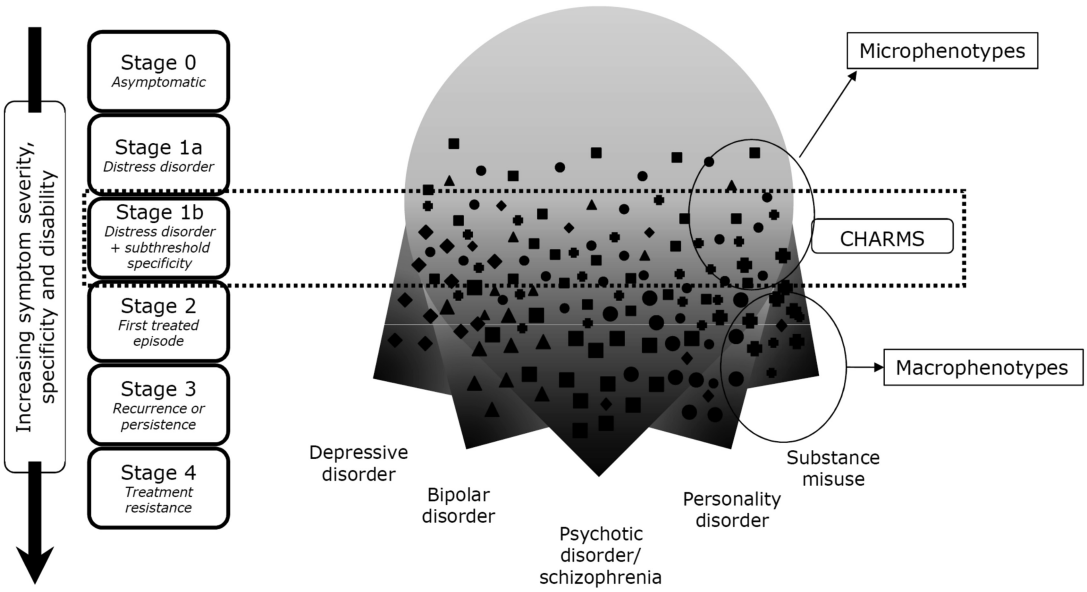

Figure 2 New transdiagnostic Clinical High At Risk Mental State (CHARMS) paradigm in the context of clinical staging. The shapes represent different types of symptoms

tient progress[73]. Interventions can also be tailored at critical decision points according to response or other patient characteristics, such as specific biomarker changes or comorbidity, and also patient preference.

Our group is currently conducting a SMART trial in a UHR sample[23], and plans to follow this with a further trial involving a combination of psychosocial and biological approaches targeting disorder progression, in the broader pluripotent at risk group (Stage 1b, identified using CHARMS criteria). Timing, personalization through biological and psychological markers, sequencing and admixture, and proportionality to stage are the key guiding principles.

New approaches to model and predict evolution of mental disorder

The model of the onset of mental disorder involving symptoms that ebb and flow, and consolidate or recede across stages, as described above, suggests the utility of approaching psychopathology as an evolving, complex system implying a combination of intra-individual and contextual factors interacting over time[74]. While it is useful to impose categories on this system for clinical decision-making, modelling change in psychopathology and predicting its evolution might more effectively be achieved using dynamic, time-dependent approaches.

Although searching for particular static factors that signal risk for future disorder (as with the Huntingtin gene mutation in Huntington's disease) may play a role, modelling risk for mental disorder may also require capturing factors (and their possible interaction) over time, i.e., must be dynamic in nature and able to incorporate fluctuations in key variables[16,40,50].

The traditional approach in psychiatric prediction studies, notably psychosis prediction, is to assess a range of variables (clinical, neurocognitive, neurobiological, genetic, etc.) upon entry to a mental health service and to determine whether these variables predict disorder onset (in the case of UHR research, first episode psychosis) or an increase/remission in symptom severity. This methodology rests on the notion that a single sampling of cross-sectional data can accurately predict the outcome of interest. The highly dynamic and changeable nature of psychopathology and the heterogeneous nature of early symptoms and symptom trajectories (see above) indicates the need for more dynamic models of prediction[24,74]. Such models of dynamic change have predominantly emerged from disciplines outside of psychiatry and therefore cross-disciplinary fertilization is important for progress in the field.

An example is dynamical systems theory, with origins in mathematics and physics, which seeks to describe the behaviour of complex dynamical systems such as the climate, ecosystems and financial markets[75]. Increasingly, mental health has been conceptualized in these terms, i.e., as a system with many elements which interact with each other over time (as in network theory[76], see below). The architecture of such a system reflects how it will change over time[77]: in a system with loosely connected, heterogeneous elements, change occurs gradually in response to changing conditions, whereas a system with highly interconnected, homogenous elements may initially resist change but then reach a critical threshold or "tipping point" towards another state.

In the context of psychopathology, these two "system states" may correspond to "healthy" and "disordered"/"ill" states[78,79]. Tipping points tend to be preceded by early warning signs, such as the phenomenon of "critical slowing down", which refers to the system taking increasingly longer to return to its previous state after a perturbation/stressor[80,81]. There is emerging evidence, using simulation data and fine-grained longitudinal time series data collected using ecological momentary assessment, that transitions in mental health (at this stage, depression and bipolar disorder) are preceded by critical slowing down[78,79].

A conceptually related approach is the "network perspective" of psychopathology, which has gained traction in recent years. This approach conceptualizes mental disorder not as the consequence of an underlying latent variable (a "common cause"), but as a result of a dynamic interplay of symptoms[82-84], with symptoms actively influencing/causing each other, rather than being the passive expression of an underlying disease process. Within the context of pluripotency during early psychopathology, it has been proposed that the way in which networked symptoms influence each other during early stages of mental ill health may be less concentrated and stable than in later stages[76].

Preliminary empirical work is consistent with this proposal, positioning network dynamics theory within the clinical staging framework and suggesting that, with increased clinical stage severity, symptomatology becomes more specialized and differentiated, giving rise to diagnostic specificity associated with greater inter- and intra-mental state connection strength, and greater inter- and intra-mental state connection variability[85]. Empirical investigations into the predictive potential of dynamic symptom networks for the onset and progression of psychosis are currently underway[86].

Another dynamic prediction approach, more agnostic with regard to theoretical principles, is joint modelling. This is a statistical method that combines multilevel modelling (using repeated clinical assessments) with survival analysis (allowing for the time-to-event nature of determining outcome in prediction studies)[87-89]. The approach can be used to identify symptom trajectories (e.g., persistence of negative symptoms, intensification of general psychopathology) that predict outcome, taking into account censored data and time to follow-up (as in survival analysis).

Importantly, it allows for the generation of a risk calculator that can be updated over time based on repeated assessments (using clinical or other information), a more refined method of predicting outcome than the existing risk calculators[90,91]. Initial work using this approach with data from our recent UHR intervention trial[19] shows that dynamic prediction using joint modelling produces much stronger predictive models, particularly positive predictive values, than using baseline data alone[89]. This

approach could equally be applied to transdiagnostic outcomes within a broader risk group, such as a CHARMS cohort.

We have recently argued that such concepts and analytic approaches may be useful for predicting onset of more severe stages of disorder transdiagnostically, as they take the evolving clinical picture into account[74]. They offer a means of modelling and predicting how mental disorder may evolve across clinical stages, capturing how and why microphenotypes disperse, cohere, sustain, expand or entrench. They may also guide the identification of "dynamic signatures" for risk of *particular* disorders[74] (e.g., critical slowing down may prove to be a more reliable indicator of imminent onset of depression than of psychotic disorder). As indicated above, "broadening" and "narrowing" the approach to risk identification and predictive factors are not mutually exclusive.

Importantly, these new prediction approaches link well with the process that occurs in real-world clinical decision making[92]. Clinical decision making regarding possible treatment changes and prognostic judgements is generally "adaptive" in nature – it reacts to and is updated based on gathering further clinical information and the unfolding symptomatology of the patient, rather than relying solely on the profile of the patient's first clinical presentation[23]. Using the conceptual and analytic approaches outlined here may provide an empirically-based and rigorous guide for making decisions regarding treatment modification in response to the evolution of a patient's clinical profile over time. In this way, they may help refine treatment decision making and possibly be incorporated into adaptive clinical trial designs, described above, which are currently generally based purely on a category of response/non-response at the end of a pre-specified time period[93].

CONCLUSIONS

The "at risk mental state" for psychosis approach has been a highly productive research paradigm over the last 25 years. However, the limitations of current risk identification approaches, the diffuse and unstable symptom patterns in early stages, and their pluripotent, transdiagnostic trajectories all indicate the need to develop a new strategy. The approach we have recently adopted has been guided by the clinical staging model and adapts the original "at risk mental state" model to encompass a broader range of inputs and output target syndromes. This approach is supported by a number of novel modelling and prediction strategies, such as dynamical systems theory, network theory, and joint modelling.

A holistic strategy can be developed that applies these new prediction approaches, as well as machine learning and iterative probabilistic multimodal models, to a blend of subjective psychological data, physical disturbances and biomarkers acquired through fine-grained sequential or longitudinal assessments. This strategy will ultimately enhance our understanding and ability to predict the onset, early course and evolution of mental ill health, further opening pathways for preventive interventions.

REFERENCES

1. McGorry PD, Edwards J, Mihalopoulos C et al. EPPIC: an evolving system of early detection and optimal management. Schizophr Bull 1996;22:305-26.
2. McGorry PD, Copolov DL, Singh BS. Current concepts in functional psychosis. The case for a loosening of associations. Schizophr Res 1990;3: 221-34.
3. Henry LP, Harris MG, Amminger GP et al. Early Psychosis Prevention and Intervention Centre long-term follow-up study of first-episode psychosis: methodology and baseline characteristics. Early Interv Psychiatry 2007;1: 49-60.
4. Fusar-Poli P, Bonoldi I, Yung AR et al. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Arch Gen Psychiatry 2012;69:220-9.
5. Cannon TD. How schizophrenia develops: cognitive and brain mechanisms underlying onset of psychosis. Trends Cogn Sci 2015;19:744-56.
6. Wood SJ, Yung AR, McGorry PD et al. Neuroimaging and treatment evidence for clinical staging in psychotic disorders: from the at-risk mental state to chronic schizophrenia. Biol Psychiatry 2011;70:619-25.
7. van Os J, Guloksuz S. A critique of the "ultra-high risk" and "transition" paradigm. World Psychiatry 2017;16:200-6.
8. Cannon TD, Chung Y, He G et al. Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. Biol Psychiatry 2015;77:147-57.
9. McGorry P, Keshavan M, Goldstone S et al. Biomarkers and clinical staging in psychiatry. World Psychiatry 2014;13:211-23.
10. Pantelis C, Velakoulis D, McGorry PD et al. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. Lancet 2003;361:281-8.
11. Fusar-Poli P, McGuire P, Borgwardt S. Mapping prodromal psychosis: a critical review of neuroimaging studies. Eur Psychiatry 2012;27:181-91.
12. Cotter J, Drake RJ, Bucci S et al. What drives poor functioning in the at-risk mental state? A systematic review. Schizophr Res 2014;159:267-77.
13. Yung AR, Cotter J, Wood SJ et al. Childhood maltreatment and transition to psychotic disorder independently predict long-term functioning in young people at ultra-high risk for psychosis. Psychol Med 2015;45:3453-65.
14. Yung AR, Nelson B. The ultra-high risk concept – a review. Can J Psychiatry 2013;58:5-12.
15. Ruhrmann S, Schultze-Lutter F, Salokangas RK et al. Prediction of psychosis in adolescents and young adults at high risk: results from the Prospective European Prediction of Psychosis Study. Arch Gen Psychiatry 2010;67: 241-51.
16. McGorry P, Nelson B. Why we need a transdiagnostic staging approach to emerging psychopathology, early diagnosis, and treatment. JAMA Psychiatry 2016;73:191-2.
17. Fusar-Poli P, Carpenter WT, Woods SW et al. Attenuated psychosis syndrome: ready for DSM-5.1? Annu Rev Clin Psychol 2014;10:155-92.
18. van der Gaag M, Smit F, Bechdolf A et al. Preventing a first episode of psychosis: meta-analysis of randomized controlled prevention trials of 12 month and longer-term follow-ups. Schizophr Res 2013;149:56-62.
19. McGorry PD, Nelson B, Markulev C et al. Effect of omega-3 polyunsaturated fatty acids in young people at ultrahigh risk for psychotic disorders: the NEURAPRO randomized clinical trial. JAMA Psychiatry 2017;74:19-27.
20. Stafford MR, Jackson H, Mayo-Wilson E et al. Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013;346:f185.
21. Yung AR. Treatment of people at ultra-high risk for psychosis. World Psychiatry 2017;16:207-8.
22. Yung AR, Nelson B, Thompson A et al. The psychosis threshold in ultra high risk (prodromal) research: is it valid? Schizophr Res 2010;120:1-6.
23. Nelson B, Amminger GP, Yuen HP et al. Staged Treatment in Early Psychosis: a sequential multiple assignment randomised trial of interventions for ultra high risk of psychosis patients. Early Interv Psychiatry (in press).
24. Polari A, Lavoie S, Yuen HP et al. Clinical trajectories in the ultra-high risk for psychosis population. Schizophr Res (in press).
25. Fusar-Poli P, Rocchetti M, Sardella A et al. Disorder, not just state of risk: meta-analysis of functioning and quality of life in people at high risk of psychosis. Br J Psychiatry 2015;207:198-206.
26. Carpenter WT, Regier D, Tandon R. Misunderstandings about attenuated psychosis syndrome in the DSM-5. Schizophr Res 2014;152:303.
27. Lin A, Wood SJ, Nelson B et al. Outcomes of nontransitioned cases in a sample at ultra-high risk for psychosis. Am J Psychiatry 2015;172:249-58.

28. Rutigliano G, Valmaggia L, Landi P et al. Persistence or recurrence of non-psychotic comorbid mental disorders associated with 6-year poor functional outcomes in patients at ultra high risk for psychosis. J Affect Disord 2016;203:101-10.
29. Cuijpers P. Examining the effects of prevention programs on the incidence of new cases of mental disorders: the lack of statistical power. Am J Psychiatry 2003;160:1385-91.
30. Lee TY, Lee J, Kim M et al. Can we predict psychosis outside the clinical high-risk state? A systematic review of non-psychotic risk syndromes for mental disorders. Schizophr Bull 2018;44:276-85.
31. Nelson B, Yuen HP, Wood SJ et al. Long-term follow-up of a group at ultra high risk ("prodromal") for psychosis: the PACE 400 study. JAMA Psychiatry 2013;70:793-802.
32. Wigman JT, van Nierop M, Vollebergh WA et al. Evidence that psychotic symptoms are prevalent in disorders of anxiety and depression, impacting on illness onset, risk, and severity – implications for diagnosis and ultra-high risk research. Schizophr Bull 2012;38:247-57.
33. Lim J, Rekhi G, Rapisarda A et al. Impact of psychiatric comorbidity in individuals at ultra high risk of psychosis – findings from the Longitudinal Youth at Risk Study (LYRIKS). Schizophr Res 2015;164:8-14.
34. Shah JL, Crawford A, Mustafa SS et al. Is the clinical high-risk state a valid concept? Retrospective examination in a first-episode psychosis sample. Psychiatr Serv 2017;68:1046-52.
35. Frances A. Saving normal: an insider's revolt against out-of-control psychiatric diagnosis, DSM-5, big pharma, and the medicalization of ordinary life. New York: Morrow, 2013.
36. Greenberg G. The book of woe: the DSM and the unmaking of psychiatry. New York: Blue Rider, 2013.
37. Scott RD. The treatment barrier: part 2. The patient as an unrecognized agent. Br J Med Psychol 1973;46:57-67.
38. Eaton WW. The sociology of mental disorders, 3rd ed. Westport: Praeger, 2001.
39. McGorry P, van Os J. Redeeming diagnosis in psychiatry: timing versus specificity. Lancet 2013;381:343-5.
40. van Os J. The dynamics of subthreshold psychopathology: implications for diagnosis and treatment. Am J Psychiatry 2013;170:695-98.
41. van Os J, Linscott RJ, Myin-Germeys I et al. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. Psychol Med 2009; 39:179-95.
42. Kelleher I, Corcoran P, Keeley H et al. Psychotic symptoms and population risk for suicide attempt: a prospective cohort study. JAMA Psychiatry 2013; 70:940-8.
43. Kelleher I, Devlin N, Wigman JT et al. Psychotic experiences in a mental health clinic sample: implications for suicidality, multimorbidity and functioning. Psychol Med 2013;44:1615-24.
44. Kelleher I, Keeley H, Corcoran P et al. Clinicopathological significance of psychotic experiences in non-psychotic young people: evidence from four population-based studies. Br J Psychiatry 2012;201:26-32.
45. Kirmayer L, Lemelson R, Cummings C (eds). Revisioning psychiatry: integrating biological, clinical and cultural perspectives. Cambridge: Cambridge University Press, 2015.
46. Pilowsky I. Abnormal illness behaviour. Br J Med Psychol 1969;42:347-51.
47. Pilowsky I. A general classification of abnormal illness behaviours. Br J Med Psychol 1978;51:131-7.
48. Parnas J. The RDoC program: psychiatry without psyche? World Psychiatry 2014;13:46-7.
49. Borsboom D, Cramer A, Kalis A. Brain disorders? Not really... Why network structures block reductionism in psychopathology research. Behav Brain Sci (in press).
50. Nelson B, Hartmann JA, Parnas J. Detail, dynamics and depth: useful correctives for some current research trends. Br J Psychiatry (in press).
51. Lahey BB, Krueger RF, Rathouz PJ et al. Validity and utility of the general factor of psychopathology. World Psychiatry 2017;16:142-3.
52. Sanislow CA. Updating the Research Domain Criteria. World Psychiatry 2016;15:222-3.
53. Kotov R, Krueger RF, Watson D. A paradigm shift in psychiatric classification: the Hierarchical Taxonomy Of Psychopathology (HiTOP). World Psychiatry 2018;17:24-5.
54. McGorry PD. Issues for DSM-V: clinical staging: a heuristic pathway to valid nosology and safer, more effective treatment in psychiatry. Am J Psychiatry 2007;164:859-60.
55. McGorry PD. Early clinical phenotypes, clinical staging, and strategic biomarker research: building blocks for personalized psychiatry. Biol Psychiatry 2013;74:394-5.
56. Ajnakina O. First episode psychosis: looking backwards and forwards. London: King's College London, 2017.
57. Yung AR, Yuen HP, Berger G et al. Declining transition rate in ultra high risk (prodromal) services: dilution or reduction of risk? Schizophr Bull 2007;33:673-81.
58. Valmaggia LR, Byrne M, Day F et al. Duration of untreated psychosis and need for admission in patients who engage with mental health services in the prodromal phase. Br J Psychiatry 2015;207:130-4.
59. Valmaggia LR, McCrone P, Knapp M et al. Economic impact of early intervention in people at high risk of psychosis. Psychol Med 2009;39:1617-26.
60. Phillips LJ, Cotton S, Mihalopoulos C et al. Cost implications of specific and non-specific treatment for young persons at ultra high risk of developing a first episode of psychosis. Early Interv Psychiatry 2009;3:28-34.
61. Savill M, D'Ambrosio J, Cannon TD et al. Psychosis risk screening in different populations using the Prodromal Questionnaire: a systematic review. Early Interv Psychiatry 2018;12:3-14.
62. Rietdijk J, Klaassen R, Ising H et al. Detection of people at risk of developing a first psychosis: comparison of two recruitment strategies. Acta Psychiatr Scand 2012;126:21-30.
63. PSYSCAN. Translating neuroimaging findings from research into clinical practice. http://psyscan.eu.
64. PRONIA. Personalised prognostic tools for early psychosis management. https://www.pronia.eu.
65. Clark SR, Baune BT, Schubert KO et al. Prediction of transition from ultra-high risk to first-episode psychosis using a probabilistic model combining history, clinical assessment and fatty-acid biomarkers. Transl Psychiatry 2016;6:e897.
66. Rickwood DJ, Telford NR, Mazzer KR et al. The services provided to young people through the headspace centres across Australia. Med J Aust 2015; 202:533-6.
67. Rickwood DJ, Telford NR, Parker AG et al. headspace - Australia's innovation in youth mental health: who are the clients and why are they presenting? Med J Aust 2014;200:108-11.
68. Purcell R, Jorm AF, Hickie IB et al. Demographic and clinical characteristics of young people seeking help at youth mental health services: baseline findings of the Transitions Study. Early Interv Psychiatry 2015;9:487-97.
69. Shah JL. Sub-threshold mental illness in adolescents: within and beyond DSM's boundaries. Soc Psychiatry Psychiatr Epidemiol 2015;50:675-7.
70. Hartmann JA, Nelson B, Spooner R et al. Broad clinical high-risk mental state (CHARMS): methodology of a cohort study validating criteria for pluripotent risk. Early Interv Psychiatry (in press).
71. van Os J, Murray RM. Can we identify and treat "schizophrenia light" to prevent true psychotic illness? BMJ 2013;346:f304.
72. Power L, Polari AR, Yung AR et al. Distress in relation to attenuated psychotic symptoms in the ultra-high-risk population is not associated with increased risk of psychotic disorder. Early Interv Psychiatry 2015;10:258-62.
73. Lei H, Nahum-Shani I, Lynch K et al. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol 2012;8:21-48.
74. Nelson B, McGorry PD, Wichers M et al. Moving from static to dynamic models of the onset of mental disorder. JAMA Psychiatry 2017;74:528-34.
75. Scheffer M. Critical transitions in nature and society. Princeton: Princeton University Press, 2009.
76. Borsboom D. A network theory of mental disorders. World Psychiatry 2017; 16:5-13.
77. Scheffer M, Carpenter SR, Lenton TM et al. Anticipating critical transitions. Science 2012;338:344-8.
78. van de Leemput IA, Wichers M, Cramer AO et al. Critical slowing down as early warning for the onset and termination of depression. Proc Natl Acad Sci USA 2014;111:87-92.
79. Bayani A, Hadaeghi F, Jafari S et al. Critical slowing down as an early warning of transitions in episodes of bipolar disorder: a simulation study based on a computational model of circadian activity rhythms. Chronobiol Int 2017;34:235-45.
80. Scheffer M. Complex systems: foreseeing tipping points. Nature 2010;467: 411-2.
81. Scheffer M, Bascompte J, Brock WA et al. Early-warning signals for critical transitions. Nature 2009;461:53-9.
82. Borsboom D, Cramer AO. Network analysis: an integrative approach to the structure of psychopathology. Annu Rev Clin Psychol 2013;9:91-121.

83. van Borkulo C, Boschloo L, Borsboom D et al. Association of symptom network structure with the course of longitudinal depression. JAMA Psychiatry 2015;72:1219-26.
84. Isvoranu AM, Borsboom D, van Os J et al. A network approach to environmental impact in psychotic disorder: brief theoretical framework. Schizophr Bull 2016;42:870-3.
85. Wigman JTW, van Os J, Thiery E et al. Psychiatric diagnosis revisited: towards a system of staging and profiling combining nomothetic and idiographic parameters of momentary mental states. PLoS One 2013;8: e59559.
86. Booij SH, Wichers M, de Jonge P et al. Study protocol for a prospective cohort study examining the predictive potential of dynamic symptom networks for the onset and progression of psychosis: the Mapping Individual Routes of Risk and Resilience (Mirorr) study. BMJ Open 2018;8: e019059.
87. Yuen HP, Mackinnon A. Performance of joint modelling of time-to-event data with time-dependent predictors: an assessment based on transition to psychosis data. PeerJ 2016;4:e2582.
88. Yuen HP, Mackinnon A, Nelson B. A new method for analysing transition to psychosis: joint modelling of time-to-event outcome with time-dependent predictors. Int J Methods Psychiatry Res (in press).
89. Yuen HP, Mackinnon A, Hartmann J et al. A novel approach for developing prediction models of transition to psychosis: dynamic prediction using joint modelling. Submitted for publication.
90. Cannon TD, Yu C, Addington J et al. An individualized risk calculator for research in prodromal psychosis. Am J Psychiatry 2016;173:980-8.
91. Fusar-Poli P, Rutigliano G, Stahl D et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. JAMA Psychiatry 2017;74:493-500.
92. Chitty RN. Why clinicians are natural bayesians: is there a bayesian doctor in the house? BMJ 2005;330:1390.
93. Bhatt DL, Mehta C. Adaptive designs for clinical trials. N Engl J Med 2016; 375:65-74.

DOI:10.1002/wps.20514

Robustness and replicability of psychopathology networks

Network approaches to psychopathology hold that mental disorders arise from the interplay between symptoms in a network structure[1,2]. In the past few years, statistical techniques that estimate networks were developed and applied to many disorders[3]. As empirical findings start to accumulate, the question arising is which of these findings are robust and replicable. Here we evaluate the state of psychopathological network research based on three methodological criteria: model quality, precision, and replicability.

Model quality. One important quality of statistical modeling techniques is their ability to recover the "true" model that generated the data, a necessary prerequisite for justifying inferences based on network models. This is evaluated through: a) mathematical analysis: prove that a technique will recover the generating network model from data (e.g., by showing that it converges to the true model as sample size increases); and b) simulation studies: evaluate a technique's performance under various circumstances (e.g., for different network structures, sample sizes, and parameter settings).

Current state-of-the-art network techniques (i.e., pairwise Markov random fields[4]) have been vetted through mathematical proofs and simulation studies[5,6]: they efficiently recover the "true" model underlying the data. In general, such techniques minimize the false positive rate at the expense of statistical power. As a result, these techniques are more likely to omit "true" network connections, than to include spurious connections[5,6]. In sum, these techniques are vetted, conservative tools for estimating psychopathology network structures.

Precision and robustness. When a researcher has estimated a network from empirical data using vetted methodology, the question is to what extent the parameter estimates are precise: how robust are the results? For instance, if the relationship between self-worth and suicidal thoughts seems stronger than that between sleep difficulty and suicidal thoughts, it is necessary to investigate if model parameters are estimated with sufficient precision to justify this inference. If not, the result may not replicate in other samples.

Precision of network parameter estimates can vary considerably depending on factors such as sample size, network size, and network structure. Therefore, these factors must be assessed and reported on a case-by-case basis, by evaluating the statistical precision of parameter estimates (e.g., with confidence intervals) and the robustness of the model as a whole (e.g., investigating network structures in subsamples).

Dedicated freeware methodology for doing this recently became available[4], which allows researchers to report confidence intervals for estimated network parameters as an integral part of their results. This practice was quickly embraced by the majority of the network community, who now publish their work including detailed robustness checks. Naturally, results of such analyses should constrain the researcher's conclusions proportionately to their content: stronger claims (e.g., "insomnia is the most central node in the depression network") require stronger evidence than weaker claims (e.g., "insomnia is connected to the depression network").

Replicability. When network analysis seems to warrant an empirical conclusion (e.g., a particular symptom is highly central, or one network is more densely connected than another), the next question is whether the relevant phenomenon can be replicated in other samples. Ideally, replication research differs from the original study only in features that are deemed irrelevant to the phenomenon under investigation (e.g., by using a different sample from the same population). However, as is often the case in replication research, it is sometimes unclear whether differences between a study and its purported replication are relevant or not. For instance, if a network is first estimated on a community sample, and a replication is attempted in a patient sample, it may be unrealistic to assume that the same network holds in both populations. In such cases, studies probe not only a finding's replicability, but also its generalizability. Consequently, if inconsistent findings arise, this may either be because the phenomenon is unstable or illusory (i.e., the finding is not replicable) or because of substantively meaningful differences between studies (i.e., the finding is not generalizable to the context of the new study). In contrast, if an empirical phenomenon is observed consistently across studies, this provides compound evidence for both its replicability and generalizability[7].

Several recent empirical studies have evaluated the replicability of networks. The general picture which emerges is that network structures replicate and generalize well. For example, networks of major depression and generalized anxiety disorder symptoms are nearly identical in the US and Australia; post-traumatic stress disorder (PTSD) networks are similar across different populations and sources of trauma; and major depression networks are invariant across environmental and genetic risk factors (e.g., age of onset)[7,8].

Although network structures appear replicable and generalizable, detailed inferences based on them may be more susceptible to variation across studies. For example, the centrality of nodes seems to vary across PTSD networks, and a reported difference in network density between remitted and persistent major depression cases in adults was not fully replicated in an adolescent sample[8]. Future research should critically interrogate such findings to determine if inconsistency between studies is best characterized as a failure to replicate or a failure to generalize across contexts.

In conclusion, the model quality of network analysis techniques is good, while precision and robustness can now adequately be assessed with freely available methodological tools. Burgeoning replication research suggests that the structure of networks is typically consistent across studies, while stronger inferences based on these structures (e.g., centrality) have occasionally yielded mixed results.

Network analysis is a promising approach that may lead to significant improvements in research on and treatment of psychopathology[9], but researchers should be careful not to overstate causal conclusions based on network analysis as long as the causal interpretation of models has not been thoroughly investigated. The assessment of network robustness and replicability is an important step in this process and should be an important research focus in the next few years.

Denny Borsboom[1], Donald J. Robinaugh[1], The Psychosystems Group[1], Mijke Rhemtulla[2], Angélique O.J. Cramer[3]
[1]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands; [2]Department of Psychology, University of California, Davis, CA, USA; [3]Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands

The Psychosystems Group includes Sacha Epskamp, Eiko I. Fried, Lourens J. Waldorp, Claudia D. van Borkulo, and Han L.J. van der Maas.

1. Borsboom D. World Psychiatry 2017;16:5-13.
2. Cramer AOJ, van Borkulo CD, Giltay EJ et al. PLoS One 2016;11:e0167490.
3. Fried EI, van Borkulo CD, Cramer AOJ et al. Soc Psychiatry Psychiatr Epidemiol 2017;52:1-10.
4. Epskamp S, Borsboom D, Fried EI. Behav Res Methods 2018;50:195-212.
5. Bühlmann P, van de Geer S. Statistics for high-dimensional data: methods, theory and applications. Berlin: Springer, 2011.
6. Van Borkulo CD, Borsboom D, Epskamp S et al. Sci Rep 2014;4:5918.
7. Borsboom D, Fried EI, Epskamp S et al. J Abnorm Psychol 2017;126:989-99.
8. Fried EI. http://psych-networks.com/7-new-papers-network-replicability/.
9. McNally RJ. Behav Res Ther 2016;86:95-104.

DOI:10.1002/wps.20515

Accelerated biological aging in serious mental disorders

Individuals with serious mental disorders (SMDs) die at an earlier average age, even after controlling for suicide[1]. They are also at increased risk for developing somatic diseases that are typically associated with advanced age, such as cardiovascular diseases, metabolic syndrome, immune dysregulation and dementia[1].

The causes of this are likely multi-factorial, including genetic predisposition, biological changes set in motion by early life adversity, and lifestyle factors. Lifestyle factors, while obviously important, do not fully explain the increased mortality and morbidity in these individuals, and consequently, "accelerated biological aging" is increasingly being seen as an intrinsic factor in SMDs, at least in some individuals[2,3].

To the extent this hypothesis is true, the scope of pathophysiology in these illnesses would broaden considerably, and they might no longer be framed as only "mental disorders" or even brain diseases, but rather as whole-body, multi-system illnesses (or at least as illnesses with substantial somatic comorbidity), of which the psychiatric presentation is just the most readily observable pathology[3]. Understanding the mediators of such potential acceleration of aging should expand preventative and therapeutic opportunities to improve physical as well as mental health in affected individuals.

The notion of accelerated biological aging in SMDs is supported by reports of acceleration of certain biomarkers of age, such as leukocyte telomere length[2] and epigenetic age[4]. However, data on these biomarkers remain relatively sparse to this point, and several questions arise: a) Do these markers measure aging *per se*, or just the presence of factors that themselves mediate aging? b) Are these markers causally related to SMDs or just correlated with them? c) Is accelerated aging specific to particular psychiatric diagnoses or to certain physiological perturbations that traverse diagnostic boundaries? d) Do different aging biomarkers reflect the same or different underlying aging processes? Here I briefly review recent data pertinent to these questions.

Leukocyte telomere length and epigenetic age both significantly track chronological age, with correlation coefficients of −0.38 to −0.51 (for the former) and 0.96 (for the latter). Both of these markers significantly predict disease and mortality, strengthening the view that they are measurable markers of the aging process and of rates of aging. However, leukocyte telomere length and epigenetic age are independent predictors of chronological age and mortality risk[5]. Therefore, while they both measure processes that evolve with aging or are associated with aging, the specific processes are different, and their underlying mediators likely differ.

Telomere shortening can occur in response to inflammation, oxidative stress, stress hormones and other factors[2,3]. As such, it may signal the cumulative presence of a toxic cellular environment, rather than directly informing on the aging process itself. Indeed, leukocyte telomere length is often found to be inversely correlated with circulating inflammatory and oxidative stress factors[2,3]. Another major determinant of telomere shortening is a cell's mitotic history, since telomeres fail to fully replicate after each cell division, unless acted upon by the intracellular enzyme telomerase.

When cells reach a critically short telomere length, they may undergo replicative senescence, apoptosis, genomic instability or oncogenic transformation[2]. This can be especially problematic in tissues whose mitotic capacity is necessary for cellular replacement, such as hematopoietic stem cells and – of particular relevance to psychiatry – neuronal stem cells in the dentate gyrus of the hippocampus. Of great concern (and also of great preventative opportunity), early life stress, even *in utero*, has been associated with shortened leukocyte telomere length in newborns and in adults.

Telomeres in SMDs may progressively shorten with illness chronicity and/or severity, but, interestingly, even never-depressed girls at high genetic risk for developing depression already have short telomeres compared to girls at low genetic

risk[6], suggesting a genetic (or epigenetic) link to telomere shortening even before illness onset, and raising the possibility that short telomeres are a *risk factor* for developing certain SMDs.

Telomerase has non-canonical effects unrelated to telomere lengthening, and may, in fact, have direct anti-aging, neurotrophic and antidepressant effects[7]. Human studies are just beginning to explore whether telomerase activation provides such benefits in humans or mediates the therapeutic effects of certain psychotropic drugs[8].

Epigenetic age is a more recent candidate marker of the aging process, owing largely to the discovery of an "epigenetic clock" by Horvath[4] that remarkably tracks chronological age in humans and also predicts all-cause mortality. This "clock" is based on progressive age-related changes in methylation of 5'-C-phosphate-G-3' (CpG) sites at specific DNA loci. While largely pre-programmed, methylation of these sites is also influenced by the environment.

Considerably less work has been reported on epigenetic age in SMDs compared to leukocyte telomere length, and the findings are not entirely consistent. However, as was the case with telomere shortening, epigenetic age is apparently accelerated following life stress, perhaps even prenatal stress, and this might be related to glucocorticoid effects on methylation[9].

Although not strictly a biomarker of aging, another set of relevant subcellular biomarkers involves the mitochondria, which are related to the aging process, are affected by stress and glucocorticoids, interact with telomere length and telomerase, and are likely dysregulated in SMDs. Mitochondrial pathology may be assessed by mitochondrial mutations, inefficient energy generation, increased reactive oxygen species generation and altered mitochondrial DNA copy number. Although widely studied in somatic diseases, the characterization of mitochondria in SMDs is in its infancy, some results are conflicting, and the relationship of mitochondria to epigenetic age and leukocyte telomere length is actively being investigated[10].

In summary, the landscape of psychiatric illness is changing, with a new focus on subcellular components and processes in addition to neurotransmitters. All or nearly all of the biomarkers discussed here lack diagnostic specificity, but may, rather, be specific to particular biochemical disturbances (e.g., glucocorticoids, inflammation, oxidative stress) that contribute to psychiatric illnesses. Such studies are reinforcing the concept that purely phenomenologic diagnoses may obscure biological underpinnings of psychiatric pathology.

The biomarkers discussed here are usually studied in peripheral leukocytes or in saliva. While the relationship between these peripheral markers and brain markers is uncertain, they likely have import in their own right, and may contribute to somatic as well as psychiatric presentations.

A key area of investigation is to determine whether biological aging in SMDs can be decelerated with appropriate treatment. When psychiatrists and other clinicians view mental illnesses as whole body diseases, the focus will change from specific behaviors to systemic, whole body pathologies, and personalized medicine will increasingly match target-based therapies to specific biological indicators.

Owen M. Wolkowitz
University of California San Francisco School of Medicine, San Francisco, CA, USA

1. Liu NH, Daumit GL, Dua T et al. World Psychiatry 2017;16:30-40.
2. Lindqvist D, Epel ES, Mellon SH et al. Neurosci Biobehav Rev 2015;55:333-64.
3. Wolkowitz OM, Reus VI, Mellon SH. Dialogues Clin Neurosci 2011;13:25-39.
4. Horvath S. Genome Biol 2013;14:R115.
5. Marioni RE, Harris SE, Shah S et al. Int J Epidemiol 2018;47:356.
6. Gotlib IH, LeMoult J, Colich NL et al. Mol Psychiatry 2015;20:615-20.
7. Jaskelioff M, Muller FL, Paik JH et al. Nature 2011;469:102-6.
8. Bersani FS, Lindqvist D, Mellon SH et al. Drug Discov Today 2015;20:1305-9.
9. Zannas AS, Arloth J, Carrillo-Roa T et al. Genome Biol 2015;16:266.
10. Picard M. J Aging Res 2011;2011:814096.

DOI:10.1002/wps.20546

Physician-assisted death in psychiatry

Physician-assisted death (PAD) – a term encompassing both the prescription and administration of life-ending medications – originated as a response to the extreme physical suffering of many people with terminal conditions such as cancer. Given that their lives would soon end in any case, allowing them to avoid the pain associated with their final days was seen as consistent with physicians' obligations to relieve suffering. Persuaded by such arguments, at least six countries and five US states have legalized some form of PAD[1]. However, Belgium, the Netherlands and Luxembourg have gone a step further, eliminating the requirement for a terminal illness, thus making people with psychiatric disorders eligible for physician assistance in ending their lives – a move currently under consideration in Canada as well[1,2].

Supporters of PAD for psychiatric disorders argue that denying access to relief from suffering to persons with such conditions is discriminatory, reflecting a failure to recognize the real pain associated with depression, chronic psychotic disorders, and other psychiatric illnesses[3]. In Belgium and the Netherlands, clinics have been established to facilitate access to PAD, and the use of the procedure is increasing. The most recent Dutch data (from 2015) indicate that approximately 4.5% of all deaths are due to PAD, with psychiatric disorders accounting for 3% of the total[4]. Belgium has a similar proportion of PAD deaths involving persons whose suffering is primarily attributable to psychiatric disorders[2].

In light of the growing use of PAD for psychiatric indications, it is worth considering why other jurisdictions may want

to exercise caution about embracing this trend. Although advocates for psychiatric PAD often take treatment-resistant depression as the model disorder for which termination of life may be indicated, it seems clear that PAD is being used for many other disorders as well. A study by Kim et al[5], based on a sample of 66 reports filed with the Dutch entity charged with overseeing PAD, found that 49 cases involved depression, but six were reported to have substance abuse, four neurocognitive impairment, and two autism spectrum disorder. A review of 100 persons requesting PAD in Belgium[6] reported that 90% had multiple psychiatric conditions, with 58% suffering from mood disorders, at least 12% from Asperger's syndrome, 10% from eating disorders, and 7% from dissociative disorders.

Both the Dutch and Belgian studies reported that about half of patients requesting PAD had personality disorders, including 27% with borderline personality disorder in Belgium. The substantial presence of comorbid personality disorders, often highly reactive to life stresses, especially interpersonal conflict, raises the question of whether PAD may be sought impulsively, as a response to social distress and disappointment. Along those lines, in 56% of the Dutch cases, social isolation or loneliness was reported[5]. Indeed, a recent study by one of the leading advocates of psychiatric PAD in Belgium, examining the explanations by 26 patients of their requests, reported frequent comments related to social isolation, interpersonal conflict, and socio-economic stresses – all potentially remediable and none usually considered good reasons for ending one's life[2].

Difficulty in applying the core eligibility criteria for PAD to psychiatric disorders may contribute to its use in questionable cases[8]. The Belgian statute, for example, requires that persons receiving PAD have "unbearable and untreatable" disorders[6]. Whether a condition is unbearable, is not easily susceptible to objective determination; there seems to be little alternative to taking the patient's assertion at face value. However, depression and other psychiatric disorders are often associated with hopelessness and helplessness that heighten subjective distress. Thus, the perceived intolerability of suffering may itself be a symptom of the underlying disorder, rather than reflecting an independent judgment of the patient. In any case, the criterion offers no real basis on which a psychiatrist can judge the reasonableness of a person's request for PAD.

Most of the work of determining whether an applicant qualifies for PAD, then, must be done on the basis of the requirement that the psychiatric disorder be "untreatable" (or in the Dutch law, that there be "no prospect of improvement"). Few patients will have tried every possible pharmacological, psychotherapeutic, or other treatment option (e.g., electroconvulsive therapy), and it is always difficult to judge whether some as-yet-untried approach might be helpful. However, PAD laws generally also stipulate that only treatments acceptable to the person seeking PAD should be considered in determining treatability. Thus, untreatability also becomes a subjective determination made by the person requesting PAD, who – perhaps in the grip of depressive hopelessness – can simply conclude that nothing is likely to work and thus no untried options are acceptable.

Although patients must be competent to request PAD, even the most skilled of psychiatric evaluators will find it difficult to ascertain the extent to which the patient is making a judgment independent of the influence of the psychiatric disorder itself. This is particularly true for depression, in which the desire to end one's life is a common manifestation of the disorder. Other than for flagrant psychosis, which seems barely represented in the cases reported to date, the competence requirement will provide little check on the use of PAD in psychiatry.

Jurisdictions considering adoption of PAD for psychiatric disorders would be well-advised also to consider the potential, less tangible impacts of legalization. Psychiatrists and other treaters may perceive PAD laws as offering "permission" to give up on treating difficult cases. It is not unimaginable that we will see frustrated psychiatrists and families under stress suggesting PAD to problematic patients as their only option. Likewise of concern is the implicit message communicated to patients when PAD becomes available, i.e., "there are hopeless conditions in psychiatry, and you may have one". Finally, one cannot ignore the temptation for countries with inadequate psychiatric care systems to look to PAD as a substitute for investment in appropriate treatment, especially for more challenging cases.

Taken as a whole, there appear to be ample reasons to conclude that adoption of PAD for psychiatric disorders is likely to yield more harm than good, a judgment reflected in the American Psychiatric Association's position that "a psychiatrist should not prescribe or administer any intervention to a non-terminally ill person for the purpose of causing death"[9].

Paul S. Appelbaum
Department of Psychiatry, Columbia University, and New York State Psychiatric Institute, New York, NY, USA

1. Appelbaum PS. Psychiatr Serv 2017;68:315-7.
2. Verhofstadt M, Thienpont L, Peters GY. Br J Psychiatry 2017;211:238-45.
3. Schuklenk U, van de Vathorst S. J Med Ethics 2015;41:577-83.
4. van der Heide A, van Delden JJM, Onwuteaka-Philipsen BD. N Engl J Med 2017;377:492-4.
5. Kim SYH, De Vries RG, Peteet JR. JAMA Psychiatry 2016;73:362-8.
6. Thienpont L, Verhofstadt M, Van Loon T et al. BMJ Open 2015;5:e007454.
7. Miller FG, Appelbaum PS. N Engl J Med 2018;378:883-5.
8. Aviv R. The death treatment: when should people with a non-terminal illness be helped to die? The New Yorker, June 22, 2015.
9. American Psychiatric Association. Position statement on medical euthanasia. https://www.psychiatry.org/psychiatrists/search-directories-databases/policy-finder?g=0cad609b-9bbe-4482-a945-21d1a7bb5dd3&Page=3.

DOI:10.1002/wps.20548

The nascent empirical literature on psychopathology and terrorism

The current status of empirical knowledge regarding the relationship between psychopathology and violent radicalization has undoubtedly improved from the initial forays into the study.

Work during the 1970s and 1980s focused upon personality traits and disorders, especially three that are found within DSM cluster B: borderline, narcissistic and antisocial. Poor research designs and a lack of valid empirics ultimately undermined such arguments. Various studies supporting psychopathic and personality-level explanations were conducted in the absence of rigorous clinical diagnostic procedures. Instead, they relied upon autobiographies, biographies, second-hand case studies, media interviews and willful misreadings of actual empirical work.

In the absence of rigorous clinical and empirical procedures, the reductionist view, where terrorists are characterized as suffering from some mental disorder purely on the nature of the attack behavior, ignores the highly complex neurological, psychological and sociological processes whereby actors become desensitized to violence, and subsequently suffer psychological consequences as a result of terrorist engagement.

Despite these methodological problems, the appeal of such efforts remains influential within the literature beyond their zenith in the 1970s and 1980s. For example, studies continue to hypothesize that terrorists are driven by envy, an urge to punish and retaliate, and a lack of empathy[1].

Following movements in wider psychiatric research, the study of the terrorist has also recently become more disaggregated, with empirical analyses focusing upon specific terrorist subsets (e.g., lone-actors, foreign fighters) rather than aggregate depictions (i.e., the general terrorist). Such analyses identify a mid-way point between the initial attributional studies that sought causation in psychopathology and social explanations which overlook the potential of psychopathology.

Such studies have found evidence for the presence of mental and personality disorders with various degrees of methodological sophistication. Some simply report aggregate prevalence rates of mental disorder diagnoses. Others disaggregate across mental disorders and compare to the societal base rate. One study of 140 Dutch foreign fighters and attempted foreign fighters found that 6% had diagnosed disorders. These included psychotic, narcissistic, attention-deficit/hyperactivity, schizophrenia, autism spectrum, and post-traumatic stress disorders. An additional 20% displayed indications of other undiagnosed mental health problems[2].

An investigation examining 153 lone-actor terrorists also noted a diverse range of disorders, including traumatic brain injury, drug dependence, schizophrenia, schizoaffective disorder, delusional disorder, psychotic disorder, depression, bipolar disorder, unspecified anxiety disorder, dissociative disorder, obsessive-compulsive disorder, post-traumatic stress disorder, unspecified sleep disorder, unspecified personality disorder, and autism spectrum disorders. The authors noted that schizophrenia, delusional disorder and autism spectrum disorders were more prevalent than in the general population[3].

Other studies examine statistical associations between disorder prevalence and specific behaviours and experiences. One investigation identified that lone-actors with a mental disorder are more likely to express violent desires, seek legitimization for their intended actions, stockpile weapons, train, carry out a successful attack, kill and injure, discriminate in their targeting, and claim responsibility[4].

The study of psychopathology and terrorism has traditionally focused upon those who conducted, or at least attempted to conduct, violence. Those studies that instead focus upon individuals who hold attitudinal affinity with such cases are growing. These studies further highlight the importance of examining personality alongside several other personal, situational and attitudinal measures.

A study of 52 teenagers in Gaza highlighted that depressive symptoms were common amongst supporters of "religio-political aggression"[5]. One investigation developed a radicalization scale that asked 16 questions regarding sympathies for violent protest and terrorism. Of the 608 UK-based participants, those most sympathetic were significantly more likely to also self-report depression and to see religion as important. Condemnation of violent protest and terrorism was associated with a greater number of social contacts, less social capital, and an unavailability for work due to housekeeping or disability. There was no significant difference in terms of generalized anxiety scores[6].

A European investigation deployed an extremist attitudes scale to 1,288 adolescents in Switzerland. Personal strain (which included personal stressors, negative life events and prior stays at a psychiatric hospital) was associated with significantly higher support for violent extremism, although this effect largely disappeared once other social and individual variables were included in the analysis. Those with poor coping skills were significantly more likely to support violent extremism. Self-reported low self-control had no impact upon violent extremism[7].

The above investigations have value, as they identify disorders and symptoms which often co-occur with specific experiences. However, "detailed research would be needed to further clarify the precise nature and role (if any) of mental health problems in the development of violent activity"[8]. In many cases, active symptoms may be present, but completely unrelated. Additionally, even symptoms of disorders that are associated with an increased risk of violence (e.g., substance use and active psychosis) may never give rise to an act of violence until they are combined with environmental factors that favor violence, in the context of a situational trigger.

Although this perspective is theoretically coherent, research is yet to empirically determine at which point the experience

of psychiatric symptoms is relevant to violent radicalization. Depending on circumstance, it may be a catalyst, an inhibitory factor, and even a consequence. To improve this knowledge gap, and move forward from unfounded causal assumptions, research must look to multiple avenues.

This may include, but is not necessarily limited to: a) sequence modelling to understand when the onset of disorders typically occurs in an individual's move to radicalization and violent action; b) clinical interviews with those at risk of radicalization as well as imprisoned terrorists; c) an exploration of how prominent symptoms were at the time of the violence and their relevance in decision-making; d) evaluations of psychologically-oriented interventions countering violent extremism; e) investigating the impact of living a terrorist lifestyle upon psychological functioning; and f) examinations of whether and how the presence of psychopathology impacts recruitment into terrorist co-offending networks.

Emily Corner[1], Paul Gill[2]
[1]Centre for Social Research and Methods, Australian National University, Canberra, Australia; [2]Department of Security and Crime Science, University College London, London, UK

1. Marazziti D, Veltri A, Piccinni A. CNS Spectr (in press).
2. Weenink A. Perspect Terror 2015;9:17-32.
3. Corner E, Gill P, Mason O. Stud Conflict Terror 2016;39:560-8.
4. Corner E, Gill P. Law Hum Behav 2015;39:23-34.
5. Victoroff J, Quota S, Adelman JR et al. Aggress Behav 2010;36:219-31.
6. Bhui K, Everitt B, Jones E. PLoS One 2014;9:e105918.
7. Nivette A, Eisner M, Ribeaud D. J Res Crime Delinquency 2017;54:755-90.
8. Horgan J. The psychology of terrorism, 2nd ed. Oxon: Routledge, 2014.

DOI:10.1002/wps.20547

What is the risk-benefit ratio of long-term antipsychotic treatment in people with schizophrenia?

Christoph U. Correll[1-4], Jose M. Rubio[1-3], John M. Kane[1-3]

[1]Zucker Hillside Hospital, Psychiatry Research, Northwell Health, Glen Oaks, NY, USA; [2]Hofstra Northwell School of Medicine, Hempstead, NY, USA; [3]Feinstein Institute for Medical Research, Manhasset, NY, USA; [4]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin, Berlin, Germany

The long-term benefit-to-risk ratio of sustained antipsychotic treatment for schizophrenia has recently been questioned. In this paper, we critically examine the literature on the long-term efficacy and effectiveness of this treatment. We also review the evidence on the undesired effects, the impact on physical morbidity and mortality, as well as the neurobiological correlates of chronic exposure to antipsychotics. Finally, we summarize factors that affect the risk-benefit ratio. There is consistent evidence supporting the efficacy of antipsychotics in the short term and mid term following stabilization of acute psychotic symptoms. There is insufficient evidence supporting the notion that this effect changes in the long term. Most, but not all, of the long-term cohort studies find a decrease in efficacy during chronic treatment with antipsychotics. However, these results are inconclusive, given the extensive risk of bias, including increasing non-adherence. On the other hand, long-term studies based on national registries, which have lower risk of bias, find an advantage in terms of effectiveness during sustained antipsychotic treatment. Sustained antipsychotic treatment has been also consistently associated with lower mortality in people with schizophrenia compared to no antipsychotic treatment. Nevertheless, chronic antipsychotic use is associated with metabolic disturbance and tardive dyskinesia. The latter is the clearest undesired clinical consequence of brain functioning as a potential result of chronic antipsychotic exposure, likely from dopaminergic hypersensitivity, without otherwise clear evidence of other irreversible neurobiological changes. Adjunctive psychosocial interventions seem critical for achieving recovery. However, overall, the current literature does not support the safe reduction of antipsychotic dosages by 50% or more in stabilized individuals receiving adjunctive psychosocial interventions. In conclusion, the critical appraisal of the literature indicates that, although chronic antipsychotic use can be associated with undesirable neurologic and metabolic side effects, the evidence supporting its long-term efficacy and effectiveness, including impact on life expectancy, outweighs the evidence against this practice, overall indicating a favorable benefit-to-risk ratio. However, the finding that a minority of individuals diagnosed initially with schizophrenia appear to be relapse free for long periods, despite absence of sustained antipsychotic treatment, calls for further research on patient-level predictors of positive outcomes in people with an initial psychotic presentation.

Key words: Long-term antipsychotic treatment, schizophrenia, benefit-to-risk ratio, efficacy, effectiveness, physical morbidity, mortality, metabolic disturbance, tardive dyskinesia, psychosocial interventions, non-adherence, dopaminergic hypersensitivity

(World Psychiatry 2018;17:149–160)

Schizophrenia is a disorder characterized by acute episodes often followed by symptom improvement[1]. Most guidelines recommend at least 1-2 years of antipsychotic treatment after symptom remission of an acute episode[2-5]. Of those discontinuing antipsychotic treatment, up to 75% have a relapse within 12 to 18 months[6,7]. Meta-analyses of 26 to 52 week studies comparing second-generation antipsychotics vs. placebo in the prevention of relapse found a very favorable number-needed-to-treat (NNT) of 3-5[8,9].

Risks of acute antipsychotic treatment, compared with placebo, mostly include weight gain, metabolic disturbance, QTc prolongation, neurologic adverse effects and sedation[10]. It is generally accepted that, given the usually moderate magnitude of these potential side effects and the availability of strategies to manage them, as well as the efficacy of antipsychotics in preventing relapse, antipsychotics have a favorable risk-benefit balance during the first 1-2 years following an acute psychotic episode[2-5,11].

Clinical guidelines do not provide systematic recommendations for treatment continuation or discontinuation beyond 1-2 years, yet they warn about the risks of relapse associated with treatment discontinuation[2-5,11]. The effects of antipsychotic treatment beyond the first 2 years of treatment are not well understood, given the lack of double-blind, placebo-controlled randomized trials (RCTs)[9].

There has been an emerging body of literature on the long-term effects of antipsychotics questioning their necessity[12-15]. Long-term animal studies of antipsychotic exposure[16], naturalistic cohorts[14,15], and treatment discontinuation studies[13] have been cited by some authors who claim that antipsychotics do not improve outcomes in the long term, and that there may even be iatrogenic adverse consequences of long-term antipsychotic treatment[17]. Others suggest that there is insufficient evidence supporting iatrogenic effects[18]. Such debate, and the uncertainty in the interpretation of long-term studies, with inherent biases[12,19], results in unclear recommendations for clinicians.

In this paper, we review the literature on the potential risks and benefits of long-term antipsychotic treatment, summarizing the evidence of efficacy, effectiveness, tolerability, physical morbidity and mortality, as well as functional and structural brain changes associated with that treatment. Additionally, we review the role of interventions to optimize such risk-benefit ratio.

EFFICACY, EFFECTIVENESS AND TOLERABILITY

The longer the study, the more likely that systematic error accumulates over time and biases the results. Measurements tend to prioritize feasibility over reliability; the intervention is less controlled due to

greater influence of environmental factors; and there is greater chance of systematic or non-random drop-outs differing between the arms of the trial.

Hence, the interpretation of the results should consider how each one of these potential biases affects the study. Interpretation should also consider the literature, not isolated studies. Here, we summarize the available data separately for different methodological approaches, as all have their own strengths and limitations[20-22].

Treatment adherence and long-acting injectable antipsychotic studies

The longer the treatment, the greater the chance of insufficient adherence[9,23,24]. Data from administrative claims in the US suggest that, in clinical practice, patients with psychosis treated in an outpatient setting fill their prescriptions an average of 40-60% of the days prescribed[25]. Adherence studies find that poor mid-term adherence ranges from 11.6% based on self-report to 58.4% in studies using serum concentration[23]. In addition to high rates of insufficient adherence[24], we lack practical/reliable measures of exposure[26].

In a systematic review and meta-analysis of longitudinal studies examining relapse and its risk factors in patients following stabilization after a first psychotic episode, non-adherence was found to be the greatest predictor of relapse among twenty variables in seven long-term studies, increasing the chance of relapse by 400%[27]. Individuals in another study with non-adherence for >1 month of an 18-month follow-up had a five-fold greater chance of relapse than individuals with continuous treatment[28].

Poor adherence was also found to explain up to 36% of the effect of cannabis on the number of relapses[29]. Individuals with suboptimal adherence were found to have greater body mass index and were less likely to live in independent housing than individuals with continuous adherence over 18 months. The magnitude of these risk factors was small to moderate, with a 2% greater likelihood of being non-adherent for each point of increase in body mass index, and a 25% greater likelihood of being adherent in individuals living independently. In this study, no other undesired outcomes were associated with adherence status[30].

Long-acting injectable (LAI) formulations have also provided meaningful data. When LAIs and oral formulations were compared in RCTs, no overall difference was found regarding relapse prevention in the mid term after stabilization[31]. This is not surprising, given that the control groups taking oral medication in these RCTs tend to include patients with better treatment adherence and lower illness severity. Non-adherence levels did not differ across ten meta-analyzed trials with adherence data ($p=0.27$)[31].

When the same question was addressed by meta-analyzing mirror-image studies, where each research participant acts as his/her own control, LAI treatment phases, compared to those with oral antipsychotics, were associated with a significantly 57% lower risk of a next hospitalization and a 62% reduced risk of number of hospitalizations[32]. This is not simply the result of the order of the oral and LAI phases, as two trials confirmed that the reverse switch (i.e. from an LAI to an oral antipsychotic) was associated with poorer outcomes for the oral phase[33,34].

The finding of greater effectiveness of LAIs in mirror image studies was replicated in a meta-analysis of cohort studies, where the number of hospitalizations was reduced by 15% (14 studies; 60,260 person-years), despite greater illness severity in the LAI cohorts than the oral antipsychotic treatment cohorts ($p=0.014$)[35]. Results were particularly apparent in Scandinavian registries, that have fully generalizable national samples. In a Finnish national cohort, individuals treated naturalistically with LAIs after their first hospitalization for a schizophrenia episode had one third the risk of re-hospitalization than individuals on oral counterparts of the same antipsychotics[36]. This was replicated in a Swedish cohort including all phases of illness, following patients for a median of 6.9 years. Six of the top eight antipsychotic monotherapies that were significantly superior regarding hospitalization risk compared to not receiving any antipsychotic (hazard ratios, HRs=0.51-0.64) were LAIs (with the two oral antipsychotics being clozapine and olanzapine)[37].

In a meta-analysis that compared adverse effects with LAIs vs. the same oral antipsychotics across sixteen RCTs with a mean duration of one year, those preparations did not differ regarding 115 (96.6%) of the 119 reported adverse effects[38]. LAIs were more likely to present with akinesia, low-density lipoprotein cholesterol change and anxiety, whereas oral antipsychotics were associated with greater hyperprolactinemia. Furthermore, there were no differences regarding treatment discontinuation due to side effects and mortality[38]. Little is known, however, about differences in adverse events beyond one year of treatment.

Overall, assuming that the main advantage of LAI over oral antipsychotics is lower risk of non-adherence, this literature supports the relationship between suboptimal adherence in the long term and greater risk of relapse[27,39], while differences in adverse effects are small within the time span of one year.

Placebo-controlled antipsychotic maintenance treatment studies

Methodologically, placebo-controlled maintenance RCTs have the advantage of minimizing systematic differences between groups, yet their time frame is only mid-term (i.e., 1-3 years following stabilization), and their results assume full long-term adherence with antipsychotics (which is known to decrease over time[24]). Increasing non-adherence even in RCTs could lead to finding lower effect sizes in studies of longer duration.

A meta-analysis of 65 placebo-controlled maintenance RCTs found an overall NNT of 3 favoring antipsychotics over placebo in preventing relapse, but overall treatment effects tended to decrease as a function of study duration[9]. The proportion of individuals unimproved/worse was lower on antipsychotics, but this

difference decreased over time and was non-significant in the longer-term studies.

Supporting the hypothesis that increasing non-adherence on antipsychotics could decrease antipsychotic maintenance efficacy, the authors found a significantly greater relapse preventive effect (p=0.03) in studies comparing LAIs vs. placebo (HR=0.31) than oral medications vs. placebo (HR=0.46). In LAI studies, non-adherence could be identified and non-adherent patients were discontinued or excluded from the analyses[9].

The number of patients with at least one adverse effect did not differ between antipsychotics and placebo, and did not increase over time for individuals on antipsychotics. No differences were observed in sedation, although weight gain and at least one movement disorder were significantly more frequent during antipsychotic treatment[9].

Long-term cohort studies

Few placebo-controlled RCTs of antipsychotics last >3 years, with most lasting ≤1 year[9]. Most data beyond this initial period are derived from non-randomized, non-controlled cohort and register studies. These have the advantage of providing long-term data, not requiring consent and being highly representative of the overall population. However, given the lack of randomization and controlled intervention, subgroups are subject to various types of selection biases, and conclusions are tentative.

Non-randomized cohort studies often found that, at follow-up, individuals on antipsychotics had equal or greater illness severity compared with those off antipsychotics. For example, in the Suffolk county cohort, 175 individuals with schizophrenia showed a clinical decline over the 20-year follow-up period[40]. This decline occurred despite high and constant rates of antipsychotic prescription (86.9% at baseline and 81.8% 20 years later), and antipsychotic use was associated with worse Global Assessment of Functioning (GAF) scores and negative symptoms, yet lower disorganization and excitement[40]. In the Chicago cohort, which followed 70 individuals with schizophrenia from early illness for 20 years, 8% of the 15 unmedicated individuals had some degree of psychotic symptoms, versus 68% of the 25 individuals treated continuously with antipsychotics[14]. In the Northern Finland 1966 Birth Cohort, which followed patients for almost 20 years, those who were off antipsychotics were more often in remission, and no differences in remission rates between treatment groups were found[41,42]. Similarly, the OPUS cohort in Denmark found that, among the 90% of the individuals who did not have sustained remission 10 years after their first episode, more were on than off antipsychotics[43,44].

Nevertheless, in those non-randomized, uncontrolled studies, adherence levels to antipsychotic treatment are unknown, and most importantly, there is a high risk of confounding by indication and reverse causation, in that greater illness severity could be the cause of continued antipsychotic treatment, rather than being the effect. Interestingly, different results were found in a retrospective cohort study of individuals with schizophrenia whose access to antipsychotic treatment had been restricted. In this cohort from rural China, those who had access to antipsychotics did substantially better after 14 years than those without access[45].

Thus, despite the pattern of patients with worse outcomes being overrepresented in the treatment groups of several cohort studies, the interpretation regarding cause and effect is difficult, and reverse causation cannot be excluded.

On the other hand, results from large, national samples analyzed with statistical methods to adjust for baseline differences support the notion that treatment failure and hospitalization[37], as well as mortality risk from suicide[46,47], are significantly greater in patients not receiving antipsychotics than in those who are.

Dose-reduction and dose-discontinuation studies

Dose-reduction and dose-discontinuation studies (DRDD) evaluate outcomes associated with these treatment strategies compared with long-term continuation of antipsychotic treatment. DRDD studies often have the advantage of a longer time span than antipsychotic maintenance trials, yet with greater degree of randomization and control than naturalistic cohort studies.

Wunderink et al[13] conducted the study with the longest follow-up period to date, consisting of two phases. In the first phase, 131 individuals with a first episode of psychosis were allocated to 2 years of either symptom guided DRDD or treatment continuation[48]. The initial goal of stopping antipsychotic treatment in the DRDD group was changed to dose reduction only, due to too many relapses after antipsychotic discontinuation. In the second phase, 103 individuals were evaluated once after 5 years of uncontrolled community treatment[13]. In the initial RCT, the DRDD group had twice as many relapses as the maintenance group (43% vs. 21%, p=0.011), although about 20% were able to successfully stop the medication without relapses. There were no differences in symptom severity, both groups having low Positive and Negative Syndrome Scale (PANSS) scores throughout[48]. At 5 years, there were no differences in relapse rates or symptom severity. However, recovery rates were twice as likely in the initial dose-reduction group (40.4% vs. 17.6%, p=0.004), driven not by symptomatic remission (69.2% vs. 66.7%, p=0.79), but by functional remission (46.2% vs. 19.6%, p=0.01), and 8 of the 11 patients off antipsychotics for 2 years were in the original dose-reduction condition. These results have been cited as important evidence that antipsychotics could postpone rather than prevent relapse, while impacting negatively on functional recovery in the long-term[12,14,15,17,19].

These findings should be interpreted with caution. As the authors acknowledge, the participants had very low symptom severity. Their conclusions might not apply to more severely ill patients. Also, the difference in antipsychotic exposure between the two groups was only questionably clinically meaningful (1.4 mg/day of haloperidol equivalents), without significant differences in months per

patient without antipsychotic prescription. Less than 50% of the sample approached for the original RCT agreed to participate, and only 43.7% of the patients at baseline were diagnosed with schizophrenia[48]. Therefore, it remains possible that the results were related to factors other than the 2-year intervention (i.e., DRDD or antipsychotic maintenance dose continuation), which was followed by 5 years of uncontrolled community care, especially given the small dose differences between treatment arms at 7 years. The lack of blinded assessment and reverse causation could also have influenced the results.

Antipsychotic dose reduction vs. standard maintenance dose has also been examined in other studies with shorter follow-up. In a meta-analysis of 13 trials with follow-up of 24 and 104 weeks (11 trials lasting $\geq$1 year), Uchida et al[49] found no differences between low antipsychotic dose (50-100% of the defined daily doses[50]) and standard antipsychotic dose, with respect to overall treatment failure (p=0.53) or hospitalization (p=0.40). Yet, very low dose (<50% of the defined daily doses[50]) were associated with greater risk of hospitalization (p=0.002) and relapse (p=0.0004). In a pilot study, cognitive symptoms were significantly improved when the antipsychotic dose was reduced to 50% of the defined daily dose[51].

A more recent uncontrolled discontinuation study with an intermediate follow-up period found greater rates of symptom recurrence and lower functional status in 46 individuals who had recovered from a single psychotic episode and who had opted to being treated with DRDD compared to 22 patients who had opted for continuation of antipsychotic treatment for 3 years[52].

Comments

There is a trade-off of strengths and weaknesses between study designs, with generally greater chance of bias in longer-term studies and, especially, uncontrolled studies in which more symptomatic and impaired patients are more likely to receive long-term antipsychotic treatment. There is consistent evidence, though, supporting the efficacy of antipsychotics in preventing relapse in the mid term (i.e., 1-3 years) following stabilization. These data come from studies of adherence, trials of LAIs, national registries, placebo-controlled maintenance trials and DRDD trials.

Most, but not all, of the studies with follow-up >3 years reported worse outcomes associated with continued antipsychotic use. However, these results are inconclusive, given small and selective patient samples and extensive risk of bias[13-15]. Conversely, long-term register studies of much larger and representative national cohorts of patients diagnosed with schizophrenia confirmed significantly less treatment failure and suicide-related mortality in antipsychotic-treated patients compared to those not treated with antipsychotics[37,46,47].

In conclusion, there is a strong evidence supporting mid-term efficacy, and a lack of convincing evidence against long-term efficacy of antipsychotic treatment.

PHYSICAL MORBIDITY AND MORTALITY

Schizophrenia is associated with a well-established excess of physical morbidity and premature mortality, while antipsychotics are associated with cardiovascular risk factors[53-60].

Individuals with schizophrenia have a greater prevalence of sedentary lifestyle, obesity, cardiovascular illness, diabetes, nicotine smoking and tobacco-related disorders, sexually transmitted diseases, obstetric complications, and altered pain sensitivity[61,62], while also having lower rates of health care services utilization and medical treatment for such conditions, which results in large unaddressed gaps in medical care[63]. While it is unclear the role that differences in health care systems play in physical morbidity in schizophrenia, given the limited availability of comparable data from a variety of countries[61], it seems clear that this morbidity plays an important role in reducing the life expectancy of individuals with schizophrenia across different settings.

A recent systematic review and meta-analysis including 11 studies from various countries found a weighted mean decrement in life expectancy of 14.5 years in patients with schizophrenia, with significant variations depending on gender and country[64]. While overall life expectancy has recently increased in developed countries, it is concerning that patients with schizophrenia appear not to have benefited from such improvements, so that the mortality gap affecting these patients has increased[65]. The drivers of this excess mortality seem to be poor physical health and decreased health care service utilization in patients with schizophrenia[66,67].

In the US, natural causes account for a vast majority of deaths, with only 1/7 related to unnatural causes (accidents, suicide or homicide). Chronic medical illness associated with smoking, obesity and a sedentary lifestyle account for most of the variance in premature mortality. These results seem to vary across countries, likely reflecting public health characteristics. A 10-year longitudinal study in Ethiopia found that premature mortality was double in patients with schizophrenia, with infectious diseases accounting for almost half of the causes of premature death, and with a greater role of suicide in premature mortality[68,69]. A similar pattern has been found in other developing countries[70,71].

The metabolic and cardiovascular side effects of long-term antipsychotic treatment have been a source of concern as possible contributors to the increase of physical morbidity and premature mortality, especially in developed countries where most of the mortality in schizophrenia is related to consequences of metabolic disturbance and cardiovascular disease[55,56,72]. While the metabolic consequences of long-term antipsychotic treatment are widely appreciated[53,54,57,58,60], the understanding of their contribution to morbidity and mortality in schizophrenia has evolved over the last several years.

There has been a growing literature identifying health care service utilization

patterns in schizophrenia associated with worse outcomes. In a national Swedish cohort, individuals with schizophrenia were less likely to have received a diagnosis of cancer or ischemic heart disease at the moment of dying of these causes[73]. These data suggest poor prevention and early treatment of medical conditions. In another sample, individuals with schizophrenia diagnosed with cardiovascular illness were less likely to use lipid-lowering and anti-hypertensive medication, which was altogether associated with worse outcomes[74]. To what extent antipsychotic treatment moderates the association between schizophrenia and poor health care utilization is not yet well understood.

The role of antipsychotics in reducing premature mortality in schizophrenia has been better characterized. Despite antipsychotic treatment elevating cardiovascular risk factors, long-term treatment is consistently associated with lower mortality rates compared to no long-term treatment[46,47,75-77], but still higher rates than in individuals without schizophrenia[46].

National registries constitute the best approach to study the relationship between long-term antipsychotic treatment and all-cause mortality as well as mortality related to cardiovascular illness, given the availability of cumulative dose data. In a seminal study, Tiihonen et al[47] found that, compared to individuals with schizophrenia not receiving antipsychotic treatment, those with longer antipsychotic treatment had greater decrements in premature mortality, including from cardiovascular causes[47]. Given the possible survivor bias, the same group studied the role of cumulative antipsychotic dose over a 5-year period in influencing mortality in schizophrenia adjusting for an extensive number of variables. They found in a separate sample that all – low, moderate and high – antipsychotic cumulative doses were associated with lower mortality rates than no antipsychotic use. Patients with schizophrenia with low and moderate – but not high – cumulative doses of antipsychotics had lower rates of mortality due to cardiovascular disease, whereas those with high – but not moderate or low – doses had low mortality rates due to suicide[46].

Beyond these individual findings, a recent meta-analysis found a consistent association of antipsychotic use and decrement in all-cause mortality, with some evidence of a dose effect[75]. The seeming disconnect between adverse antipsychotic cardiovascular effects in short- and longer-term studies and reduced (or, at least, not elevated) all-cause and cardiovascular illness-related mortality in long-term database studies may be explained by a beneficial link between improved psychiatric symptom control and improved healthy lifestyle behaviors as well as access to medical care[78].

Despite being consistent, these register-based findings should not be interpreted as clearly establishing a causal relationship between long-term antipsychotic treatment and reduced all-cause mortality, given the limitations of observational studies. However, national registries, despite their exposure to potentially unmeasured confounders, currently constitute the most adequate method to assess the long-term effects of antipsychotics on morbidity and mortality. Future research should improve their design by adjusting analyses for relevant potential confounders that have not been measured (e.g., body mass index, metabolic values, psychiatric illness symptom severity, and functionality).

Comments

Individuals with schizophrenia have significantly greater physical morbidity and premature mortality than the general population. While this finding is related to unhealthier lifestyle and lower health care service utilization, the role of antipsychotics is less clear. Long-term antipsychotic treatment is associated with significantly greater rates of metabolic and cardiovascular risk factors and disease, yet patients treated with antipsychotics over the long-term seem to have significantly lower mortality rates, including death due to cardiovascular disease, at low and moderate doses, compared to individuals with schizophrenia not receiving antipsychotics. This finding has been replicated with large effect sizes in various national registries, adjusting for an extensive number of potential confounders, and with some evidence suggesting a time and dose effect.

Though these data are limited by their observational nature, they are consistent enough to provide support for a favorable risk-benefit balance for the long-term use of antipsychotics in schizophrenia in reducing mortality.

BRAIN STRUCTURE AND FUNCTIONING

Schizophrenia has been associated with various brain volumetric abnormalities since the emergence of neuroimaging[79]. However, the nature and clinical relevance of these findings still remain unclear[80], and even less so the role of antipsychotics[18]. The cortical and subcortical regions found to have lower volume in schizophrenia have most frequently been the anterior cingulate cortex, insula, hippocampus, and thalamus[81,82], although several other areas have been implicated, with variability across studies probably due to methodological differences.

Never treated patients with chronic schizophrenia show a significantly accelerated decline in prefrontal and temporal cortical thickness[83], suggesting a neurodegenerative illness course. Reduced hippocampal and thalamic volumes have been observed in individuals at high risk of developing psychosis[84]. High-risk individuals who transitioned to psychosis presented with further progression of the whole brain volume reduction, even before antipsychotic treatment[85], and reductions in brain regions, such as the anterior cingulate, have been identified as potential biomarkers indicative of greater risk of transition to psychosis[86]. Despite grey matter reduction being a consistent finding, what this means at the neuropathological level is unclear[87-91].

Brain tissue loss is a non-specific finding, observed with antipsychotic exposure[92], changes in body weight[93], alcohol use[94,95], and steroid use[96]. Volumet-

ric changes in drug-naïve patients do not seem to be correlated with clinical impairment or duration of illness, not supporting a neurodegenerative hypothesis[83-86,97]. A more recent perspective is that volumetric reductions reflect a reduction of neuropil[80], and that volumetric variations can be heterogeneous in schizophrenia, although decrements in specific regions, such as the anterior cingulate cortex, might be more homogeneous and therefore more specific to that disorder[98].

A generalized decrement of grey matter volume associated with antipsychotic treatment duration and cumulative doses has been repeatedly reported[92,99]. However, these studies are limited by the fact that the duration and cumulative dose of antipsychotics can be a marker of illness severity or illness duration, making it difficult to distinguish a reduction due to illness severity, illness duration or antipsychotic exposure. In a meta-analysis of longitudinal studies, the grey matter decrement was directly related to the cumulative dose of first-generation antipsychotics during the window of observation, whereas the opposite was true for second-generation antipsychotics[97]. This finding is difficult to interpret and, as acknowledged by the authors, may in part be due to confounders, such as weight gain associated with second-generation antipsychotics.

Other findings contradict the notion that antipsychotics cause a decrement in grey matter in schizophrenia. The ENIGMA neuroimaging consortium found that, among 2,028 patients, antipsychotic-naïve individuals had greater volumetric deficits in the hippocampus compared with antipsychotic-treated ones[100], whereas thalamus and basal ganglia volume deficits in untreated patients have been found to be corrected with antipsychotic treatment[92,100]. A longitudinal study comparing grey matter volumes before and after initiation of antipsychotic treatment in first-episode patients found that antipsychotics minimized these decrements, particularly in the striatum[101]. Another study of patients who were stabilized on antipsychotic treatment and allocated to either antipsychotic maintenance or antipsychotic withdrawal found that after one year there were no differences in volumetric parameters between the two groups[102].

Brain volume reductions need to be interpreted within the context of the effects of untreated psychosis and of clinical outcome findings. The reanalysis of a study that had raised considerable concern about the potential dose-dependent adverse effect of antipsychotic treatment on brain tissue loss[103] revealed that the duration of psychosis had a 3-fold greater detrimental effect on total brain and frontal lobe grey matter loss compared to the duration of antipsychotic treatment[104]. Furthermore, brain volumetric changes do not seem to correlate with poor clinical response or outcomes. In patients treated with clozapine, both a grey matter decrement and a clinical improvement have been reported[105], whereas in other studies the opposite was found[106].

Moreover, measuring volumetric brain changes during antipsychotic treatment without assessing functional brain status confuses the discussion. A cross-sectional study in 23 antipsychotic-treated and 21 untreated first-episode patients found significant cortical thinning within the former group in the dorsolateral prefrontal and temporal cortex. However, the medicated patient group showed significantly higher dorsolateral prefrontal cortex activation and significantly better cognitive performance than the unmedicated group[107].

Thus, the evidence does not seem to support a causal or detrimental relationship between long-term antipsychotic use and clinically relevant brain volumetric changes, with some data even suggesting that brain volume reductions could be associated with better brain network integration.

Contrary to the ambiguous literature on structural changes with chronic treatment, findings on functional changes have been more consistent. Long-term antipsychotic treatment has been associated with an increase in the number and affinity of dopamine D2 receptors, which results in a state of dopaminergic supersensitivity, and has been replicated in animal[16,108] and human models[109]. Tardive dyskinesia is a clinical consequence of long-term antipsychotic use that has been associated with dopaminergic supersensitivity[110], but also other possible mechanisms[111], and with greater risk in genetically vulnerable populations[112].

The estimated risk of tardive dyskinesia with first-generation antipsychotics is 3-5% per year of exposure (at least for the first 5 years)[113], being lower with second-generation antipsychotics[114]. Early parkinsonism and higher antipsychotic doses have been associated with this side effect[115]. A recent meta-analysis estimated a global mean prevalence of 25% in patients with schizophrenia treated with antipsychotics, with great variability depending on geographical and treatment-related factors[115].

Some studies reported that patients with tardive dyskinesia are at greater risk of rebound psychosis upon antipsychotic withdrawal[116], development of treatment resistance[117], and physical morbidity and mortality[118], although these results have not been consistently replicated[119]. The degree to which chronic antipsychotic exposure plays a role in these potential outcomes associated with tardive dyskinesia (i.e., whether, beyond causing that side effect, chronic antipsychotic treatment has a causal role in these outcomes) is not well understood[120].

Second-generation antipsychotics should be first-line maintenance treatment agents to decrease the risk of tardive dyskinesia. Two agents, valbenazine and deutetrabenazine, have been recently approved in the US for the treatment of this side effect of antipsychotic treatment, having shown moderate to high efficacy[121,122].

Following the hypothesized mechanism underlying tardive dyskinesia, dopamine supersensitivity related psychosis either during antipsychotic treatment or upon antipsychotic discontinuation has been a theoretical concern[117,123]. The hypothesis is that chronic dopaminergic blockade resulting in dopamine D2 receptor upregulation and dopaminergic hypersensitivity in the mesolimbic pathway may increase the risk of relapse and reduce antipsychotic efficacy in the long term.

Dopamine supersensitivity psychosis was first described in a series of ten case reports of patients who had abrupt onset of psychosis upon the discontinuation of antipsychotic treatment[124]. The existence of this phenomenon has been controversial and only supported by small studies[125]. Nevertheless, there has been a recent resurgent interest in dopamine supersensitivity as a potential cause of the emergence of treatment resistance [123,124,126,127]. However, a meta-analysis of RCTs found no differences in relapse rates between abrupt and gradual antipsychotic withdrawal or between different antipsychotic doses prior to discontinuation[9]. Moreover, if dopamine hypersensitivity were a major reason for the lack of long-term efficacy, then the partial D2 agonist aripiprazole, which has not been associated with upregulation of dopamine D2 receptors, at least in adult animal models[128], should be associated with significantly lower relapse rates than full dopamine D2 antagonists, but there are no data to support this[129,130].

Comments

Overall, tardive dyskinesia is the clearest adverse clinical consequence in brain functioning of long-term antipsychotic treatment, which may be related to dopamine supersensitivity in a subgroup of vulnerable individuals. This risk should be evaluated when considering long-term antipsychotic treatment, and preventive strategies utilized. In addition, patients should be examined before initiating treatment to determine the presence of preexisting abnormal involuntary movements.

Other effects of long-term antipsychotic treatment on brain structure and function, particularly neuropathological changes and the risk of dopamine supersensitivity psychosis, are insufficiently substantiated. The current literature does not provide consistent evidence to support irreversible functional and structural brain changes as a consequence of long-term antipsychotic treatment other than tardive dyskinesia.

THE ROLE OF PSYCHOSOCIAL STRATEGIES IN MODIFYING THE RISK-BENEFIT RATIO OF ANTIPSYCHOTICS

While symptom reduction and response, as well as relapse prevention, are relevant outcomes, functional recovery is a preeminent goal of treatment in schizophrenia[39]. Unfortunately, when using criteria based on both clinical and social domains, recovery rates in schizophrenia have remained low, with a meta-analytically derived median of 13.5% across five decades, without improvement over time (although only two studies contributed data to the last decade)[131]. While, in an aforementioned meta-analysis[9], antipsychotic maintenance treatment was superior to placebo in preventing relapse with an NNT = 3, employment rates did not differ, pointing toward the need for psychosocial interventions to achieve improved functional outcomes.

A recent meta-analysis found a significant small to medium association between clinical outcomes and personal recovery, but psychotic symptoms – which are the main target of antipsychotic medications – showed a smaller correlation than affective symptoms with personal recovery[132]. These data underscore that antipsychotics alone are insufficient and that adjunctive multimodal psychosocial treatments are needed to help stabilized patients achieve personal recovery goals[133].

The Schizophrenia Patient Outcomes Research Team (PORT)[134] reviewed the evidence supporting a wide variety of psychosocial interventions for the long-term treatment of schizophrenia. The committee recommended eight psychosocial interventions with various indications and for different populations. Of these, cognitive behavioral therapy (CBT) was specifically recommended, with evidence supporting its efficacy in reducing positive, negative and overall symptoms in individuals treated with antipsychotic drugs[135]. While one of the goals of CBT is psychoeducation on antipsychotic drug adherence, the efficacy of CBT in improving this outcome has been inconclusive[136].

Interestingly, the evidence supporting the efficacy of CBT in reducing psychotic symptoms in individuals not taking antipsychotic medication[137], or individuals whose symptoms fail to respond to antipsychotic treatment[138,139], has been more consistent. This finding suggests that the impact of CBT goes beyond improving adherence with antipsychotic medications, having an antipsychotic effect on its own. However, to our knowledge, there have not been head-to-head comparisons of CBT with long-term antipsychotic dose reduction strategies that would provide data about CBT as a partial or total substitution for long-term antipsychotic treatment[139].

Family-based psychosocial treatments were another of the interventions recommended by the Schizophrenia PORT, with evidence for reducing relapses and rehospitalizations, and improving treatment adherence[134]. These interventions are based on psychoeducation, and are not generally conceived as partial or total alternatives to antipsychotics, but rather as augmentation. In a large Chinese study that randomized first-episode patients to antipsychotic treatment alone or augmented with family interventions for one year, those in the augmentation arm were less likely to discontinue antipsychotics, showed greater improvements in insight, social functioning and activities of daily living, as well as access to employment or education[140]. These results have been substantially replicated[141]. In a trial that compared family interventions augmenting regular or reduced antipsychotic dose, those treated with low-dose antipsychotics and family therapy were more likely to relapse than those with family therapy and regular antipsychotic dose[142].

More recently, the Recovery After an Initial Schizophrenia Episode - Early Treatment Program (RAISE-ETP) study tested the feasibility and effectiveness of the integration of various psychosocial and pharmacological interventions in the treatment of 404 first psychotic episode patients in 34 community clinics across the US[133]. This study compared coordinated specialty care (which included CBT-based psychotherapy, family education and support, supported education and/or employ-

ment, and guided pharmacotherapy) with treatment as usual, showing superiority of the former in improving quality of life, increasing time in education or at work, and reducing symptom severity[133]. Because pharmacotherapy also differed between the two compared conditions, it is difficult to draw firm conclusions regarding effects of specific modalities. However, it seems unlikely that the psychosocial interventions included in coordinated specialty care could serve as substitute to medications, rather than as an effective augmentation strategy, given the lack of differences in the antipsychotic dose used between the two arms[143].

While psychosocial interventions seem effective augmenting strategies, rather than partial or total alternatives to antipsychotics, they can help improve the long-term risk-benefit ratio of antipsychotics by improving symptomatic and psychosocial outcomes and by reducing the risk of cardiometabolic side effects. A meta-analysis of various non-pharmacological interventions, ranging from healthy lifestyle and behavioral interventions to CBT-based psychotherapies, demonstrated their effectiveness in significantly reducing body weight, body mass index and serum lipids associated with antipsychotic use[144]. Some of these advantages persisted over time. Unfortunately, challenges in engagement limit the effectiveness of these interventions[145,146].

Comments

Psychosocial interventions are effective augmentation strategies for the treatment of schizophrenia, particularly CBT-based interventions, which seem to have antipsychotic effects independent of improving antipsychotic adherence. These interventions can be effectively implemented beyond academic centers.

Evidence suggests that psychosocial interventions can improve the long-term risk-benefit ratio of antipsychotics by improving functional, recovery-focused outcomes and by decreasing the burden associated with antipsychotic treatment, rather than by necessarily allowing a decrease in antipsychotic doses.

INDIVIDUAL DIFFERENCES IN THE RISK-BENEFIT RATIO OF LONG-TERM ANTIPSYCHOTIC TREATMENT

While the diagnosis of schizophrenia has been associated with poor outcome and need for long-term antipsychotic treatment, the heterogeneity in response and illness course has resulted in calls to broaden the view towards a psychosis syndrome with variable outcome patterns[147,148]. Some studies suggest that a minority of patients could potentially discontinue antipsychotic treatment without risk of relapse. The literature indicates that this would apply to between 4% and 30% of the patients that are stabilized after an acute episode[43,48,52,149,150].

This variable range likely reflects heterogeneity in the studied populations, criteria for diagnosis and relapse, duration of follow-up, and exposure to non-pharmacologic interventions. Therefore, we need better epidemiological data and predictors of successful antipsychotic discontinuation in patients presenting with a psychotic syndrome consistent with a diagnosis of schizophrenia. Some studies have identified abrupt onset and older age, female gender, higher GAF scores, working, having a partner, living independently and the absence of substance abuse as significant predictors of better outcomes[43,149], whereas others have not been able to find any significant predictors[52].

A more consistent observation, however, is that previous successful antipsychotic withdrawal predicts successful withdrawal during follow-up[13,43,48,149]. This finding indicates that a minority of individuals with a psychotic syndrome fulfilling criteria for schizophrenia can successfully discontinue antipsychotic treatment, and the risk of relapse probably decreases as they move past a critical high-risk period for relapse. However, to date, there is no reliable evidence-based method to identify such individuals.

This question, however, may benefit from research that is being conducted aimed at patient-level prediction of treatment response. A wide range of predictors have been recently identified, involving genetic[151] and neuroimaging[152-154] perspectives. Also, individual risk scores based on clinical variables have been developed to predict transition from clinical high risk for psychosis to suprathreshold psychosis[155], and future research could develop similar models to predict treatment response. At present, despite some promising findings, the field is not ready to apply patient-level predictors of antipsychotic response in real-world care[156]. Future research should equally address the development of prediction models for successful treatment discontinuation.

Comments

To date there is no evidence-based strategy that enables us to identify individuals who would benefit from antipsychotic dose reduction or discontinuation with minimal increase in relapse risk. Future research should capitalize on the recent advances in patient-level predictors of treatment response in order to identify these low-risk individuals.

CONCLUSIONS AND RECOMMENDATIONS

Overall, antipsychotic maintenance treatment should be recommended for the mid term (i.e., 1-3 years), since there is strong evidence supporting efficacy of antipsychotics in reducing relapses over this time frame. Data on long-term outcomes are more equivocal and, although the effect of antipsychotics seems to decrease over time, this could be an artifact of long-term study designs. Increasing non-adherence and reverse causation may play a significant role in the observed time trends, while alternative hypotheses, including dopamine supersensitivity psychosis, are less well substantiated.

Additionally, mortality and neuropathological findings do not support an accrual of damage from cumulative antipsychotic dose and duration (with the exception of tardive dyskinesia). On the contrary, long-term antipsychotic main-

tenance treatment has consistently been associated with lower all-cause and specific-cause mortality compared to antipsychotic discontinuation in large national and representative samples of patients with schizophrenia.

Despite lack of long-term randomized, placebo-controlled trials and residual uncertainty regarding a subgroup of patients who fulfill criteria for schizophrenia and who may only suffer one single psychotic episode, it is reasonable to recommend antipsychotic treatment in the long term (i.e., >3 years), although with several additional suggestions. Continued antipsychotic treatment with ≥50% of the standard defined daily dose should be implemented (going below such doses increases the risk of relapse). LAIs should be prioritized to minimize breaks in treatment adherence, or to at least make them known, allowing for additional interventions to continue adequate treatment. Second-generation antipsychotics should be preferred over first-generation ones to minimize the risk of tardive dyskinesia. Psychosocial interventions, particularly CBT and family-based interventions, are useful as augmentation, even when there are residual or treatment resistant symptoms, yet these therapies are not a substitute for antipsychotic treatment. Some behavioral interventions can also be used to reduce some of the negative impacts of continued antipsychotic treatment (i.e., metabolic side effects).

In patients who have achieved successful antipsychotic discontinuation for <1 year, close monitoring is recommended, keeping in mind that only a minority of patients can successfully discontinue antipsychotics. There are no evidence-based methods to identify individuals who may be managed successfully with antipsychotic doses <50% of standard antipsychotic doses, or who can safely discontinue antipsychotics. Therefore, the recommendation to continue long-term treatment applies to patients in general. While it is recognized that shared decision making is relevant, clinicians should use the available evidence and discuss the risks of the illness and relapse-related biopsychosocial cost versus the risks of antipsychotic treatment, and clearly present the probability of relapse when stopping or continuing antipsychotic treatment. While the uncertainty is largest after the first episode of psychosis, following a second episode the arguments for antipsychotic maintenance treatment are even greater.

Future research should include predictive models of successful treatment discontinuation in addition to prediction of treatment response.

ACKNOWLEDGEMENT

The first two authors contributed equally to this paper.

REFERENCES

1. Kahn RS, Sommer IE, Murray RM et al. Schizophrenia. Nat Rev Dis Primer 2015;1:15067.
2. Lehman AF, Lieberman JA, Dixon LB et al. Practice guideline for the treatment of patients with schizophrenia, 2nd ed. Am J Psychiatry 2004;161(Suppl. 2):1-56.
3. Buchanan RW, Kreyenbuhl J, Kelly DL et al. The 2009 schizophrenia PORT psychopharmacological treatment recommendations and summary statements. Schizophr Bull 2010;36:71-93.
4. Crockford D, Addington D. Canadian schizophrenia guidelines: schizophrenia and other psychotic disorders with coexisting substance use disorders. Can J Psychiatry 2017;62:624-34.
5. Galletly C, Castle D, Dark F et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. Aust N Z J Psychiatry 2016;50:410-72.
6. Davis JM, Matalon L, Watanabe MD et al. Depot antipsychotic drugs. Place in therapy. Drugs 1994;47:741-73.
7. Kissling W. The current unsatisfactory state of relapse prevention in schizophrenic psychoses – suggestions for improvement. Clin Neuropharmacol 1991;14(Suppl. 2):S33-44.
8. Leucht S, Barnes TRE, Kissling W et al. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. Am J Psychiatry 2003;160:1209-22.
9. Leucht S, Tardy M, Komossa K et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. Lancet 2012;379:2063-71.
10. Leucht S, Cipriani A, Spineli L et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet 2013;382:951-62.
11. Takeuchi H, Suzuki T, Uchida H et al. Antipsychotic treatment for schizophrenia in the maintenance phase: a systematic review of the guidelines and algorithms. Schizophr Res 2012;134:219-25.
12. Murray RM, Quattrone D, Natesan S et al. Should psychiatrists be more cautious about the long-term prophylactic use of antipsychotics? Br J Psychiatry 2016;209:361-5.
13. Wunderink L, Nieboer RM, Wiersma D et al. Recovery in remitted first-episode psychosis at 7 years of follow-up of an early dose reduction/discontinuation or maintenance treatment strategy: long-term follow-up of a 2-year randomized clinical trial. JAMA Psychiatry 2013;70:913-20.
14. Harrow M, Jobe TH, Faull RN. Does treatment of schizophrenia with antipsychotic medications eliminate or reduce psychosis? A 20-year multi-follow-up study. Psychol Med 2014;44:3007-16.
15. Harrow M, Jobe TH, Faull RN et al. A 20-year multi-followup longitudinal study assessing whether antipsychotic medications contribute to work functioning in schizophrenia. Psychiatry Res 2017;256:267-74.
16. Samaha A-N, Seeman P, Stewart J et al. "Breakthrough" dopamine supersensitivity during ongoing antipsychotic treatment leads to treatment failure over time. J Neurosci 2007;27:2979-86.
17. Moncrieff J. Antipsychotic maintenance treatment: time to rethink? PLoS Med 2015;12:e1001861.
18. Goff DC, Falkai P, Fleischhacker WW et al. The long-term effects of antipsychotic medication on clinical course in schizophrenia. Am J Psychiatry 2017;174:840-9.
19. Gøtzsche PC, Young AH, Crace J. Does long term use of psychiatric drugs cause more harm than good? BMJ 2015;350:h2435.
20. Correll CU, Kishimoto T, Nielsen J et al. Quantifying clinical relevance in the treatment of schizophrenia. Clin Ther 2011;33:B16-39.
21. Kane JM, Kishimoto T, Correll CU. Assessing the comparative effectiveness of long-acting injectable vs. oral antipsychotic medications in the prevention of relapse provides a case study in comparative effectiveness research in psychiatry. J Clin Epidemiol 2013;66(Suppl. 8):S37-41.
22. Correll CU, Kishimoto T, Kane JM. Randomized controlled trials in schizophrenia: opportunities, limitations, and trial design alternatives. Dialogues Clin Neurosci 2011;13:155-72.
23. Kane JM, Kishimoto T, Correll CU. Non-adherence to medication in patients with psychotic disorders: epidemiology, contributing factors and management strategies. World Psychiatry 2013;12:216-26.
24. Valenstein M, Ganoczy D, McCarthy JF et al. Antipsychotic adherence over time among patients receiving treatment for schizophrenia: a retrospective review. J Clin Psychiatry 2006;67:1542-50.
25. Rajagopalan K, Wade S, Meyer N et al. Real-world adherence assessment of lurasidone and other oral atypical antipsychotics among patients with schizophrenia: an administrative claims analysis. Curr Med Res Opin 2017;33:813-20.
26. Lopez LV, Shaikh A, Merson J et al. Accuracy of clinician assessments of medication status in the emergency setting: a comparison of clinician assessment of antipsychotic usage and plasma level determination. J Clin Psychopharmacol 2017;37:310-4.

27. Alvarez-Jimenez M, Priede A, Hetrick SE et al. Risk factors for relapse following treatment for first episode psychosis: a systematic review and meta-analysis of longitudinal studies. Schizophr Res 2012;139:116-28.
28. Winton-Brown TT, Elanjithara T, Power P et al. Five-fold increased risk of relapse following breaks in antipsychotic treatment of first episode psychosis. Schizophr Res 2017;179:50-6.
29. Schoeler T, Petros N, Di Forti M et al. Poor medication adherence and risk of relapse associated with continued cannabis use in patients with first-episode psychosis: a prospective analysis. Lancet Psychiatry 2017;4:627-33.
30. Novick D, Haro JM, Suarez D et al. Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia. Psychiatry Res 2010;176:109-13.
31. Kishimoto T, Robenzadeh A, Leucht C et al. Long-acting injectable vs oral antipsychotics for relapse prevention in schizophrenia: a meta-analysis of randomized trials. Schizophr Bull 2014;40:192-213.
32. Kishimoto T, Nitta M, Borenstein M et al. Long-acting injectable versus oral antipsychotics in schizophrenia: a systematic review and meta-analysis of mirror-image studies. J Clin Psychiatry 2013;74:957-65.
33. Barnes TRE, Drake RJ, Dunn G et al. Effect of prior treatment with antipsychotic long-acting injection on randomised clinical trial treatment outcomes. Br J Psychiatry 2013;203:215-20.
34. Voss EA, Ryan PB, Stang PE et al. Switching from risperidone long-acting injectable to paliperidone long-acting injectable or oral antipsychotics: analysis of a Medicaid claims database. Int Clin Psychopharmacol 2015;30:151-7.
35. Kishimoto T, Hagi K, Nitta M et al. Effectiveness of long-acting injectable vs oral antipsychotics in patients with schizophrenia: a meta-analysis of prospective and retrospective cohort studies. Schizophr Bull (in press).
36. Tiihonen J, Haukka J, Taylor M et al. A nationwide cohort study of oral and depot antipsychotics after first hospitalization for schizophrenia. Am J Psychiatry 2011;168:603-9.
37. Tiihonen J, Mittendorfer-Rutz E, Majak M et al. Real-world effectiveness of antipsychotic treatments in a nationwide cohort of 29823 patients with schizophrenia. JAMA Psychiatry 2017;74:686-93.
38. Misawa F, Kishimoto T, Hagi K et al. Safety and tolerability of long-acting injectable versus oral antipsychotics: a meta-analysis of randomized controlled studies comparing the same antipsychotics. Schizophr Res 2016; 176:220-30.
39. Carbon M, Correll CU. Clinical predictors of therapeutic response to antipsychotics in schizophrenia. Dialogues Clin Neurosci 2014; 16:505-24.
40. Kotov R, Fochtmann L, Li K et al. Declining clinical course of psychotic disorders over the two decades following first hospitalization: evidence from the Suffolk County Mental Health Project. Am J Psychiatry 2017;174:1064-74.
41. Husa AP, Rannikko I, Moilanen J et al. Lifetime use of antipsychotic medication and its relation to change of verbal learning and memory in midlife schizophrenia – an observational 9-year follow-up study. Schizophr Res 2014;158:134-41.
42. Moilanen J, Haapea M, Miettunen J et al. Characteristics of subjects with schizophrenia spectrum disorder with and without antipsychotic medication – A 10-year follow-up of the Northern Finland 1966 Birth Cohort study. Eur Psychiatry 2013;28:53-8.
43. Gotfredsen DR, Wils RS, Hjorthøj C et al. Stability and development of psychotic symptoms and the use of antipsychotic medication – long-term follow-up. Psychol Med 2017;47: 2118-29.
44. Wils RS, Gotfredsen DR, Hjorthøj C et al. Antipsychotic medication and remission of psychotic symptoms 10 years after a first-episode psychosis. Schizophr Res 2017;182:42-8.
45. Ran M-S, Weng X, Chan CL-W et al. Different outcomes of never-treated and treated patients with schizophrenia: 14-year follow-up study in rural China. Br J Psychiatry 2015;207: 495-500.
46. Tiihonen J, Mittendorfer-Rutz E, Torniainen M et al. Mortality and cumulative exposure to antipsychotics, antidepressants, and benzodiazepines in patients with schizophrenia: an observational follow-up study. Am J Psychiatry 2016;173:600-6.
47. Tiihonen J, Lönnqvist J, Wahlbeck K et al. 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). Lancet 2009;374:620-7.
48. Wunderink L, Nienhuis FJ, Sytema S et al. Guided discontinuation versus maintenance treatment in remitted first-episode psychosis: relapse rates and functional outcome. J Clin Psychiatry 2007;68:654-61.
49. Uchida H, Suzuki T, Takeuchi H et al. Low dose vs standard dose of antipsychotics for relapse prevention in schizophrenia: meta-analysis. Schizophr Bull 2011;37:788-99.
50. World Health Organization. Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2017. http://www.whocc.no/atcddd/.
51. Takeuchi H, Suzuki T, Remington G et al. Effects of risperidone and olanzapine dose reduction on cognitive function in stable patients with schizophrenia: an open-label, randomized, controlled, pilot study. Schizophr Bull 2013;39:993-8.
52. Mayoral-van Son J, de la Foz VO, Martinez-Garcia O et al. Clinical outcome after antipsychotic treatment discontinuation in functionally recovered first-episode nonaffective psychosis individuals: a 3-year naturalistic follow-up study. J Clin Psychiatry 2016;77:492-500.
53. De Hert M, Detraux J, van Winkel R et al. Metabolic and cardiovascular adverse effects associated with antipsychotic drugs. Nat Rev Endocrinol 2011;8:114-26.
54. De Hert M, Correll CU, Bobes J et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. World Psychiatry 2011; 10:52-77.
55. Correll CU, Solmi M, Veronese N et al. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. World Psychiatry 2017; 16:163-80.
56. Stubbs B, Koyanagi A, Veronese N et al. Physical multimorbidity and psychosis: comprehensive cross sectional analysis including 242,952 people across 48 low- and middle-income countries. BMC Med 2016;14:189.
57. Vancampfort D, Correll CU, Galling B et al. Diabetes mellitus in people with schizophrenia, bipolar disorder and major depressive disorder: a systematic review and large scale meta-analysis. World Psychiatry 2016;15:166-74.
58. Vancampfort D, Stubbs B, Mitchell AJ et al. Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. World Psychiatry 2015;14:339-47.
59. Correll CU, Detraux J, De Lepeleire J et al. Effects of antipsychotics, antidepressants and mood stabilizers on risk for physical diseases in people with schizophrenia, depression and bipolar disorder. World Psychiatry 2015;14:119-36.
60. Vancampfort D, Wampers M, Mitchell AJ et al. A meta-analysis of cardio-metabolic abnormalities in drug naïve, first-episode and multi-episode patients with schizophrenia versus general population controls. World Psychiatry 2013;12:240-50.
61. Vancampfort D, Firth J, Schuch FB et al. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. World Psychiatry 2017;16:308-15.
62. Leucht S, Burkard T, Henderson J et al. Physical illness and schizophrenia: a review of the literature. Acta Psychiatr Scand 2007;116:317-33.
63. Liu NH, Daumit GL, Dua T et al. Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. World Psychiatry 2017;16:30-40.
64. Hjorthøj C, Stürup AE, McGrath JJ. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. Lancet Psychiatry 2017;4:295-301.
65. Saha S, Chant D, McGrath J. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? Arch Gen Psychiatry 2007;64:1123-31.
66. Olfson M, Gerhard T, Huang C et al. Premature mortality among adults with schizophrenia in the United States. JAMA Psychiatry 2015;72:1172-81.
67. Rubio JM, Correll CU. Duration and relevance of untreated psychiatric disorders, 1: Psychotic disorders. J Clin Psychiatry 2017;78:358-9.
68. Fekadu A, Medhin G, Kebede D et al. Excess mortality in severe mental illness: 10-year population-based cohort study in rural Ethiopia. Br J Psychiatry 2015;206:289-96.
69. Ran MS, Chan CL, Chen EY et al. Differences in mortality and suicidal behaviour between treated and never-treated people with schizophrenia in rural China. Br J Psychiatry 2009; 195:126-31.
70. Charlson FJ, Baxter AJ, Dua T et al. Excess mortality from mental, neurological, and substance use disorders in the Global Burden of Disease Study 2010. In: Patel V, Chisholm D, Dua T et al (eds). Mental, neurological, and substance use disorders: disease control priorities, 3rd ed. Washington: International Bank for Reconstruction and Development/ World Bank, 2016.

71. Ponnudurai R, Jayakar J, Sathiya Sekaran B. Assessment of mortality and marital status of schizophrenic patients over a period of 13 years. Indian J Psychiatry 2006;48:84-7.
72. Casey DE, Haupt DW, Newcomer JW et al. Antipsychotic-induced weight gain and metabolic abnormalities: implications for increased mortality in patients with schizophrenia. J Clin Psychiatry 2004;65(Suppl. 7):4-18.
73. Crump C, Winkleby MA, Sundquist K et al. Comorbidities and mortality in persons with schizophrenia: a Swedish national cohort study. Am J Psychiatry 2013;170:324-33.
74. Lahti M, Tiihonen J, Wildgust H et al. Cardiovascular morbidity, mortality and pharmacotherapy in patients with schizophrenia. Psychol Med 2012;42:2275-85.
75. Vermeulen J, van Rooijen G, Doedens P et al. Antipsychotic medication and long-term mortality risk in patients with schizophrenia; a systematic review and meta-analysis. Psychol Med 2017;47:2217-28.
76. Torniainen M, Mittendorfer-Rutz E, Tanskanen A et al. Antipsychotic treatment and mortality in schizophrenia. Schizophr Bull 2015;41:656-63.
77. Tiihonen J, Mittendorfer-Rutz E, Alexanderson K et al. Antipsychotics and mortality in a nationwide cohort of 29,823 patients with schizophrenia. Presented at the 30th Congress of the European College of Neuropsychopharmacology, Paris, September 2017.
78. Rubio JM, Correll CU. Reduced all-cause mortality with antipsychotics and antidepressants compared to increased all-cause mortality with benzodiazepines in patients with schizophrenia observed in naturalistic treatment settings. Evid Based Ment Health 2017;20:e6.
79. Johnstone EC, Crow TJ, Frith CD et al. Cerebral ventricular size and cognitive impairment in chronic schizophrenia. Lancet 1976;2:924-6.
80. Bakhshi K, Chance SA. The neuropathology of schizophrenia: a selective review of past studies and emerging themes in brain structure and cytoarchitecture. Neuroscience 2015;303:82-102.
81. Honea R, Crow TJ, Passingham D et al. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. Am J Psychiatry 2005;162:2233-45.
82. Crow TJ, Chance SA, Priddle TH et al. Laterality interacts with sex across the schizophrenia/bipolarity continuum: an interpretation of meta-analyses of structural MRI. Psychiatry Res 2013;210:1232-44.
83. Zhang W, Deng W, Yao L et al. Brain structural abnormalities in a group of never-medicated patients with long-term schizophrenia. Am J Psychiatry 2015;172:995-1003.
84. Harrisberger F, Buechler R, Smieskova R et al. Alterations in the hippocampus and thalamus in individuals at high risk for psychosis. NPJ Schizophr 2016;2:16033.
85. McIntosh AM, Owens DC, Moorhead WJ et al. Longitudinal volume reductions in people at high genetic risk of schizophrenia as they develop psychosis. Biol Psychiatry 2011;69:953-8.
86. Takayanagi Y, Kulason S, Sasabayashi D et al. Reduced thickness of the anterior cingulate cortex in individuals with an at-risk mental state who later develop psychosis. Schizophr Bull 2017;43:907-13.
87. Weinberger DR, McClure RK. Neurotoxicity, neuroplasticity, and magnetic resonance imaging morphometry: what is happening in the schizophrenic brain? Arch Gen Psychiatry 2002;59:553-8.
88. Mathalon DH, Rapoport JL, Davis KL et al. Neurotoxicity, neuroplasticity, and magnetic resonance imaging morphometry. Arch Gen Psychiatry 2003;60:846-8.
89. Pakkenberg B. Total nerve cell number in neocortex in chronic schizophrenics and controls estimated using optical disectors. Biol Psychiatry 1993;34:768-72.
90. Selemon LD, Rajkowska G, Goldman-Rakic PS. Elevated neuronal density in prefrontal area 46 in brains from schizophrenic patients: application of a three-dimensional, stereologic counting method. J Comp Neurol 1998;392:402-12.
91. Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. Brain J Neurol 1999;122(Pt. 4):593-624.
92. Haijma SV, Van Haren N, Cahn W et al. Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. Schizophr Bull 2013;39:1129-38.
93. Swayze VW, Andersen A, Arndt S et al. Reversibility of brain tissue loss in anorexia nervosa assessed with a computerized Talairach 3-D proportional grid. Psychol Med 1996;26:381-90.
94. Pfefferbaum A, Sullivan EV, Mathalon DH et al. Longitudinal changes in magnetic resonance imaging brain volumes in abstinent and relapsed alcoholics. Alcohol Clin Exp Res 1995;19:1177-91.
95. Schroth G, Naegele T, Klose U et al. Reversible brain shrinkage in abstinent alcoholics, measured by MRI. Neuroradiology 1988;30:385-9.
96. Gordon N. Apparent cerebral atrophy in patients on treatment with steroids. Dev Med Child Neurol 1980;22:502-6.
97. Vita A, De Peri L, Deste G et al. The effect of antipsychotic treatment on cortical gray matter changes in schizophrenia: does the class matter? A meta-analysis and meta-regression of longitudinal magnetic resonance imaging studies. Biol Psychiatry 2015;78:403-12.
98. Brugger S, Howes OD. Heterogeneity and homogeneity of regional brain structure in schizophrenia. JAMA Psychiatry 2017;74:1104-11.
99. Fusar-Poli P, Smieskova R, Kempton MJ et al. Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. Neurosci Biobehav Rev 2013;37:1680-91.
100. van Erp TGM, Hibar DP, Rasmussen JM et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. Mol Psychiatry 2016;21:547-53.
101. Leung M, Cheung C, Yu K et al. Gray matter in first-episode schizophrenia before and after antipsychotic drug treatment. Anatomical likelihood estimation meta-analyses with sample size weighting. Schizophr Bull 2011;37:199-211.
102. Boonstra G, van Haren NEM, Schnack HG et al. Brain volume changes after withdrawal of atypical antipsychotics in patients with first-episode schizophrenia. J Clin Psychopharmacol 2011;31:146-53.
103. Ho BC, Andreasen NC, Ziebell S et al. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. Arch Gen Psychiatry 2011;68:128-37.
104. Andreasen NC, Liu D, Ziebell S et al. Relapse duration, treatment intensity, and brain tissue loss in schizophrenia: a prospective longitudinal MRI study. Am J Psychiatry 2013;170:609-15.
105. Ahmed M, Cannon DM, Scanlon C et al. Progressive brain atrophy and cortical thinning in schizophrenia after commencing clozapine treatment. Neuropsychopharmacology 2015;40:2409-17.
106. Lieberman J, Chakos M, Wu H et al. Longitudinal study of brain morphology in first episode schizophrenia. Biol Psychiatry 2001;49:487-99.
107. Lesh TA, Tanase C, Geib BR et al. A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. JAMA Psychiatry 2015;72:226-34.
108. Burt DR, Creese I, Snyder SH. Antischizophrenic drugs: chronic treatment elevates dopamine receptor binding in brain. Science 1977;196:326-8.
109. Silvestri S, Seeman MV, Negrete JC et al. Increased dopamine D2 receptor binding after long-term treatment with antipsychotics in humans: a clinical PET study. Psychopharmacology 2000;152:174-80.
110. Klawans HL, Goetz CG, Perlik S. Tardive dyskinesia: review and update. Am J Psychiatry 1980;137:900-8.
111. Lohr JB, Kuczenski R, Niculescu AB. Oxidative mechanisms and tardive dyskinesia. CNS Drugs 2003;17:47-62.
112. Bakker PR, van Harten PN, van Os J. Antipsychotic-induced tardive dyskinesia and polymorphic variations in COMT, DRD2, CYP1A2 and MnSOD genes: a meta-analysis of pharmacogenetic interactions. Mol Psychiatry 2008;13:544-56.
113. Kane JM, Woerner M, Weinhold P et al. A prospective study of tardive dyskinesia development: preliminary results. J Clin Psychopharmacol 1982;2:345-9.
114. Correll CU, Leucht S, Kane JM. Lower risk for tardive dyskinesia associated with second-generation antipsychotics: a systematic review of 1-year studies. Am J Psychiatry 2004;161:414-25.
115. Carbon M, Hsieh C-H, Kane JM et al. Tardive dyskinesia prevalence in the period of second-generation antipsychotic use: a meta-analysis. J Clin Psychiatry 2017;78:e264-78.
116. Lieberman JA, Alvir J, Geisler S et al. Methylphenidate response, psychopathology and tardive dyskinesia as predictors of relapse in schizophrenia. Neuropsychopharmacology 1994;11:107-18.
117. Yamanaka H, Kanahara N, Suzuki T et al. Impact of dopamine supersensitivity psychosis in treatment-resistant schizophrenia: an analysis of multi-factors predicting long-term prognosis. Schizophr Res 2016;170:252-8.
118. Youssef HA, Waddington JL. Morbidity and mortality in tardive dyskinesia: associations in chronic schizophrenia. Acta Psychiatr Scand 1987;75:74-7.
119. Apud JA, Egan MF, Wyatt RJ. Neuroleptic withdrawal in treatment-resistant patients with

schizophrenia: tardive dyskinesia is not associated with supersensitive psychosis. Schizophr Res 2003;63:151-60.
120. Kane JM. Tardive dyskinesia circa 2006. Am J Psychiatry 2006;163:1316-8.
121. Kane JM, Correll CU, Liang GS et al. Efficacy of valbenazine (nbi-98854) in treating subjects with tardive dyskinesia and schizophrenia or schizoaffective disorder. Psychopharmacol Bull 2017;47:69-76.
122. Fernandez HH, Factor SA, Hauser RA et al. Randomized controlled trial of deutetrabenazine for tardive dyskinesia: the ARM-TD study. Neurology 2017;88:2003-10.
123. Kimura H, Kanahara N, Komatsu N et al. A prospective comparative study of risperidone long-acting injectable for treatment-resistant schizophrenia with dopamine supersensitivity psychosis. Schizophr Res 2014;155:52-8.
124. Chouinard G, Jones BD. Neuroleptic-induced supersensitivity psychosis: clinical and pharmacologic characteristics. Am J Psychiatry 1980;137:16-21.
125. Moncrieff J. Does antipsychotic withdrawal provoke psychosis? Review of the literature on rapid onset psychosis (supersensitivity psychosis) and withdrawal-related relapse. Acta Psychiatr Scand 2006;114:3-13.
126. Nakata Y, Kanahara N, Iyo M. Dopamine supersensitivity psychosis in schizophrenia: concepts and implications in clinical practice. J Psychopharmacol 2017;31:1511-8.
127. Yin J, Barr AM, Ramos-Miguel A et al. Antipsychotic induced dopamine supersensitivity psychosis: a comprehensive review. Curr Neuropharmacol 2017;15:174-83.
128. Koener B, Goursaud S, Van De Stadt M et al. Pharmacological blockade of dopamine D2 receptors by aripiprazole is not associated with striatal sensitization. Naunyn-Schmiedeberg's Arch Pharmacol 2011;383:65-77.
129. Kane JM, Osuntokun O, Kryzhanovskaya LA et al. A 28-week, randomized, double-blind study of olanzapine versus aripiprazole in the treatment of schizophrenia. J Clin Psychiatry 2009;70:572-81.
130. Stauffer V, Ascher-Svanum H, Liu L et al. Maintenance of response with atypical antipsychotics in the treatment of schizophrenia: a post-hoc analysis of 5 double-blind, randomized clinical trials. BMC Psychiatry 2009;9:13.
131. Jääskeläinen E, Juola P, Hirvonen N et al. A systematic review and meta-analysis of recovery in schizophrenia. Schizophr Bull 2013;39:1296-306.
132. Van Eck RM, Burger TJ, Vellinga A et al. The relationship between clinical and personal recovery in patients with schizophrenia spectrum disorders: a systematic review and meta-analysis. Schizophr Bull (in press).
133. Kane JM, Robinson DG, Schooler NR et al. Comprehensive versus usual community care for first-episode psychosis: 2-year outcomes from the NIMH RAISE Early Treatment Program. Am J Psychiatry 2016;173:362-72.
134. Dixon LB, Dickerson F, Bellack AS et al. The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. Schizophr Bull 2010;36:48-70.
135. Wykes T, Steel C, Everitt B et al. Cognitive behavior therapy for schizophrenia: effect sizes, clinical models, and methodological rigor. Schizophr Bull 2008;34:523-37.
136. Turkington D, Kingdon D, Weiden PJ. Cognitive behavior therapy for schizophrenia. Am J Psychiatry 2006;163:365-73.
137. Morrison AP, Turkington D, Pyle M et al. Cognitive therapy for people with schizophrenia spectrum disorders not taking antipsychotic drugs: a single-blind randomised controlled trial. Lancet 2014;383:1395-403.
138. Sensky T, Turkington D, Kingdon D et al. A randomized controlled trial of cognitive-behavioral therapy for persistent symptoms in schizophrenia resistant to medication. Arch Gen Psychiatry 2000;57:165-72.
139. Burns AMN, Erickson DH, Brenner CA. Cognitive-behavioral therapy for medication-resistant psychosis: a meta-analytic review. Psychiatr Serv 2014;65:874-80.
140. Guo X, Zhai J, Liu Z et al. Effect of antipsychotic medication alone vs combined with psychosocial intervention on outcomes of early-stage schizophrenia: a randomized, 1-year study. Arch Gen Psychiatry 2010;67:895-904.
141. Pilling S, Bebbington P, Kuipers E et al. Psychological treatments in schizophrenia: I. Meta-analysis of family intervention and cognitive behaviour therapy. Psychol Med 2002;32:763-82.
142. Schooler NR, Keith SJ, Severe JB et al. Relapse and rehospitalization during maintenance treatment of schizophrenia. The effects of dose reduction and family treatment. Arch Gen Psychiatry 1997;54:453-63.
143. Robinson DG, Schooler NR, Correll CU et al. Psychopharmacological treatment in the RAISE-ETP study: outcomes of a manual and computer decision support system based intervention. Am J Psychiatry 2018;175:169-79.
144. Caemmerer J, Correll CU, Maayan L. Acute and maintenance effects of non-pharmacologic interventions for antipsychotic associated weight gain and metabolic abnormalities: a meta-analytic comparison of randomized controlled trials. Schizophr Res 2012;140:159-68.
145. Speyer H, Nørgaard HCB, Birk M et al. The CHANGE trial: no superiority of lifestyle coaching plus care coordination plus treatment as usual compared to treatment as usual alone in reducing risk of cardiovascular disease in adults with schizophrenia spectrum disorders and abdominal obesity. World Psychiatry 2016;15:155-65.
146. Jakobsen AS, Speyer H, Nørgaard HCB et al. Effect of lifestyle coaching versus care coordination versus treatment as usual in people with severe mental illness and overweight: two-years follow-up of the randomized CHANGE trial. PLoS One 2017;12:e0185881.
147. Guloksuz S, van Os J. The slow death of the concept of schizophrenia and the painful birth of the psychosis spectrum. Psychol Med 2018;48:229-44.
148. van Os J, Guloksuz S. A critique of the "ultra-high risk" and "transition" paradigm. World Psychiatry 2017;16:200-6.
149. Nishikawa T, Hayashi T, Koga I et al. Neuroleptic withdrawal with remitted schizophrenics: a naturalistic follow-up study. Psychiatry 2007;70:68-79.
150. Gitlin M, Nuechterlein K, Subotnik KL et al. Clinical outcome following neuroleptic discontinuation in patients with remitted recent-onset schizophrenia. Am J Psychiatry 2001;158:1835-42.
151. Zhang J-P, Lencz T, Malhotra AK. D2 receptor genetic variation and clinical response to antipsychotic drug treatment: a meta-analysis. Am J Psychiatry 2010;167:763-72.
152. Sarpal DK, Robinson DG, Lencz T et al. Antipsychotic treatment and functional connectivity of the striatum in first-episode schizophrenia. JAMA Psychiatry 2015;72:5.
153. Sarpal DK, Argyelan M, Robinson DG et al. Baseline striatal functional connectivity as a predictor of response to antipsychotic drug treatment. Am J Psychiatry 2016;173:69-77.
154. Hadley JA, Nenert R, Kraguljac NV et al. Ventral tegmental area/midbrain functional connectivity and response to antipsychotic medication in schizophrenia. Neuropsychopharmacology 2014;39:1020-30.
155. Carrión RE, Cornblatt BA, Burton CZ et al. Personalized prediction of psychosis: external validation of the NAPLS-2 psychosis risk calculator with the EDIPPP project. Am J Psychiatry 2016;173:989-96.
156. Prata D, Mechelli A, Kapur S. Clinically meaningful biomarkers for psychosis: a systematic and quantitative review. Neurosci Biobehav Rev 2014;45:134-141.

DOI:10.1002/wps.20516

Increasing expectations and knowledge require a more subtle use of prophylactic antipsychotics

Those who comment on the use of antipsychotics in 2018 face two challenges. The first stems from rising expectations. The move from incarceration of psychotic people in asylums to care in the community has transformed the lives of many in Western European countries. Undoubtedly, antipsychotics played a major role in facilitating this. Furthermore, there is overwhelming evidence that antipsychotics are essential in acute psychosis and that many patients will benefit from taking them for a period thereafter.

However, as care has improved, so expectations of recovery have increased. This has been accompanied by calls for patients and relatives to have a greater voice in planning care. In some countries, their representatives have been incorporated into policy making[1]. In others, patients have been relegated to shouting their criticisms from offstage. One example of the latter is the website "Mad in America" (https://www.madinamerica.com); a brief look should cause any psychiatrist to reflect on why antipsychotics attract such opprobrium from many of those they are intended to help.

It is in this context that the prescription of antipsychotics for prevention of recurrence, rather than treatment of active symptoms, should be considered. Drugs intended to be taken prophylactically need to be extremely safe and tolerable; witness the arguments concerning the pros and cons of statins. In recent years, concern has been raised about the risk/benefit ratio of prophylactic antipsychotics[2]. The paper by Correll et al[3] is the second of two responses from the psychopharmacological establishment, and takes a less dogmatic approach than its predecessor[4].

Correll et al accept that most antipsychotics increase the risk of obesity and the metabolic syndrome. Their review addresses, but fails to resolve, the paradox that we clinicians commonly see the adverse effects of antipsychotics on the physical health of our patients, yet mortality appears to be lower in those patients who take rather than do not take antipsychotics (at least for those who receive low or moderate doses)[5]. Fortunately, we are now better able to minimize metabolic effects by prescribing antipsychotics with lower propensity to weight gain.

The coverage of the worrying issue of the effects of antipsychotics on brain structure is less satisfactory. Correll et al too readily dismiss the evidence that prolonged antipsychotic use is associated with decreased grey matter and fail to cite the monkey and rodent studies in which administration of antipsychotics causes brain volume losses[4]. This is an unresolved issue that deserves intensive investigation rather than bland reassurances.

The second challenge to traditional practice comes from the explosion in knowledge about psychosis since the prophylactic use of antipsychotics was introduced in the 1970s. Then schizophrenia was considered a discrete neurodegenerative disease. Now we know that schizophrenia is the severe end of a continuum of psychosis, and that the final common pathway underlying positive symptoms is dopamine dysregulation[6]. We used to think that dopamine blockade addressed the locus of abnormality in the D2 receptor, but it is now clear that the primary problem in most patients is presynaptic: they synthesize excessive striatal dopamine. Antipsychotics block the effect of the released dopamine and thus diminish aberrant perceptions secondary to increased salience. They do little for established delusions; neither do they help negative symptoms or cognitive dysfunction: indeed, there is considerable evidence that high doses impair both of these.

We now also know that many of the environmental risk factors for psychosis (e.g., child abuse, migration) increase striatal dopamine synthesis, and that people with schizophrenia show greater dopamine release in response to everyday hassles[6,7]. Furthermore, as people become ill, the stress caused by the psychosis itself (e.g., by beliefs that they may be harmed) and its consequences (e.g., compulsory hospitalization) likely result in further release of dopamine, and thus more abnormal salience and worsening psychosis[6].

Care must therefore include efforts to minimize stresses and find an appropriate social niche for patients, to facilitate decrease in dopamine synthesis. Psychological treatments, including cognitive-behavioural and cognitive remediation therapies, possibly the newer avatar therapy, and last but not least physical exercise, should also be available. Unfortunately, too often sufferers languish in hostels full of drug takers in the worst parts of town; or they are homeless or in jail, situations that we psychiatrists would have difficulty in coping with, let alone those with increased sensitivity to stress.

Advances in understanding the genetic architecture of schizophrenia have demonstrated shared genes with bipolar disorder and also with depression, post-traumatic stress disorder and anxiety disorders[8]. This should not surprise clinicians who have noted that many people with a diagnosis of schizophrenia also suffer from mood swings, or find anxiety and depression as disabling as positive symptoms. Mood stabilizers, antidepressants and psychological treatments can ameliorate these and lessen the drive to psychosis.

Given that hard evidence for the benefits of antipsychotics extends only to about two years, what to do in the long-term requires further research. Much effort has been expended by pharmaceutical companies emphasizing the importance of adherence to antipsychotics. Less attention has focused on the value of moderate and rational prescribing. As a result, many patients receive excessive doses of D2 blockers for prolonged periods. Correll et al note that many recovered patients are able to remain well on doses

of antipsychotics smaller than that which was needed in the acute episode (though they say not less than 50%). Guidelines need to be developed on when and how slowly to reduce antipsychotics, and in whom it is appropriate to eventually stop them.

In their initial trial of the prophylactic use of antipsychotics, Leff and Wing[9] reported that these were helpful to patients with a moderate, but not those with a very good, outlook. Similarly, Correll et al accept that a significant minority of people who receive the diagnosis of schizophrenia (perhaps up to 20%) will be able to come off the drugs without disadvantage, probably because they have milder illnesses.

Leff and Wing[9] also noted that those with a very poor outcome do not benefit from continued antipsychotics. The reason that such individuals are treatment resistant is because they do not synthesize excessive striatal dopamine[6]. There appear to be two types of treatment resistance[10]. First, those who have never responded to antipsychotics and whose psychosis may not involve dopamine dysregulation. Second, those who once responded to D2 blockers but have lost this ability, possibly due to the development of dopamine supersensitivity. Correll et al ignore the evidence that prolonged administration of antipsychotics to animals cause an increase in D2 receptor numbers, and that the resultant dopamine supersensitivity causes antipsychotics to lose their efficacy[2]. They do, however, cite reports that partial dopamine agonists may have less propensity to cause dopamine supersensitivity. Once again this is an issue that demands further investigation.

Finally, we psychiatrists need to reach out to our patients and to those groups critical of antipsychotic prescribing. Doctors and patients may have different priorities; patients may put more emphasis on remaining slim rather than having voices totally eradicated, or may consider it more important to be alert enough to work rather than to have conventional thoughts. In the absence of such conversations, patients may become disillusioned with psychiatry and rely on alternatives such as the Hearing Voices Network or therapies without any evidence base.

Robin M. Murray, Marta Di Forti
Institute of Psychiatry, Psychology and Neuroscience, King's College, London, UK

1. Schizophrenia Commission. The abandoned illness. https://www.rethink.org/about-us/the-schizophrenia-commission.
2. Murray RM, Quattrone D, Natesan S et al. Br J Psychiatry 2016;209:361-5.
3. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
4. Goff DC, Falkai P, Fleischhacker WW et al. Am J Psychiatry 2017;174:840-9.
5. Torniainen M, Mittendorfer-Rutz E, Tanskanen A et al. Schizophr Bull 2015;41:656-63.
6. Howes OD, Murray RM. Lancet 2014;383:1677-87.
7. Klippel A, Myin-Germeys I, Chavez-Baldini U et al. Schizophr Bull 2017;43:302-31.
8. Duncan LE, Shen H, Ballon JS et al. Schizophr Bull (in press).
9. Leff JP, Wing JK. BMJ 1971;3:599-604.
10. Lally J, Ajnakina O, Di Forti M et al. Psychol Med 2016;46:3231-40.

DOI:10.1002/wps.20517

Long-term antipsychotic treatment of schizophrenia: does it help or hurt over a 20-year period?

Correll et al[1] argue for a positive view of the risk-benefit ratio for long-term continuous antipsychotic treatment of schizophrenia. They claim that studies of long-term outcome which show negative results are not convincing because of confounding factors. Their chief argument is that in "non-randomized, uncontrolled studies... there is a high risk of confounding by indication and reverse causation, in that greater illness severity could be the cause of continued antipsychotic treatment, rather than being the effect"[1]. The other argument is that long-term continuous use of antipsychotics does not involve significant morbidity from dopamine supersensitivity psychosis. Here we provide evidence which severely questions both of these conclusions, showing that they overestimate the benefits and underestimate the risks of long-term antipsychotic treatment.

There are at least eight studies assessing whether schizophrenia patients improve when treated longer than two-three years with antipsychotic medications. These studies have been conducted by eight different investigator groups. They include those by Wunderink et al in the Netherlands[2], our own Chicago Followup Study[3], the Suffolk County study of Kotov et al[4] in the US, and the long-term data provided by the Danish OPUS trial[5], the AESOP-10 study in England[6], the Finnish Birth Cohort Study[7], the Alberta Hospital Follow-Up Study in Western Canada[8], and the international follow-up study by Harrison et al[9]. These research programs included samples studied from 7 to 20 years. Unlike short-term studies, none of them showed positive long-term results.

Correll et al quote for support a study by Ran et al[10] favoring long-term use of antipsychotics for schizophrenia in China. However, there are many weaknesses in that study. In particular, the untreated group was selected from much older unmarried chronic rural uneducated patients, while the treated group consisted of younger married educated urban patients, some of whom had received only one short period of medication over the 14 year period, rather than being continuously medicated.

As we have noted, one argument used to explain the negative results of long-term antipsychotic treatment is that the schizophrenia patients on antipsychotics for a prolonged period are more severely ill than those not on antipsychotics. However, there are no clear features on which everyone would agree distinguishing "more severely ill people with schizophrenia". Nor is it always clear what "severity" means in relation to schizophrenia. One

frequently used criterion for severity refers to more blatant psychotic illness. However, some episodes of blatant psychosis clear up quickly and thus these psychotic patients may not be more severely ill in every respect.

Another potential criterion for severity in people with schizophrenia involves those whose disorder is more likely to be sustained over a longer period of time, or who have a poorer long-term prognosis. To control for this possible confounder, we have utilized the prognostic indices outlined by Vaillant, Stephens and Zigler. These were collected in our studies at index hospitalization. Later we compared the long-term outcome of poor-prognosis schizophrenia patients medicated with antipsychotics for 15-20 years to that of poor-prognosis patients not prescribed antipsychotics for 15-20 years. We also compared a good-prognosis sample of patients prescribed antipsychotics for 15-20 years with a good-prognosis sample of patients not prescribed antipsychotics for 15-20 years. In both comparisons, those patients not on antipsychotics for 15-20 years had fewer symptoms and better outcomes after the first 2-3 years[3].

An additional limitation of Correll et al's paper is that they do not fully address the evidence on dopamine supersensitivity psychosis from animals and from humans. They limit their discussion to short-term studies of psychotic relapse and the potential loss of antipsychotic efficacy, while ignoring the serious risk for the syndrome resulting from continuous long-term antipsychotic treatment.

The clinical picture of dopamine supersensitivity psychosis is well defined and occurs with increasing frequency after two to three years of continuous antipsychotic maintenance use. Studies indicate that the syndrome manifests in 70% of patients with treatment resistant schizophrenia[11]. Other studies show that the switch to aripiprazole, mentioned by the authors, may actually unmask and intensify psychotic symptoms previously suppressed by stronger D2 antagonists[12]. While long-term continuous use of antipsychotics may induce the syndrome, these medications also block psychotic symptoms, which therefore remain largely unrecognized until the "breakthrough" of more severe symptoms occurs and leads to treatment resistance.

While several research groups have described dopamine supersensitivity psychosis as a serious risk of long-term continuous use of antipsychotics, there has been a systematic failure to incorporate this finding into the risk-benefit ratio for continuous use of antipsychotics. The same applies to the possible negative impact of long-term antipsychotic treatment on work functioning[3]: the block of dopamine receptors may indeed reduce drive and motivation.

Unfortunately, views about the long-term efficacy of antipsychotics are often based on the results from short-term (0-2 years) evaluations. As we have highlighted, there are at least eight major studies which fail to find better outcomes for schizophrenia patients treated on a long-term basis with antipsychotics. These negative results from multiple large well-documented long-term studies are a clear warning sign.

Martin Harrow, Thomas H. Jobe
Department of Psychiatry, University of Illinois at Chicago, College of Medicine, Chicago, IL, USA

1. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
2. Wunderink L, Nieboer RM, Wiersma D et al. JAMA Psychiatry 2013;70:913-20.
3. Harrow M, Jobe TH, Faull RN et al. Psychiatry Res 2017;256:267-74.
4. Kotov R, Fochtmann L, Li K et al. Am J Psychiatry 2017;174:1064-74.
5. Wils RS, Gotfredsen DR, Hjorthøj C et al. Schizophr Res 2017;182:42-8.
6. Morgan C, Lappin J, Heslin M et al. Psychol Med 2014;44:2713-26.
7. Moilanen J, Haapea M, Miettunen J et al. Eur Psychiatry 2013;28:53-8.
8. Bland RC, Parker JH. Arch Gen Psychiatry 1978;35:72-7.
9. Harrison G, Hopper KI, Craig T et al. Br J Psychiatry 2001;178:506-17.
10. Ran MS, Weng X, Chan CL et al. Br J Psychiatry 2015;207:495-500.
11. Suzuki T, Kanahara N, Yamanaka H et al. Psychiatry Res 2015;227:278-82.
12. Takase M, Kanahara N, Oda Y et al. J Psychopharmacol 2015;29:383-9.

DOI:10.1002/wps.20518

Disease modifying effects of antipsychotic drugs in schizophrenia: a clinical and neurobiological perspective

Only in psychiatry would the benefits of one of the great pharmacological breakthroughs in the history of medicine be questioned over a half century after its introduction to clinical practice. When H. Laborit, a French Naval Surgeon stationed in Tunisia, serendipitously realized that chlorpromazine, a compound synthesized by the chemist P. Charpentier, could be used for the treatment of schizophrenia, and brought it to the attention of J. Delay and P. Deniker, psychiatrists at St. Anne's Hospital, a chain of events ensued that changed the course of psychiatry and ushered in the age of psychopharmacology[1]. The advent of this antipsychotic prototype was of comparable significance to other therapeutic milestones like the discovery of insulin, antibiotics and L-dopa.

In the ensuing years, numerous studies by eminent researchers in many countries documented the therapeutic efficacy of chlorpromazine, and the other antipsychotics that followed, in relieving the acute psychotic symptoms of schizophrenia and preventing their recurrence[2]. And while neurological side effects were prevalent, and in many cases problematic, in most instances they could be managed with dose adjustment or adjunctive medications. Second generation ("atypical") medications in turn provided comparable or (in clozapine's case) superior efficacy, and fewer neurological but more metabolic side effects. However, in both cases, the therapeutic benefits of antipsychotics, when used properly, more than offset their side effects[3].

In addition to symptom suppression, longer term studies of patients in their first episode or early stages of illness suggested that antipsychotic drugs, by virtue of their ability to limit the duration and number of psychotic episodes, could impact the clinical deterioration which Kraepelin considered the defining feature of what he termed dementia praecox[4]. In other words, antipsychotics might not just be symptom suppressing, but could mitigate the progression of schizophrenia. If confirmed, this would mean that psychiatry had treatments that could modify the course of the illness, something that had not been achieved with other brain diseases, such as Alzheimer's, Parkinson's and Huntington's.

The evidence for this aspirational therapeutic effect is somewhat circumstantial, but nevertheless compelling, and includes the following.

Treatment studies of first episode patients have consistently found associations between the duration of psychosis prior to treatment and outcome[5]. Specifically, these studies have found that longer periods of active psychotic symptoms prior to first treatment were associated with poorer outcome. What is remarkable is that this relationship was present for outcomes measured in multiple ways, including the time to or level of recovery from the first episode, the time to or likelihood of relapsing after recovery from the first episode, and long-term outcomes measured globally for up to five years after entering treatment for a first episode. Moreover, maintenance treatment studies have demonstrated the prophylactic effect of antipsychotic drugs in preventing relapse; treatment, then, may be responsible for mitigating the course of the illness and producing more favorable outcomes.

Furthermore, numerous investigations of brain morphology (post-mortem and neuroimaging) have demonstrated structural abnormalities in various anatomic regions in schizophrenia patients compared to control subjects. These abnormalities primarily involve volume reductions of gray matter in soft tissue structures (e.g., hippocampus, temporal and frontal cortices, superior temporal gyrus, thalamus) and volume enlargements of fluid containing structures (e.g., ventricular system, subarachnoid space); but they also include shape anomalies and neurodevelopmental anomalies like cavum septum pellucidum, callosal agenesis and gray matter heterotopias. To the extent that some of these pathomorphologic features represent an atrophic process associated with illness progression, they are a target for therapeutic intervention. Various studies have demonstrated gray matter volume changes consistent with progression in specific anatomic regions, and an association between cumulative intake of atypical antipsychotic medication and less pronounced cortical thinning has been reported[6]. While the correlations of treatment and volume change cannot be confirmed as neuroprotective or disease modifying, they are certainly consistent with that interpretation.

Finally, since the introduction of antipsychotic medications into clinical practice, the frequency of the phenomenologic subtypes has changed. Historically, it was postulated that the less severe forms of schizophrenia were characterized by formed delusions, hallucinations and affective symptoms, and paranoid subtype diagnoses, while the more malignant forms exhibited negative, disorganized and motor symptoms and received hebephrenic and catatonic diagnoses. If there is indeed a continuum of severity in illness subtypes, a unidirectional pattern of change in patients' symptoms and diagnoses should reflect progression of the illness. Studies which have found an association between longer periods of untreated psychosis and a greater number of exacerbations and greater likelihood of developing negative, hebephrenic and catatonic symptoms are consistent with this interpretation. However, since antipsychotics came into use, the proportion of patients with predominant negative symptoms and hebephrenic and catatonic symptoms has decreased[4].

Given the obvious acute and prophylactic benefits of antipsychotics, and the possibility that they may be disease modifying, it is hard to understand why there would still be questions as to their effectiveness. In fact, I cannot think of another medication class in other disease areas which has faced similar challenges to its effectiveness after longstanding use and voluminous supportive evidence. Classic "debunking" studies like the Cardiac Arrhythmia Suppression Trial (CAST)[7] and the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT)[8] were either rigorous tests of clinical lore or comparative effectiveness studies. Given the number and consistency of studies, and numerous meta-analyses, I wonder why reviews like that of Correll et al[9] still need to be written for antipsychotics.

It is my contention that the enduring skepticism and distorted views of the clinical effects of antipsychotic drugs are mostly due to the stigma of mental illness and prejudice toward psychiatry, the medical specialty which is focused on their study and care[10]. The stigma historically associated with mental illness is currently perpetuated by lay and professional groups, who oppose the use and deny the efficacy of medication on ideological grounds. They are anti-psychiatry or anti-medical in their ideological orientation, and motivated by biased beliefs. Some lay persons challenge the notion of mental illness, the validity underpinning psychiatric nosology and the evidence supporting the therapeutic basis of psychotropic medications. Some professionals are motivated by factional disputes in mental health care between medical and psychosocial approaches. The latter seek to deny or diminish the evidence that mental disorders have biological bases and are effectively treated with somatic (medications, brain stimulation) forms of treatment, in favor of psychological explanations and psychotherapeutic approaches.

It is certainly appropriate, indeed warranted, to require hard evidence for the efficacy and safety of medical treatments as justification for their clinical use, but it is prejudicial and disingenuous to keep moving the threshold of proof higher and higher because of dogmatically held

views. While we seek and hope for future scientific breakthroughs that will yield better drugs and even greater therapeutic advances, we must recognize and be grateful for what we have, and put them to the best use for our patients[11].

Jeffrey A. Lieberman
Department of Psychiatry, College of Physicians & Surgeons, Columbia University, New York State Psychiatric Institute, New York, NY, USA

1. Dunitz M. Ten years that changed the face of mental illness. London: Livery House, 1999.
2. Davis JM, Andriukaitis S. J Clin Psychopharmacol 1986;6:2S-10S.
3. Lieberman JA, Stroup TS, McEvoy JP et al. N Engl J Med 2005;353:1209-23.
4. Wyatt RJ. Schizophr Bull 1991;17:325-51.
5. Penttila M, Jaaskelainen E, Hirvonen N et al. Br J Psychiatry 2014;205:88-94.
6. van Haren NE, Schnack HG, Cahn W et al. Arch Gen Psychiatry 2011;68:871-80.
7. Echt DS, Liebson PR, Mitchell LB et al. N Engl J Med 1991;324:781-8.
8. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. JAMA 2002;288:2981-97.
9. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
10. Lieberman JA. Shrinks: the untold story of psychiatry. Boston: Little, Brown and Company, 2015.
11. Tiihonen J, Tanskanen A, Taipale H. Am J Psychiatry (in press).

DOI:10.1002/wps.20543

"Will I need to take these medications for the rest of my life?"

Correll et al[1] respond to a growing body of literature that calls into question the long-term use of antipsychotic medications in the treatment of schizophrenia. This recent literature has vexed clinicians who very commonly prescribe antipsychotics on a long-term basis, and who may have held a sense of certainty in the necessity of the therapy.

To address this issue, Correll et al characterize the balance between risks and benefits of long-term antipsychotic treatment. They place past evidence of poor outcomes associated with long-term antipsychotic use in the context of many other benefits (such as that on mortality and relapse prevention), and stratify the literature according to possible bias in each research method. Ultimately, they give an analysis of the benefits and risks of long-term antipsychotic treatment that favors treatment.

In this commentary we focus on applying these principles to working with individuals, particularly people who recently developed schizophrenia. We highlight challenges that will be faced by nearly every clinician who manages this disorder.

First, many – perhaps most – recent onset patients will stop their medication at one time or another. First episode studies have reported up to a 37.1% non-adherence rate[2] and other studies which include longer observation periods report even higher rates. One naturalistic study in Finland reported non-adherence in 58.4% of its sample, which was confirmed by measuring serum concentration[3].

Second, the relationship between clinicians and patients with schizophrenia is often skewed toward the patient feeling controlled by others, particularly prescribers or family members. For most other illnesses, patients accept treatment because it makes them feel better or because it protects them from something they wish to avoid. This is often not so in schizophrenia. For young patients with the illness, particularly those who enter a stable remission following a psychotic episode, the most impassioned psychoeducational approaches to improving adherence may not instill a belief that they need to continue their medication.

In addition, nearly all patients will ask the question "Will I need to take these medications for the rest of my life?". There is only one honest answer to this question, which is "Probably, but I can't be certain". Many individuals believe that they will be the *exceptional* patient who will do well off medications. Correll et al cite that perhaps 4-30% of patients stabilized after an acute episode may discontinue antipsychotics without risk of relapse. They add that, currently, we do not have a clinically reliable means of predicting which patients will have this maverick response to antipsychotic discontinuation. A challenge then remains: how to help individuals with recent-onset schizophrenia to make decisions according to an optimal balance of clinical benefit and personal autonomy.

We propose that a reasonable goal during these early years is to assist patients in taking some ownership of their illness and its management. In doing so, one might change the clinician-patient relationship from one in which the patient may feel controlled by the clinician to one in which the two work collaboratively. A poor relationship with a provider, and the experience of coercion, have been shown to be predictors of negative attitudes towards treatment in those receiving antipsychotics[4]. We emphasize the importance of changing this relationship.

For many, a discussion of the benefits and risks described by Correll et al, combined with the memory of a painful psychotic experience, will suffice. Others may still be skeptical of their need for long-term medication. Prescribers should emphasize the importance of remaining on medications for the first one to two years as well as the potential risks of discontinuation, which includes high rates of relapse[1,5]. However, if the patient is committed to stopping medication, we concur with the recommendation[5] that a trial of dosage reduction with possible discontinuation may be carried out with medical supervision and concurrent psychosocial interventions, in a select population. Clinicians may choose to perform a longer and gentler dose-reduction schedule if they sense a higher risk of relapse.

Dose reduction can be characterized as a learning opportunity for the benefit of both the patient and the prescriber. It may yield important data on the patient's ability to tolerate a period of time on a lower dose of antipsychotic medication, or off of it altogether. Although there are clearly risks associated with this approach, earlier studies[6] found that careful monitoring of patients for prodromal symptoms can

substantially reduce the risk of severe psychotic relapse.

There are, of course, factors that may predict a more successful discontinuation trial. In a recent review[5], several such factors were listed: lack of schizophrenia diagnosis, better premorbid social and occupational functioning, good social support, shorter duration of illness, and shorter duration of untreated psychosis. These factors may help identify the better candidates for discontinuation. Timing, as well, is an important component, as it appears that patients who achieve remission for three months in the first two years of illness have a better clinical prognosis[7]. This better prognosis is felt by some to indicate a higher likelihood of tolerating dose reduction and discontinuation[5].

We support the conclusions outlined in the paper by Correll et al, and we believe that the current literature undermines the clinical certainty of antipsychotic medications in the long-term treatment of schizophrenia. While not a certainty, long-term antipsychotic treatment is a very common outcome for people with schizophrenia. We encourage a sense of curiosity about the *possibility* of dose reduction and discontinuation in appropriate patients.

This open-mindedness will strengthen the therapeutic bond between provider and patient, and might likely lead to better clinical outcomes. In her book *The Center Cannot Hold*[8], E.R. Saks, a Professor in the Gould School of Law at the University of Southern California, describes how experiencing a different sense of reality on and off medications was a revelation which led her to accept that she had a mental illness. She observed that the more she accepted her illness, the less the illness defined her.

Stephen R. Marder, Michael F. Zito
Desert Pacific Mental Illness Research, Education, and Clinical Center, Semel Institute for Neuroscience at UCLA, Los Angeles, CA, USA

1. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
2. Perkins DO, Gu H, Weiden PJ et al. J Clin Psychiatry 2008;69:106-13.
3. Jonsdottir HCB, Opjordsmoen S, Birkenaes AB et al. J Clin Psychopharmacol 2010;30:169-75.
4. Day JC, Bentall RP, Roberts C et al. Arch Gen Psychiatry 2005;62:717-24.
5. Alvarez-Jimenez M, O'Donoghue B, Thompson A et al. CNS Drugs 2016;30:357-68.
6. Marder SR, Wirshing WC, Van Putten T et al. Arch Gen Psychiatry 1994;51:280-7.
7. Cassidy CM, Norman R, Manchanda R et al. Schizophr Bull 2010;36:1001-8.
8. Saks ER. The center cannot hold: my journey through madness. New York: Hyperion, 2007.

DOI:10.1002/wps.20519

Is there compelling evidence that schizophrenia long-term treatment guidelines should be changed?

For decades psychiatrists have tried to stop the "revolving door" phenomenon, meaning that patients with schizophrenia stop their antipsychotic treatment, relapse and have to be treated again, often leading to long rehospitalizations with a worse outcome than that of the previous episode. New findings on the long-term outcome of schizophrenia[1], potential brain volume loss[2], dopamine supersensitivity effects of antipsychotics[3], and a study on "guided discontinuation" in first-episode patients[4] challenge the usefulness of these attempts.

K. Popper would say that scientific progress is made if existing knowledge is challenged by new evidence. The question, however, is whether the new evidence is compelling enough to justify major changes of treatment guidelines. The issue is complex, because several different lines of evidence must be considered together. In brief, my current understanding of this literature is the following.

Recent studies showed that the long-term course of schizophrenia is not always chronic[1]. This finding is not new, as the classic follow-up studies, some from the pre-antipsychotic era, reported that a substantial proportion of patients do quite well after years to decades (see for example Ciompi[5]). These findings have not attracted enough attention. They are very good news, because in my opinion it is a major role of a doctor to give hope. Our clinical perception may be biased, because we usually see patients with a poor outcome. The ones with a good outcome may not come back to the clinic and may disappear from the system. Moreover, several other studies have found that approximately 20% of patients with a first episode of schizophrenia will not have a second one, implying that every fifth patient may be treated unnecessarily on the long term, but unfortunately we cannot predict in advance who these patients may be.

The literature on the long-term outcome of schizophrenia is, however, extremely heterogeneous and difficult to interpret, in particular because different definitions of outcome and patient populations are used, limiting comparability. But, according to a methodologically sound review of follow-up studies, only 13.5% of patients recover in the long term[6]. The majority of national registry studies demonstrate that the mortality of untreated patients with schizophrenia is higher than that of those receiving antipsychotics[7]. In the long term, the multiple side effects of antipsychotics probably contribute to the excess mortality of schizophrenia, but they must be weighed against suicidality and self-neglect when patients are psychotic. Comparably benign antipsychotics are now available (for example, partial dopamine agonists) and many side effects are dose-related.

The potential brain volume loss associated with antipsychotics worries me. As I am not a brain imaging researcher, it is difficult for me to understand whether the magnitude of this brain volume loss is relevant, and what is the relative contribution of the progress of the disease and of antipsychotic treatment to this loss. One study questioned clinical relevance, because the cognition of treated patients was better than that of untreated

patients despite more brain volume loss in the former group[8].

That antipsychotic drugs can lead to dopamine supersensitivity effects has been hypothesized since the 1970s[3], and there is important evidence from animal studies supporting this point of view. We addressed this issue from various angles in a comprehensive meta-analysis[9], but did not find evidence to substantiate these effects in patients. I believe that these effects exist but, given the large effect size for relapse prevention of antipsychotics compared to placebo[9], they do not explain all the superiority of antipsychotics over placebo. There are patients who need them continuously for symptom suppression.

Wunderink et al[4] found a better functional outcome at seven year follow-up in first-episode patients in whom gradual withdrawal and dose reduction of antipsychotics had been tried. It should be emphasized that these were *first-episode* patients (the data do not apply to chronic patients) *in remission* (not all first-episode patients reach a remission) and that in the initial randomized phase complete withdrawal was only possible in 20% of the patients (remarkably, again the magic 20%). These findings are interesting, but the follow-up phase of the study was naturalistic, making it hard to believe that the better outcome was caused by the withdrawal attempts. Replications with better methodology are needed; if they turn out positive, guideline changes would indeed be necessary.

The main counterargument to these new findings is the high risk for relapse whenever patients stop their antipsychotics. Within one year, antipsychotics reduced the relapse risk from 64 to 27% in chronic patients and from 61 to 26% in first-episode patients[9]. Another systematic review in first-episode patients even reported that 77% of untreated patients compared to 3% of treated patients will relapse[10]. Very long-term placebo-controlled randomized studies, say with 5 to 10 year follow-up, are basically impossible for methodological reasons. But, in withdrawal studies, chronic patients who had been stabilized for up to 3-6 years with antipsychotics before randomization still benefitted from these medications, since their relapse risk was higher when they stopped the drugs compared to staying on them[9].

We found that the effect size of antipsychotics for relapse prevention is one of the largest of all medical drug treatments (keeping the limitations of such comparisons in mind)[11]. For example, antihypertensive drugs reduce the risk for cardiovascular events only from 18.1 to 14.1%[11], but their use is not questioned.

In acute treatment, psychiatrists often use non-evidence-based, highly irrational polypharmacy and high doses of antipsychotics. They should be avoided whenever possible. If such strategies have still been used, the first step in the maintenance phase is to reduce excessive doses (for example, to the middle of the range in the labels) and polypharmacy. This applies to *all* patients. Then, a distinction must be made between *chronic* and *first-episode* patients. Among the latter, 20% will not have a second episode, and some might actually not have schizophrenia. For example, it can be very hard to differentiate schizophrenia from drug-induced psychotic disorder. First-episode patients benefit as much from relapse prevention as multiple-episode patients, and to treat them for at least 1-2 years is evidence-based[9], but remitted patients may want to find out whether they are among those who do not need long-term treatment. As relapses often occur only months after stopping antipsychotics, and if potential dopamine supersensitivity effects are taken seriously, any dose reduction must be made extremely carefully (an influential consensus conference recommended to reduce doses by no more than 20% every six months[12]). In those patients who have already had multiple episodes and who often have already tried several times to reduce or stop their treatment, further down titration may not be recommended, in particular if there are no important side effects.

Finally, the clinical circumstances must be considered. If the acute episode was short and mild or the diagnosis unclear, or if there are important side effects, one might be more ready to attempt careful dose reduction. If in the acute phase there was a suicide attempt or an aggressive act (with potentially dire consequences for the patient, such as forensic treatment), one would be very hesitant. And, if patients have clearly improved but are still symptomatic despite treatment, attempts to reduce doses also appear counterintuitive. As – due to the subjectivity of psychiatric outcomes – there is room for interpretation, in the future the evidence will have to be presented such that patients can decide themselves.

Stefan Leucht
Department of Psychiatry and Psychotherapy, Technische Universität München Klinikum rechts der Isar, Munich, Germany

1. Morgan C, Lappin J, Heslin M et al. Psychol Med 2014;44:2713-26.
2. Huhtaniska S, Jääskeläinen E, Hirvonen N et al. Hum Psychopharmacol Clin Exp 2017;32: e2574.
3. Chouinard G, Jones BD. Am J Psychiatry 1980; 137:16-21.
4. Wunderink L, Nieboer RM, Wiersma D et al. JAMA Psychiatry 2013;70:913-20.
5. Ciompi L. Schizophr Bull 1980;6:606-18.
6. Jääskeläinen E, Juola P, Hirvonen N et al. Schizophr Bull 2013;39:1296-306.
7. Vermeulen J, van Rooijen G, Doedens P et al. Psychol Med 2017;47:2217-28.
8. Lesh TA, Tanase C, Geib BR et al. JAMA Psychiatry 2015;72:226-34.
9. Leucht S, Tardy M, Komossa K et al. Lancet 2012;379:2063-71.
10. Zipursky RB, Menezes NM, Streiner DL. Schizophr Res 2014;152:408-14.
11. Leucht S, Hierl S, Dold M et al. Br J Psychiatry 2012;200:97-106.
12. Kissling W, Kane JM, Barnes TR et al. In: Kissling W (ed). Guidelines for neuroleptic relapse prevention in schizophrenia. Heidelberg: Springer, 1991:155-63.

DOI:10.1002/wps.20520

Antipsychotic maintenance treatment in schizophrenia and the importance of preventing relapse

The paper by Correll et al[1] critically reviews the literature pertaining to maintenance antipsychotic treatment in schizophrenia. This is a highly important, but poorly understood topic. The paucity of well-conducted long-term studies makes it difficult to draw firm conclusions regarding the risk to benefit ratio of ongoing antipsychotic treatment. However, this paper provides a comprehensive overview of the pros and cons of that treatment. Clinicians would do well to read the paper carefully.

Despite its demonstrated benefits, it is well recognized that long-term antipsychotic treatment is associated with substantial safety risks, adverse effects and inconveniences. For these reasons, patients and clinicians continue to entertain the possibility of stopping treatment at some stage. While the option of successfully discontinuing antipsychotics once a favourable response has been achieved would be highly desirable, the reality is that no current strategies can realistically be expected to achieve this goal. Despite our best efforts, the illness remains often characterized by chronicity, recurrence of psychotic symptoms when treatment is discontinued, and enduring deficits with negative effects on functionality, autonomy and independent living, as well as quality of life[2].

There are several important aspects concerning the nature of relapse events that clinicians and patients should be aware of when considering antipsychotic treatment discontinuation. First, relapse rates are higher than usually recognized when antipsychotics are discontinued, even after a single episode of psychosis. A recent systematic review reported a weighted mean one-year recurrence rate of 77%, and by two years the risk of recurrence had increased to over 90%[3].

Second, there are no clinically useful predictors of which individuals are likely to successfully discontinue antipsychotic treatment. Indeed, one study in a small sample found that, counterintuitively, patients who respond most favourably to treatment might be at particular risk of relapse[4].

Third, there are no reliable warning signs of imminent relapse, and early rescue medication interventions may not effectively prevent full-blown illness recurrence[5]. Evidence suggests that, once a first psychotic episode has occurred, there is a reduced threshold for illness recurrence. Unlike the first episode, where the onset of illness is frequently gradual and prodromal symptoms emerge over months and even years, the second and subsequent episodes tend to occur abruptly, with no reliable early warning signs and with rapid return of symptom severity levels similar to those of the previous episode[6]. Consequently, treatment discontinuation, even with careful follow-up and immediate re-initiation of treatment, runs the risk of exposing patients to the consequences of full-blown psychosis. This means that the often cited strategy of "targeted discontinuation" – i.e., carefully monitoring patients while treatment is reduced and discontinued, with immediate re-introduction of treatment at the first signs of recurrence – may not be effective.

Fourth, a longer period of treatment prior to discontinuation does not reduce the risk of relapse. Studies in which treatment was continued for two years before discontinuation reported similar relapse rates to those in which patients were treated for six months before discontinuation[7]. Although longer term discontinuation studies have not been conducted, there is no reason to believe that treating patients for a longer period will reduce their chance of illness recurrence once medication is discontinued.

Finally, no discontinuation strategies have been demonstrated to improve the chance of successfully stopping antipsychotic treatment. As pointed out by Correll et al[1], while psychosocial interventions are effective adjunctive therapies, they cannot be regarded as an alternative to antipsychotic medication. Furthermore, other approaches – such as gradual dose reduction followed by discontinuation of antipsychotic treatment – have not been successful.

There are serious psychosocial risks associated with illness recurrence. For example, there is a risk of self-harm and harm to others. In addition, relapses may disrupt friendships and relationships, and impact negatively on education and employment. They may also restrict autonomy, contribute to stigma, and cause patients and their families immense distress. Furthermore, relapses add hugely to the overall economic burden of treating schizophrenia.

In addition to these negative psychosocial consequences of relapse, there may be an additional risk of biological harm. While the treatment response when antipsychotics are re-initiated after relapse is variable, some patients exhibit protracted impairment of response and, importantly, treatment failure emerges in a subgroup of about one in six patients. Treatment failure occurs irrespective of whether it is the first or a subsequent relapse, and even when treatment is re-introduced immediately after the first signs of illness recurrence[8].

Given all of these potential hazards associated with illness recurrence, together with the clear-cut evidence for efficacy of antipsychotics in relapse prevention studies[9], it is understandable that clinicians continue to prioritize relapse prevention via continuous antipsychotic treatment as a treatment goal. This is despite the substantial adverse effect burden associated with antipsychotic medication. This burden can be reduced by judicious selection of the best tolerated antipsychotic according to the individual patient's profile, and at the lowest effective dose. There is also a need for the development of new antipsychotic medications that are better tolerated and at the same time more effective in providing

uninterrupted treatment, including long-acting injectable formulations.

Finally, there is an urgent need for further studies aimed at better identifying individuals who are more likely to successfully discontinue treatment, as well as at characterizing clinically useful early warning signs of impending relapse and developing treatment strategies more likely to result in successful discontinuation.

In the meantime, recommending ongoing maintenance treatment with the safest and best tolerated antipsychotic at the lowest therapeutic dose is the best option for achieving optimal outcomes.

Robin Emsley
Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

1. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
2. Andreasen NC, Carpenter WT Jr, Kane JM et al. Am J Psychiatry 2005;162:441-9.
3. Zipursky RB, Menezes NM, Streiner DL. Schizophr Res 2014;152:408-14.
4. Gaebel W, Riesbeck M, Wolwer W et al. Schizophr Res 2016;170:168-76.
5. Schooler NR, Keith SJ, Severe JB et al. Arch Gen Psychiatry 1997;54:453-63.
6. Emsley R, Chiliza B, Asmal L et al. BMC Psychiatry 2013;13:50.
7. Emsley R, Nuamah I, Hough D et al. Schizophr Res 2012;138:29-34.
8. Emsley R, Chiliza B, Asmal L. Schizophr Res 2013;148:117-21.
9. Leucht S, Tardy M, Komossa K et al. Cochrane Database Syst Rev 2012;5:CD008016.

DOI:10.1002/wps.20521

The long-term treatment of schizophrenia with antipsychotics: a perennial debate

A number of thoughts come to mind, when revisiting the recent debate around the benefit/risk ratio of antipsychotic maintenance treatment in schizophrenia patients.

First, it appears difficult to explain why one of the best documented findings in psychiatric treatment research, namely the fact that continuous intake of antipsychotics prevents relapses with a number-needed-to-treat of 3[1], a success rate which must be seen with envy from the perspective of other medical specialties, is questioned on the basis of a handful of studies of suboptimal methodological rigor[2].

Second, one wonders why historical hypothetical constructs, such as "supersensitivity psychosis"[3], which have not proven reasonably valid ever since they were originally put forward, experience a sudden renaissance.

Third, it is interesting to note how renowned clinician researchers, when reviewing the topic based on the same datasets, come to, at least subliminally, divergent conclusions, advocating the judicious use of antipsychotics on one hand[4] and providing cautionary criticism on the other[5].

Last, I find it disconcerting that rigorously designed state of the art clinical trials, fulfilling both the demands of academic psychopharmacology and the rules and guidelines of registration agencies, are still discussed with a scepticism of an almost paranoid quality just because they are "industry sponsored".

Let me set the record straight: I am absolutely in favour of iconoclastic paradigm shattering, if it is evidence based. This is one of the guiding principles of scientific research, to either replicate or falsify. However, in my humble view, I fail to find substantial evidence from a significant number of clinical trials which convincingly puts the principle fact that antipsychotics prevent relapse in schizophrenia into question. Needless to say, this does not obviate the necessity to adjust the finer details of antipsychotic relapse prevention. More recently, treatment expectations have extended beyond the mere prevention of the recurrence of psychotic symptoms. This takes me from my more general points to issues which more specifically address Correll et al's review[6]. While the authors provide a thoughtful, balanced and clinically most useful discussion of the topic, two issues, in my mind, deserve additional attention. One of them deals with outcome assessment and the other with psychosocial outcomes.

I would like to elaborate on assessment methodology from three perspectives: diagnosis, safety monitoring and quantifying psychosocial outcomes. With respect to the first, it needs to be acknowledged that schizophrenia is still a somewhat elusive concept. Despite the efforts of the DSM-5 and the forthcoming ICD-11, the heterogeneity of the syndrome, both with respect to psychopathological presentation and neurobiological underpinnings, has left us with a certain degree of diagnostic uncertainty. Obviously, this inhomogeneity also impinges upon the quality of clinical trials, leaving us with a considerable degree of variance, even when looking at clearly defined outcome measures such as symptom recurrence. This implies that, as evidenced in basically all other fields of medicine, we are left with group findings based on mean values, which allow us only limited predictions of individual outcomes. Although personalized or precision medicine is on everybody's wish list in our field as well, it has not yet become a clinical reality, although there is some light at the end of the tunnel[7].

A problem which appears somewhat easier to solve is that of reliably assessing safety and tolerability. Many clinical trialists still rely on spontaneous reporting of adverse events. This is notoriously unreliable, especially in a disorder with well-known communicative and cognitive impairments. Standardized rating scales for all adverse events, such as those available for extrapyramidal motor side effects, need to be implemented into clinical trials, especially into phase II and phase III studies. The discrepancy between rating scale based and subjectively assessed adverse events has been well documented for motor side effects[8].

In this context, I would also like to underscore the difficulties in reliably documenting psychosocial outcomes, starting by pointing to the difficulties in assessing quality of life from either a subjective or an objective perspective in patients suffering from reality distortion. This becomes even more challenging when considering the influence of sociocultural and geopolitical diversity in larger scale multicenter, often international, clinical trials. The same holds true when considering other relevant psychosocial outcomes, such as employment rates, which differ tremendously based on regional specifics. Even within the same country, recruiting patients from diverse socioeconomic backgrounds renders the interpretation of the obtained results very difficult.

Lastly, I would like to underscore the importance of stigma and discrimination from two different points of view, namely those caused by a psychotic relapse and by the side effects of medication. Starting with the latter, those of us with enough experience in the field to still remember heavily parkinsonized and akathisic patients on antipsychotics do appreciate the fact that these side effects, albeit not totally eliminated, are, in the true sense of the word, considerably less visible with new generation antipsychotics. Apart from the subjective discomfort that patients with motor side effects experience, this also considerably lessens the stigma caused by medication, as patients are less obviously "disturbed" in their motor appearance.

On a different but related note, stigma and discrimination can also be among the sequelae of psychotic symptoms, and the negative impact that unusual, odd and sometimes dangerous behaviour can have on psychosocial (re-)integration cannot be appreciated enough. It has been well documented that reducing antipsychotic dose below a critical level, or discontinuing medication altogether, enhances the risk for residual symptoms and/or relapse[9]. In an ideal world, society may find a certain level of symptom acceptable, if the patient does not subjectively suffer, yet, unfortunately, we do not live in this ideal world, and symptoms such as those experienced by schizophrenia patients still lead to a considerable amount of stigma and discrimination, which must not be underestimated.

All in all, I fully agree with Correll et al that the bulk of the available evidence still supports the judicious evidence-based use of maintenance antipsychotic treatment in most patients suffering from schizophrenia. Involving patients and, if available, significant others in treatment considerations is a *conditio sine qua non*. In addition, regular risk/benefit assessments, as well as medication adjustments based on a monitoring of symptom and safety/tolerability levels, are an obvious requirement. Although we may not yet have the tools to provide predictive personalized medicine, individualized care based on these considerations allows to optimize management options for every person affected with this serious mental disorder.

W. Wolfgang Fleischhacker
Division of Psychiatry I, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University Innsbruck, Innsbruck, Austria

1. Leucht S, Tardy M, Komossa K et al. Lancet 2012;379:2063-71.
2. Wunderink L, Nieboer RM, Wiersma D et al. JAMA Psychiatry 2013;70:913-20.
3. Chouinard G, Jones BD. Am J Psychiatry 1980; 137:16-21.
4. Goff DC, Falkai P, Fleischhacker WW et al. Am J Psychiatry 2017;174:840-9.
5. Murray RM, Quattrone D, Natesan S et al. Br J Psychiatry 2016;209:361-5.
6. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
7. DeLisi LE, Fleischhacker WW. Curr Opin Psychiatry 2016;29:187-9.
8. Widschwendter CG, Karayal ON, Kolluri S et al. Int J Neuropsychopharmacol 2015;18(12).
9. Uchida H, Suzuki T, Takeuchi H et al. Schizophr Bull 2011;37:788-99.

DOI:10.1002/wps.20542

Antipsychotic drugs: challenges and future directions

Some sixty years on from the first use of chlorpromazine to treat schizophrenia, it is worth reflecting on where we have come from. Back in the 1950s, it was not known that dopamine was a neurotransmitter, how antipsychotics worked, what symptoms they worked on, or indeed if they worked at all[1]. Now we know that dopamine is a neurotransmitter, antipsychotics are all dopamine receptor blockers and, as Correll et al[2] nicely review, large randomized, double-blind placebo-controlled trials have unequivocally demonstrated that they work both to treat acute psychotic episodes and to reduce relapse rates over the short to medium term.

Recent meta-analytic data generated from over sixty years of placebo-controlled trials estimate the standardized mean difference (SMD) between antipsychotics and placebo to be 0.38, with a greater effect seen on positive symptoms (SMD= 0.45) than negative symptoms (SMD= 0.35), quality of life (SMD=0.35) or depression (SMD=0.27)[3]. Such effect sizes are comparable to or larger than those found for treatments used for many common physical health conditions, including angiotensin-converting enzyme (ACE) inhibitors for reducing cardiac events and mortality due to hypertension (SMD= 0.16) and statins for reducing the risk of cardiac disease and stroke (SMD=0.15)[4]. Clearly, we have come a long way from the 1950s, but, despite these robust data on antipsychotics, many fundamental gaps in knowledge remain.

One glaring gap highlighted in this Forum is that as of yet we are unable to say conclusively what the optimum length of treatment with antipsychotic medication should be, once a patient has recovered from an acute episode. In current practice, many patients are treated with antipsychotic medication long-term if not lifelong, in an attempt to prevent the frequency and severity of relapses that can be so disruptive to a person's life.

Where patients are symptom free but experiencing side effects, such as weight

gain, that may shorten life as well as affect its quality, the risk-benefit balance for relapse prevention is finely poised. Yet, as Correll et al highlight, there is little evidence from randomized, double blind controlled studies to support prophylactic treatment beyond two-three years. Whilst some naturalistic studies do provide support for treatment beyond this term, the inherent limitations of these designs mean that the question remains unresolved, and guidelines cannot be conclusive.

This is a challenge to the field which needs to be met. We will need longer and, crucially, larger randomized controlled studies. This will not be easy, but other fields have risen to the challenge. For instance, in the case of the examples discussed above, statins and ACE inhibitors, there are now a number of randomized placebo-controlled trials with several thousand patients. These studies are roughly two orders of magnitude larger and five to ten times longer than the typical long-term randomized controlled study in schizophrenia. These large sample sizes give the power to have extended follow-up and account for treatment changes and drop-out. It is likely that we will need new ways of working, including international academic consortia as well as partnership with the pharmaceutical industry and governments, to achieve such large-scale studies.

Correll et al also highlight heterogeneity in schizophrenia, something that is increasingly becoming apparent in the neurobiology underlying the disorder as well as its clinical manifestations, course and treatment response[5].

Treatment resistance is probably the most clinically important manifestation of heterogeneity in patients with schizophrenia, and remains a major issue that continues to provoke debate over its pathophysiology, diagnosis and clinical management[6]. About a third of patients are thought to have treatment resistant illnesses, and around 15% show treatment resistance from illness onset[7]. Moreover, we have no way to identify the individuals whose illness will benefit from antipsychotic treatment.

Thus, large numbers of patients currently receive antipsychotic treatment although their illness is unlikely to respond to dopamine antagonists. The solutions to this will likely be found in part through identifying biomarkers that allow disease stratification, for example by the likelihood of response to dopamine receptor antagonists and, in the future, novel non-dopamine receptor blocking medication.

As both trial data and clinical experience show, current antipsychotic treatment works most effectively in reducing the positive symptoms of schizophrenia, whereas the negative and cognitive symptoms often remain problematic. Cognitive symptoms in particular are associated with poor functional outcomes in schizophrenia[8], yet our current treatments do nothing for them. In fact, there is evidence to suggest that dopamine antagonists may cause secondary negative and cognitive symptoms in people with schizophrenia[9]. Put simply, taking an antipsychotic may be unpleasant for some patients, and lead to secondary symptoms. This highlights the third challenge to the field: the need to develop treatments that are more than just antipsychotic and that patients are happy to take in the long term if necessary.

The final challenge is that our current antipsychotic medications are not disease modifying. Pre-synaptic striatal dopamine dysfunction is thought to drive the symptoms of schizophrenia[10], yet all of our current antipsychotic drugs act post-synaptically. Thus, they block the consequences of pre-synaptic dopamine dysfunction but do not address the underlying dopamine dysfunction, which remains present even after long-term treatment. This provides a neurobiological explanation for why patients may relapse on stopping antipsychotic treatment.

Targeting the upstream abnormality and/or the factors that lead to it is an alternative approach that could both be better tolerated and more effective in the long term. Broadly speaking, the glutamatergic and GABAergic systems have excitatory and inhibitory effects, respectively, on the dopamine system. Genetic studies measuring copy number variants in patients with schizophrenia[11] suggest that abnormalities in both neurotransmitter systems may be critical to the upstream regulation of dopamine. Findings like these suggest that targeting GABA and glutamate control of subcortical dopamine function could modify the pathophysiology, and potentially even be disease modifying. The interaction between psychosocial factors and biological changes[12] also highlights the potential for psychological treatments to be disease modifying.

It is clear that we have come a long way from the 1950s in terms of both understanding of the pathophysiology of schizophrenia and its treatment, and this has thrown up new questions and issues. Antipsychotic drugs are likely to remain a crucial part of our therapeutic arsenal for years to come, so it behoves us to address the questions that remain.

Oliver H. Howes[1-3], Stephen J. Kaar[1]
[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [2]Medical Research Council London Institute of Medical Sciences, London, UK; [3]Institute of Clinical Sciences, Imperial College, Hammersmith Hospital Campus, London, UK

1. Howes O, McCutcheon R, Stone J. J Psychopharmacol 2015;29:97-115.
2. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
3. Leucht S, Leucht C, Huhn M et al. Am J Psychiatry 2017;174:927-42.
4. Leucht S, Hierl S, Kissling W et al. Br J Psychiatry 2012;200:97-106.
5. Brugger SP, Howes OD. JAMA Psychiatry 2017; 74:1104-11.
6. Howes OD, McCutcheon R, Agid O et al. Am J Psychiatry 2016;174:216-29.
7. Lally J, Ajnakina O, Di Forti M et al. Psychol Med 2016;46:3231-40.
8. Green MF. J Clin Psychiatry 2016;77(Suppl. 2): 8-11.
9. Kirschner M, Aleman A, Kaiser S. Schizophr Res 2017;186:29-38.
10. Demjaha A, Murray RM, McGuire PK et al. Am J Psychiatry 2012;169:1203-10.
11. Pocklington AJ, Rees E, Walters JT et al. Neuron 2015;86:1203-14.
12. Howes OD, Murray RM. Lancet 2014;383:1677-87.

DOI:10.1002/wps.20522

Under-utilized opportunities to optimize medication management in long-term treatment of schizophrenia

Millions of people use antipsychotic medications and thousands of clinicians prescribe and monitor this treatment every day worldwide. In their review, Correll et al[1] highlight key issues regarding the long-term use of these medications. Herein we further discuss such issues, based on additional literature and data from Finnish cohort studies[2-6].

Three cornerstones in the long-term medication of schizophrenia are evidence-based use of antipsychotics, adjuvant psychosocial therapies, and optimal medication management. Correll et al review key findings and problems with the first and second of these cornerstones. Could good medication management alleviate these problems?

Medication management is defined as a process aimed to facilitate safe and effective use of medications and optimize therapeutic outcomes[7]. Multiple models of medication management have been put forward[2,7], but no systematic reviews, meta-analyses or universal treatment recommendations are available. Incomplete evidence on the content and cost-effectiveness of optimal antipsychotic medication management must be balanced with the development of organization-specific practices.

Proper approaches to medication management include shared decision making in prescription, follow-up, and monitoring at regular intervals[7]. In addition, careful documentation of response, continuity, and coordination of care should be ensured by a well-trained multidisciplinary team. However, in clinical practice, medication management is often suboptimal[6,7]. Schizophrenia patients with impaired cognition or motivation and/or poor financial resources have an elevated risk of inadequate medication management.

Important principles include avoiding maximal doses and polypharmacy, in favor of using the lowest possible effective and tolerated dose, choosing an antipsychotic associated with minimal side effects, and using adjuvant psychosocial interventions. For instance, maximal psychosocial and medication management interventions reduced the mean dose of antipsychotics from 370 to 160 mg/day as chlorpromazine equivalents in a Finnish therapeutic community ward of patients with acute psychosis[2]. However, currently used doses may sometimes be too high, as a reflection of insufficient or missing psychosocial therapies or poor medication management.

Current clinical practice guidelines are non-specific with respect to optimal doses, dose tapering, low-dose maintenance and discontinuation of antipsychotics. Guidelines do not specify how to go about tapering (i.e., at what point in the clinical course of illness, over what time period). This uncertainty may induce clinicians to set the bar high in dose reduction or withdrawal owing to potential risks. In practice, changes in antipsychotic dosing are done by testing and monitoring clinical response in the individual patient. This testing presupposes good medication management.

A vital long-term aim in medication management is to minimize unwanted drug effects such as tardive dyskinesia, weight gain or metabolic disturbances. Adverse effects attributable to chronic antipsychotic exposure are often cumulative over a period of years. A meta-analysis found associations between long-term antipsychotic use and brain volume changes[8]. Antipsychotics may also impact on brain plasticity and cognitive functioning[5]. Brain effects appear to be dose-dependent: high cumulative doses are related to brain alterations[3] and cognitive decline[5]. Paying attention to side effects, and adjusting and trying to find the lowest effective and tolerated dose could also decrease the dramatically high prevalence of medication non-adherence in schizophrenia[9].

As stressed by Correll et al, there are major methodological challenges when studying long-term antipsychotic use. Scientific evidence on dose reduction or medication discontinuation is primarily based on observational studies, which are subject to potential biases. Randomized controlled trials (RCTs) can help determine only the short-term efficacy and adverse effects. Such trials tend to be reductionistic when analyzing the complex interactions between brain, environment, drug effects and care. RCTs tend not to detect different subgroup effect sizes and long-term advantages and harms.

Non-adherence and attrition are also alarming problems in medication studies. Effective medication management may reduce them. In the Northern Finland Birth Cohort 1966 Study[3-6], we initially had a low participation rate of patients during the 9-year follow-up (44%). With maximal management efforts, including home visits, the participation rate increased in subsequent follow-ups (67%).

There are no major breakthroughs in the efficacy of antipsychotic treatment in sight. Current antipsychotics diminish illness expression, but do not restore lost complex brain functions. Many patients (and clinicians) do not utilize these medications optimally, even though their efficacy is quite high. Improving medication management and thus the risk-benefit ratio of antipsychotics is a realistic goal in the near future.

In summary, current care guidelines and practice standards advise us on how to use antipsychotics, most definitively at the group level and during the first years of illness. Long-term use and medication management skills and services are inadequately studied. It is important to learn what not to do when aiming at long-term improvement in medication management. Do not leave the patient alone with the medication. Do not forget the intellectual power of and need for psychoeducation and social support to patients and relatives. Do not remain silent or uninformed on patients' drug attitudes, treatment adherence, and negative experiences. Do not use only your brain, but also your heart and empathy.

Do not assume (even though you are a well-trained and experienced clinician) sole responsibility for long-term care and medication, but involve psychiatric and somatic teams, ensure continuity of care and organizational support.

The efficacy and risk-benefit balance of antipsychotics are not fixed. They can be improved through optimal medication management, particularly after the first years of the disease.

Matti Isohanni[1], Jouko Miettunen[1,2], Erika Jääskeläinen[1-3], Jani Moilanen[1], Anja Hulkko[1,2], Sanna Huhtaniska[1,2]

[1]Center for Life Course Health Research, University of Oulu, Oulu, Finland; [2]Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland; [3]Department of Psychiatry, Oulu University Hospital, Oulu, Finland

1. Correll CU, Rubio JM, Kane JM. World Psychiatry 2018;17:149-60.
2. Isohanni M. Acta Universitatis Ouluensis 1983; D111:5.
3. Veijola J, Guo JY, Moilanen JS et al. PLoS One 2014;9:e101689.
4. Moilanen J, Haapea M, Miettunen J et al. Eur Psychiatry 2013;28:53-8.
5. Husa AP, Rannikko I, Moilanen J et al. Schizophr Res 2014;158:134-41.
6. Nykänen S, Puska V, Tolonen J-P et al. Psychiatry Res 2016;235:160-8.
7. Howard PB, El-Mallakh P, Miller AL. Psychiatr Serv 2009;60:929-35.
8. Huhtaniska S, Jääskeläinen E, Hirvonen N et al. Hum Psychopharmacol Clin Exp 2017;32:e2574.
9. García S, Martínez-Cengotitabengoa M, López-Zurbano S et al. J Clin Psychopharmacol 2016; 36:355-71.

DOI:10.1002/wps.20523

The ICD-11 developmental field study of reliability of diagnoses of high-burden mental disorders: results among adult patients in mental health settings of 13 countries

Geoffrey M. Reed[1,2*], Pratap Sharan[3*], Tahilia J. Rebello[1,4], Jared W. Keeley[5], María Elena Medina-Mora[2], Oye Gureje[6], José Luis Ayuso-Mateos[7], Shigenobu Kanba[8], Brigitte Khoury[9], Cary S. Kogan[10], Valery N. Krasnov[11], Mario Maj[12], Jair de Jesus Mari[13], Dan J. Stein[14], Min Zhao[15], Tsuyoshi Akiyama[16], Howard F. Andrews[4,17], Elson Asevedo[13], Majda Cheour[18], Tecelli Domínguez-Martínez[2,19], Joseph El-Khoury[9], Andrea Fiorillo[12], Jean Grenier[20], Nitin Gupta[21], Lola Kola[6], Maya Kulygina[11], Itziar Leal-Leturia[7], Mario Luciano[12], Bulumko Lusu[14], J. Nicolas I. Martínez-López[2], Chihiro Matsumoto[22], Lucky Umukoro Onofa[23], Sabrina Paterniti[24], Shivani Purnima[3], Rebeca Robles[2], Manoj K. Sahu[25], Goodman Sibeko[14], Na Zhong[15], Michael B. First[1,4], Wolfgang Gaebel[26], Anne M. Lovell[27], Toshimasa Maruta[28], Michael C. Roberts[29], Kathleen M. Pike[1]

[1]Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, USA; [2]National Institute of Psychiatry Ramón de la Fuente Muñiz, Mexico City, Mexico; [3]Department of Psychiatry, All India Institute of Medical Sciences, New Delhi, India; [4]New York State Psychiatric Institute, New York, NY, USA; [5]Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA; [6]Department of Psychiatry, University of Ibadan, Nigeria; [7]Department of Psychiatry, Universidad Autonoma de Madrid, IIS-P and Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain; [8]Department of Neuropsychiatry, Kyushu University, Fukuoka City, Japan; [9]Department of Psychiatry, American University of Beirut Medical Center, Beirut, Lebanon; [10]School of Psychology, University of Ottawa, Ottawa, Ontario, Canada; [11]Moscow Research Institute of Psychiatry, National Medical Research Centre for Psychiatry and Narcology, Moscow, Russian Federation; [12]Department of Psychiatry, University of Campania "L. Vanvitelli", Naples, Italy; [13]Department of Psychiatry, Universidade Federal de São Paulo, São Paulo, Brazil; [14]Department of Psychiatry, University of Cape Town and South African Medical Research Council Unit on Risk and Resilience in Mental Disorders, Cape Town, South Africa; [15]Shanghai Mental Health Center and Department of Psychiatry, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China; [16]NTT Medical Center Tokyo, Tokyo, Japan; [17]Departments of Biostatistics and Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, USA; [18]Department of Psychiatry, Tunis Al Manar University and Al Razi Hospital, Tunis, Tunisia; [19]Cátedras CONACYT, National Council for Science and Technology, Mexico City, Mexico; [20]Institut du Savoir Montfort - Hôpital Montfort & Université d'Ottawa, Ottawa, Ontario, Canada; [21]Department of Psychiatry, Government Medical College and Hospital, Chandigarh, India; [22]Japanese Society of Psychiatry and Neurology, Tokyo, Japan; [23]Federal Neuropsychiatric Hospital Aro, Abeokuta, Nigeria; [24]Institute of Mental Health Research, Royal Ottawa Mental Health Centre, and Department of Psychiatry, University of Ottawa, Ontario, Canada; [25]Pt. Jawahar Lal Nehru Memorial Medical College, Raipur, Chhattisgarh, India; [26]Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany; [27]Institut National de la Santé et de la Recherche Médicale U988, Paris, France; [28]Health Management Center, Seitoku University, Matsudo City, Japan; [29]Office of Graduate Studies and Clinical Child Psychology Program, University of Kansas, Lawrence, KS, USA

*Joint first authors

Reliable, clinically useful, and globally applicable diagnostic classification of mental disorders is an essential foundation for global mental health. The World Health Organization (WHO) is nearing completion of the 11th revision of the International Classification of Diseases and Related Health Problems (ICD-11). The present study assessed inter-diagnostician reliability of mental disorders accounting for the greatest proportion of global disease burden and the highest levels of service utilization – schizophrenia and other primary psychotic disorders, mood disorders, anxiety and fear-related disorders, and disorders specifically associated with stress – among adult patients presenting for treatment at 28 participating centers in 13 countries. A concurrent joint-rater design was used, focusing specifically on whether two clinicians, relying on the same clinical information, agreed on the diagnosis when separately applying the ICD-11 diagnostic guidelines. A total of 1,806 patients were assessed by 339 clinicians in the local language. Intraclass kappa coefficients for diagnoses weighted by site and study prevalence ranged from 0.45 (dysthymic disorder) to 0.88 (social anxiety disorder) and would be considered moderate to almost perfect for all diagnoses. Overall, the reliability of the ICD-11 diagnostic guidelines was superior to that previously reported for equivalent ICD-10 guidelines. These data provide support for the suitability of the ICD-11 diagnostic guidelines for implementation at a global level. The findings will inform further revision of the ICD-11 diagnostic guidelines prior to their publication and the development of programs to support professional training and implementation of the ICD-11 by WHO member states.

Key words: International Classification of Diseases, ICD-11, diagnosis, mental disorders, reliability, schizophrenia, mood disorders, anxiety disorders, disorders specifically associated with stress

(World Psychiatry 2018;17:174–186)

A classification system that ensures satisfactorily reliable, clinically useful, and globally applicable diagnosis of mental disorders provides an essential foundation for global mental health. Such a system facilitates efficient identification of people with the greatest mental health needs when they seek health care and supports access to appropriate and cost-effective treatment[1].

Classification systems form the interface between health encounters and health information, and are an important foundation for decisions related to health policy and resource allocation at system, national and global levels. A classification that is too cumbersome to use at the encounter level or does not provide clinically useful information to the treating health professional will not be used as intended, cannot provide valid aggregate data, and will fail to support good clinical practice, research, and policy making[2].

The World Health Organization (WHO) is nearing completion of the 11th revision of the International Classification of Diseases and Related Health Problems (ICD-11), to be released for use by WHO member states in 2018. The revision has provided a major opportunity to bring the ICD classification of mental and behavioural disorders in line with current empirical evidence and clinical practice.

To achieve these aims, the WHO Department of Mental Health and Substance Abuse appointed a series of Working Groups to focus on different disorder areas, and these groups

have conducted comprehensive reviews of available evidence, on which their recommendations are based[3-8]. In addition, the Department of Mental Health and Substance Abuse has undertaken a systematic and comprehensive program of formative and evaluative field studies focused particularly on the global applicability and clinical utility of the revised Clinical Descriptions and Diagnostic Guidelines (CDDG) for ICD-11 Mental, Behavioural and Neurodevelopmental Disorders. These field studies are substantially different from previous efforts in their use of innovative methodologies to investigate the application of the guidelines in the context of clinical decision making[9].

The effectiveness of the ICD-11 CDDG in producing more consistent clinical judgments as compared to the ICD-10 CDDG[10] is currently being tested in a series of Internet-based, multilingual case controlled field studies using standardized case material in the form of vignettes, as these allow for experimental manipulation of the clinical information in a way that isolates the effects of the classification system on diagnostic decision making[11,12]. The use of the written vignettes offers many advantages in terms of standardization and experimental control.

As an important next step in evaluating the CDDG, studies of their implementation in clinical settings provide a fuller approximation of the subtleties of assessment, interpretation and decision making involved in making diagnoses in real patients. Accordingly, ecological implementation field studies (EIFS) are being conducted in clinical settings in a range of countries to investigate the diagnostic reliability and clinical utility of the proposed ICD-11 CDDG. The EIFS centers are located in countries that diverge widely in terms of languages, culture, and resource levels. The initial results of these studies are described in this paper.

The proposed structure and content of the ICD-11 CDDG were designed to enhance their clinical utility, validity and reliability[13]. The WHO has focused on improving clinical utility in the ICD-11 because it is critical to the WHO's public health goals related to reducing the global burden of mental disorders[1]. The WHO defines clinical utility for classificatory systems as including their value in communicating among stakeholders, their implementation characteristics in clinical practice (e.g., goodness of fit, time required to use them), and their usefulness in making clinical management decisions[14].

Thus, clinical utility, validity and reliability are distinct but overlapping constructs[15]. An example of the relationship between reliability and clinical utility of diagnoses was provided by the ICD-10 CDDG field trials[16], which showed that diagnoses with lower reliability were accompanied by lower-than-average ratings of clinical utility (e.g., diagnostic fit, confidence in diagnosis, ease of use, and adequacy of description). Similarly, aspects of the validity of diagnostic constructs also relate to their inherent clinical usefulness in the care of patients, for example in predicting treatment response or course of illness[17].

The reliability of mental disorders diagnoses has been a focus of attention in the revision processes of both the ICD and the Diagnostic and Statistical Manual of Mental Disorders (DSM) of the American Psychiatric Association since the 1970s[18]. Both classificatory systems adopted a descriptive approach to providing diagnostic guidance[19-21], in part based on studies suggesting that deficiencies in pre-DSM-III classification systems were major sources of unreliability[22-24].

In general, studies of the reliability of diagnostic classifications following the publication of DSM-III documented improved results[18]. However, the lower diagnostic reliability documented in the DSM-5 field trials[25] compared to previous field trials has highlighted the profound influence of methodology on estimates of diagnostic reliability[26]. That is, reliability is not solely a property of the classification, but also a product of the method used to estimate it. This makes comparisons across studies with different methodologies quite difficult.

The current study has used a naturalistic, joint-rater design to estimate inter-diagnostician reliability. Unlike some previous studies of the reliability of classification systems[27,28], structured interviews, which could be expected to increase reliability substantially[29-31], were not used. No instruction or training was provided regarding how clinician raters should perform the diagnostic interview, and clinician raters received relatively minimal training on the new ICD-11 guidelines. The attempt was therefore to approximate the conditions under which the guidelines will be applied in clinical settings after their publication.

The joint-rater design was used in order to minimize information variance and to focus specifically on the question of whether two clinicians, relying on the same clinical information, agree on the diagnoses to be assigned to the patient when separately applying the ICD-11 diagnostic guidelines.

Similar to the naturalistic design of the current ICD-11 EIFS, the developers of the field studies for DSM-III, ICD-10 and DSM-5 also chose not to employ structured diagnostic interviews because they are not commonly used in general clinical settings[16,25,32,33]. The DSM-III and ICD-10 CDDG field trials demonstrated good diagnostic reliability for most major classes of disorders. However, reliability estimates were likely inflated in the case of DSM-III by presentation of estimates only for disorder groupings (rather than individual diagnoses)[32] and, in the case of the ICD-10 CDDG, by the use of case conferences – in which one diagnostician interviewed the patient and then presented the case to other assessors – for establishing inter-diagnostician reliability[16].

The DSM-5 field trials also used a naturalistic design, employing two diagnosticians to assess inter-rater agreement on diagnoses and computing reliability at the level of individual diagnoses[34]. However, those field trials used a sequential, test-retest design (two diagnosticians interviewing the participant at different time points) to establish inter-clinician reliability, rather than the concurrent, joint-rater design (two clinicians interviewing the participant together) employed in the ICD-11 EIFS. The DSM-5 design, therefore, did not control for information variance and so would almost certainly yield lower reliabilities[26,35]. Therefore, reliability estimates of the recent DSM-5 field trials and the current ICD-11 EIFS are not comparable.

Arguably, the DSM-5 design is a test of the diagnostic reliability of psychiatric diagnoses more generally and not specifically of the new diagnostic manual.

The ICD-11 EIFS were designed as developmental studies with the goal of using the results in the final revision of the guidelines, rather than solely as evaluative field studies, which aim to assess what users can expect in terms of the classification's psychometric properties after the classification has been completed[36]. The concurrent joint-rater reliability design was preferred for the EIFS because it made it possible to focus on variation in the application or interpretation of the diagnostic guidelines, controlling for variance due to patient factors (e.g., giving different histories to the diagnosticians) and extraneous clinician factors (e.g., variations in the thoroughness of the interview).

The concurrent joint-rater design employed in ICD-11 EIFS focused specifically on the role of the diagnostic guidelines themselves as a source of unreliability. In a developmental field study, identification of high levels of clinician-criterion incongruity should prompt changes to the diagnostic guidelines, whereas clinician errors are likely better addressed through training on the use of the classification and clinical interviewing.

The reliability arm of EIFS described in this paper specifically targeted four groups of disorders among adult patients: schizophrenia and other primary psychotic disorders, mood disorders (including both depressive and bipolar disorders), anxiety and fear-related disorders, and disorders specifically associated with stress. These diagnoses account for the greatest proportion of global disease burden among mental disorders[37] and the highest levels of service utilization in mental health settings.

This paper describes the EIFS results concerning reliability of the proposed ICD-11 CDDG among adult patients in 13 countries.

METHODS

Study design and procedures

Two study protocols were implemented to assess the reliability of the proposed ICD-11 diagnostic guidelines. Protocol 1 tested the reliability of the guidelines for schizophrenia and other primary psychotic disorders and for mood disorders, while Protocol 2 tested the guidelines for mood disorders, anxiety and fear-related disorders, and disorders specifically associated with stress.

Adult (≥18 years of age) patients exhibiting any psychotic symptoms and presenting for care at the participating field study center were eligible for Protocol 1. Adult (≥18 years of age) patients exhibiting mood symptoms, anxiety symptoms, or stress-related symptoms but no psychotic symptoms and presenting for care at the participating field study center were eligible for Protocol 2. These requirements were intended to produce an enriched sample that was likely to have at least one of the conditions being tested, but whose diagnostic status was not determined in advance.

Exclusion criteria for both protocols were the following: communication difficulty sufficient to interfere with participation in the diagnostic interview (e.g., lack of proficiency in the language of the clinicians at the site); cognitive dysfunction to an extent that would interfere with participation in the diagnostic interview; current incapacitation due to severe physical illness or pain; current substance intoxication or withdrawal or serious medication side effects; and current imminent risk of harm to self or other. These criteria essentially functioned to allow any consenting patient exhibiting the index symptoms to be recruited, unless they could not reasonably be expected to participate in the diagnostic interview.

Protocols were implemented at 28 sites in 13 countries. Additional site information is presented in Table 1.

The local language was always used for the diagnostic interviews. The ICD-11 guidelines, training materials, and all material for the study were developed in English. Materials were then translated into four other languages – Chinese, Japanese, Russian and Spanish – with the collaboration of field study centers, using a thorough forward and back-translation process. In other sites, the English guidelines and training materials were used even though the interviews were conducted in other languages, again replicating the circumstances under which the ICD-11 will be implemented.

All sites obtained ethical clearance from their institutional review boards prior to study implementation. Research teams defined local procedures for obtaining consent and for reporting and addressing any adverse events that might be experienced by participants who were being interviewed as part of the study (e.g., inability to complete the interview due to high levels of symptoms or distress). Participants were assigned unique identification numbers, and no confidential or identifying information was reported to anyone outside the site.

A site director was responsible at each site for recruiting clinician raters. According to the practice standards of their countries, all clinician raters were qualified to make mental disorders diagnoses independently as a part of their scope of practice. Advanced residents in psychiatry (following completion of first two years of residency) could function as interviewers but were always paired with a fully qualified individual. Training was organized either at the level of the site or for multiple sites within a given country.

Clinician raters were provided with the ICD-11 diagnostic guidelines being tested and were asked to review them prior to the training session. The training session reviewed central features of the ICD-11 diagnostic guidelines in those areas covered by the protocols and their differences with ICD-10. The sessions used a standard set of slides developed by the WHO. Interactive exercises provided an opportunity for practice in applying the guidelines to case vignettes. The only difference between Protocol 1 and Protocol 2 was that, for the former,

Table 1 Participating country and study site information

Country	Protocol(s) implemented	N. sites	Site names	N. raters
Brazil	1	1	Universidade Federal de São Paulo	21
Canada	2	1	Royal Ottawa Mental Health Centre/University of Ottawa Institute of Mental Health Research	7
China	1 and 2	1	Shanghai Mental Health Center	25
India	1 and 2	3	All India Institute of Medical Sciences, New Delhi	44
			Government Medical College Hospital, Chandigarh	
			Pandit Jawaharlal Nehru Memorial Medical College, Raipur	
Italy	1	1	University of Campania "L. Vanvitelli", Naples	14
Japan	1 and 2	11	Kyushu University	90
			Hokkaido University	
			University of Occupational & Environmental Health, Kitakyushu	
			Tokyo Medical Dental University	
			Tokyo Metropolitan Matsuzawa Hospital	
			Nihon University School of Medicine, Tokyo	
			Nagoya University	
			Hizen National Psychiatric Center, Yoshinogari	
			NTT Medical Center Tokyo	
			Tokyo University	
			Tokushima University	
Lebanon	1 and 2	2	American University of Beirut	14
			Hôpital Psychiatrique De La Croix, Jal El Dib	
Mexico	1 and 2	1	National Institute of Psychiatry Ramón de la Fuente Muñiz, Mexico City	23
Nigeria	1	2	University College Hospital, Ibadan	32
			Federal Neuropsychiatric Hospital, Aro, Abeokuta	
Russia	1	2	Moscow Research Institute of Psychiatry	41
			First Saint Petersburg City Mental Hospital	
South Africa	1 and 2	1	Valkenberg Psychiatric Hospital, Cape Town	10
Spain	1 and 2	1	Hospital Universitario La Princesa, Madrid	6
Tunisia	1 and 2	1	Razi Hospital, Tunis	12

clinician raters were informed that they were required to assess for schizophrenia and other primary psychotic disorders and for mood disorders, as well as for any other area they deemed relevant in arriving at a diagnostic formulation, while for the latter they were required to assess for mood disorders, anxiety and fear-related disorders, and disorders specifically associated with stress. No other instruction was given about how to approach the interview, and it was left to the judgment of the clinician raters to determine how best to perform the assessment, according to their professional training and usual practice, as will be the case when the ICD-11 is implemented.

Training sessions lasted for approximately two hours per protocol (i.e., approximately four hours for sites that were doing both Protocol 1 and Protocol 2). Training sessions were therefore not dissimilar to those that clinicians might realistically receive when the ICD-11 is implemented in their countries. The sessions also covered the study flow and data collection procedures. Post-training and prior to start of data collection, clinician raters registered to participate using an online registration system, providing demographic information as well as details regarding their clinical experience (see Table 2).

A broader group of clinicians at each study site were given information on the study inclusion and exclusion criteria and referral procedures, and asked to refer qualifying patients to either Protocol 1 or Protocol 2. At most sites, clinician raters who conducted joint-rater interviews were also part of the pool of referring clinicians, in which case they were not permitted to

interview any patient they had referred. Referring clinicians were invited to participate in the training sessions for interviewers, though this was not mandatory.

Upon referral, a research coordinator explained the study to referred patients and obtained their informed consent. Following informed consent, patients were interviewed by two clinicians who had no prior clinical contact with the patient. One clinician rater served as the primary interviewer and the second as an observer. The observer was allowed to ask additional follow-up questions at the end of the interview. Clinician pairings were varied as much as possible given constraints of availability and scheduling, and participating clinicians alternated as primary interviewer and observer.

The clinician raters were instructed to set aside 60-90 min for the joint-rater interview. They were asked to approach assessments as they would in routine practice. The range and length of the diagnostic interviews were therefore substantially consistent with usual practice in participating mental health centers.

Based on the interview, and in some cases additional supplementary material provided to both clinicians (e.g., patient file excluding current or prior psychiatric diagnoses and psychotropic medication prescriptions, interviews with family members), clinician raters independently arrived at a diagnostic formulation consisting of up to three diagnoses. Diagnoses were non-hierarchical (i.e., not specified as primary, secondary or tertiary) and could fall within any mental, behavioural or neurodevelopmental disorder diagnostic grouping. Participating clinicians could also specify a non-mental or behavioural disorder diagnosis, or no diagnosis. For diagnoses included in Protocol 1 and Protocol 2, additional detailed questions were asked about symptom presentation and clinical utility of the guidelines.

Following the interview, both clinician raters independently provided data based on the interview using a secure web-based data collection system. Participating clinicians were instructed to record their data within 24 hours. Information provided included each clinician rater's diagnostic formulation, and ratings of the presence or absence of each element of any disorder from the diagnostic groupings that were the focus of Protocol 1 or Protocol 2. Data provided by each clinician also included responses to detailed questions about the clinical utility of the diagnostic guidelines as applied to that particular patient.

Participants

A total of 339 clinicians from the 28 study sites in 13 countries (see Table 2) served as clinician raters for Protocol 1 and/or Protocol 2. The mean age of clinician raters was 37.2 ± 8.3 years, and the ages were comparable across countries. There was a slight majority of male clinician raters in the global sample (56.6%). The overwhelming majority of clinician raters in the study were psychiatrists (93.2%), with a small representation of psychologists (3.8%), nurses (1.5%) and other medical professionals (1.5%). The average clinical experience of the clinician raters was 7.6 ± 7.5 years.

As shown in Table 3, 1,806 patients were recruited into the study for Protocol 1 (N=1,041) or Protocol 2 (N=765). The average age of participating patients was 39.9 ± 13.7 years, and ages were comparable across countries. The global sample had an equal gender distribution. The marital status of the majority of patients in the global sample was single (54.9%); 33.1% were married/cohabitating, 9.8% were separated/divorced and 2.2% were widowed. More than half of the patients in the global sample were unemployed (55.9%) and only 22.3% of the patients had full time employment. A slight majority of recruited patients in the global sample were inpatients (55.0%) and the remainder were nearly all outpatients (44.4%). The small remaining proportion (0.6%) were enrolled in other types of programs such as partial day hospitalization.

Data collection, management and processing

Data reported by clinician interviewers were securely collected using the Electronic Field Study System (EFSS), a web-based data collection system developed using Qualtrics™ (Provo, UT, USA) and made available in five study languages. Clinicians logged onto the EFSS using a unique password to report all study data.

Data from the sites were stored and managed centrally by the Data Coordinating Center (DCC) at Columbia University. Data quality was established through continuous monitoring of the data collection procedures by local research staff at each site and through use of programming functions within Qualtrics™, such as forced response and content validation options. This provided a mechanism for collecting data in a standardized, uniform format from all sites. Site-based research teams kept records of any errors in data entry that were passed on to the DCC for correction.

Data analysis

The main analysis of the study addressed the reliability of diagnoses included in Protocols 1 and 2. Data from both protocols were combined in the current analyses. Diagnostic reliability was estimated based on agreements between clinician raters irrespective of whether the diagnosis was listed first, second or third. For example, if for a particular patient one clinician rater diagnosed single episode depressive disorder, panic disorder, and agoraphobia, and the other clinician rater diagnosed agoraphobia and single episode depressive disorder, both clinician raters would have agreement on the diagnosis of single episode depressive disorder and agoraphobia, but disagreement for panic disorder.

Only diagnoses that occurred at least 30 times across the study were included in these analyses, as diagnoses assigned less frequently were not considered to have sufficient stability for the present evaluation. To estimate diagnostic reliability,

Table 2 Demographics of clinician raters by country

	Total (N=339)	Brazil (N=21)	Canada (N=7)	China (N=25)	India (N=44)	Italy (N=14)	Japan (N=90)	Lebanon (N=14)	Mexico (N=23)	Nigeria (N=32)	Russia (N=41)	South Africa (N=10)	Spain (N=6)	Tunisia (N=12)
Age, years (mean±SD)	37.2 ± 8.3	35.5 ± 8.6	44.4 ± 13.8	32.6 ± 4.8	34.1 ± 7.4	39.8 ± 6.2	38.9 ± 7.7	36.1 ± 8.6	37.6 ± 7.9	37.8 ± 6.0	39.2 ± 11.1	35.5 ± 7.0	32.0 ± 5.9	38.3± 9.3
Gender, N (%)														
Male	192 (56.6)	10 (47.6)	1 (14.3)	5 (20.0)	29 (65.9)	9 (64.3)	72 (80.0)	7 (50.0)	12 (52.2)	25 (78.1)	14 (34.1)	5 (50.0)	2 (33.3)	1 (8.3)
Female	145 (42.8)	11 (52.4)	6 (85.7)	20 (80.0)	15 (34.1)	5 (35.7)	16 (17.8)	7 (50.0)	11 (47.8)	7 (21.9)	27 (65.9)	5 (50.0)	4 (66.7)	11 (91.7)
Clinical profession, N (%)														
Psychiatry	316 (93.2)	21 (100)	2 (28.6)	25 (100)	44 (100)	14 (100)	88 (97.8)	11 (78.6)	22 (95.7)	32 (100)	39 (95.1)	3 (30.0)	5 (83.3)	10 (83.3)
Psychology	13 (3.8)	0	5 (71.4)	0	0	0	0	3 (21.4)	1 (4.3)	0	1 (2.4)	1 (10.0)	1 (16.7)	1 (8.3)
Nursing	5 (1.5)	0	0	0	0	0	0	0	0	0	0	5 (50.0)	0	0
Other medical	5 (1.5)	0	0	0	0	0	2 (2.2)	0	0	0	1 (2.4)	1 (10.0)	0	1 (8.3)
Years of experience (mean±SD)	7.6 ± 7.5	6.6 ± 7.4	13.3 ± 11.9	4.2 ± 3.9	5.4 ± 6.4	7.7 ± 7.2	7.1 ± 6.7	7.8 ± 7.4	9.2 ± 8.2	5.8 ± 4.1	13.6 ± 10.3	6.4 ± 3.9	6.0 ± 4.8	6.5 ± 6.6

Table 3 Demographics of patients by country

	Total (N=1,806)	Brazil (N=100)	Canada (N=53)	China (N=203)	India (N=209)	Italy (N=100)	Japan (N=168)	Lebanon (N=103)	Mexico (N=153)	Nigeria (N=132)	Russia (N=104)	South Africa (N=208)	Spain (N=70)	Tunisia (N=203)
Age, years (mean±SD)	39.9 ± 13.7	32.9 ± 9.6	39.8 ± 14.2	43.9 ± 15.6	36.5 ± 11.4	41.4 ± 11.2	47.0 ± 15.1	36.4 ± 12.5	38.1 ± 13.0	37.5 ± 12.2	36.3 ± 11.7	35.1 ± 11.0	52.0 ± 16.2	43.2 ± 12.6
Gender, N (%)														
Male	908 (50.3)	62 (62.0)	19 (35.8)	123 (60.6)	120 (57.4)	50 (50.0)	72 (42.9)	38 (36.9)	48 (31.4)	65 (49.2)	44 (42.3)	133 (63.9)	26 (37.1)	108 (53.2)
Female	897 (49.7)	38 (38.0)	33 (62.3)	80 (39.4)	89 (42.6)	50 (50.0)	96 (57.1)	65 (63.1)	105 (68.6)	67 (50.8)	60 (57.7)	75 (36.1)	44 (62.9)	95 (46.8)
Relationship status, N (%)														
Single	992 (54.9)	81 (81.0)	22 (41.5)	110 (54.2)	66 (31.6)	71 (71.0)	77 (45.8)	68 (66.0)	91 (59.5)	68 (51.5)	65 (62.5)	167 (80.3)	28 (40.0)	78 (38.4)
Married/ cohabitating	597 (33.1)	12 (12.0)	17 (32.1)	75 (36.9)	133 (63.6)	19 (19.0)	64 (38.1)	20 (19.4)	42 (27.5)	41 (31.1)	22 (21.2)	25 (12.0)	28 (40.0)	99 (48.8)
Separated/ divorced	177 (9.8)	6 (6.0)	13 (24.5)	15 (7.4)	4 (1.9)	7 (7.0)	21 (12.5)	15 (14.6)	20 (13.1)	18 (13.6)	13 (12.5)	13 (6.3)	9 (12.9)	23 (11.3)
Widowed	40 (2.2)	1 (1.0)	1 (1.9)	3 (1.5)	6 (2.9)	3 (3.0)	6 (3.6)	0	0	5 (3.8)	4 (3.8)	3 (1.4)	5 (7.1)	3 (1.5)
Employment, N (%)														
Full time	403 (22.3)	4 (4.0)	14 (26.4)	47 (23.2)	69 (33.0)	11 (11.0)	26 (15.5)	16 (15.5)	17 (11.1)	41 (31.1)	22 (21.2)	22 (10.6)	26 (37.1)	88 (43.3)
Part time	142 (7.9)	5 (5.0)	6 (11.3)	3 (1.5)	12 (5.7)	9 (9.0)	14 (8.3)	11 (10.7)	31 (20.3)	11 (8.3)	6 (5.8)	8 (3.8)	3 (4.3)	23 (11.3)
Unemployed	1009 (55.9)	76 (76.0)	30 (56.6)	80 (39.4)	110 (52.6)	74 (74.0)	109 (64.9)	66 (64.1)	79 (51.6)	64 (48.5)	53 (51.0)	167 (80.3)	20 (28.6)	81 (39.9)
Student	136 (7.5)	6 (6.0)	4 (7.5)	15 (7.4)	15 (7.2)	4 (4.0)	10 (6.0)	15 (14.6)	30 (19.6)	10 (7.6)	7 (6.7)	12 (5.8)	2 (2.9)	6 (3.0)
Retired	152 (8.4)	10 (10.0)	1 (1.9)	62 (30.5)	3 (1.4)	2 (2.0)	15 (8.9)	0	5 (3.3)	8 (6.1)	18 (17.3)	0	22 (31.4)	6 (3.0)
Treatment setting, N (%)														
Outpatient	801 (44.4)	82 (82.0)	53 (100)	0	122 (58.4)	67 (67.0)	48 (28.6)	14 (13.6)	135 (88.2)	84 (63.6)	4 (3.8)	0	49 (70.0)	143 (70.4)
Inpatient	994 (55.0)	18 (18.0)	0	203 (100)	87 (41.6)	33 (33.0)	120 (71.4)	89 (86.4)	17 (11.1)	48 (36.4)	91 (87.5)	207 (99.5)	21 (30.0)	60 (29.6)
Other	11 (0.6)	0	0	0	0	0	0	0	1 (0.7)	0	9 (8.7)	1 (0.5)	0	0

Table 4 Concurrent reliability of ICD-11 diagnoses

	Number of diagnoses (N)	Joint rater agreement (intraclass kappa)	Standard error	Bootstrapped 95% CI
Schizophrenia	725	0.87	0.012	0.84-0.89
Schizoaffective disorder	189	0.66	0.035	0.58-0.72
Acute and transient psychotic disorder	40	0.45	0.087	0.27-0.60
Delusional disorder	30	0.69	0.084	0.51-0.84
Bipolar I disorder	351	0.84	0.017	0.81-0.87
Bipolar II disorder	95	0.62	0.048	0.52-0.70
Single episode depressive disorder	191	0.64	0.035	0.57-0.77
Recurrent depressive disorder	267	0.74	0.025	0.69-0.79
Dysthymic disorder	57	0.45	0.073	0.28-0.58
Generalized anxiety disorder	129	0.62	0.044	0.53-0.70
Panic disorder	59	0.57	0.069	0.42-0.69
Agoraphobia	46	0.62	0.072	0.47-0.75
Social anxiety disorder	38	0.88	0.045	0.78-0.95
Post-traumatic stress disorder	51	0.49	0.076	0.33-0.64
Complex post-traumatic stress disorder	45	0.56	0.077	0.40-0.71
Adjustment disorder	82	0.73	0.046	0.63-0.81

intraclass kappa coefficients for diagnoses weighted by site and study prevalence were calculated. Bootstrapped 95% confidence intervals for kappa, based upon 1,000 resamples, were then calculated. All analyses were conducted using SPSS.

Landis and Koch[38] adjectives were used to describe ranges of reliability values for kappa: slight (from 0 to 0.20), fair (from 0.21 to 0.40), moderate (from 0.41 to 0.60), substantial (from 0.61 to 0.80), and almost perfect (from 0.81 to 1.0).

RESULTS

Estimates of joint-rater agreement are shown in Table 4, along with bootstrapped 95% confidence intervals. The point estimate of kappa ranged from 0.45 (dysthymic disorder) to 0.88 (social anxiety disorder) and would be considered moderate to almost perfect according to Landis and Koch adjectives for all diagnoses for which it was calculated.

The kappa estimates were almost perfect for schizophrenia (0.87) and bipolar I disorder (0.84); substantial for schizoaffective disorder (0.66), delusional disorder (0.69), bipolar II disorder (0.62), single episode depressive disorder (0.64), recurrent depressive disorder (0.74), generalized anxiety disorder (0.62), agoraphobia (0.62), and adjustment disorder (0.73); and moderate for acute and transient psychotic disorder (0.45), dysthymic disorder (0.45), panic disorder (0.57), post-traumatic stress disorder (0.49), and the newly introduced diagnosis of complex post-traumatic stress disorder (0.56).

In general, point estimates of kappa were lower for disorders for which smaller samples were obtained. The higher number of diagnoses of primary psychotic and mood disorders reflects the type of settings (55% inpatient) and the nature of the centers (tertiary and secondary care) involved in the reliability arm of EIFS.

The estimates of kappa were precise for all diagnoses for which it was calculated (confidence interval <0.5; standard error <0.1). The lower bound estimates of the confidence interval for kappa were higher than 0.4 (fair reliability) for 13 of the 16 disorders. However, the lower bound estimates were only in the fair range (from 0.2 to 0.4) for acute and transient psychotic disorder (0.27), dysthymic disorder (0.28), and post-traumatic stress disorder (0.33). All diagnoses with lower bound confidence interval estimates of kappa (<0.4) were made less often, suggesting that higher reliability for these disorders might accrue in samples of larger sizes.

Table 5 provides a comparison of the results of joint-rater agreement in the current study of the ICD-11 CDDG with the results of the ICD-10 CDDG field trial[16]. This comparison is intended to be illustrative rather than exact because of major differences in study methodologies. Unlike the ICD-11 EIFS, which used two raters for face-to-face joint rater interviews, the ICD-10 field study used case conferences, in which one rater conducted a face-to-face interview and then presented the case to other raters as a basis for establishing inter-rater reliability. The case conference methodology is likely to produce more consensus-based results, in which reliability would be correspondingly higher. Further, though most ICD-10 diagnoses correspond closely to proposed ICD-11 diagnoses, they are not identical.

While statistical comparisons of the two studies are not justified, in 10 of the 14 possible comparisons between the ICD-

Table 5 Comparison of reliability estimates in ICD-11 CDDG EIFS and ICD-10 CDDG field trials

ICD-11 EIFS		ICD-10 CDDG field trail	
	Kappa (N)		**Kappa (N)**
Schizophrenia	0.87 (725)	F20 Schizophrenia	0.81 (490)
Schizoaffective disorder	0.66 (189)	F36 Schizoaffective disorder	0.48 (148)
Acute and transient psychotic disorder	0.45 (40)	F23 Acute/transient psychotic disorders	0.65 (146)
Delusional disorder	0.69 (30)	F22.0 Delusional disorder	0.62 (83)
Bipolar I disorder	0.84 (351)	F30 Manic episode	0.69 (53)
		F31 Bipolar affective disorders	0.81 (259)
Single episode depressive disorder	0.64 (191)	F32 Depressive episode	0.66 (353)
Recurrent depressive disorder	0.74 (267)	F33 Recurrent depressive disorders	0.69 (302)
Dysthymic disorder	0.45 (57)	F34.1 Dysthymia	0.36 (101)
Generalized anxiety disorder	0.62 (129)	F41.1 Generalized anxiety disorder	0.48 (67)
Panic disorder	0.57 (59)	F41.0 Panic disorder	0.74 (31)
Agoraphobia	0.62 (46)	F40.0 Agoraphobia	0.51 (22)
Social anxiety disorder	0.88 (38)	F40.1 Social phobias	0.41 (22)
Post-traumatic stress disorder	0.49 (51)	F43.1 Post-traumatic stress disorder	0.62 (23)
Adjustment disorder	0.73 (82)	F43.2 Adjustment disorder	0.54 (107)

CDDG – Clinical Descriptions and Diagnostic Guidelines, EIFS – Ecological Implementation Field Studies

11 CDDG EIFS and the ICD-10 CDDG field study, the kappa values were higher for ICD-11. These differences tended to be modest.

DISCUSSION

The 11th revision of the Mental, Behavioural and Neurodevelopmental Disorders chapter of the ICD has proposed substantive changes to the conceptualization of many disorders, which may impact their reliability, validity and clinical utility. Field studies that assess how well the proposed changes perform in the hands of the intended users are crucial to this revision process. Accordingly, the EIFS for proposed ICD-11 CDDG were conducted in a broad spectrum of secondary and tertiary mental health care settings across countries with varied languages, cultures, and resource levels.

The results of the ICD-11 EIFS show that all common and high-burden disorders in the adult population covered in the current study were diagnosed with at least satisfactory – and in most cases excellent – reliability by a sample of clinician raters that included advanced trainees in psychiatry as well as more experienced clinicians. This suggests that the proposed ICD-11 CDDG are suitable for use at a global level and that their satisfactory implementation extends beyond application to written vignettes to application to real patients in clinical settings.

Reliability of diagnosis impacts clinical communication, generalizability of the guidelines across patient populations, and tailoring of treatments according to diagnosis, in addition to the selection of samples for research. The DSM-III had introduced fully operationalized diagnostic criteria in the classification of mental disorders as a way of improving diagnostic reliability[30,31]. The ICD-11 CDDG were designed to align with the overarching principles of categorization emerging from earlier studies analyzing how clinicians naturally organize clinical conditions[2]. ICD-11 disorders are presented in terms of the essential features that clinicians could reasonably expect to find in all cases, in an effort to communicate the essence of the disorder, with greater flexibility for clinical and cultural variation[13]. The ICD-11 CDDG avoid fully operationalized criteria characterized by precise cutoffs and symptom counts, unless these are specifically empirically supported. The present results challenge the assumption that the more clinician-friendly, less concretely algorithmic, and less precisely specified approach adopted for the ICD-11 CDDG is inherently less reliable.

The reliability coefficients observed in this study were based on routine clinical assessments (lasting about one hour) using open form interviews by clinicians with diverse training and experience. The results were similar to those achieved by diagnostic assessments using more complex and time consuming structured instruments[26,39,40]. These results suggest that the use of more uniform procedures by clinicians based on a brief training may yield adequate reliability for commonly diagnosed mental disorders in clinical settings. A hypothesis that would be well worth testing – given the resources that are devoted to the refinement of diagnostic criteria – is that further gains could be obtained by focusing greater attention on

appropriate training in diagnostic skills and interviewing techniques[41], rather than on introducing greater precision in the strict operationalization of diagnostic guidelines.

In general, the reliability of diagnoses in ICD-11 CDDG was superior to that of diagnoses in ICD-10 CDDG[16], though strict comparisons are not appropriate due to differences in methodology of these field studies. Similar comparisons with the studies of ICD-10 Diagnostic Criteria for Research[33] and the DSM-III[32] were not performed because of even greater methodological differences. The ICD-10 Diagnostic Criteria for Research field trial involved the use of a structured diagnostic instrument that covered the diagnostic criteria for assessment[33]. The published results of DSM-III field trial provided kappa values for disorder groupings rather than for specific disorders[32], which would tend to maximize reliability results because disagreements within a grouping are substantially more likely than disagreements concerning disorders from different groupings.

Changes to the ICD-11 CDDG relative to the ICD-10 CDDG were proposed by expert working groups based on the available scientific evidence and with explicit attention to additional sources of information related to clinical utility and global applicability. In no case were changes proposed solely to improve reliability, though the more consistent presentation of information in the ICD-11 CDDG as compared to the ICD-10 CDDG[13] likely helped in this regard. However, had the outcome of these changes been an overall reduction in reliability of the ICD-11 CDDG relative to the ICD-10 CDDG, this would have been cause for concern.

The reliability of ICD-11 CDDG generalized anxiety disorder, agoraphobia, social anxiety disorder, and adjustment disorder seems to have improved relative to the ICD-10 CDDG. This is reassuring, because the reliability of milder disorders compared to more severe disorders (e.g., schizophrenia and bipolar disorder) was lower in ICD-10 field trials[16,33]. Data from DSM-5 field trials suggest that disorders that are more broadly defined have higher reliability[25]. A number of hierarchical exclusion rules have been eliminated for anxiety and fear-related disorders in ICD-11 CDDG because they lacked specific empirical support[7]. Similarly, the subtypes of adjustment disorder have been eliminated from ICD-11 CDDG because they lacked evidence for validity or clinical utility[5].

The conceptualization of generalized anxiety disorder has been broadened in ICD-11 CDDG to include worry as an alternative essential feature to generalized apprehension and accompanying physiological symptoms[7], based in part on studies that show that worry is a central characteristic of the disorder[42]. Agoraphobia is reconceptualized to include a broader array of feared stimuli (fear of situations, fear of specific negative outcomes) and behaviours manifested in response to these stimuli (avoidance or entering the situations under specific conditions or enduring the situation with intense fear/anxiety), partly to allow for situations that may be more representative of those reported in low- and middle-income countries[43]. The ICD-11 conceptualization of social anxiety disorder has broadened the concept of ways in which the person could fear being negatively evaluated by others to include cultural variants of the disorder (i.e., fears of humiliation, embarrassment, rejection, or being offensive) as well as the range of behaviours in response to social stimuli[44,45]. It is possible that the greater attention to the cognitive and behavioural components of anxiety disorders and their contextual and cultural features in the ICD-11 CDDG as compared to the ICD-10 CDDG[7] helped to improve the reliability of these diagnoses.

Changes made in the diagnostic guidelines for adjustment disorder based on an earlier case-controlled study of disorders specifically associated with stress[9], particularly in providing additional guidance on differentiation from normal stress reactions, likely improved its diagnostic reliability in the current study.

Schizoaffective disorder is not a rare diagnosis in clinical populations, and its reliability is subject to ongoing discussion[46]. Jager et al[47] reviewed six studies and reported kappa scores between 0.08 and 0.63, concluding that only one study showed good agreement. In a meta-analysis of studies on sequential reliability (test-retest) of schizoaffective disorders, Santelmann et al[46] documented a mean difference of approximately 0.2 for kappa between schizoaffective disorder and other diagnoses such as schizophrenia, bipolar disorder and unipolar depression. The improved reliability of ICD-11 schizoaffective disorder in comparison to ICD-10 CDDG diagnosis may be related to the decision, in the proposed ICD-11 CDDG, to even more clearly apply the diagnostic requirements to the current episode rather than to the longitudinal presentation of the illness[3]. This is different from the longitudinal approach historically and currently taken by the DSM, on which most previous studies have been based[46,47]. Again, the purpose of the changes made for ICD-11 was to increase the clinical utility of the categories, and to the extent possible their validity, but it is reassuring that improved reliability appears to have been an outcome of these changes.

Some areas of the classification merit further consideration based on these results. The ICD-11 CDDG diagnoses of acute and transient psychotic disorder, panic disorder, and posttraumatic stress disorder seemed to have lower reliability than the equivalent categories in the ICD-10 CDDG, though it was not considered appropriate to analyze these differences statistically. However, these differences are modest in size (in all cases <0.2), and the reliability estimates for the ICD-11 CDDG in these categories are still in the moderate range.

However, unlike the categories discussed previously that were broadened in the ICD-11 CDDG, the description of each of these disorders has been narrowed in terms of their essential features. ICD-11 acute and transient psychotic disorder now exclusively comprises acute psychoses with "polymorphic" presentation[3], which is not strictly comparable to the broader concept tested in the ICD-10 field trial[16]. The reliability of acute and transient psychotic disorder with polymorphic presentation in ICD-10 Diagnostic Criteria for Research field trial[33] was similar to that in the present study. Nevertheless, based on these

results, the description of acute and transient psychotic disorder has been revised for the final version of the guidelines to define this aspect of the disorder more explicitly and to provide additional guidance on how to differentiate it from other conditions.

The proposed ICD-11 CDDG for panic disorder now require a clear discrimination between panic attacks of unexpected nature and panic attacks occurring in relation to symptoms of specific mental disorders (i.e., phobic disorders, some obsessive-compulsive disorders, and disorders specifically associated with stress). If panic attacks can be explained as due to symptoms of other specific mental disorders, a "with panic attacks" qualifier should be used rather than an additional co-occurring diagnosis of panic disorder. If some panic attacks over the course of the disorder have been unexpected and not exclusively in response to stimuli associated with the focus of apprehension related to the relevant disorder, a separate diagnosis of panic disorder should be assigned. In such cases, it is not necessary to apply the "with panic attacks" qualifier[7]. The lower kappa value for the ICD-11 CDDG as compared to the ICD-10 CDDG for panic disorder suggests that clinicians may have found it difficult to differentiate between expected and unexpected panic attacks or have been unclear about when to use the "with panic attacks" qualifier and when instead apply an additional diagnosis of panic disorder. This provides an example of an apparent trade-off between validity and reliability. Based on the results of this study, the final version of the ICD-11 CDDG will contain more detailed guidance on how to differentiate between unexpected and expected panic attacks and on how to decide whether applying the "with panic attacks" qualifier or a co-occurring panic disorder diagnosis. Increased emphasis on this issue in training programs as a part of ICD-11 implementation may also be helpful.

Though post-traumatic stress disorder is a well-recognized clinical entity, it has been criticized for the broad composition of its symptom clusters and high levels of co-occurrence with other disorders. Studies have also suggested that the threshold for an ICD-10 diagnosis of the disorder is relatively low[48,49]. The ICD-11 CDDG for post-traumatic stress disorder diagnosis are conceptually narrower than the ICD-10 ones, and now require the presence of re-experiencing of intrusive symptoms in the "here and now", as opposed to only experiencing intrusive memories of the traumatic event, as well as the presence of functional impairment[5]. This model has garnered increasing empirical support[50]. However, an earlier Internet-based study on disorders specifically associated with stress[9] showed that clinicians did not consistently apply the proposed ICD-11 guidelines regarding the required element of re-experiencing of the traumatic event(s). The subsequent version of the ICD-11 CDDG used in the present study provided additional clarification regarding re-experiencing in PTSD. However, the application of some of the changes introduced in the ICD-11 for post-traumatic stress disorder still appears to be difficult for practicing clinicians. Further exploration of the discrepancies between clinician raters at the level of specific symptoms may cast additional light on this issue. A specific focus on the new conceptualization of post-traumatic stress disorder as a part of ICD-11 training programs will also likely to be needed.

Some of the limitations of the ICD-11 EIFS need to be acknowledged. First, it bears repeating that the joint-rater (concurrent) method of testing reliability, which constrains the information provided to the two diagnosticians to be identical, usually generates higher kappa values compared to those obtained when separate interviews are conducted[26,51]. Second, the present study was conducted in multiple centers in diverse countries, including a very high proportion of low- and middle-income countries, but participating clinicians cannot be considered to be a globally representative sample of diagnosing mental health professionals. Participating institutions were typically high-status secondary or tertiary care centers, where the training of clinicians in diagnostic classification and interviewing is likely to meet the highest national standards. Clinician interviewers in the study would also have had some specific interest in diagnostic classification and in learning about the ICD-11. It can therefore be assumed that the reliabilities obtained in the study are higher than those that will be obtained in usual practice across all settings where the ICD-11 CDDG will be implemented. However, these problems are inherent in any field study, unless they can be overcome by a level of resources substantially in excess of those available for the EIFS.

Moreover, because the study sites were large academic settings that would tend to serve patients with moderate to severe mental health problems, the results may not be generalizable to patients with milder disorders seen in community settings. Mitigating this concern somewhat is the fact that ICD-11 CDDG include specific guidance on delineation of disorders from normal variation and have raised diagnostic thresholds for some of the conditions tested in EIFS (e.g., disorders specifically associated with stress).

Finally, the current study assessed only a relatively small proportion of the wide range of mental disorder diagnoses that may be applied to adult patients, focusing on those that are responsible for the highest level of disease burden and account for the greatest proportions of mental health services in participating centers. A much broader range of diagnostic categories is being addressed via Internet-based studies[9,12] and the overall consistency between the results of the two types of studies is reassuring in this regard.

CONCLUSIONS

As a developmental field study[36], the ICD-11 CDDG EIFS has been designed to provide information regarding the source of diagnostic disagreements through assessment of each element of the diagnostic guidelines for those disorders included in the protocols. This study has provided additional data for the WHO to use in improving the diagnostic guidelines prior to their publication. The WHO will also use the data in the development of training manuals and training courses for

clinicians in order to support member states in their implementation of ICD-11, with specific attention to the low- and middle-income countries in which the overwhelming majority of the world's population live.

The primary conclusion of this multi-country study is that the proposed ICD-11 CDDG can be interpreted in a consistent manner by diagnosing mental health professionals in a wide range of countries. The global applicability of the ICD-11 CDDG conceptualization of commonly diagnosed mental disorders is supported by the assessment of reliability of these guidelines in diverse settings (across 28 sites in 13 countries and in five languages) using a naturalistic field study design and a training approach that can easily be replicated for ICD-11 implementation. In the limited number of conditions that fell short, the findings will inform further revision prior to publication of the ICD-11.

The magnitude of this collaboration, the inclusion of clinicians in practice around the globe, the administration of the study in multiple languages, and the completion of this research in time to have the findings inform the final guidelines are major strengths of the ICD-11 research program. In addition to the specific value of this study in shaping the ICD-11, the EIFS and the WHO's Global Clinical Practice Network[52] for Internet-based ICD-11 field studies (http://gcp.network) have galvanized interest among clinicians around the world to participate in ongoing research that will continue to improve many dimensions of clinical understanding of mental illness and mental health service delivery.

ACKNOWLEDGEMENTS

The opinions contained in the paper are those of its authors and, except as specifically stated, are not intended to represent the official policies or positions of the World Health Organization. Funding was received for national activities related to this project in the following countries: Brazil - Conselho Nacional de Desenvolvimento Científico e Tecnológico; Canada - University Medical Research Fund, Royal's University of Ottawa Institute of Mental Health Research; Japan - Japanese Society of Psychiatry and Neurology, and Japan Agency for Medical Research and Development; Mexico - Consejo Nacional de Ciencia y Tecnología. Additional support for data collection in Brazil, Lebanon, Nigeria, South Africa and Tunisia was provided by the Columbia University Global Mental Health Program. Otherwise, this project was funded by in-kind contributions of the participating institutions. The authors express their gratitude to the following individuals who contributed substantially to the conduct of this research: Gustavo M. Barros, Ary Gadelha, Michel Haddad, Nuno H.P. Santos (Brazil); Huajian Ma, Zhen Wang, Jingjing Huang (China); Huma Kamal, Nidhi Malhotra (India); Gaia Sampogna, Lucia Del Gaudio, Giuseppe Piegari, Francesco Perris, Luca Steardo Jr (Italy); Tomofumi Miura, Itta Namamura, Kiyokazu Atake, Ayako Endo, Yuki Kako, Shinichi Kishi, Michihiko Koeda, Shinsuke Kondo, Akeo Kurumaji, Shusuke Numata, Naoya Oribe, Futoshi Suzuki, Masashi Yagi (Japan); Sariah Daouk, Chadia Haddad, François Kazour, Nicole Khauli (Lebanon); Francisco Juárez, Alejandra González, Omar Hernández, Carolina Muñoz (Mexico); Mayokun Odunleye (Nigeria); Tatiana Kiska, Oleg Limankin, Pavel Ponizovsky (Russian Federation); Roxanne James, Christine Lochner, Adele Pretorius (South Africa); Carolina Ávila, Cora Fernández, Julián Gómez, Ana Izquierdo, Beatriz Vicario, Rubén Vicente (Spain); Rahma Damak (Tunisia).

REFERENCES

1. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. World Psychiatry 2011;10:86-92.
2. Reed GM, Roberts MC, Keeley J et al. Mental health professionals' natural taxonomies of mental disorders: implications for the clinical utility of the ICD-11 and the DSM-5. J Clin Psychol 2013;69:1191-212.
3. Gaebel W. Status of psychotic disorders in ICD-11. Schizophr Bull 2012;38: 895-8.
4. Maj M, Reed GM. The ICD-11 classification of mood and anxiety disorders: background and options. World Psychiatry 2012;11(Suppl. 1).
5. Maercker A, Brewin CR, Bryant RA et al. Diagnosis and classification of disorders specifically associated with stress: proposals for ICD-11. World Psychiatry 2013;12:198-206.
6. Stein DJ, Kogan CS, Atmaca M et al. The classification of obsessive-compulsive and related disorders in the ICD-11. J Affect Disord 2016;190: 663-74.
7. Kogan CS, Stein DJ, Maj M et al. The classification of anxiety and fear-related disorders in the ICD-11. Depress Anxiety 2016;33:1141-54.
8. Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence and effect of personality disorder. Lancet 2015;385:717-26.
9. Keeley JW, Reed GM, Roberts MC et al. Disorders specifically associated with stress: a case-controlled field study for ICD-11 Mental and Behavioural Disorders. Int J Clin Health Psychol 2016;16:109-27.
10. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: World Health Organization, 1992.
11. Evans SC, Roberts MC, Keeley JW et al. Using vignette methodologies for study clinicians' decision-making: validity, utility, and application in ICD-11 field studies. Int J Clin Health Psychol 2015;15:160-70.
12. Keeley JW, Reed GM, Roberts MC et al. Developing a science of clinical utility in diagnostic classification systems: field study strategies for ICD-11 mental and behavioural disorders. Am Psychol 2016;71:3-16.
13. First MB, Reed GM, Hyman SE et al. The development of the ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. World Psychiatry 2015;14:82-90.
14. Reed GM. Toward ICD-11: improving the clinical utility of WHO's international classification of mental disorders. Prof Psychol Res Pr 2010;41:457-64.
15. Reed GM, First MB, Medina-Mora ME et al. Draft diagnostic guidelines for ICD-11 mental and behavioural disorders available for review and comment. World Psychiatry 2016;15:112-3.
16. Sartorius N, Kaelber CT, Cooper JE et al. Progress toward achieving a common language in psychiatry. Results from the field trial of the clinical guidelines accompanying the WHO classification of mental and behavioral disorders in ICD-10. Arch Gen Psychiatry 1993;50:115-24.
17. Mullins-Sweatt SN, Widiger TA. Clinical utility and DSM-5. Psychol Assess 2009;21:302-12.
18. Blashfield RK, Keeley JW, Flanagan EH et al. The cycle of classification: DSM-I through DSM-5. Annu Rev Clin Psychol 2014;10:25-51.
19. Stengel E. Classification of mental disorders. Bull World Health Organ 1959;21:601-63
20. Spitzer R, Sheehy M, Endicott J. DSM-III: Guiding principles. In: Rakoff V, Stancer H, Kedward H (eds). Psychiatric diagnosis. New York: Brunner/Mazel, 1977:1-24.
21. World Health Organization. Glossary of mental disorders and guide to their classification. Geneva: World Health Organization, 1974.
22. Beck AT, Ward CH, Mendelson M et al. Reliability of psychiatric diagnosis 2: a study of consistency of clinical judgments and ratings. Am J Psychiatry 1962;119:351-7.
23. Ward CH, Beck AT, Mendelson M et al. The psychiatric nomenclature: reasons for diagnostic disagreement. Arch Gen Psychiatry 1962;7:198-205.
24. Harvey PD, Heaton RK, Carpenter WT Jr et al. Diagnosis of schizophrenia: consistency across information sources and stability of the condition. Schizophr Res 2012;140:9-14.
25. Regier DA, Narrow WE, Clarke DE et al. DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnosis. Am J Psychiatry 2013;170:59-70.
26. Chmielewski M, Clark LA, Bagby RM et al. Method matters: understanding diagnostic reliability in DSM-IV and DSM-5. J Abnorm Psychol 2015;124: 764-9.
27. Lahey BB, Applegate B, Barkley RA et al. DSM-IV field trials for oppositional defiant disorder and conduct disorder in children and adolescents. Am J Psychiatry 1994;151:1163-71.

28. Rounsaville BJ, Kosten TR, Williams JB et al. A field trial of DSM-III-R psychoactive substance dependence disorders. Am J Psychiatry 1987;144:351-5.
29. Brittain PJ, Stahl D, Rucker J et al. A review of the reliability and validity of OPCRIT in relation to its use for the routine clinical assessment of mental health patients. Int J Methods Psychiatr Res 2013;22:110-37.
30. Aboraya A, Rankin E, France C et al. The reliability of psychiatric diagnosis revisited: the clinician's guide to improve the reliability of psychiatric diagnosis. Psychiatry 2006;3:41-50.
31. First MB. Counterpoint – there isn't enough evidence available to speculate on the reliability of diagnoses in clinical settings. Psychiatry 2007;4: 24-5.
32. Spitzer R, Forman J, Nee J. DSM-III field trials, I: initial interrater diagnostic reliability. Am J Psychiatry 1979;136:815-7.
33. Sartorius N, Ustün TB, Korten A et al. Progress toward achieving a common language in psychiatry, II: Results from the international field trials of the ICD-10 diagnostic criteria for research for mental and behavioral disorders. Am J Psychiatry 1995;152:1427-37.
34. Clarke DE, Narrow WE, Regier DA et al. DSM-5 field trials in the United States and Canada, Part I: study design, sampling strategy, implementation, and analytic approaches. Am J Psychiatry 2013;170:43-58.
35. Feinn R, Gelernter J, Cubells JF et al. Sources of unreliability in the diagnosis of substance dependence. J Stud Alcohol Drugs 2009;70:475-81.
36. First MB. The importance of developmental field trials in the revision of psychiatric classifications. Lancet Psychiatry 2016;3:579-84.
37. Whiteford HA, Degenhardt L, Rehm J et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 2013;382:1575-86.
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
39. Andrews G, Peters L, Guzman A-M et al. A comparison of two structured diagnostic interviews: CIDI and SCAN. Aust N Z J Psychiatry 1995;29:124-32.
40. Lobbestael J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). Clin Psychol Psychother 2011;18:75-9.
41. Bruehl S, Ohrbach R, Sharma S et al. Approaches to demonstrating the reliability and validity of core diagnostic criteria for chronic pain. J Pain 2016;17(Suppl. 9):T118-31.
42. Olatunji BO, Broman-Fulks JJ, Bergman SM et al. A taxometric investigation of the latent structure of worry: dimensionality and associations with depression, anxiety, and stress. Behav Ther 2010;41:212-28.
43. Stein DJ. Agoraphobia and panic disorder: options for ICD-11. World Psychiatry 2012;11(Suppl. 1):89-93.
44. Emmelkamp PMG. Specific and social phobias in ICD-11. World Psychiatry 2012;11(Suppl. 1):94-9.
45. Kinoshita Y, Chen J, Rapee RM et al. Cross-cultural study of conviction subtype Taijin Kyofu: proposal and reliability of Nagoya-Osaka diagnostic criteria for social anxiety disorder. J Nerv Ment Dis 2008;196:307-13.
46. Santelmann H, Franklin J, Bußhoff J, et al. Test-retest reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression – a systematic review and metaanalysis. Bipolar Disord 2015;17:753-68.
47. Jager M, Haack S, Becker T et al. Schizoaffective disorder – an ongoing challenge for psychiatric nosology. Eur Psychiatry 2011;26:159-65.
48. Brewin CR, Fuchkan N, Huntley Z et al. Outreach and screening following the 2005 London bombings: usage and outcomes. Psychol Med 2010;40: 2049-57.
49. Peters L, Slade T, Andrews G. A comparison of ICD-10 and DSM-IV criteria for posttraumatic stress disorder. J Trauma Stress 1999;12:335-43.
50. Brewin CR, Cloitre M, Hyland P et al. A review of current evidence regarding the ICD-11 proposals for diagnosing PTSD and complex PTSD. Clin Psychol Rev 2017;58:1-15.
51. Kraemer HC. The reliability of clinical diagnoses: state of the art. Annu Rev Clin Psychol 2014;10:111-30.
52. Reed GM, Rebello TJ, Pike KM et al. WHO's Global Clinical Practice Network for mental health. Lancet Psychiatry 2015;2:379-80.

DOI:10.1002/wps.20524

Do mental health professionals use diagnostic classifications the way we think they do? A global survey

Michael B. First[1], Tahilia J. Rebello[2], Jared W. Keeley[3], Rachna Bhargava[4], Yunfei Dai[5], Maya Kulygina[6], Chihiro Matsumoto[7], Rebeca Robles[8], Anne-Claire Stona[9], Geoffrey M. Reed[2,10]

[1]Department of Psychiatry, Columbia University, College of Physicians and Surgeons and New York State Psychiatric Institute, New York, NY, USA; [2]Global Mental Health Program, Columbia University, College of Physicians and Surgeons and Research Foundation for Mental Hygiene, New York, NY, USA; [3]Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA; [4]All India Institute of Medical Sciences, New Delhi, India; [5]Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China; [6]Moscow Research Institute of Psychiatry, National Medical Research Centre for Psychiatry and Narcology, Moscow, Russian Federation; [7]Japanese Society of Psychiatry and Neurology, Tokyo, Japan; [8]National Institute of Psychiatry Ramón de la Fuente Muñiz, Mexico City, Mexico; [9]Ministry for Solidarity and Health, Paris, France; [10]World Health Organization, Geneva, Switzerland

We report on a global survey of diagnosing mental health professionals, primarily psychiatrists, conducted as a part of the development of the ICD-11 mental and behavioural disorders classification. The survey assessed these professionals' use of various components of the ICD-10 and the DSM, their attitudes concerning the utility of these systems, and usage of "residual" (i.e., "other" or "unspecified") categories. In previous surveys, most mental health professionals reported they often use a formal classification system in everyday clinical work, but very little is known about precisely how they are using those systems. For example, it has been suggested that most clinicians employ only the diagnostic labels or codes from the ICD-10 in order to meet administrative requirements. The present survey was conducted with clinicians who were members of the Global Clinical Practice Network (GCPN), established by the World Health Organization as a tool for global participation in ICD-11 field studies. A total of 1,764 GCPN members from 92 countries completed the survey, with 1,335 answering the questions with reference to the ICD-10 and 429 to the DSM (DSM-IV, DSM-IV-TR or DSM-5). The most frequent reported use of the classification systems was for administrative or billing purposes, with 68.1% reporting often or routinely using them for that purpose. A bit more than half (57.4%) of respondents reported often or routinely going through diagnostic guidelines or criteria systematically to determine whether they apply to individual patients. Although ICD-10 users were more likely than DSM-5 users to utilize the classification for administrative purposes, other differences were either slight or not significant. Both classifications were rated to be most useful for assigning a diagnosis, communicating with other health care professionals and teaching, and least useful for treatment selection and determining prognosis. ICD-10 was rated more useful than DSM-5 for administrative purposes. A majority of clinicians reported using "residual" categories at least sometimes, with around 12% of ICD-10 users and 19% of DSM users employing them often or routinely, most commonly for clinical presentations that do not conform to a specific diagnostic category or when there is insufficient information to make a more specific diagnosis. These results provide the most comprehensive available information about the use of diagnostic classifications of mental disorders in ordinary clinical practice.

Key words: Classifications of mental disorders, ICD-11, ICD-10, DSM-5, Global Clinical Practice Network, ordinary clinical practice, psychiatric diagnosis, use for administrative purposes, clinical utility, residual diagnostic categories

(World Psychiatry 2018;17:187–195)

For the past ten years, the World Health Organization (WHO) has been revising the Mental and Behavioural Disorders chapter as part of the development of the 11th edition of the International Classification of Diseases and Related Health Problems (ICD-11). A major focus of the proposed changes for ICD-11 has been to improve the clinical utility of the classification for use by frontline mental health professionals, including psychiatrists[1].

A first step in efforts to improve the clinical utility of a classification is to collect baseline information about how the classification is currently being used[2,3]. Two surveys were conducted by the WHO at the outset of the ICD-11 revision process in order to determine psychiatrists'[4] and psychologists'[5] attitudes towards, and usage of, mental health classifications. Survey questions were primarily directed towards assessing respondents' views about the classification of mental disorders, covering topics such as their opinions about the main purposes of a classification system, the ideal number of diagnostic categories, the desired level of flexibility in application of the criteria, the best way to address concepts of severity and functional status, whether disorders should be rated dimensionally or categorically, and whether the current classification system is difficult to apply cross-culturally. The minority of questions focusing on classification usage included how often a formal classification was used in day-to-day clinical work, which classification was used most, and which diagnostic categories were most used in daily clinical practice. For those diagnoses used at least once per week, respondents provided ratings of their ease of use and goodness of fit.

In those surveys, 79% of participating psychiatrists and 60% of participating psychologists reported that they "often" or "almost always" use a formal classification system as part of their everyday clinical work, with an additional 14% of psychiatrists and 18% of psychologists indicating that they "sometimes" use one. However, these results do not tell us precisely *how* clinicians use formal classification systems in their practices. For example, such use might involve employing only the diagnostic labels or the diagnostic codes, using diagnostic prototypes embodied in the classification's definitions, or applying diagnostic guidelines or criteria[6]. In fact, conventional wisdom about psychiatrists' use of classification systems suggests that ICD is often used only as a coding system to meet

administrative requirements, so that the impact of revisions of that system may be lower than usually realized.

In fact, very little is known about the global implementation of either the ICD or DSM classifications by psychiatrists and other mental health professionals in their clinical practices. Most of the limited information we have about clinicians' reported clinical usage of psychiatric classifications comes from surveys[7-11]. Those surveys, however, have focused almost entirely on respondents' attitudes and preferences about classification systems, rather than on collecting information about their usage. For example, a survey developed by Mellsop et al in 2006[12], administered to psychiatrists in seven different countries[13,14] included only one general question about usage, asking respondents how often (i.e., routinely, sometimes, or never) they used each of the five DSM-IV axes, ICD-10, and the International Classification of Functioning (ICF) in their clinical practice. An exception was a 1991 survey of American child psychiatrists attending a national meeting, which included several specific questions regarding how respondents used the DSM-III-R criteria for childhood disorders[15]. The survey found that, depending on the diagnosis, 47 to 66% of the respondents reported that they generally assessed all applicable DSM-III-R criteria when making a diagnosis and that 28 to 49% often referred to the manual before assigning a diagnosis.

The present paper reports on a detailed survey of global mental health professionals' actual usage of the ICD-10 and the most recent two editions of the DSM. Its purpose was to shed light on clinicians' use of the various components of ICD and DSM (i.e., diagnostic codes, diagnostic guidelines/criteria, descriptive text) as a part of their routine clinical practice, as well as to compare clinicians' patterns of use of these two classifications. The survey also queried clinicians about their attitudes concerning the utility of the ICD and DSM for various purposes (e.g., communication, treatment selection), as well as collecting information about their usage of the "residual" categories (i.e., other specified, unspecified, not otherwise specified).

Both the ICD and DSM include such "residual categories" for use in situations where the patients' clinical presentation does not meet the definitional requirements for any specific disorder or when there is insufficient information available for the clinician to make a specific diagnosis (for example, in an emergency department setting). However, it has been suggested that the relatively high rates of the use of these categories may in fact be an indirect clue that clinicians find the ICD/DSM categories difficult to use or not accurately descriptive of their patients[16], or that providing specific diagnostic information in a patient's medical record may be harmful to the patient (e.g., stigmatizing).

The survey was conducted with clinicians registered as a part of the Global Clinical Practice Network (GCPN)[17,18], which was established by the WHO Department of Mental Health and Substance Abuse for the purpose of direct participation by clinicians around the world in field studies related to the development of the ICD-11 chapter on Mental and Behavioural Disorders. GCPN members are mental health professionals who have completed their training and are qualified to practice in their country of residence (https://gcp.network). The GCPN consists now of more than 14,000 mental health professionals from 154 countries, more than half of whom are psychiatrists. Although the network was initially established for the purpose of conducting ICD-11 field studies via the Internet[19], it also provides an opportunity to survey clinicians from a range of professional backgrounds and from all over the world on other related topics. The present survey of classification usage is the first of the network to study a topic other than the new ICD-11 diagnostic guidelines.

METHODS

Participants

Participants were recruited from the GCPN. At the time of study sample selection, there were 11,707 registered GCPN members from 139 countries, across nine registration languages. It was determined that the study would be conducted in six languages: Chinese, English, French, Japanese, Spanish and Russian. Language selection was based on an adequate number of GCPN members who are proficient in that language and the availability of appropriate translation resources.

The Internet-based survey used in the study was programmed in the six languages using the Qualtrics survey software. The survey contained questions related to classification use followed by a separate module on technology use. The original survey was developed in English, assessed for global applicability and relevance (e.g., examples used in questions), and then translated into the other five languages using a rigorous process including validation by bilingual content experts[19].

Survey invitations were sent directly via Qualtrics to 9,792 registered GCPN members who, based on information provided at the time of registration, were proficient in one of the study languages and were actively providing clinical services or direct clinical supervision. Reminder e-mails were sent two and four weeks after the first invitation to all those who had not yet completed the survey. Data collection in each language was closed two months after the initial invitation.

Of the 9,792 GCPN members invited to participate, 2,960 (30.2%) clicked on the embedded link in the survey invitation and explicitly agreed to participate in the study by confirming consent in the first survey question. This participation rate is comparable to the diagnostic field studies conducted using the GCPN.

At the beginning of the survey, consenting participants were asked to state whether they: a) were currently providing direct mental health services to patients for at least one hour per week, b) were usually responsible for assigning a mental disorder diagnosis to patients, and c) had often or routinely used the ICD-10 or a version of the DSM (DSM-5, DSM-IV or

DSM-IV-TR) during the past year, which was determined based on their response to a four-point scale (never/rarely, sometimes, often, routinely). Individuals who did not meet the above criteria were not asked the remaining questions about classification use, but instead taken directly to the technology module.

Participants who indicated that they had often or routinely used only one of the target classification systems during the past year were asked a series of detailed follow-up questions for that particular classification system. Participants who indicated that they had often or routinely used both the ICD-10 and some edition of the DSM during the last year ("mixed users") were asked to indicate those purposes for which they used the ICD and DSM classifications often or routinely. If they indicated that they used either the ICD or DSM for clinical purposes, they were then asked to answer detailed questions regarding that system. Participants who indicated that they used both the ICD-10 and some edition of the DSM for clinical purposes were assigned to answer detailed questions about ICD-10. Participants who indicated that they used the DSM for clinical purposes were instructed to answer detailed questions on the version of the DSM they were using, with preference given to the DSM-5 if they were using more than one version. Meaningful differences in patterns of use were not found between users of the DSM-5, DSM-IV or DSM-IV-TR, so DSM users were combined in the analyses.

Survey content

Once assigned to either the ICD-10 or a DSM use module, participants were asked to indicate which version(s) they had used over the past year in either printed or electronic format. A unique feature of the ICD-10 is the existence of different versions[20], including the Clinical Descriptions and Diagnostic Guidelines[21], intended for use by mental health professionals in clinical practice, the Diagnostic Criteria for Research[22], and the statistical version of the classification, used for the collection and reporting of health information by WHO member states, which contains only brief definitions of mental and behavioural disorders[23]. ICD-10 users were asked to indicate whether they had used each of these three versions within the past year and, if they had not used a particular version, whether they had ever seen it. Participants were presented with samples of each system so that they would report as accurately as possible regarding their use of specific versions. Both ICD-10 and DSM users were asked questions about their use of printed and electronic formats of the relevant system.

Participants were then asked to provide detailed information about how frequently they used the assigned diagnostic system in specific ways (e.g., systematically going through the diagnostic guidelines or criteria, reviewing other parts of available text in addition to diagnostic guidelines or criteria) and for specific purposes (e.g., administrative and billing uses, communicating with patients and family members). Frequency of usage was assessed using a four-point scale (never/rarely, sometimes, often, routinely) for the initial diagnostic phase and during the ongoing treatment of patients. Participants were also asked to rate the utility of the relevant diagnostic system for specific purposes (e.g., selecting a treatment, assessing probable prognosis) using the following four-point scale: not at all/slightly useful, moderately useful, very useful, extremely useful.

Finally, participants were asked how often they used "other specified" or "unspecified" categories in the ICD-10 or "not otherwise specified" categories in the DSM, and offered a range of reasons for using these categories. In order to determine whether the respondents were using these categories appropriately, the range of reasons included some that would be considered legitimate (e.g., for presentations that do not conform to any specified category) and others that would be more questionable (e.g., to prevent more specific diagnostic information from being entered into the patient's record).

RESULTS

Sample demographics

After excluding participants who did not meet the eligibility requirements for the survey and 13 individuals who agreed to participate and met the eligibility requirements but did not provide sufficient data for analysis, the final sample for the study consisted of 1,764 GCPN members. As shown in Table 1, these included 1,335 participants who were assigned the ICD-10 version of the survey (75.7% of the sample) and 429 participants who were assigned the DSM version of the survey (24.3%).

Demographic characteristics of the sample are shown in Table 1. Nearly two-thirds of the sample (63.4%) were male and three-quarters (74.6%) were physicians, approximately 90% of whom were psychiatrists. Participants were from 92 countries, and 39.5% were practicing in low- or middle-income countries. All global regions were represented; while the representation of regions appears imbalanced, it closely resembles the representation of mental health professionals across the world[24]. A substantial majority (64.1%) completed the survey in a language other than English. The average age of participants was 46.2 ± 11.3 years, with a mean of 16.1 ± 10.5 years of experience following completion of their training.

Classification usage

When asked about which version of the classification they were using during the past year, almost three-quarters of the ICD-10 users (73.8%) reported that they had used the Clinical Descriptions and Diagnostic Guidelines, while 26.4% had used the Diagnostic Criteria for Research, and 32.0% had used the statistical version (percentages are non-exclusive). Only 7.6% of ICD-10 users indicated that they had never seen the Clinical Descriptions and Diagnostic Guidelines. A majority of the DSM users (86.9%) reported using the full version (i.e., diag-

Table 1 Demographic characteristics of the sample

	Completed ICD-10 version of survey (N=1,335)	Completed DSM version of survey (N=429)	Total (N=1,764)
Age at time of network registration, years (mean±SD)	45.4 ± 10.7	48.6 ± 12.6	46.2 ± 11.3
Years of experience after training completion (mean±SD)	15.6 ± 10.1	17.5± 11.7	16.1 ± 10.5
Gender, N (%)			
Male	877 (65.7)	242 (56.4)	1,119 (63.4)
Female	457 (34.2)	187 (43.6)	644 (36.5)
Other or not available	1 (<0.1)	0	1 (<0.1)
Professional discipline, N (%)			
Medicine	1,102 (82.5)	214 (49.9)	1,316 (74.6)
Psychology	198 (14.8)	144 (33.6)	342 (19.4)
Nursing	2 (0.1)	1 (0.2)	3 (0.2)
Social work	4 (0.3)	27 (6.3)	31 (1.8)
Counseling	13 (1.0)	23 (5.4)	36 (2.0)
Sex therapy	0	4 (0.9)	4 (0.2)
Speech therapy	1 (<0.1)	0	1 (<0.1)
Occupational therapy	14 (1.0)	0	1 (<0.1)
Other		16 (3.7)	30 (1.7)
Country income level, N (%)			
High	781 (58.5)	284 (66.2)	1,065 (60.4)
Upper-middle	404 (30.3)	106 (24.7)	510 (28.9)
Lower-middle	136 (10.2)	27 (6.3)	163 (9.2)
Low	14 (1.0)	11 (2.6)	25 (1.4)
Language of participation, N (%)			
Chinese	254 (19.0)	10 (2.3)	264 (15.0)
English	429 (32.1)	204 (47.6)	633 (35.9)
French	144 (10.8)	59 (13.8)	203 (11.5)
Japanese	137 (10.3)	63 (14.7)	200 (11.3)
Russian	229 (17.2)	0	229 (13.0)
Spanish	142 (10.6)	93 (21.7)	235 (13.3)
WHO global region, N (%)			
Africa	27 (2.0)	15 (3.5)	42 (2.4)
Americas - North	11 (0.8)	131 (30.5)	142 (8.0)
Americas - South	128 (9.6)	78 (18.2)	206 (11.7)
Eastern Mediterranean	24 (1.8)	31 (7.2)	55 (3.1)
Europe	644 (48.2)	71 (16.6)	715 (40.5)
South-East Asia	92 (6.9)	15 (3.5)	107 (6.1)
Western Pacific - Asia	395 (29.6)	74 (17.2)	469 (26.6)
Western Pacific - Oceania	14 (1.0)	14 (3.3)	28 (1.6)

nostic criteria plus descriptive text), 47.3% used the version containing only the diagnostic criteria and 11.7% used a listing of DSM disorders and codes without diagnostic criteria (percentages are non-exclusive).

Despite the availability of electronic sources of the diagnostic codes and diagnostic guidelines or criteria, respondents primarily relied on printed versions as their sources. With respect to obtaining diagnostic codes, while 93.3% of ICD-10 users and 84.6% of DSM users reported obtaining them from printed versions, only 44.1% of ICD-10 users and 30.5% of DSM-5 users obtained them from electronic sources (e.g., WHO or American Psychiatric Association websites, drop-down

Table 2 Overall classification usage pattern (N=1,764)

	Never/Rarely	Sometimes	Often	Routinely
Frequency of use of diagnostic codes for administrative/billing purposes (%)				
Initial diagnosis	18.4	13.5	19.2	48.9
Ongoing treatment	18.4	18.5	22.6	40.5
Frequency of systematically going through diagnostic guidelines/criteria to determine whether they apply to individual cases (%)				
Initial diagnosis	5.2	37.4	33.9	23.5
Ongoing treatment	8.8	43.2	31.3	16.7
Frequency of making diagnosis without referring to diagnostic guidelines/criteria (%)				
Initial diagnosis	18.2	32.0	36.8	13.0
Ongoing treatment	17.2	32.8	35.9	14.1
Frequency of referring to relevant additional text sections outside diagnostic guidelines/criteria (%)				
Initial diagnosis	16.8	46.9	26.1	10.1
Ongoing treatment	21.4	50.5	21.3	6.9
Frequency of using the diagnostic system to communicate or share information with patient and/or family (%)				
Initial diagnosis	25.9	39.2	21.1	13.8
Ongoing treatment	26.6	41.0	20.9	11.5

menus in electronic health records or other software). The breakdown of sources for obtaining diagnostic guidelines or criteria was similarity tilted towards printed versions, with 92.5% of ICD-10 users and 84.8% of DSM users obtaining them from hardcopy sources and around 35% of both ICD-10 and DSM-5 users obtaining them electronically (percentages are non-exclusive).

The usage pattern for various components of the diagnostic classifications (i.e., diagnostic codes, diagnostic guidelines/criteria, descriptive text) is presented in Table 2. Clinicians reported using diagnostic classifications most often for administrative or billing purposes, with 68.1% reporting that they used them often or routinely for the initial evaluation.

With respect to diagnostic practices, the survey asked respondents to indicate how often they systematically go through diagnostic guidelines or criteria to determine whether they apply to individual patients, as well as how often they make a diagnosis without referring to guidelines or criteria. These were not presented as mutually exclusive questions. A bit more than half (57.4%) of the respondents reported going through the diagnostic guidelines or criteria often or routinely during the initial assessment of individual patients, dropping to 48.0% during ongoing treatment. Approximately half of the clinicians reported often or routinely making a diagnosis without referring to the diagnostic guidelines or criteria, which was essentially the same during the initial diagnostic assessment and during ongoing treatment (49.8% and 50.0%, respectively). Usage of the additional text sections was much less common, with only 36.2% reporting that they referred to the text often or routinely during the initial evaluation, and only 28.2% during ongoing treatment.

Usage of the classification system for the purpose of communicating or sharing information with the patient or family was not frequent, with 34.9% using it often or routinely for that purpose during the initial evaluation and 32.4% during ongoing treatment.

In order to facilitate the comparison of usage patterns and utility ratings among ICD-10 and DSM users, Likert scale frequency tables were converted to standard weighted frequencies. This was done by assigning a value of 1 to never/rarely, 2 to sometimes, 3 to often, and 4 to routinely, and multiplying the frequency of each response option by its point value. The resulting scores were then transformed into a standard weighted frequency by summing all the values for a question, subtracting that value from the minimum possible sum, and dividing the total by the range of possible scores. Using this method, the resulting values range from 0 to 1, are roughly on the same scale, and the magnitude of each individual response is taken into account.

Comparative usage patterns for ICD-10 and DSM users during initial diagnosis and ongoing treatment are shown in Table 3. ICD-10 users were more likely than DSM users to use it for administrative and billing purposes, especially during initial diagnosis. DSM users were less likely to indicate that they make diagnoses without referring to the diagnostic guidelines or criteria and more likely to indicate that they go through the diagnostic guidelines or criteria systematically to determine whether they apply to individual cases, but these differences were small in absolute terms. For both systems, participants indicated that they were more likely to go through the guidelines or criteria and to refer to additional text sections during the initial diagnostic assessment than during ongoing treatment.

Table 3 Comparison of usage patterns for ICD-10 and DSM users (standard weighted frequencies)

	ICD-10 (N=1,335)	DSM (N=429)	χ^2 (df=3)
Frequency of use of diagnostic codes for administrative/billing purposes			
Initial diagnosis	.7021	.5369	58.83***
Ongoing treatment	.6557	.4965	57.41***
χ^2 (df=3)	28.66***	0.14	
Frequency of systematically going through diagnostic guidelines/criteria to determine whether they apply to individual cases			
Initial diagnosis	.5643	.6511	32.79***
Ongoing treatment	.5101	.5478	7.18
χ^2 (df=3)	24.28***	27.46***	
Frequency of making diagnosis without referring to diagnostic guidelines/ criteria			
Initial diagnosis	.4961	.4390	11.68**
Ongoing treatment	.4979	.4639	4.33
χ^2 (df=3)	1.61	1.99	
Frequency of referring to relevant additional text sections outside diagnostic guidelines/criteria			
Initial diagnosis	.4215	.4646	8.63*
Ongoing treatment	.3810	.3722	1.98
χ^2 (df=3)	14.55**	23.77***	
Frequency of using the diagnostic system to communicate or share information with patient and/or family			
Initial diagnosis	.4010	.4343	4.93
Ongoing treatment	.3868	.4017	2.50
χ^2 (df=3)	3.49	2.15	

*p < 0.05, **p <0 .01, ***p <0.001

Utility of the classifications

Participants' ratings of the utility of the ICD-10 and DSM during the past year for a variety of different purposes are shown in Table 4. Both systems received the highest ratings of utility for meeting administrative requirements, assigning a diagnosis, communicating with other health care professionals, and teaching trainees or students, and the lowest ratings for selecting a treatment and assessing probable prognosis.

ICD-10 users rated that system as more useful for meeting administrative requirements as compared to ratings of the DSM by DSM users, and the DSM was judged by its users as slightly more useful for educating the patient and/or family about the diagnosis, although this latter difference was small in absolute terms. Otherwise, utility ratings by ICD-10 and DSM users were similar.

Results concerning "mixed users", i.e., those who reported that they often or routinely used both the ICD-10 and some editions of the DSM, are shown in Table 5. In our experience, there is widespread confusion among US and Canadian professionals about whether they are using the ICD-10 or the DSM for making diagnoses or for administrative purposes, due to the existence of US and Canadian clinical ICD modifications. For this reason, 55 survey participants from the US and Canada were not included in the analysis.

Table 4 Utility of ICD-10 and DSM in the past year (standard weighted frequencies)

	ICD-10 (N=1,335)	DSM (N=429)	χ^2 (df=3)
Meeting administrative requirements	.7486	.5066	236.71***
Assigning a diagnosis	.6777	.6589	4.85
Selecting a treatment	.3658	.3388	4.99
Educating patient and/or family about diagnosis	.3910	.4406	11.32*
Assessing probable prognosis	.3870	.3916	0.95
Communicating with other health care professionals	.6449	.6426	0.77
Teaching trainees or students	.6275	.6535	3.52

*p < 0.05, ***p < 0.001

Mixed users were substantially more likely to report using the ICD-10 (70.7%) than the DSM (21.0%) for fulfilling administrative requirements. However, they were equally likely to report using the ICD-10 and the DSM for assigning diagnoses in clinical practice. Mixed users more frequently use the DSM for research and education.

Usage of "residual" categories

A total of 67.5% of ICD-10 users and 72.7% of DSM users indicated that they at least sometimes employed "residual" categories, with 11.6% of ICD-10 users and 19.3% of DSM users reporting that they employed these categories often or routinely. The reasons that participants endorsed for using these categories, expressed as standard weighted frequencies to facilitate comparisons, are shown in Table 6.

The most commonly endorsed reasons for both ICD-10 and DSM users were clinical presentations that do not conform to any specific diagnostic category and insufficient information to make a more specific diagnosis. There were no significant

Table 5 Usage of ICD-10 and DSM by "mixed users" (N=605)

	ICD-10	DSM	χ^2 (df=1)
Fulfilling administrative requirements, N (%)	428 (70.7)	127(21.0)	115.27***
Assigning diagnoses in clinical practice, N (%)	445 (73.6)	458 (75.7)	0.11
Research, N (%)	246 (40.7)	475 (78.5)	46.23***
Education, N (%)	347 (57.4)	498 (82.3)	15.97***

***p < 0.001

Table 6 ICD-10 and DSM users' reasons for employing "residual" categories (standard weighted frequencies)

	ICD-10 (N=916)	DSM (N=317)	χ^2 (df=3)
Because the patient's presentation does not conform to any of the specific categories	.6001	.6090	1.61
Because there is insufficient information to make a more specific diagnosis	.4554	.4733	4.56
To indicate that it cannot be determined whether the symptoms are due to a primary condition or are secondary	.2882	.2863	0.11
Because the patient meets the requirements for more than one diagnosis in a grouping	.2926	.2265	17.64***
To prevent more specific diagnostic information from being entered into the patient's record	.1850	.1795	0.66
Because making a more specific diagnosis is not useful for patient care	.1709	.1934	2.75

***$p < 0.001$

differences between the responses of ICD-10 and DSM users except that the former were somewhat more likely to indicate that they employ these categories when the patient meets the diagnostic requirements for multiple categories in a grouping.

DISCUSSION

Clinicians make mental and behavioural disorder diagnoses in everyday clinical practice for a variety of reasons: a) a diagnosis is generally required in order to meet administrative requirements; b) diagnostic labels provide a convenient shorthand for communicating the patient's clinical presentation to other clinicians; c) a diagnosis is often important for determining the patient's prognosis and selection of treatment; and d) the diagnosis can facilitate the education of the patient and family about the illness. Diagnostic classification systems provide clinicians with tools intended to meet these needs: diagnostic codes for meeting administrative requirements, diagnostic guidelines or criteria to facilitate accurate and reliable diagnoses, and accompanying text to facilitate differential diagnosis and the appreciation of the role of developmental and culture-related features in the clinical presentation. However, the extent to which clinicians make use of these elements of diagnostic systems in clinical practice is unknown[3].

Several aspects of the results of this survey of GCPN users confirm conventional wisdom about patterns of classification usage. In particular, the most frequently reported use of a classification system is to obtain diagnostic codes for administrative or billing purposes. This almost certainly reflects the fact that the provision of a diagnostic code is a requirement for clinical encounters in most countries. Nonetheless, 18.4% of respondents reported that they rarely or never use a classification for that purpose, which likely reflects the fact that in some practice settings the responsibility for looking up the appropriate diagnostic code is not the clinician's but instead falls on non-clinical personnel (e.g., medical billing and coding specialists).

A majority of GCPN clinicians (57.4%) reported that they often or routinely go through diagnostic guidelines or criteria systematically during the process of making an initial diagnosis, which is at variance with the widespread belief that clinicians only use the classification, in particular the ICD-10, for the purpose of obtaining diagnostic codes. Only 5.2% of GCPN clinicians reported that they never or rarely go through the diagnostic guidelines or criteria systematically during the initial diagnostic process. The practice of making a diagnosis without referring to the diagnostic guidelines or criteria was a bit less common, with just less than half of GCPN clinicians reporting often or routinely doing this during the initial evaluation. The use of the classification for ancillary purposes was less frequent, with only 34.9% reporting often or routinely using it to communicate or share diagnostic information with patients and their families.

A comparison of usage patterns between ICD and DSM users reveals that the ICD classification is used much more commonly among this sample for administrative and billing purposes as compared to the DSM classification. This is unsurprising, because the ICD is required for administrative use in most countries in which documentation of diagnoses for clinical encounters is needed. The only other significant difference between ICD and DSM users is the pattern of usage of the diagnostic guidelines or criteria, with DSM users being more likely to go through the diagnostic criteria to determine whether they apply as compared to the ICD users, who were correspondingly more likely to make psychiatric diagnoses without referring to the diagnostic guidelines, although these differences were small in magnitude. This may reflect a difference in the perceived utility of the ICD diagnostic guidelines vs. the DSM criteria, but it could also reflect the greater complexity of the DSM criteria, which makes them more difficult to recall as compared to the ICD guidelines. Slightly greater usage of the DSM additional text as compared to the ICD-10 text likely reflects the much more extensive text sections in the DSM. In recognition of the unevenness of the ICD-10 text, the newly developed ICD-11 text is more extensive and follows a uniform template from disorder to disorder[20].

Clinicians' ratings of the utility of the ICD and DSM classifications for various purposes were highest for applications such as meeting administrative requirements, assigning a diagnosis, communicating with other health care professionals, and teaching trainees or students, and lowest for selecting a treatment and assessing probable prognosis. This result likely re-

flects long-identified weaknesses of descriptive categorical classification systems[25,26], namely the diagnostic heterogeneity of the categories and the lack of a one-to-one relationship between diagnostic categories and treatment options. Several previous surveys of clinicians' attitudes towards mental health classification (including the WPA-WHO study undertaken early in the development of the ICD-11[4]) included a question which asked respondents to indicate the single most important purpose of a diagnostic classification. In each of these surveys, the two top-prioritized purposes were to facilitate communication among clinicians and to inform treatment decisions. From this perspective, the ICD and DSM classifications get a mixed grade: their utility for communication with other health care professionals was one of the three use types in the top tier of ratings, whereas utility for selecting treatment was one of three use types in the bottom tier. Clinicians also do not consider the classifications to be particularly useful in communicating with the patient or family, although the DSM was rated slightly higher in this regard than the ICD.

Finally, the question on the use of the "residual" (other specified, unspecified, and not otherwise specified) categories indicated that these categories are employed relatively often (more so by DSM than ICD-10 users), although a substantial minority (around 32% of ICD-10 users and 27% of DSM users) reported that they rarely or never employed them. The survey results suggest that, for the most part, clinicians are using these categories appropriately: the top three most commonly endorsed reasons were those that would be considered to be legitimate (i.e., presentations that do not conform to specific diagnoses, insufficient information to make a more specific diagnosis, and inability to determine whether symptoms are primary or secondary).

Although the higher usage of these residual categories by DSM users might suggest that the ICD-10 classification has better diagnostic coverage (i.e., that the ICD-10 categories are overall more broadly defined and more likely to cover patient presentations in clinical settings than the more narrow DSM categories), there was no difference in frequency between the two classification systems in respondents' answers with respect to the reason that best corresponds to differences in coverage (i.e., "because the patient's presentation does not conform to any of the specific categories"). The only reason for using residual categories that was given more frequently by ICD-10 users was to indicate that the patient's presentation met the requirements for more than one diagnosis in a grouping, which is an inappropriate use of these categories, given that the convention in ICD-10 (and DSM) is to give multiple comorbid diagnoses in such cases.

The main strengths of this study are the inclusion of survey questions specifically focusing on classification usage rather than just on attitudes about usage, and its diverse sample, which included clinicians from a wide variety of geographical locations, languages, and country income levels. All participants were individuals who indicated that they customarily assigned diagnoses in clinical practice.

The main limitation of the survey is that the sample is not representative of the whole population of mental health clinicians in terms of their level of interest in diagnosis and classification, given that GCPN members joined the network specifically to participate in studies of diagnostic classification and thus were likely to be more interested in diagnostic and classification issues and more proficient in the use of classification systems than the average clinician. Thus, the relatively high frequency of systematically reviewing diagnostic guidelines or criteria in order to determine whether they apply to individual cases may not generalize to a population of clinicians with a wider range of levels of interest in diagnosis and classification. However, it should be noted that this generalizability problem is inherent to all surveys, even those that randomly select participants, since response rates are traditionally low, and people who agree to participate are those most interested in the topic covered by the survey.

Additionally, some answers in the present survey may have been subject to a social desirability bias, as clinicians could have wanted to present their diagnostic practice in the best possible light.

CONCLUSIONS

If it were the case, as suggested by conventional wisdom, that clinicians' use of the ICD and DSM classifications is largely confined to the diagnostic labels and codes, then current efforts to improve the clinical utility of the ICD diagnostic guidelines and DSM diagnostic criteria would have a limited impact on clinical practice. Although the survey sample of GCPN members was likely self-selected to use diagnostic guidelines or criteria more often than the average clinician, the results of this survey suggest that clinicians do use the diagnostic guidelines and criteria in routine clinical practice and that efforts to revise and update them is likely to have an impact on that practice.

Because of limitations in using self-report methodology to examine actual behaviour, it would be useful to employ additional methodologies in the future[2,3], such as direct observation of clinician's classification usage in clinical settings. Such research would help not only to improve classification systems, but also to enhance the function of classification as the interface between clinical practice and health information.

REFERENCES

1. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. World Psychiatry 2011;10:86-93.
2. First M, Pincus H, Levine J et al. Clinical utility as a criterion for revising psychiatric diagnoses. Am J Psychiatry 2004;161:946-54.
3. First MB, Bhat V, Adler D et al. How do clinicians actually use the DSM in clinical practice and why we need to know more. J Nerv Ment Dis 2014; 202:841-4.

4. Reed GM, Correia JM, Esparza P et al. The WPA-WHO global survey of psychiatrists' attitudes towards mental disorders classification. World Psychiatry 2011;10:118-31.
5. Evans SC, Reed GM, Roberts MC et al. Psychologists' perspectives on the diagnostic classification of mental disorders: results from the WHO-IUPsyS Global Survey. Int J Psychology 2013;48:177-93.
6. Maj M. The media campaign on the DSM-5: recurring comments and lessons for the future of diagnosis in psychiatric practice. Epidemiol Psychiatr Sci 2015;24:197-202.
7. Jampala V, Sierles F, Taylor M. Consumers' views of DSM-III: attitudes and practices of U.S. psychiatrists and 1984 graduating psychiatric residents. Am J Psychiatry 1986;143:148-53.
8. Jampala V, Sierles F, Taylor M. The use of DSM-III in the United States: a case of not going by the book. Compr Psychiatry 1988;29:39-47.
9. Mezzich JE. International surveys on the use of ICD-10 and related diagnostic systems. Psychopathology 2002;35:72-5.
10. Someya T, Takahashi M, Takahashi S. Is DSM widely accepted by Japanese clinicians? Psychiatry Clin Neurosci 2001;55:437-50.
11. Zimmerman M, Jampala VC, Sierles FS et al. DSM-III and DSM-III-R: what are American psychiatrists using and why? Compr Psychiatry 1993; 34:365-74.
12. Mellsop GW, Dutu G, Robinson G. New Zealand psychiatrists views on global features of ICD-10 and DSM-IV. Aust N Z J Psychiatry 2007;41:157-65.
13. Mellsop GW, Banzato C, Shinfuku N et al. An international study of the views of psychiatrists on present and preferred characteristics of classifications of psychiatric disorders. Int J Ment Health 2007;36:17-25.
14. Suzuki Y, Takahashi T, Nagamine M et al. Comparison of psychiatrists' views on classification of mental disorders in four East Asian countries/areas. Asian J Psychiatry 2010;3:20-5.
15. Setterberg SR, Ernst M, Rao U et al. Child psychiatrists' views of DSM-III-R: a survey of usage and opinions. J Am Acad Child Adolesc Psychiatry 1991;30:652-8.
16. Reed GM. Toward ICD-11: improving the clinical utility of WHO's International Classification of Mental Disorders. Prof Psychol Res Pr 2010;41:457-64.
17. Reed GM, Rebello TJ, Pike KM et al. WHO's Global Clinical Practice Network for mental health. Lancet Psychiatry 2015;2:379-80.
18. Reed GM, First MB, Medina-Mora ME et al. Draft diagnostic guidelines for ICD-11 mental and behavioural disorders available for review and comment. World Psychiatry 2016;15:112-3.
19. Keeley JW, Reed GM, Roberts MC et al. Developing a science of clinical utility in diagnostic classification systems: field study strategies for ICD-11 Mental and Behavioural Disorders. Am Psychol 2016;71:3-16.
20. First MB, Reed GM, Hyman SE et al. The development of the ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. World Psychiatry 2015;14:82-90.
21. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: World Health Organization, 1992.
22. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders - Diagnostic Criteria for Research. Geneva: World Health Organization, 1993.
23. World Health Organization. ICD-10 Version: 2016 Chapter V Mental and Behavioural Disorders. http://apps.who.int/classifications/icd10.
24. World Health Organization. Mental Health Atlas. Geneva: World Health Organization, 2014.
25. Widiger T, Clark LA. Towards DSM-IV and the classification of psychopathology. Psychol Bull 2000;126:946-63.
26. Jablensky A. Psychiatric classification: validity and utility. World Psychiatry 2016;15:26-31.

DOI:10.1002/wps.20525

Lack of evidence to favor specific preventive interventions in psychosis: a network meta-analysis

Cathy Davies[1], Andrea Cipriani[2], John P.A. Ioannidis[3-7], Joaquim Radua[1,8,9], Daniel Stahl[10], Umberto Provenzani[1,11], Philip McGuire[12,13], Paolo Fusar-Poli[1,11,13,14]

[1]Early Psychosis: Interventions & Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [2]Department of Psychiatry, University of Oxford, and Oxford Health NHS Foundation Trust, Oxford, UK; [3]Department of Medicine, Stanford Prevention Research Center, Stanford, CA, USA; [4]Department of Health Research and Policy, Stanford University School of Medicine, Stanford, CA, USA; [5]Department of Biomedical Data Science, Stanford University School of Medicine, Stanford, CA, USA; [6]Meta-Research Innovation Center at Stanford, Stanford University, Stanford, CA, USA; [7]Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, USA; [8]FIDMAG Germanes Hospitalàries, CIBERSAM, Sant Boi de Llobregat, Spain; [9]Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Stockholm, Sweden; [10]Biostatistics Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [11]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy; [12]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [13]National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre, London, UK; [14]OASIS Service, South London and Maudsley NHS Foundation Trust, London, UK

Preventing psychosis in patients at clinical high risk may be a promising avenue for pre-emptively ameliorating outcomes of the most severe psychiatric disorder. However, information on how each preventive intervention fares against other currently available treatment options remains unavailable. The aim of the current study was to quantify the consistency and magnitude of effects of specific preventive interventions for psychosis, comparing different treatments in a network meta-analysis. PsycINFO, Web of Science, Cochrane Central Register of Controlled Trials, and unpublished/grey literature were searched up to July 18, 2017, to identify randomized controlled trials conducted in individuals at clinical high risk for psychosis, comparing different types of intervention and reporting transition to psychosis. Two reviewers independently extracted data. Data were synthesized using network meta-analyses. The primary outcome was transition to psychosis at different time points and the secondary outcome was treatment acceptability (dropout due to any cause). Effect sizes were reported as odds ratios and 95% confidence intervals (CIs). Sixteen studies (2,035 patients, 57% male, mean age 20.1 years) reported on risk of transition. The treatments tested were needs-based interventions (NBI); omega-3 + NBI; ziprasidone + NBI; olanzapine + NBI; aripiprazole + NBI; integrated psychological interventions; family therapy + NBI; D-serine + NBI; cognitive behavioural therapy, French & Morrison protocol (CBT-F) + NBI; CBT-F + risperidone + NBI; and cognitive behavioural therapy, van der Gaag protocol (CBT-V) + CBT-F + NBI. The network meta-analysis showed no evidence of significantly superior efficacy of any one intervention over the others at 6 and 12 months (insufficient data were available after 12 months). Similarly, there was no evidence for intervention differences in acceptability at either time point. Tests for inconsistency were non-significant and sensitivity analyses controlling for different clustering of interventions and biases did not materially affect the interpretation of the results. In summary, this study indicates that, to date, there is no evidence that any specific intervention is particularly effective over the others in preventing transition to psychosis. Further experimental research is needed.

Key words: Psychosis, risk, prevention, needs-based interventions, cognitive behavioural therapy, antipsychotics, omega-3, integrated psychological interventions, family therapy, network meta-analysis, guidelines

(World Psychiatry 2018;17:196–209)

Individuals at clinical high risk for psychosis (CHR-P)[1] present with attenuated psychotic symptoms, impairments of social, emotional and cognitive functioning[2], and help-seeking behaviour[3]. They have around 20% risk of developing psychosis (but not any other non-psychotic disorder[4,5]) over a two-year period[6].

Primary indicated prevention in CHR-P individuals has the unique potential to alter the course of the disorder[7] and improve clinical outcomes[8]. Current international guidelines – such as those of the National Institute for Health and Care Excellence (NICE) and the European Psychiatric Association (EPA) – recommend that CHR-P individuals be primarily offered cognitive behavioural therapy (CBT) with or without family interventions[9,10]. However, while prophylactic treatment with antipsychotics is altogether prohibited by NICE guidelines[9], the EPA allows its use in the case of severe and progressive symptomatology[10].

The evidence supporting these partially conflicting recommendations is relatively unclear[11], despite several pairwise meta-analyses having been published to date[10,12-18]. For example, earlier meta-analyses concluded that no reliable recommendations with respect to specific interventions could be made, because studies were too heterogeneous[12], with comparable efficacy across different treatments[16] or no effects at all[17]. The most recent meta-analysis concluded that both CBT and antipsychotics are effective[13]. The other meta-analyses were affected by mistakes[19] or methodological limitations, such as the use of overall effect sizes computed across heterogeneous interventions of questionable clinical interpretability[10,12,18], inclusion of patients not assessed with standard CHR-P instruments (e.g., with schizotypal disorders[20])[12,13,15,18], inclusion of non-randomized and uncontrolled trials[10], pooling of time-dependent outcomes[21] in the same group (e.g., 6 and 12 months[18]) or no time stratification at all[13], or poor meta-analytical approaches[13]. Meta-analyses have acquired a major influence on clinical practice and guidelines[22], so they can be particularly harmful if they are of suboptimal quality.

Another problem is that the included trials involved a variety of specific interventions[12], which were inconsistently clustered in pairwise comparisons. For example, although CBT is

an umbrella term for a plethora of heterogeneous strategies[23], different CBT protocols have been lumped together, and the specific efficacy of each defining element or specific protocol remains unclear[24].

The objective of this network meta-analysis (NMA) was to summarize the available evidence about the specific efficacy of different preventive interventions in CHR-P individuals. NMA offers additional benefits over standard pairwise analyses in that the comparative efficacy of specific interventions can be estimated and ranked, even when two treatments have never been compared directly head-to-head[25]. Furthermore, since NMA can improve the precision of estimates by allowing integration of both direct and indirect treatment effect estimates[26], it is recommended over pairwise meta-analyses by the World Health Organization as a basis for clinical guidelines[27]. Therefore, NMA should be considered the highest level of evidence in CHR-P treatment guidelines[28].

METHODS

The protocol for this study was registered on PROSPERO (CRD42017069550). The study was conducted in accordance with the PRISMA statement[29].

Interventions included

We included all randomized controlled trials (RCTs) of pharmacological and/or non-pharmacological interventions for CHR-P individuals. We were *a priori* interested in the following non-pharmacological interventions: CBT (various protocols), psychoeducation, family therapy, supportive counselling, needs-based interventions (NBI), and integrated psychological therapies. We were also interested in the following pharmacological interventions: antipsychotics (olanzapine, risperidone, ziprasidone, aripiprazole) and novel/experimental pharmacotherapies (omega-3 fatty acids and D-serine). As indicated in the protocol, additional interventions emerging from the literature search were also considered (e.g., glycine and cognitive remediation).

The definition of the exact types of interventions is essential to reduce heterogeneity and produce robust informative results of direct clinical significance. As such, we first took each trial and carefully identified the treatment components that were characterizing each specific intervention, as detailed below.

Needs-based interventions (NBI)

Since CHR-P patients recruited in clinical trials are help-seeking adolescents and young adults accessing clinical services, randomizing them to no treatment is not considered a reasonable or ethical option[30]. Defining "treatment as usual" in these samples is also challenging, because treatment is not standardized and largely depends on local service configurations and the availability of specific resources or competences.

We therefore used the most established and original definition of NBI employed by the founders of the CHR-P paradigm, which focuses on the presenting symptoms and problems already manifest[31]. In accordance with this definition[32], NBI may include any of the following components: a) supportive psychotherapy primarily focusing on pertinent issues such as social relationships and vocational or family problems; b) case management, providing psychosocial assistance with accommodation, education or employment; c) brief family psychoeducation and support; d) medications other than antipsychotics; and e) clinical monitoring and crisis management [31,33].

Cognitive behavioural therapy, French & Morrison protocol (CBT-F)

The CBT-F protocol[34] is based on the principles developed by Beck[35]. The intervention is formulation-driven, problem-focused and time-limited, with manualized strategies selected on the basis of the patient's prioritized problem. The key components include building engagement, collaborative goal-setting and formulation, normalizing experiences, evaluating appraisals and core beliefs, and behavioural experiments[34,36].

Cognitive behavioural therapy, van der Gaag protocol (CBT-V)

The protocol developed by van der Gaag et al[37] essentially includes the French & Morrison protocol[34], but with two additional components. These comprise psychoeducation about dopamine system supersensitivity and training/behavioural experiments on cognitive biases that may contribute to paranoia[38]. Further behavioural goals include sustaining school and work attendance, enhancing social relationships, and reducing cannabis use[37].

Integrated psychological interventions, Bechdolf protocol (IPI)

The protocol developed by Bechdolf et al[39] contains a number of components, including individual CBT-F[34], manualized group social skills training, computerized cognitive remediation to address thought and perception deficits, and manualized psychoeducational multi-family group sessions[39,40].

Family-focused therapy, Miklowitz protocol (FFT)

A family-focused therapy (FFT) protocol, initially designed for those with or at risk of bipolar disorder, was adapted by Miklowitz et al[41] for the CHR-P population. The key components include psychoeducation and development of a prevention plan with the patient and family, sessions where the patient and family practice skills for better com-

munication, and sessions focusing on enhancing problem solving skills[41].

Pharmacological interventions

Pharmacological interventions included currently licensed medications, novel or experimental pharmacotherapies, and nutritional supplements.

Placebo

The placebo designation was reserved for placebo pills administered as pharmacological control conditions. Placebos were designed to match the active drug intervention in appearance but without the pharmacological compound of interest.

Nodes for the network meta-analysis

The specific interventions listed above were pooled into "nodes" for the network meta-analysis. Nodes were defined by the linear combination of any of the above specific interventions. Each individual pharmacological treatment was assigned to its own node. As indicated in the protocol, different dosages of the same drug/molecule were classed under the same node. Placebo was initially considered as a separate node from NBI. However, in line with the protocol, sensitivity analyses investigated the effect of alternate clustering of nodes (see statistical analysis).

Search strategy and selection criteria

We performed a multi-step literature search using the following keywords: (risk OR prodromal OR prodrom* OR ultra high risk OR clinical high risk OR high risk OR genetic high risk OR at risk mental state OR risk of progression OR progression to first-episode OR prodromally symptomatic OR basic symptoms) AND (psychosis) AND (RCT OR randomized controlled trial OR placebo controlled trial OR trial).

First, systematic searches were conducted in the Web of Science (which includes Web of Science Core Collection, BIOSIS Citation Index, KCI - Korean Journal Database, MEDLINE, Russian Science Citation Index, and SciELO Citation Index), the Cochrane Central Register of Controlled Trials, and Ovid/PsychINFO databases, until July 18, 2017, with no restrictions on language or publication date.

Second, we used Scopus/Web of Science to search reference lists of retrieved articles and previously conducted systematic reviews and meta-analyses. We manually searched for published and unpublished data in relevant conference proceedings, trial registries and drug-approval agencies. In addition, we contacted study authors for supplemental data and searched the OpenGrey database for grey literature.

Abstracts identified by this process were then screened, and full-text articles were retrieved for further inspection against the inclusion and exclusion criteria (as detailed *a priori* in the protocol). The literature search, study selection and data extraction were conducted by two authors (CD, UP) independently. During all stages, in the case of disagreement, consensus was reached through discussion with a third author (PFP).

Studies were eligible for inclusion when the following criteria were fulfilled: a) original articles, abstracts or pilot studies; b) RCTs (including cluster randomized trials, but excluding cross-over studies); c) designed as blinded (either single- or double-blind); d) conducted in CHR-P individuals as established by validated assessments, i.e. Comprehensive Assessment of At-Risk Mental States (CAARMS)[42], Structured Interview for Psychosis-risk Syndromes (SIPS)[43,44], Positive and Negative Syndrome Scale (PANSS)[45], Brief Psychiatric Rating Scale (BPRS)[46], or Early Recognition Inventory (ERIraos)[47]; e) comparing specific preventive interventions as defined above; and f) sample size of 10 or greater[48].

The exclusion criteria were: a) reviews/non-original data; b) studies lacking at least two compared groups; c) studies of first-episode psychosis or other non-CHR-P groups; d) lack of data needed for meta-analytical computation of the primary (transition) outcome (authors were contacted and asked to provide summary data); e) lack of proper randomization (quasi-randomization, observational naturalistic studies); f) samples size < 10; and g) articles presenting overlapping, redundant data (for a particular outcome at the same time point). Specifically, in the case of overlapping samples, we used the largest one. Studies that were designed as blinded but could not maintain blinding during follow-up (e.g., for psychological interventions) were not excluded.

Outcome measures and data extraction

The primary outcome was transition to psychosis. Due to the variable effect of time on transition risk[6,21], we stratified outcomes and analyses into 6 and 12 month follow-up time points. Sample sizes were based on the numbers randomized to each arm, to prevent artificial inflation of transition risk[6,49]. Participants who dropped out of individual studies after randomization were classified as non-transitions[6,10,14,50].

Where studies did not report sufficient data to extract the primary outcome, we contacted the relevant authors. In the case of non-response or where studies presented data graphically, numerical data were digitally extracted from the Kaplan-Meier plots using a previously validated procedure[51,52], as defined in the protocol.

The secondary outcome was the acceptability of interventions (discontinuation due to any cause), indexed as the number of participants who dropped out of each arm for any reason following randomization, over the number randomized [53-55].

In addition, we extracted the following information for each study: first author and year of publication, country, types of outcomes, intervention and control descriptions, study design,

quality assessment (see below), intervention period and follow-up duration, study arm details (sample size, mean age, percent male), and diagnostic tools used for CHR-P diagnosis and determining transition to psychosis.

Quality of the evidence

Risk of bias

The Cochrane Risk of Bias tool[56] was used to assess and classify the risk of bias in each of the included studies, as per criteria defined *a priori*. A judgement was made about whether each study had a high, low or unclear risk of bias in each of the following six domains: random sequence generation, allocation concealment, blinding of participants and study personnel, blinding of outcome assessments, incomplete outcome data, and selective outcome reporting.

The overall risk of bias was classified as low if none of the above domains was rated as high risk and three or less were rated as unclear risk. It was classified as moderate if one domain was rated as high risk, or none rated as high risk but four or more rated as unclear risk. All other studies were classified as having a high risk of bias[57].

To represent the quality of evidence associated with comparisons in the network meta-analysis, we used coloured edges in the network plots, as recommended[58].

GRADE

We assessed the certainty of evidence contributing to network estimates of the primary outcome using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework[59]. The GRADE method characterizes the quality of a body of evidence on the basis of six factors: study limitations, imprecision, heterogeneity, inconsistency, indirectness, and publication bias[59].

We tabulated the findings for the above six factors to aid in the decision-making process for the downgrading of evidence. If one of the factors was present for a comparison, then the overall confidence rating for that comparison was considered for downgrading by one or two levels (as appropriate). Each comparison started as high quality/confidence (as based on RCTs), and was downgraded to moderate, low or very low, depending on the presence, severity and potential impact of the aforementioned factors. These represented the final judgements about the certainty of the evidence[59,60].

Statistical analysis

Frequentist NMAs were conducted for transition and acceptability outcomes using the *network* package in STATA (version SE 14.2; StataCorp). First, a network plot was constructed for each outcome[61] to ensure that nodes of the network were sufficiently connected[58]. We then performed a NMA assuming consistency and a common heterogeneity across all comparisons in the network. This allowed us to derive a single summary treatment effect (odds ratio, OR) for every possible pairwise comparison of treatments, which takes account of all evidence from the network of trials, including both direct and indirect comparisons. Correlations in effect sizes induced by multi-arm trials[62] were accounted for[58,63]. The resulting relative ORs with 95% confidence intervals (CIs) for each pair of treatments were reported in league tables[64].

The interventions were then ranked by the surface under the cumulative ranking curve (SUCRA), which accounts for the location as well as the variance of all relative treatment effects[65]. SUCRA is a numeric presentation of the overall ranking and provides a single number (from 0 to 100%) associated with each intervention[66]. The higher the SUCRA value, and the closer to 100%, the higher the likelihood that an intervention is in the top rank, and vice versa[66]. Cluster ranking methods[58,65] – using both transition and acceptability SUCRA values – were used to order the treatments in league tables, in line with recent guidance which requires interpretation of SUCRA only in the context of NMA uncertainty, rather than at face value[66]. Statistical significance was set at $p < 0.05$.

We assessed the assumption of consistency by calculating, for each closed loop, an inconsistency factor (differences between direct and indirect evidence) along with 95% CIs and associated p values. We plotted the results graphically as the ratio of ORs (RORs) and 95% CIs for each loop[64]. Inconsistency was defined as disagreement between direct and indirect evidence, with 95% CIs for RORs excluding 1.

Given the low power of the loop-specific approach and its focus on local inconsistency (between direct and indirect evidence), we also tested a full design-by-treatment model[62] for the primary outcome to evaluate inconsistency more globally, including between trials with different designs (e.g., two-arm vs. multi-arm). A NMA under the inconsistency model was applied and a χ^2 test was used to infer about the statistical significance of all possible inconsistencies in the networks[67].

The transitivity assumption was examined by assessing the distributions of potential effect modifiers for every comparison in the network, including percentage of males[68], age[69], percentage exposed to antipsychotic medications at baseline[70], type of blinding and publication year[6]. The presence of small-study effects was assessed by visual inspection of comparison-adjusted funnel plots[59].

To evaluate the impact of study quality and our data analysis procedures, we conducted sensitivity analyses for the primary outcome restricted to: a) studies with a low risk of bias for the blinding of outcome assessments; b) studies whose data were not digitally extracted (e.g., from Kaplan-Meier plots); and c) published data only. We also repeated the analyses after applying alternate clustering of the following nodes: a) pooling NBI and placebo; b) pooling different CBT protocols; c) pooling different types of antipsychotic molecules, and d) separating the different NBI components (i.e., supportive therapy vs. clinical monitoring vs. other). Finally, network meta-regressions

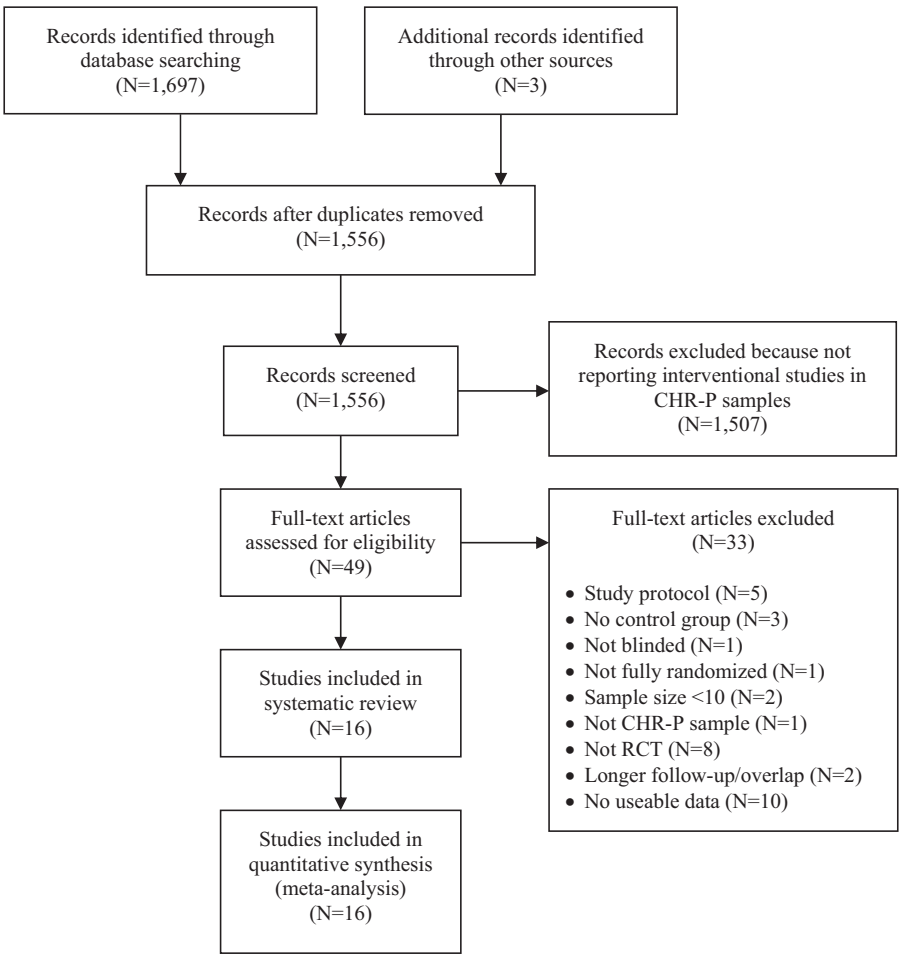

Figure 1 PRISMA flow chart of the study selection process. CHR-P – clinical high risk for psychosis, RCT – randomized controlled trial

were planned in the case of substantial heterogeneity and at least ten studies[71] to test the impact of different CHR-P diagnostic instruments/criteria.

RESULTS

Characteristics of the included studies

1,556 references were found in the literature search, most of which were not reporting RCTs in CHR-P individuals; 49 were fully screened for the inclusion and exclusion criteria, resulting in a final sample of 16 studies (Figure 1). There were only five, four, two and three studies reporting data for 18, 24, 36 and >36 month time points, respectively, and therefore all results reported hereafter are for 6 and 12 months only.

The 16 studies used in the analyses of the primary outcome contributed data on 2,035 patients, with a mean age of 20.1 ± 2.8 years, and 57% were male (Table 1). The mean sample size was 127 (range 44-304). Six studies were conducted in North America, six in Europe, three in Australia and one was multinational. Two studies were three-arm and the rest were two-arm trials. Two studies had a treatment duration of <6 months, ten of 6 months, and four of 12 months. Of the 14 studies with available information on sponsorship/funding, three[31,75,81,82] acknowledged pharmaceutical company grants. The CAARMS and the SIPS were the most common CHR-P diagnostic instruments[44] (six and seven studies, respectively).

For the 6-month analysis of the primary outcome, these 16 studies provided data on 20 direct comparisons between 11 different treatment nodes (Figure 2). Three studies provided follow-up data only for the 6-month analysis, and therefore the 12-month analysis consisted of 13 studies (N=1,811), providing data on 17 direct comparisons between 8 different treatment nodes (Figure 2). The network plots for the acceptability outcome were the same at 12 months and similar at 6 months (integrated psychological interventions was missing).

Primary outcome: transition

Results of the NMA showed a lack of evidence for clearly superior efficacy of specific treatments in preventing transi-

Table 1 Details of included studies

Study	Study arms (N)	Network inclusion	Treatment duration (months)	Follow-up time points (months)	% male	Mean age	CHR-P criteria	Study design	Country	% exposed to antipsychotics at baseline
Addington et al[30]	CBT-F + NBI (27)	6, 12	6	6,12,18	71	20.9	SIPS	SB-RCT	Canada	0
	NBI (24)									
Amminger et al[33]	Omega-3 + NBI (41)	6, 12	3	6,12,84	33	16.4	PANSS	DB-RCT	Austria	0
	NBI (40)									
Bechdolf et al[39]	IPI (63)	6, 12	12	6,12,18, 24	63	26.0	ERIraos	SB-RCT	Germany	0
	NBI (65)									
Bechdolf et al[72]	NBI + ARI (96)	6, 12	12	6,12	66	24.4	SIPS + BS	SB-RCT	Germany	3.4
	NBI (55)									
	CBT-F + NBI (129)									
Cadenhead et al[73]	Omega-3 + NBI (65)	6, 12	6	6,12,18, 24	56	18.8	SIPS	DB-RCT	US, Canada	0
	NBI (62)									
Kantrowitz et al[74]	D-serine + NBI (20)	6	4	4	66	19.4	SIPS	DB-RCT	US	11.4
	NBI (24)									
McGlashan et al[75]	NBI + OLA (31)	6, 12	12	12, 24	65	17.7	SIPS	DB-RCT	US, Canada	10
	NBI (29)									
McGorry et al[31]	CBT-F + RIS + NBI (31)	6, 12	6	6,12,36-48	58	20.0	BPRS	SB-RCT	Australia	0
	NBI (28)									
McGorry et al[76]	Omega-3 + NBI (153)	6, 12	6	6,12	46	19.2	CAARMS	DB-RCT	Multi-national	0
	NBI (151)									
Miklowitz et al[41]	FFT + NBI (66)	6	6	6	57	17.4	SIPS	SB-RCT	US, Canada	20.9
	NBI (63)									

Table 1 Details of included studies *(continued)*

Study	Study arms (N)	Network inclusion	Treatment duration (months)	Follow-up time points (months)	% male	Mean age	CHR-P criteria	Study design	Country	% exposed to antipsychotics at baseline
Morrison et al[36]	CBT-F + NBI (37)	6, 12	6	6,12,36	69	22.0	CAARMS	SB-RCT	UK	0
	NBI (23)									
Morrison et al[77]	CBT-F + NBI (144)	6, 12	6	6,12,18,24	63	20.7	CAARMS	SB-RCT	UK	0
	NBI (144)									
Stain et al[78]	CBT-F + NBI (30)	6, 12	6	6,12	40	16.3	CAARMS	SB-RCT	Australia	0
	NBI (27)									
van der Gaag et al[37]	CBT-F + CBT-V + NBI (98)	6, 12	6	6,12,18,48	49	22.7	CAARMS	SB-RCT	The Netherlands	1.5
	CBT-F + NBI (103)									
Woods et al[79] Woods[80]	NBI + ZIP (24)	6	6	6	64	22.3	SIPS	DB-RCT	US	0
	NBI (27)									
Yung et al[81] McGorry et al[82]	CBT-F + NBI (44)	6, 12	12	6,12	39	18.1	CAARMS	SB-RCT	Australia	0
	CBT-F + RIS + NBI (43)									
	NBI (28)									

CBT – cognitive behavioural therapy, CBT-F – French & Morrison CBT protocol, NBI – needs-based interventions (including placebo), IPI – integrated psychological interventions, ARI – aripiprazole, OLA – olanzapine, RIS – risperidone, FFT – family-focused therapy, CBT-V – van der Gaag CBT protocol, ZIP – ziprasidone, SIPS – Structured Interview for Psychosis-risk Syndromes, PANSS – Positive and Negative Syndrome Scale, ERIraos – Early Recognition Inventory, BS – basic symptoms, BPRS – Brief Psychiatric Rating Scale, CAARMS – Comprehensive Assessment of At-Risk Mental States, SB-RCT – single-blind randomized controlled trial, DB-RCT – double-blind randomized controlled trial

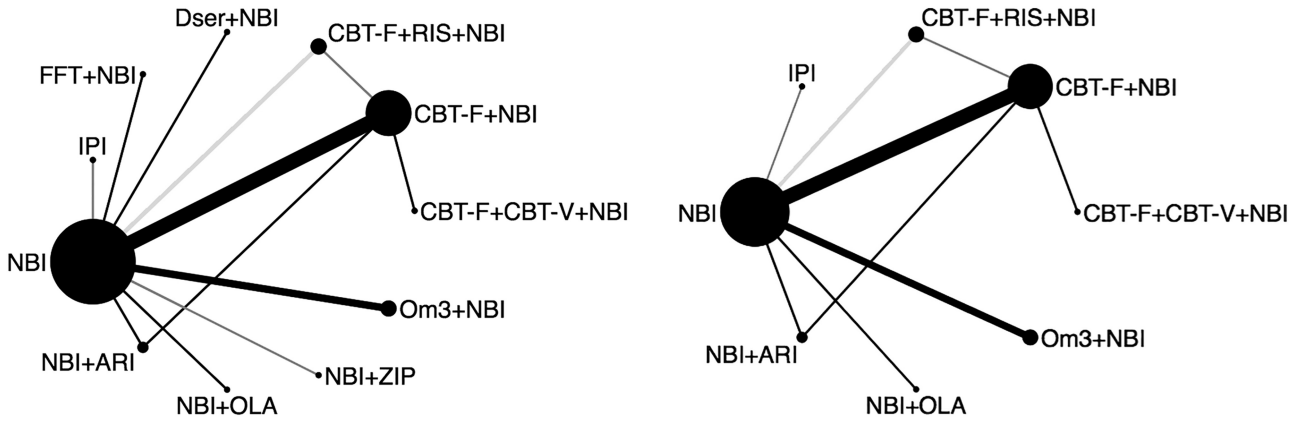

Figure 2 Network plots of direct comparisons in the network meta-analysis for transition outcome at 6 (on the left) and 12 months (on the right). The width of the lines is proportional to the number of trials comparing each pair of treatments and the size of each node is proportional to the number of studies testing the specific treatment. The color of the lines represents the comparison-specific bias level for the blinding of outcome assessments in the majority of trials (black = low risk, dark grey = unclear risk, light grey = high risk). NBI – needs-based interventions (including placebo), IPI – integrated psychological interventions, FFT – family-focused therapy, Dser – D-serine, CBT – cognitive behavioural therapy, CBT-F – French & Morrison CBT protocol, CBT-V – van der Gaag CBT protocol, RIS – risperidone, Om3 – omega-3 fatty acids, ZIP – ziprasidone, OLA – olanzapine, ARI – aripiprazole.

tion, with no significant effects of any one intervention over any others at 6 or 12 month time points (Figures 3 and 4).

Using NBI as a comparator, the OR and 95% CI for each treatment (all OR < 1 favor the given treatment) at 6 months were: 0.06 (0.00-1.90) for integrated psychological interventions; 0.17 (0.01-2.69) for family-focused therapy + NBI; 0.22 (0.02-2.17) for CBT-F + CBT-V + NBI; 0.29 (0.03-2.57) for olanzapine + NBI; 0.21 (0.04-1.08) for CBT-F + risperidone + NBI; 0.52 (0.03-10.72) for ziprasidone + NBI; 0.56 (0.03-11.51) for D-serine + NBI; 0.64 (0.15-2.68) for omega-3 + NBI; 0.73 (0.27-2.01) for CBT-F + NBI; and 0.94 (0.15-5.73) for aripiprazole + NBI.

At 12 months, ORs against the NBI comparator were: 0.04 (0.00-1.06) for integrated psychological interventions; 0.15 (0.02-1.25) for olanzapine + NBI; 0.21 (0.03-1.60) for CBT-F + CBT-V + NBI; 0.43 (0.11-1.68) for CBT-F + risperidone + NBI; 0.58 (0.23-1.47) for CBT-F + NBI; 0.64 (0.18-2.26) for omega-3 + NBI; and 1.39 (0.26-7.28) for aripiprazole + NBI.

While almost all the interventions at both time points had estimates favoring them over NBI, the differences were not beyond chance, and the 95% CIs for the NMA estimates were often very large, indicating substantial imprecision. The cluster ranking (based on SUCRA values for transition and acceptability) at 6 and 12 months is illustrated by the ordering of treatments in Tables 2 and 3.

No statistically significant inconsistency was evident at any time point, with 95% CIs for all RORs compatible with zero inconsistency (ROR=1). However, only two loops were available. Using the design-by-treatment interaction test[62], we found no evidence for significant inconsistency for 6 month (p=0.90) and 12 month (p=0.93) networks.

Only two studies had an overall low risk of bias[33,79]; five had unclear risk[72-76], and nine had high risk[30,31,36,37,39,41,77,78,81]. The edges (lines) in Figure 2 reflect the Cochrane risk of bias for the blinding of outcome assessments, estimated as the level of bias in the majority of trials and weighted according to the number of studies in each comparison[58]. The GRADE assessment highlighted low or very low confidence in almost all estimates, primarily due to study limitations (high risks of bias) and imprecision.

The numbers of studies remaining (at 6 and 12 months, respectively) after exclusion of those with a high or unclear risk of bias for the blinding of outcome assessments were 10 and 8; after exclusion of those whose data were extracted by digitizing Kaplan-Meier plots were 13 and 12; after exclusion of unpublished studies were 13 and 11. The NMA model was refitted accordingly and no differences in conclusions were observed for any OR at any time point.

Repeating the analyses treating NBI + placebo as a separate node to NBI, or separating the different NBI components, had no effect on the NMA estimates, and therefore we used the pooled NBI + placebo in the main analysis (Table 1, Figures 2-4). Similarly, pooling together different CBT protocols or different antipsychotic molecules in the same node produced no significant results. There were not enough studies to allow robust meta-regression analyses on the type of CHR-P instruments. Visual inspection of funnel plots revealed no substantive evidence of small-study effects.

Secondary outcome: acceptability

Acceptability data were available for 14 of 16 studies at 6 months (N=1,848), and 12 of 13 studies at 12 months (N=1,752). There were no significant differences in acceptability between any treatment comparisons at 6 or 12 months (Figures 3 and 4). The SUCRA cluster ranking (for transition and acceptability) is illustrated in those figures.

IPI	-	-	-	-	-	-	-	-	-	-
0.39 (0.00 to 31.26)	FFT + NBI	0.58 (0.12 to 2.73)	0.20 (0.03 to 1.18)	0.86 (0.18 to 4.15)	0.38 (0.07 to 2.01)	0.69 (0.13 to 3.85)	0.59 (0.17 to 2.03)	0.67 (0.21 to 2.14)	0.84 (0.21 to 3.38)	0.59 (0.20 to 1.70)
0.29 (0.00 to 17.13)	0.74 (0.02 to 27.15)	CBT-F + CBT-V +NBI	0.34 (0.05 to 2.09)	1.48 (0.32 to 6.85)	0.65 (0.12 to 3.56)	1.19 (0.21 to 6.82)	1.01 (0.28 to 3.65)	1.15 (0.41 to 3.21)	1.43 (0.37 to 5.60)	1.01 (0.33 to 3.08)
0.22 (0.00 to 12.43)	0.57 (0.02 to 19.56)	0.77 (0.03 to 17.86)	OLA + NBI	4.37 (0.69 to 27.70)	1.92 (0.28 to 13.20)	3.52 (0.49 to 25.15)	2.97 (0.62 to 14.36)	3.41 (0.76 to 15.42)	4.23 (0.77 to 23.21)	2.98 (0.71 to 12.54)
0.31 (0.01 to 13.33)	0.79 (0.03 to 20.17)	1.07 (0.07 to 15.83)	1.40 (0.09 to 21.37)	CBT-F + RIS + NBI	0.44 (0.08 to 2.48)	0.81 (0.14 to 4.74)	0.68 (0.18 to 2.56)	0.78 (0.25 to 2.45)	0.97 (0.23 to 4.05)	0.68 (0.21 to 2.17)
0.12 (0.00 to 11.54)	0.32 (0.01 to 19.30)	0.43 (0.01 to 18.69)	0.56 (0.01 to 23.08)	0.40 (0.01 to 12.41)	ZIP + NBI	1.83 (0.29 to 11.69)	1.55 (0.37 to 6.48)	1.77 (0.46 to 6.91)	2.20 (0.46 to 10.59)	1.55 (0.43 to 5.57)
0.12 (0.00 to 10.94)	0.30 (0.00 to 18.30)	0.40 (0.01 to 17.74)	0.52 (0.01 to 21.90)	0.38 (0.01 to 11.79)	0.94 (0.01 to 67.94)	D-serine + NBI	0.85 (0.19 to 3.74)	0.97 (0.24 to 4.00)	1.20 (0.24 to 6.09)	0.85 (0.22 to 3.24)
0.10 (0.00 to 3.98)	0.26 (0.01 to 5.94)	0.35 (0.02 to 5.11)	0.45 (0.03 to 6.17)	0.33 (0.04 to 2.93)	0.82 (0.03 to 23.12)	0.87 (0.03 to 24.80)	Om3 + NBI	1.15 (0.52 to 2.51)	1.42 (0.47 to 4.33)	1.00 (0.53 to 1.90)
0.09 (0.00 to 3.02)	0.23 (0.01 to 4.38)	0.30 (0.04 to 2.34)	0.40 (0.04 to 4.39)	0.28 (0.05 to 1.66)	0.71 (0.03 to 17.27)	0.76 (0.03 to 18.53)	0.88 (0.15 to 5.04)	CBT-F + NBI	1.24 (0.50 to 3.06)	0.87 (0.55 to 1.37)
0.07 (0.00 to 3.19)	0.18 (0.01 to 4.89)	0.24 (0.02 to 3.56)	0.31 (0.02 to 5.26)	0.22 (0.02 to 2.38)	0.56 (0.02 to 18.82)	0.59 (0.02 to 20.18)	0.68 (0.07 to 6.84)	0.78 (0.13 to 4.64)	ARI + NBI	0.70 (0.28 to 1.74)
0.06 (0.00 to 1.90)	0.17 (0.01 to 2.69)	0.22 (0.02 to 2.17)	0.29 (0.03 to 2.57)	0.21 (0.04 to 1.08)	0.52 (0.03 to 10.72)	0.56 (0.03 to 11.51)	0.64 (0.15 to 2.68)	0.73 (0.27 to 2.01)	0.94 (0.15 to 5.73)	NBI

Comparison Transition Acceptability

Figure 3 Relative effect sizes for transition to psychosis and acceptability (dropout for any reason) at 6 months, odds ratios (95% CI). Comparisons between treatments should be read from left to right, and the estimate is in the cell in common between the column-defining treatment and the row-defining treatment. Treatments are reported in descending order (from top left to bottom right) as per the cluster ranking for transition and acceptability. For transition, an OR less than 1 favors the column-defined treatment. For acceptability, an OR less than 1 favors the row-defined treatment. All 95% CIs include the null hypothesis OR = 1. Dashes (-) indicate no available NMA estimate. CBT – cognitive behavioural therapy, CBT-F – French & Morrison CBT protocol, CBT-V – van der Gaag CBT protocol, NBI – needs-based interventions (including placebo), RIS – risperidone, FFT – family-focused therapy, IPI – integrated psychological interventions, ARI – aripiprazole, OLA – olanzapine, ZIP – ziprasidone, Om3 – omega-3 fatty acids.

IPI	0.73 (0.15 to 3.48)	1.49 (0.34 to 6.54)	1.43 (0.35 to 5.89)	1.72 (0.55 to 5.43)	1.37 (0.41 to 4.54)	1.68 (0.57 to 4.91)	2.72 (0.75 to 9.83)
0.26 (0.01 to 12.94)	**OLA + NBI**	2.05 (0.45 to 9.45)	1.97 (0.46 to 8.52)	2.37 (0.71 to 7.94)	1.89 (0.54 to 6.63)	2.31 (0.74 to 7.20)	3.74 (0.98 to 14.29)
0.19 (0.00 to 9.17)	0.73 (0.04 to 13.83)	**CBT-F + CBT-V +NBI**	0.96 (0.26 to 3.51)	1.15 (0.45 to 2.93)	0.92 (0.30 to 2.86)	1.12 (0.41 to 3.11)	1.82 (0.57 to 5.84)
0.09 (0.00 to 3.23)	0.35 (0.03 to 4.35)	0.48 (0.05 to 5.03)	**CBT-F + RIS + NBI**	1.20 (0.49 to 2.95)	0.96 (0.33 to 2.74)	1.17 (0.47 to 2.93)	1.89 (0.62 to 5.76)
0.07 (0.00 to 2.06)	0.26 (0.03 to 2.60)	0.35 (0.06 to 2.20)	0.74 (0.17 to 3.22)	**CBT-F + NBI**	0.80 (0.42 to 1.52)	0.97 (0.65 to 1.47)	1.58 (0.78 to 3.18)
0.06 (0.00 to 2.07)	0.24 (0.02 to 2.74)	0.32 (0.03 to 3.52)	0.67 (0.10 to 4.24)	0.90 (0.19 to 4.28)	**Om3 + NBI**	1.22 (0.72 to 2.07)	1.98 (0.82 to 4.80)
0.04 (0.00 to 1.06)	0.15 (0.02 to 1.25)	0.21 (0.03 to 1.60)	0.43 (0.11 to 1.68)	0.58 (0.23 to 1.47)	0.64 (0.18 to 2.26)	**NBI**	1.62 (0.80 to 3.29)
0.03 (0.00 to 1.13)	0.11 (0.01 to 1.60)	0.15 (0.01 to 1.72)	0.31 (0.04 to 2.47)	0.42 (0.08 to 2.14)	0.46 (0.06 to 3.72)	0.72 (0.14 to 3.78)	**ARI + NBI**

Comparison Transition Acceptability

Figure 4 Relative effect sizes for transition to psychosis and acceptability (dropout for any reason) at 12 months, odds ratios (95% CI). Comparisons between treatments should be read from left to right, and the estimate is in the cell in common between the column-defining treatment and the row-defining treatment. Treatments are reported in descending order (from top left to bottom right) as per the cluster ranking for transition and acceptability. For transition, an OR less than 1 favors the column-defined treatment. For acceptability, an OR less than 1 favors the row-defined treatment. All 95% CIs include the null hypothesis OR = 1. CBT – cognitive behavioural therapy, CBT-F – French & Morrison CBT protocol, CBT-V – van der Gaag CBT protocol, NBI – needs-based interventions (including placebo), RIS – risperidone, IPI – integrated psychological interventions, ARI – aripiprazole, OLA – olanzapine, Om3 – omega-3 fatty acids.

DISCUSSION

This is the first network meta-analysis exploring the efficacy of specific interventions for the prevention of psychosis in CHR-P individuals. Adopting strict inclusion criteria, a total of 16 RCTs, with 2,035 patients, were included in the analyses. There were not enough studies to analyze data with a NMA approach beyond 6 and 12 month follow-ups. Two networks were established at 6 and 12 months, including 11 and 8 nodes, respectively. Network meta-analyses showed no clear evidence of superior efficacy for any specific intervention at any time point. The results were not affected by biases, inconsistency or small-study effects.

The main finding of the current study is that there is a lack of evidence to favor specific effective interventions to prevent psychosis in CHR-P individuals. Our analyses were based on a detailed protocol, which defined the exact type of interventions and nodes *a priori*. This was done with the aim of providing robust informative results of direct clinical significance. For example, deconstructing the efficacy of different types of CBT that are based on different protocols[83] seems necessary to inform accurate and evidence-based clinical guidelines for patients, clinicians and policy makers. Our NMA comparing the different CBT protocols is also timely, since authors have recently claimed that the "black box" of CBT should be unpacked into its specific therapeutic components[23,24,84-86].

In a similar fashion, our NMA represents the first attempt at deconstructing – through sensitivity analyses – the effect of different components (including placebo) that characterize NBI, which is usually employed as the control condition in this field. We also restricted our literature search to include only RCTs designed to be blinded, and studies that strictly used CHR-P assessment instruments, to minimize selection biases. Therefore, to date, our study represents the most fine-grained analysis that has deconstructed the specific effect of preventive interventions for psychosis.

Negative (non-significant) results are rarely published in psychiatric literature[87], which is affected by excess of statistical significance[88-92]. In fact, interpreting negative findings is particularly challenging, because absence of evidence is not evidence of absence[93]. In particular, when large CIs are observed (as in Figures 3 and 4), some sizeable effects may still have been missed. Nevertheless, our work represents the most powered data synthesis in this field. For example, the meta-analysis by Stafford et al[15] – on which current clinical guidelines are based – analyzed 11 studies, but one of them included an open-label trial (N=124)[94] and another did not assess participants against standard CHR-P criteria (N=79)[20], leaving nine studies (N=1,043) that are in common with the current

NMA. Since that meta-analysis, seven new trials involving 992 new CHR-P participants (an increase of more than 50%) have been published, all of which reported non-significant effects[41,72-74,76,78-80]. Since our NMA included these new data, it is more powered than previous pairwise analyses.

In the context of power considerations, indirect evidence, when combined with direct evidence through NMA, increases the power and precision of treatment effect estimates compared to pairwise analyses[26]. Furthermore, when we pooled different CBT protocols or antipsychotic molecules in the same node – thus increasing the statistical power – no significant results were still observed. Overall, the core result of our NMA is more congruent with the evidence emerging from the most recent trials, compared to previous evidence syntheses.

The current lack of evidence to support specific preventive treatments is also consistent with the fact that the three largest interventional studies in this field have all produced negative findings[95]. Earlier studies that dominated the conclusions of some previous meta-analyses (e.g., the omega-3 trial[33]) were likely false positives. There is also converging lack of significant benefits on other clinical outcomes besides transition to psychosis, such as attenuated symptom severity[14,15,96], functioning[10,14,18], depressive comorbidities[15], distress[14], and quality of life[14,15].

These findings, taken together, are particularly problematic given the conceptual concerns over the clinical validity and significance of the dichotomous concept of transition within the CHR-P paradigm[97,98]. More to the point, it is not clear whether the currently tested treatments are only delaying the onset of psychosis as opposed to altering the course of the disorder[7]. Long-term outcome trials are scarce and the results are conflicting.

The additional caveat is that the exact mechanism of action of the tested preventive treatments is – at best – poorly defined, due to lack of an established and validated pathophysiological model underlying the onset of psychosis in CHR-P samples. A lack of mechanistic models forces researchers to proceed with empirical attempts that may eventually prove unsuccessful, as has ultimately been the case for omega-3 fatty acids[76]. However, as our ability to stratify CHR-P individuals into more homogenous subtypes improves, so may our success in testing specific treatments targeted to underlying biological and psychological mechanisms[99].

Our findings may have an impact on research and clinical practice. In times of scarce resources, our NMA can help to focus the next generation of research on the most promising interventions. Although our ranking analysis should be interpreted cautiously[66,100] in the context of non-superiority of any intervention compared to any other, it suggests that CBT-F, which currently represents the most widely adopted intervention, may not be the best candidate (of relevance, the largest CBT-F trial to date provided non-significant results[77]). On the other hand, the apparent promising profile of integrated psychological approaches could be the target of future replications.

Future research in this area will need to test novel interventions that may act on underlying psychological or neurobiological processes associated with the onset of psychosis. Although there are no clinically valid CHR-P biomarkers yet available[101], several international consortia are ongoing (PRONIA[102], NAPLS[103], PSYSCAN[104]) with the aim of developing them. At the same time, it seems warranted to address the clinical heterogeneity[1,6,49,105,106] that may prevent the discovery of reliable preventive treatments, and to improve the design of the next generation of trials. For example, it is apparent that unstructured recruitment processes and risk enrichment procedures in samples undergoing CHR-P assessment have a substantial role in determining the actual level of risk for psychosis in these individuals[107-109], leading to underpowered and non-significant trials[95]. On a clinical side, individuals meeting CHR-P criteria may be informed that, at present, there is no evidence for specific treatments being more effective than any others, and current options should be carefully weighted on a personal basis depending on an individual's needs.

This study has some limitations. First, only 16 RCTs were included, reflecting the paucity of high-quality studies available in the CHR-P field. However, capitalizing on the increased power and precision of NMA[26], the Cochrane group has conducted such analysis in even smaller databases, including as few as three to seven studies[110-113]. Furthermore, sufficient data were available for 6 and 12 month networks only, which precluded insight into whether treatments may have some effectiveness in the longer term. As a result of the sparse literature, many nodes were not well connected, with the corollary of limited ability to check for inconsistency, more imprecise estimates and wide 95% CIs.

In addition, the quality of NMA rests on the quality of included studies, many of which were found to be at high or unclear risk of bias, with GRADE confidence estimates predominantly low or very low – suggesting that true effects may be substantially different from the estimates. This is particularly the case for trials including any psychological interventions. We addressed this issue through a strict and detailed assessment of biases and sensitivity analyses. Going forward, given that all comparisons in the NMA were downgraded due to study limitations (risk of bias) and imprecision, the addition of high-quality studies with adequate sample sizes is needed to improve these confidence ratings.

A final limitation is that, whilst dropout due to any cause was available from the majority of trials, this is a rather crude measure of treatment acceptability, and a more proximal index, such as specific adverse effects, may have revealed significant differences between treatments, in particular for trials of antipsychotic molecules. However, these outcomes are rarely reported in the CHR-P literature.

In conclusion, there is currently no evidence to favor specific interventions for the prevention of psychosis. Further experimental research in this field is needed.

ACKNOWLEDGEMENTS

This work was supported in part by the UK National Institute for Health Research (NIHR) Biomedical Research Centre, South London and Maudsley NHS Foundation Trust. A. Cipriani is supported by the NIHR Oxford Cognitive Health Clinical Research Facility. The funders had no influence on the design, collection, analysis and interpretation of the data, writing of the report and decision to submit this article for publication.

REFERENCES

1. Fusar-Poli P. The clinical high-risk state for psychosis (CHR-P), Version II. Schizophr Bull 2017;43:44-7.
2. Fusar-Poli P, Borgwardt S, Bechdolf A et al. The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA Psychiatry 2013;70:107-20.
3. Falkenberg I, Valmaggia L, Byrnes M et al. Why are help-seeking subjects at ultra-high risk for psychosis help-seeking? Psychiatry Res 2015;228: 808-15.
4. Webb JR, Addington J, Perkins DO et al. Specificity of incident diagnostic outcomes in patients at clinical high risk for psychosis. Schizophr Bull 2015;41:1066-75.
5. Fusar-Poli P, Rutigliano G, Stahl D et al. Long-term validity of the At Risk Mental State (ARMS) for predicting psychotic and non-psychotic mental disorders. Eur Psychiatry 2017;42:49-54.
6. Fusar-Poli P, Cappucciati M, Borgwardt S et al. Heterogeneity of psychosis risk within individuals at clinical high risk. JAMA Psychiatry 2016;73: 113-20.
7. Millan MJ, Andrieux A, Bartzokis G et al. Altering the course of schizophrenia: progress and perspectives. Nat Rev Drug Discov 2016;15:485-515.
8. Fusar-Poli P, McGorry P, Kane JM. Improving outcomes of first episode psychosis. World Psychiatry 2017;16:251-65.
9. National Institute for Health and Care Excellence. Psychosis and schizophrenia in adults: prevention and management. Guideline CG178. London: National Institute for Health and Care Excellence, 2014.
10. Schmidt SJ, Schultze-Lutter F, Schimmelmann BG et al. EPA guidance on the early intervention in clinical high risk states of psychoses. Eur Psychiatry 2015;30:388-404.
11. Taylor M, Perera U. NICE CG178 psychosis and schizophrenia in adults: treatment and management - an evidence-based guideline? Br J Psychiatry 2015;206:357-9.
12. Preti A, Cella M. Randomized-controlled trials in people at ultra high risk of psychosis: a review of treatment effectiveness. Schizophr Res 2010;123: 30-6.
13. Deas G, Kelly C, Hadjinicolaou AV et al. An update on: meta-analysis of medical and non-medical treatments of the prodromal phase of psychotic illness in at risk mental states. Psychiatr Danub 2016;28:31-8.
14. Hutton P, Taylor PJ. Cognitive behavioural therapy for psychosis prevention: a systematic review and meta-analysis. Psychol Med 2014;44:449-68.
15. Stafford MR, Jackson H, Mayo-Wilson E et al. Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013;346:f185.
16. Kelly C, Hadjinicolaou AV, Holt C et al. Meta-analysis of medical and non-medical treatments of the prodromal phase of psychotic illness in at-risk mental states. Psychiatr Danub 2010;22:56-62.
17. Marshall M, Rathbone J. Early intervention for psychosis. Cochrane Database Syst Rev 2011;6:CD004718.
18. Van Der Gaag M, Smit F, Bechdolf A et al. Preventing a first episode of psychosis: meta-analysis of randomized controlled prevention trials of 12 month and longer-term follow-ups. Schizophr Res 2013;149:56-62.
19. Stafford MR, Jackson H, Mayo-Wilson E et al. Errata: Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013; 346:f762.
20. Nordentoft M, Thorup A, Petersen L et al. Transition rates from schizotypal disorder to psychotic disorder for first-contact patients included in the OPUS trial. A randomized clinical trial of integrated treatment and standard treatment. Schizophr Res 2006;83:29-40.
21. Kempton MJ, Bonoldi I, Valmaggia L et al. Speed of psychosis progression in people at ultra-high clinical risk. JAMA Psychiatry 2015;72:622-3.
22. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016;94:485-514.
23. Hartmann JA, McGorry PD, Schmidt SJ et al. Opening the black box of cognitive-behavioural case management in clients with ultra-high risk for psychosis. Psychother Psychosom 2017;86:292-9.
24. Mueser KT, Glynn SM, Meyer-Kalos PS. What are the key ingredients of optimal psychosocial treatment for persons recovering from a first episode of psychosis? World Psychiatry 2017;16:266-7.
25. Cipriani A, Higgins JPT, Geddes JR et al. Conceptual and technical challenges in network meta-analysis. Ann Intern Med 2013;159:130-7.
26. Thorlund K, Mills EJ. Sample size and power considerations in network meta-analysis. Syst Rev 2012;1:41.
27. Kanters S, Ford N, Druyts E et al. Use of network meta-analysis in clinical guidelines. Bull World Health Organ 2016;94:782-4.
28. Leucht S, Chaimani A, Cipriani AS et al. Network meta-analyses should be the highest level of evidence in treatment guidelines. Eur Arch Psychiatry Clin Neurosci 2016;266:477-80.
29. Hutton B, Salanti G, Caldwell DM et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med 2015;162:777-84.
30. Addington J, Epstein I, Liu L et al. A randomized controlled trial of cognitive behavioral therapy for individuals at clinical high risk of psychosis. Schizophr Res 2011;125:54-61.
31. McGorry PD, Yung AR, Phillips LJ et al. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Arch Gen Psychiatry 2002;59:921-8.
32. Yung AR, McGorry PD, Francey SM et al. PACE: a specialised service for young people at risk of psychotic disorders. Med J Aust 2007;187:S43-6.
33. Amminger GP, Schäfer MR, Papageorgiou K et al. Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders. Arch Gen Psychiatry 2010;67:146-54.
34. French P, Morrison A. Early detection and cognitive therapy for people at high risk of developing psychosis. Chichester: Wiley, 2004.
35. Beck AT. Cognitive therapy and the emotional disorders. Madison: International University Press, 1976.
36. Morrison AP, French P, Walford L et al. Cognitive therapy for the prevention of psychosis in people at ultra-high risk: randomised controlled trial. Br J Psychiatry 2004;185:291-7.
37. van der Gaag M, Nieman DH, Rietdijk J et al. Cognitive behavioral therapy for subjects at ultrahigh risk for developing psychosis: a randomized controlled clinical trial. Schizophr Bull 2012;38:1180-8.
38. Rietdijk J, Dragt S, Klaassen R et al. A single blind randomized controlled trial of cognitive behavioural therapy in a help-seeking population with an at risk mental state for psychosis: the Dutch Early Detection and Intervention Evaluation (EDIE-NL) trial. Trials 2010;11:30.
39. Bechdolf A, Wagner M, Ruhrmann S et al. Preventing progression to first-episode psychosis in early initial prodromal states. Br J Psychiatry 2012; 200:22-9.
40. Bechdolf A, Puetzfeld V, Guettgemanns J et al. Cognitive behaviour therapy for people at-risk of psychosis. A treatment manual. Bern: Huber, 2010.
41. Miklowitz DJ, O'Brien MP, Schlosser DA et al. Family-focused treatment for adolescents and young adults at high risk for psychosis: results of a randomized trial. J Am Acad Child Adolesc Psychiatry 2014;53:848-58.
42. Yung AR, Yuen HP, McGorry PD et al. Mapping the onset of psychosis: the Comprehensive Assessment of At Risk Mental States (CAARMS). Aust N Z J Psychiatry 2005;39:964-71.
43. McGlashan T, Walsh B, Woods S. The psychosis-risk syndrome: handbook for diagnosis and follow-up. Oxford: Oxford University Press, 2010.
44. Fusar-Poli P, Cappucciati M, Rutigliano G et al. Towards a standard psychometric diagnostic interview for subjects at ultra high risk of psychosis: CAARMS versus SIPS. Psychiatry J 2016;2016:7146341.
45. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 1987;13:261-76.
46. Overall J, Gorham D. The Brief Psychiatric Rating Scale (BPRS): recent developments in ascertainment and scaling. Psychopharmacol Bull 1988; 24:97-9.
47. Haefner H, Bechdolf A, Klosterkotter J et al. Early detection and intervention in psychosis. A practise handbook. Stuttgart: Schattauer, 2011.
48. Cipriani A, Zhou X, Del Giovane C et al. Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. Lancet 2016;388:881-90.
49. Fusar-Poli P, Cappucciati M, Bonoldi I et al. Prognosis of brief psychotic episodes: a meta-analysis. JAMA Psychiatry 2016;73:211-20.

50. Higgins JPT, Green S (eds). Cochrane handbook for systematic reviews of interventions version 5.1.0. http://handbook.cochrane.org.
51. Guyot P, Ades A, Ouwens MJ et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol 2012;12:9.
52. Radua J, Grunze H, Amann BL. Meta-analysis of the risk of subsequent mood episodes in bipolar disorder. Psychother Psychosom 2017;86:90-8.
53. Cipriani A, Furukawa TA, Salanti G et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 2018;391:1357-66.
54. Cipriani A, Furukawa TA, Salanti G et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet 2009;373:746-58.
55. Miura T, Noma H, Furukawa TA et al. Comparative efficacy and tolerability of pharmacological treatments in the maintenance treatment of bipolar disorder: a systematic review and network meta-analysis. Lancet Psychiatry 2014;1:351-9.
56. Higgins JPT, Altman DG, Gøtzsche PC et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343: d5928.
57. Furukawa TA, Salanti G, Atkinson LZ et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ Open 2016;6:e010919.
58. Chaimani A, Higgins JPT, Mavridis D et al. Graphical tools for network meta-analysis in STATA. PLoS One 2013;8:e76654.
59. Salanti G, Del Giovane C, Chaimani A et al. Evaluating the quality of evidence from a network meta-analysis. PLoS One 2014;9:e99682.
60. Guyatt G, Oxman AD, Sultan S et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol 2013;66:151-7.
61. Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. Stata J 2015;15:905-50.
62. Higgins JPT, Jackson D, Barrett JK et al. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods 2012;3:98-110.
63. White IR. Multivariate random-effects meta-regression: updates to mvmeta. Stata J 2011;11:255-70.
64. Mavridis D, Giannatsi M, Cipriani A et al. A primer on network meta-analysis with emphasis on mental health. Evid Based Ment Health 2015; 18:40-6.
65. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64:163-71.
66. Mbuagbaw L, Rochwerg B, Jaeschke R et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev 2017;6:79.
67. Chaimani A, Salanti G, Leucht S et al. Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. Evid Based Ment Health 2017;20:88-94.
68. Wilson RP, Patel R, Bhattacharyya S. Do fewer males present to clinical high-risk services for psychosis relative to first-episode services? Early Interv Psychiatry 2016;11:429-35.
69. Schultze-Lutter F, Hubl D, Schimmelmann BG et al. Age effect on prevalence of ultra-high risk for psychosis symptoms: replication in a clinical sample of an early detection of psychosis service. Eur Child Adolesc Psychiatry 2017;26:1401-5.
70. Fusar-Poli P, Frascarelli M, Valmaggia L et al. Antidepressant, antipsychotic and psychological interventions in subjects at high clinical risk for psychosis: OASIS 6-year naturalistic study. Psychol Med 2015;45:1327-39.
71. Fusar-Poli P, Papanastasiou E, Stahl D et al. Treatments of negative symptoms in schizophrenia: meta-analysis of 168 randomized placebo-controlled trials. Schizophr Bull 2015;41:892-9.
72. Bechdolf A, Müller H, Stützer H et al. PREVENT: a randomized controlled trial for the prevention of first-episode psychosis comparing cognitive-behavior therapy (CBT), clinical management, and aripiprazole combined and clinical management and placebo combined. Schizophr Bull 2017; 43(Suppl.1):S56-7.
73. Cadenhead K, Addington J, Cannon T et al. Omega-3 fatty acid versus placebo in a clinical high-risk sample from the North American Prodrome Longitudinal Studies (NAPLS) consortium. Schizophr Bull 2017; 43(Suppl.1):S16.
74. Kantrowitz JT, Woods SW, Petkova E et al. D-serine for the treatment of negative symptoms in individuals at clinical high risk of schizophrenia: a pilot, double-blind, placebo-controlled, randomised parallel group mechanistic proof-of-concept trial. Lancet Psychiatry 2015;2:403-12.
75. McGlashan TH, Zipursky RB, Perkins D et al. Randomized, double-blind trial of olanzapine versus placebo in patients prodromally symptomatic for psychosis. Am J Psychiatry 2006;163:790-9.
76. McGorry PD, Nelson B, Markulev C et al. Effect of ω-3 polyunsaturated fatty acids in young people at ultrahigh risk for psychotic disorders. The NEURAPRO randomized controlled trial. JAMA Psychiatry 2017;74:19-27.
77. Morrison AP, French P, Stewart SL et al. Early detection and intervention evaluation for people at risk of psychosis: multisite randomised controlled trial. BMJ 2012;344:e2233.
78. Stain HJ, Bucci S, Baker AL et al. A randomised controlled trial of cognitive behaviour therapy versus non-directive reflective listening for young people at ultra high risk of developing psychosis: the detection and evaluation of psychological therapy (DEPTh) trial. Schizophr Res 2016;176: 212-9.
79. Woods S, Saksa J, Compton M et al. Effects of ziprasidone versus placebo in patients at clinical high risk for psychosis. Schizophr Bull 2017;43 (Suppl.1):S58.
80. Woods SW. Ziprasidone in the psychosis prodrome (ZIP). ClinicalTrials.gov, NCT00635700.
81. Yung AR, Phillips LJ, Nelson B et al. Randomized controlled trial of interventions for young people at ultra high risk for psychosis: 6-month analysis. J Clin Psychiatry 2011;72:430-40.
82. McGorry PD, Nelson B, Phillips LJ et al. Randomized controlled trial of interventions for young people at ultra-high risk of psychosis: twelve-month outcome. J Clin Psychiatry 2013;74:349-56.
83. Bechdolf A, Phillips LJ, Francey SM et al. Recent approaches to psychological interventions for people at risk of psychosis. Eur Arch Psychiatry Clin Neurosci 2006;256:159-73.
84. Flach C, French P, Dunn G et al. Components of therapy as mechanisms of change in cognitive therapy for people at risk of psychosis: analysis of the EDIE-2 trial. Br J Psychiatry 2015;207:123-9.
85. Leichsenring F, Steinert C. Is cognitive behavioral therapy the gold standard for psychotherapy? The need for plurality in treatment and research JAMA 2017;318:1323-4.
86. Mulder R, Murray G, Rucklidge J. Common versus specific factors in psychotherapy: opening the black box. Lancet Psychiatry 2017;4:953-62.
87. Porter RJ, Boden JM, Miskowiak K et al. Failure to publish negative results: a systematic bias in psychiatric literature. Aust N Z J Psychiatry 2017;51:212-4.
88. Flint J, Cuijpers P, Horder J et al. Is there an excess of significant findings in published studies of psychotherapy for depression? Psychol Med 2015; 45:439-46.
89. Ioannidis JPA. Excess significance bias in the literature on brain volume abnormalities. Arch Gen Psychiatry 2011;68:773-80.
90. Carvalho AF, Köhler CA, Fernandes BS et al. Bias in emerging biomarkers for bipolar disorder. Psychol Med 2016;46:2287-97.
91. David SP, Ware JJ, Chu IM et al. Potential reporting bias in fMRI studies of the brain. PLoS One 2013;8:e70104.
92. Fusar-Poli P, Radua J, Frascarelli M et al. Evidence of reporting biases in voxel-based morphometry (VBM) studies of psychiatric and neurological disorders. Hum Brain Mapp 2014;35:3052-65.
93. Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ 1995;311:485.
94. Ruhrmann S, Bechdolf A, Kühn KU et al. Acute effects of treatment for prodromal symptoms for people putatively in late initial prodromal state of psychosis. Br J Psychiatry 2007;191:88-96.
95. Fusar-Poli P. Negative psychosis prevention trials. JAMA Psychiatry 2017; 74:651.
96. Devoe DJ, Peterson A, Addington J. Negative symptom interventions in youth at risk of psychosis: a systematic review and network meta-analysis. Schizophr Bull (in press).
97. Fusar-Poli P, Van Os J. Lost in transition: setting the psychosis threshold in prodromal research. Acta Psychiatr Scand 2013;127:248-52.
98. Van Os J, Guloksuz S. A critique of the "ultra-high risk" and "transition" paradigm. World Psychiatry 2017;16:200-6.
99. Yung AR. Treatment of people at ultra-high risk for psychosis. World Psychiatry 2017;16:207-8.

100. Trinquart L, Attiche N, Bafeta A et al. Uncertainty in treatment rankings: reanalysis of network meta-analyses of randomized trials. Ann Intern Med 2016;164:666-73.
101. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 2012;17:1174-9.
102. PRONIA. Personalised prognostic tools for early psychosis management. www.pronia.eu.
103. NAPLS. The North American Prodrome Longitudinal Study. https://campuspress.yale.edu/napls/.
104. PSYSCAN. Translating neuroimaging findings from research into clinical practice. http://intranet.psyscan.eu/.
105. Fusar-Poli P, Cappucciati M, De Micheli A et al. Diagnostic and prognostic significance of brief limited intermittent psychotic symptoms (BLIPS) in individuals at ultra high risk. Schizophr Bull 2017;43:48-56.
106. Fusar-Poli P, Tantardini M, De Simone S et al. Deconstructing vulnerability for psychosis: meta-analysis of environmental risk factors for psychosis in subjects at ultra high-risk. Eur Psychiatry 2017;40:65-75.
107. Fusar-Poli P, Schultze-Lutter F, Cappucciati M et al. The dark side of the moon: meta-analytical impact of recruitment strategies on risk enrichment in the clinical high risk state for psychosis. Schizophr Bull 2016;42:732-43.
108. Fusar-Poli P, Rutigliano G, Stahl D et al. Deconstructing pretest risk enrichment to optimize prediction of psychosis in individuals at clinical high risk. JAMA Psychiatry 2016;73:1260-7.
109. Fusar-Poli P. Why ultra high risk criteria for psychosis prediction do not work well outside clinical samples and what to do about it. World Psychiatry 2017;16:212-3.
110. Roccarina D, Majumdar A, Thorburn D et al. Management of people with intermediate-stage hepatocellular carcinoma: an attempted network meta-analysis. Cochrane Database Syst Rev 2017;3:CD011649.
111. Buzzetti E, Kalafateli M, Thorburn D et al. Interventions for hereditary haemochromatosis: an attempted network meta-analysis. Cochrane Database Syst Rev 2017;3:CD011647.
112. Mantzoukis K, Rodríguez-Perálvarez M, Buzzetti E et al. Pharmacological interventions for acute hepatitis B infection: an attempted network meta-analysis. Cochrane Database Syst Rev 2017;3:CD011645.
113. Desborough M, Hadjinicolaou AV, Chaimani A et al. Alternative agents to prophylactic platelet transfusion for preventing bleeding in people with thrombocytopenia due to chronic bone marrow failure: a meta-analysis and systematic review. Cochrane Database Syst Rev 2016;10:CD012055.

DOI:10.1002/wps.20526

Psychopathology in 7-year-old children with familial high risk of developing schizophrenia spectrum psychosis or bipolar disorder – The Danish High Risk and Resilience Study - VIA 7, a population-based cohort study

Ditte Ellersgaard[1-3], Kerstin Jessica Plessen[2-4], Jens Richardt Jepsen[2,4,5], Katrine Soeborg Spang[2-4], Nicoline Hemager[1-4], Birgitte Klee Burton[2-4], Camilla Jerlang Christiani[1-3], Maja Gregersen[1,2], Anne Søndergaard[1,2], Md Jamal Uddin[1,2,6], Gry Poulsen[1,2,6], Aja Greve[2,7], Ditte Gantriis[2,7], Ole Mors[2,7], Merete Nordentoft[1-3], Anne Amalie Elgaard Thorup[2-4]

[1]Mental Health Services - Capital Region of Denmark, Mental Health Centre Copenhagen, Copenhagen, Denmark; [2]The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), Aarhus, Denmark; [3]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; [4]Mental Health Services - Capital Region of Denmark, Child and Adolescent Mental Health Centre, Copenhagen, Denmark; [5]Mental Health Services - Capital Region of Denmark, Center for Neuropsychiatric Schizophrenia Research and Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research, Copenhagen, Denmark; [6]Department of Public Health - Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark; [7]Psychosis Research Unit, Aarhus University Hospital, Aarhus, Denmark

This study aimed to compare the psychopathological profiles of children at familial high risk of schizophrenia spectrum psychosis (FHR-SZ) or bipolar disorder (FHR-BP) with population-based controls. We used Danish nationwide registers to retrieve a cohort of 522 seven-year-old children of parents with schizophrenia spectrum psychosis (N=202), bipolar disorder (N=120) or none of these disorders (N=200). Psychopathology was assessed by reports from multiple informants, including children, parents and teachers. Lifetime DSM-IV diagnoses were ascertained by blinded raters through the Schedule for Affective Disorders and Schizophrenia for School-Age Children. The dimensional assessment of psychopathology was performed by the Child Behavior Checklist, the Teacher's Report Form, a modified version of the ADHD-Rating Scale, the Test Observation Form, and the State-Trait Anxiety Inventory for Children. Current level of functioning was evaluated using the Children's Global Assessment Scale (CGAS). The prevalence of lifetime psychiatric diagnoses was significantly higher in both FHR-SZ children (38.7%, odds ratio, OR=3.5, 95% confidence interval, CI: 2.2-5.7, $p < 0.001$) and FHR-BP children (35.6%, OR=3.1, 95% CI: 1.8-5.3, $p < 0.001$) compared with controls (15.2%). FHR-SZ children displayed significantly more dimensional psychopathology on all scales and subscales compared with controls except for the Anxious subscale of the Test Observation Form. FHR-BP children showed higher levels of dimensional psychopathology on several scales and subscales compared with controls, but lower levels compared with FHR-SZ children. Level of functioning was lower in both FHR-SZ children (CGAS mean score = 68.2; 95% CI: 66.3-70.2, $p < 0.0001$) and FHR-BP children (73.7; 95% CI: 71.2-76.3, $p < 0.05$) compared with controls (77.9; 95% CI: 75.9-79.9). In conclusion, already at the age of seven, FHR-SZ and FHR-BP children show a higher prevalence of a broad spectrum of categorical and dimensional psychopathology compared with controls. These results emphasize the need for developing early intervention strategies towards this vulnerable group of children.

Key words: Schizophrenia spectrum psychosis, bipolar disorder, children at familial high risk, psychiatric diagnoses, dimensional psychopathology, level of functioning, early intervention strategies

(World Psychiatry 2018;17:210–219)

The importance of early detection and intervention for the outcome of schizophrenia has received increasing attention during the last two decades. Efforts have moved from studying treatment in first-episode psychosis towards evaluating intervention before the onset of psychosis[1]. Moreover, studies on intervention in individuals with ultra-high-risk states have provided promising results[2]. Evidence also confirms that schizophrenia is a neurodevelopmental disorder with subtle signs long before psychosis onset[3,4]. These findings suggest that intervention should begin already in the premorbid phase.

Identifying early antecedents in children and adolescents is necessary in the effort to develop primary intervention strategies for severe mental illness like schizophrenia and bipolar disorder. Additionally, differentiation between shared and distinct antecedents and risk factors in the two disorders is a prerequisite in determining whether preventive interventions should or not be illness specific[5].

Since schizophrenia and bipolar disorder are rare events in the general population, familial high risk studies of children born to parents with schizophrenia (FHR-SZ) or bipolar disorder (FHR-BP) are useful in studying trajectories towards these conditions. The offspring of parents with severe mental disorders have been reported to have elevated rates of not only the disorder of their parents but also a wide range of other mental disorders[6,7].

Studies on psychopathology in child offspring of parents with schizophrenia and bipolar disorder, as opposed to adult offspring, are vital because they provide knowledge on early developmental psychopathology long before onset of the full-blown disorders. Indeed, earlier studies have found a high prevalence of a broad spectrum of Axis I disorders and dimensional psychopathology in FHR-SZ children[8-15] as well as FHR-BP children[12,16-30]. However, many previous clinical studies have weaknesses, such as small sample sizes, use of convenience samples, inclusion of children from different age groups, or lack of a proper control group. Furthermore, studies of FHR-SZ children using comprehensive semi-structured diagnostic interviews and clinical rating scales are rare.

To investigate whether FHR-SZ and FHR-BP children are at risk of developing disorders that are specific to their respective risk profiles, or if they simply share a general proneness to psychopathology, it is necessary to study children with different

familial risk profiles simultaneously. This has only been done in very few studies[12,31].

In the present study, we aimed to characterize and compare psychopathological profiles in children born to parents with schizophrenia or bipolar disorder and population-based controls.

METHODS

Data presented are part of the Danish High Risk and Resilience Study - VIA 7, a nationwide population-based cohort study of 522 seven-year-old FHR-SZ children, FHR-BP children and controls[32].

Participants

A cohort of 522 seven-year-old (age range 6.9-8.4 years) children, born and living in Denmark, with no, one or two biological parents diagnosed with schizophrenia spectrum psychosis (defined as ICD-10 codes F20, F22 and F25, or ICD-8 codes 295, 297, 298.29, 298.39, 298.89 and 298.99) or bipolar disorder (defined as ICD-10 codes F30 and F31, or ICD-8 codes 296.19 and 296.39) was identified using the Danish Civil Registration System[33] and the Danish Psychiatric Central Research Register[34], including both inpatient and outpatient contacts.

Families in which at least one parent had been diagnosed with schizophrenia spectrum psychosis (the index parent) were matched to control families on gender, age and municipality of the child. Parents from the control group could be registered with any other psychiatric diagnoses except for schizophrenia spectrum psychosis or bipolar disorder.

Families where a parent had been diagnosed with bipolar disorder were a non-matched sample, but they were comparable to the other two groups in terms of age and gender of the children.

Procedures

The study was approved by the Danish Data Protection Agency. The Danish Ministry of Health granted permission to retrieve data from the Danish registers. The study protocol was sent to the Danish Committee on Health Research Ethics, which decided that ethical approval was not needed due to the observational nature of the study. Written informed consent was obtained from all adult participants and from the legal guardians of participating children.

A group of psychologists, medical doctors and nurses carried out the assessments after being trained in the use of all instruments. The investigators who examined the children were blinded to the illness status of the parents. The caregiver who at the present time point knew the child best was asked to provide information on the child's psychopathology.

Children's psychiatric diagnoses and level of functioning

Children's psychiatric diagnoses were ascertained through the Schedule for Affective Disorders and Schizophrenia for School-Age Children - Present and Lifetime Version (K-SADS-PL)[35]. The interview was firstly carried out with the caregiver, then with the child. Best-estimate lifetime DSM-IV-diagnoses were made based on K-SADS-PL and all other available data on the child (e.g., results of cognitive tests and psychopathology scales). Consensus diagnoses were made at conferences with a child and adolescent psychiatrist (AT). In the vast majority of cases, the K-SADS-PL interviews were video-recorded, enabling the researchers to watch parts of them if there was uncertainty regarding scores.

In K-SADS-PL, probable diagnoses are made if criteria for the core symptoms are met, all but one (or a minimum of 75%) of the remaining criteria are met, and the symptoms are causing functional impairment[35]. Both definite and probable diagnoses were included in the analysis. We excluded elimination disorders, because of their questionable clinical significance.

Current level of functioning of the child was evaluated using the Children's Global Assessment Scale (CGAS)[36], as a part of the K-SADS-PL interview.

Dimensional assessment of the children's psychopathology

The Child Behavior Checklist school-age version (CBCL) was completed by the primary caregiver[37]. The scale includes 118 problem behavior items rated on a Likert scale from zero (not true) to two (very true or often true). We used the two broad-band subscales (Internalizing and Externalizing) and the six DSM-IV oriented subscales (Affective Problems, Anxiety Problems, Somatic Problems, Attention Deficit/Hyperactivity Problems, Oppositional Defiant Problems and Conduct Problems).

The Teacher's Report Form (TRF) was completed by the child's teacher[37]. In most aspects this instrument corresponds to the CBCL and most of its items have counterparts in the CBCL.

We used a modified version of the ADHD-Rating Scale (mADHD-RS)[38-40] to assess symptoms of attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD), rated by primary caregivers and teachers. The original ADHD-Rating Scale consists of two nine-item subscales related to the core symptomatology of ADHD: Inattention and Hyperactivity/Impulsivity. The mADHD-RS includes an additional eight-item subscale for problems related to oppositional defiant disorder[38,40]. The items are rated on a four-point Likert scale from zero (never or rarely) to three (very often).

The Test Observation Form (TOF) was used to assess behavioral and emotional problems observed during a test session[41].

It consists of 125 items, scored on a four-point Likert scale. It was completed by the child examiner after testing. The TOF is subdivided into the two broad-band Internalizing and Externalizing subscales and into five empirically based subscales. We excluded the open-ended item 125, where problems not covered by the other items can be rated.

The State-Trait Anxiety Inventory for Children (STAI-CH) was used to measure the children's self-reported level of anxiety[42]. This instrument consists of two 20-item scales including both direct and reversed statements. The State-Anxiety scale was used to measure current level of anxiety at the examination, and the Trait-Anxiety scale to measure the general level of anxiety. Since the STAI-CH is constructed to be used with nine- to twelve-year-old children, it was administered verbally, and the meaning of the questions was explained if needed. The scores of each subscale range from 20 (indicating a low level of anxiety) to 60 (indicating a high level of anxiety). To make the differences in percentages comparable to the other scales, 20 points were subtracted to each score before analysis, so that the potential scores ranged from 0 to 40.

Interrater reliability

All raters attended formal courses on the use of K-SADS-PL prior to data collection. Reliability ratings were held regularly during data collection. Interrater reliability was estimated based on ten video-recorded K-SADS-PL interviews using Krippendorff's alpha with 95% bootstrap confidence intervals (CIs)[43]. The combined observed agreement of K-SADS-PL skip-out criteria across sections in the screening interview was 90.3%. Krippendorff's alpha was 0.74 (95% CI: 0.63-0.82). Because of an insufficient number of cases, it was not possible to estimate Krippendorff's alpha of skip-out criteria separately for each section of the screening interview. Observed agreement ranged from 80 to 100%, except for the post-traumatic stress disorder section, where observed agreement was 20%.

Krippendorff's alpha of CGAS was 0.87 (95% Cl: 0.70-0.92).

Statistical analyses

Differences in demographic and clinical characteristics between the three groups were analyzed by one-way analysis of variance or chi-square test, as appropriate.

Between-group differences in diagnoses were evaluated using logistic regression adjusting for the children's gender. Differences in dimensional psychopathology between the groups were analyzed using generalized linear model (GLM) with Tweedie distribution and log link function, due to non-normally distributed data. Differences in CGAS scores were analyzed using GLM with normal distribution and log link function. Analyses were adjusted for children's gender.

RESULTS

Background characteristics

A final cohort of 522 children from 506 families was retrieved from Danish national registers (Figure 1). Of these, 200 FHR-SZ children, 119 FHR-BP children and 200 controls participated with some data on psychopathology.

We found several significant differences in family characteristics and home environment between the three groups (Table 1).

Children's psychiatric diagnoses

A total of 514 children were assessed with K-SADS-PL (Table 2). The prevalence of any lifetime DSM-IV Axis I psychiatric diagnoses (excluding elimination disorders) was significantly higher in both FHR-SZ children (38.7%, odds ratio, OR=3.5, 95% CI: 2.2-5.7, $p < 0.001$) and FHR-BP children (35.6%, OR=3.1, 95% CI: 1.8-5.3, $p < 0.001$) compared with controls (15.2%).

Both familial risk groups had a higher prevalence of several psychiatric diagnoses compared with controls. However, due to the small number of children with some diagnoses, it was not possible to estimate ORs for all categories. FHR-SZ children had significantly higher ORs of anxiety disorders (OR=2.8, 95% CI: 1.2-6.1, $p < 0.05$), disruptive behavior disorders (OR=6.4, 95% CI: 1.4-29.2, $p < 0.05$), ADHD (OR=3.5, 95% CI: 1.8-6.6, $p < 0.001$), and stress and adjustment disorders (OR=3.8, 95% CI: 1.0-13.8, $p < 0.05$), compared with controls. FHR-BP children had significantly higher ORs of anxiety disorders (OR=2.8, 95% CI: 1.2-6.8, $p < 0.05$), pervasive developmental disorders (OR=3.2, 95% CI: 1.0-9.9, $p < 0.05$), and stress and adjustment disorders (OR=6.0, 95% CI: 1.6-22.2, $p < 0.01$), compared with controls.

Among cases with ADHD, FHR-BP children most often presented the predominantly inattentive type of the disorder (N=8, 72.7%), while FHR-SZ children and controls most often presented the combined or predominantly hyperactive-impulsive type (N=24, 58.5%, and N=8, 57.1%, respectively). The small number of children with ADHD did not allow calculations of the significance of these findings.

Children's level of functioning and dimensional psychopathology

FHR-SZ children had a significantly lower level of functioning (CGAS mean score=68.2, 95% CI: 66.3-70.2) compared with controls (77.9, 95% CI: 75.9-79.9, $p < 0.0001$) and with FHR-BP children (73.7, 95% CI: 71.2-76.3, $p=0.0009$) (Table 3). FHR-BP children had significantly lower levels of functioning compared with controls ($p=0.0126$).

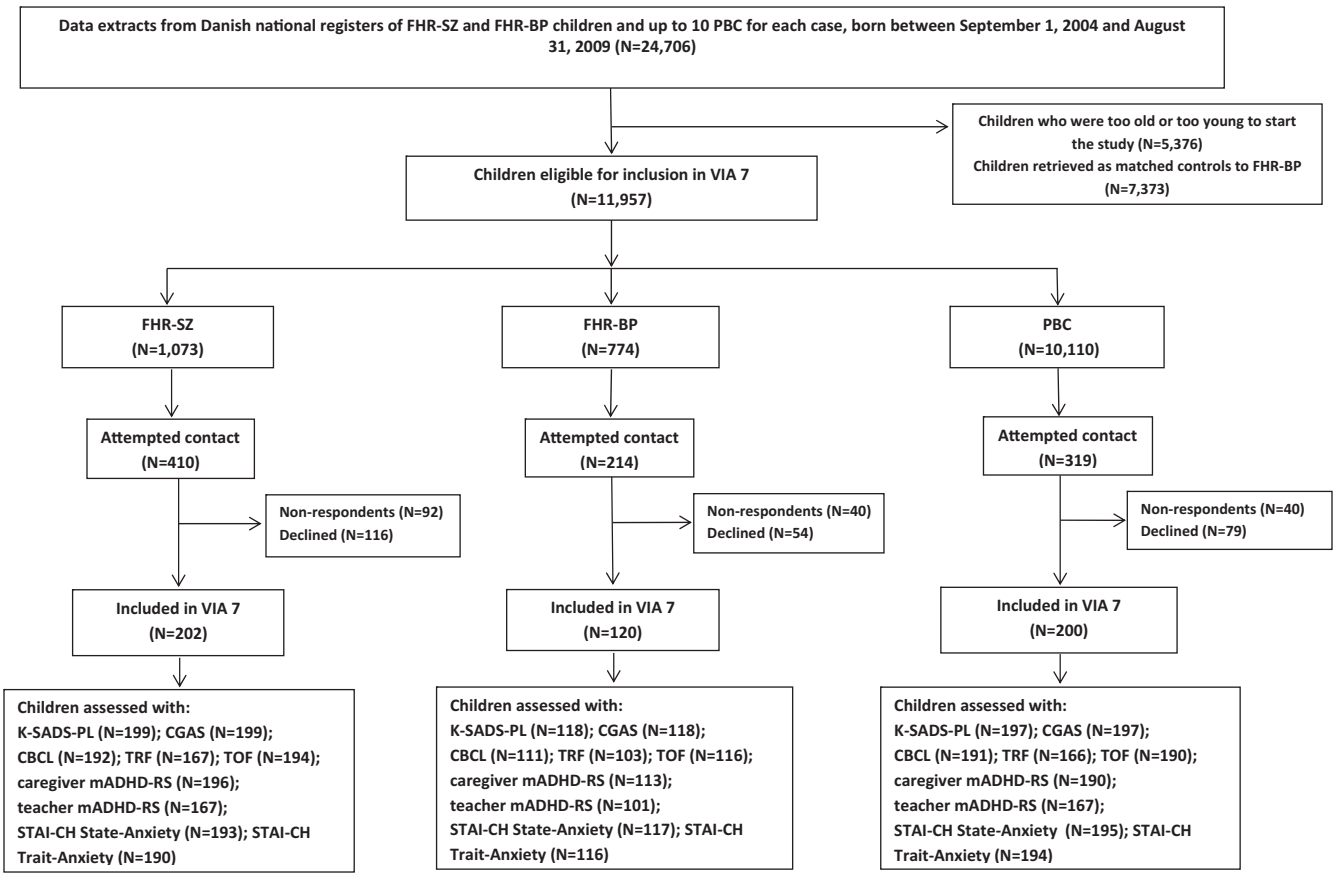

Figure 1 Flow chart of the recruitment of children in the Danish High Risk and Resilience Study - VIA 7. FHR-SZ – children of parents with schizophrenia spectrum psychosis, FHR-BP – children of parents with bipolar disorder, PBC – population-based control children of parents with no diagnoses of schizophrenia spectrum psychosis or bipolar disorder, K-SADS-PL – Schedule for Affective Disorder and Schizophrenia for School-Age Children - Present and Lifetime Version, CGAS – Children's Global Assessment Scale, CBCL – Child Behavior Checklist school-age version, TRF – Teacher's Report Form, TOF – Test Observation Form, mADHD-RS – ADHD-Rating Scale, modified version, STAI-CH – State-Trait Anxiety Inventory for Children

FHR-SZ children scored significantly higher than controls on all psychopathology scales and subscales except for the TOF Anxious subscale (Table 3; Figures 2 and 3). FHR-BP children scored significantly higher compared with controls on several psychopathology scales and subscales. However, there were no significant differences in mean scores between FHR-BP children and controls on any of the TOF subscales (Table 3; Figures 2 and 3).

FHR-SZ children had significantly higher mean scores on all the subscales of both the caregiver and teacher version of mADHD-RS compared with controls, reflecting higher levels of ADHD and oppositional defiant symptoms (Figure 4). FHR-BP children had significantly higher mean scores compared with controls on all subscales of the caregiver version of mADHD-RS except for the Hyperactivity/Impulsivity subscale. FHR-BP children and controls did not differ on the subscales of the teacher version of mADHD-RS, although the difference on the Inattention subscale and the subscale of oppositional defiant disorder problems showed a trend towards significance (Figure 4).

DISCUSSION

Main findings

The Danish High Risk and Resilience Study - VIA 7 is a nationwide cohort study of 522 seven-year-old children. It is the only population-based, representative familial high risk study, and it is the largest clinical study to date assessing psychopathology in children of parents with schizophrenia and bipolar disorder compared with controls.

We found that FHR-SZ and FHR-BP children have an equally higher prevalence of a broad spectrum of lifetime DSM-IV psychiatric diagnoses – e.g., anxiety disorders, and stress and adjustment disorders – compared with controls. Further, we found a gradient in levels of unspecific dimensional psychopathology and daily functioning between the groups, with FHR-SZ children being the most affected and controls being the least affected, whereas FHR-BP children displayed intermediate levels of psychopathology and functioning.

Table 1 Characteristics of children participating with data on psychopathology in the Danish High Risk and Resilience Study - VIA 7 and their biological parents

					Pairwise comparisons		
	FHR-SZ	FHR-BP	PBC	p	FHR-SZ vs. PBC	FHR-BP vs. PBC	FHR-BP vs. FHR-SZ
Children (N=519)	(N=200)	(N=119)	(N=200)	-	-	-	-
Female, N (%)	92 (46.0)	55 (46.2)	93 (46.5)	0.995	-	-	-
Age at inclusion, years, mean±SD	7.8 ± 0.2	7.9 ± 0.2	7.8 ± 0.2	0.096	-	-	-
Two ill parents, N (%)	8 (4.0)	1 (0.8)	-	-	-	-	-
Child's home environment							
Living with both biological parents, N (%)	80 (40.0)	62 (52.1)	169 (84.5)	<0.0001	<0.0001	<0.0001	0.035
Living out of home, N (%)	11 (5.5)	0 (0.0)	1 (0.5)	<0.001	0.003	0.440	0.009
Living with index parent, N (%)	122 (61.0)	83 (69.7)	189 (94.5)	<0.0001	<0.0001	<0.0001	0.115
Living with a single parent, N (%)	75 (37.5)	39 (32.8)	21 (10.6)	<0.0001	<0.0001	<0.0001	0.394
PSP primary caregiver, mean±SD	73.1 ± 14.0	74.5 ± 14.1	84.4 ± 9.1	<0.0001	<0.0001	<0.0001	0.346
Index parents (N=517)	(N=198)	(N=115)	(N=204)	-	-	-	-
Female, N (%)	110 (55.6)	63 (54.8)	115 (56.4)	0.962	-	-	-
Age at child's birth, years, mean±SD	30.1 ± 6.0	33.1 ± 7.0	32.8 ± 4.8	<0.0001	<0.0001	0.673	<0.0001
PSP, mean±SD	66.3 ± 15.6	68.9 ± 14.1	84.3 ± 9.9	<0.0001	<0.0001	<0.0001	0.115
Employed or studying, N (%)	92 (49.5)	60 (55.6)	185 (92.0)	<0.0001	<0.0001	<0.0001	0.313
Education							
Primary/lower secondary, N (%)	54 (30.5)	10 (9.3)	8 (4.1)				
Upper secondary, vocational, short-cycle tertiary, N (%)	75 (42.4)	44 (40.7)	95 (48.2)	<0.0001	<0.0001	0.930	<0.0001
Bachelor degree, equivalent or higher, N (%)	48 (27.1)	54 (50.0)	94 (47.7)				
Biological non-index parents (N=489)	(N=184)	(N=113)	(N=192)				
Female, N (%)	81 (44.0)	51 (45.1)	83 (43.2)	0.949	-	-	-
Age at child's birth, years, mean±SD	30.9 ± 6.4	33.1 ± 5.4	33.0 ± 4.3	<0.001	<0.001	0.856	<0.001
PSP, mean±SD	76.4 ± 14.3	81.8 ± 13.1	85.5 ± 8.4	<0.0001	<0.0001	0.013	<0.001
Employed or studying, N (%)	133 (75.6)	93 (86.1)	179 (95.2)	<0.0001	<0.0001	0.006	0.032
Education							
Primary/lower secondary, N (%)	30 (17.1)	5 (4.8)	10 (5.3)				
Upper secondary, vocational, short-cycle tertiary, N (%)	86 (49.1)	44 (41.9)	89 (47.6)	0.002	0.002	0.310	<0.001
Bachelor degree, equivalent or higher, N (%)	59 (33.7)	56 (53.3)	88 (47.1)				

Index parents refer to the biological parents with a diagnosis of schizophrenia spectrum psychosis or bipolar disorder. FHR-SZ – children with familial high risk for schizophrenia spectrum psychosis, FHR-BP – children with familial high risk for bipolar disorder, PBC – population-based controls, PSP – Personal and Social Performance Scale

Specificity of psychopathology in familial high risk children

Our findings of an elevated prevalence of psychiatric diagnoses and dimensional psychopathology in FHR-SZ and FHR-BP children are consistent with the results of earlier familial high risk studies[7-10,12,16,17,19,29,30]. Overall, both familial high risk groups in our study presented with a broad range, i.e. unspecific, categorical and dimensional psychopathology at this young age. Depressive disorders were rare in both groups, mania was absent, and only two FHR-SZ children were diagnosed with psychotic disorder not otherwise specified.

We found elevated rates of anxiety disorders as well as stress and adjustment disorders in both familial high risk groups. This is in accordance with earlier reports of anxiety disorders being common in FHR-BP children[44]. The findings support the first step of the clinical staging model suggested by Duffy et al[29], implying that anxiety and sleep disorders in childhood,

Table 2 Lifetime prevalence of DSM-IV Axis I disorders in offspring of parents with schizophrenia or bipolar disorder compared with population-based controls

	FHR-SZ (N=199)		FHR-BP (N=118)		PBC (N=197)
	N (%)	OR (95% CI)	N (%)	OR (95% CI)	N (%)
Any Axis I disorder	108 (54.3%)	2.0 (1.4-3.1)***	64 (54.2%)	2.0 (1.3-3.3)**	73 (37.1%)
Any Axis I disorder, excluding elimination disorders	77 (38.7%)	3.5 (2.2-5.7)***	42 (35.6%)	3.1 (1.8-5.3)***	30 (15.2%)
Two or more Axis I disorder, excluding elimination disorders	28 (14.1%)	4.4 (1.9-10.4)***	17 (14.4%)	4.6 (1.8-11.4)**	7 (3.6%)
Affective disorders	3 (1.5%)	-	5 (4.2%)	-	2 (1.0%)
Psychotic disorder NOS	2 (1.0%)	-	0	-	0
Anxiety disorders	23 (11.6%)	2.8 (1.2-6.1)*	14 (11.9%)	2.8 (1.2-6.8)*	9 (4.6%)
Disruptive behavior disorders	12 (6.0%)	6.4 (1.4-29.2)*	4 (3.4%)	3.5 (0.6-19.5)	2 (1.0%)
ADHD	41 (20.6%)	3.5 (1.8-6.6) ***	11 (9.3%)	1.4 (0.6-3.1)	14 (7.1%)
Pervasive developmental disorders	12 (6.0%)	2.5 (0.9-7.2)	9 (7.6%)	3.2 (1.0-9.9)*	5 (2.5%)
Post-traumatic stress disorder	4 (2.0%)	-	3 (2.5%)	-	0
Stress and adjustment disorders	11 (5.5%)	3.8 (1.0-13.8)*	10 (8.5%)	6.0 (1.6-22.2)**	3 (1.5%)
Tic disorders	7 (3.5%)	-	2 (1.7%)	-	3 (1.5%)
Elimination disorders	53 (26.6%)	1.0 (0.6-1.5)	38 (32.2%)	1.3 (0.8-2.1)	54 (27.4%)

FHR-SZ – children with familial high risk for schizophrenia spectrum psychosis, FHR-BP – children with familial high risk for bipolar disorder, PBC – population-based controls, OR – odds ratio, CI – confidence interval, NOS – not otherwise specified, ADHD – attention-deficit/hyperactivity disorder
*$p < 0.05$, **$p < 0.01$,***$p < 0.001$

as well as adjustment, mood and substance use disorders in adolescence, could represent early precursors of bipolar disorder in the offspring of parents with that disorder.

Rates of psychopathology in FHR-BP children have varied substantially in previous studies. This may be attributed to differences in parents' severity of illness, procedures for assessing offspring diagnoses, and age of the offspring[23,45]. Parents with bipolar disorder have often been recruited through inpatient and outpatient clinics, whereas they were identified through Danish registers in this study. Therefore, the group of parents in our study was likely to be more heterogeneous in terms of severity of the disorders, which may explain the lower levels of psychopathology in FHR-BP children compared with other familial high risk studies of bipolar disorder. Indeed, our findings are in line with the Dutch Bipolar Offspring Study, where most parents were recruited through a patient advocacy group[23].

Differences in psychopathological presentation between the two high risk groups

Even though evidence of the shared genetic risk factors for schizophrenia and bipolar disorder is robust, knowledge concerning common or distinct developmental psychopathology is still lacking[5]. Our findings showed that FHR-SZ and FHR-BP children both present an elevated prevalence of unspecific categorical and dimensional psychopathology, even though FHR-BP children differed less from controls than did FHR-SZ children.

Also, FHR-SZ children consistently displayed elevated levels of behavioral problems across settings, namely at home, at school and during the test session, as rated by several informants. In contrast, even though parents of FHR-BP children reported a high prevalence of behavioral and emotional problems compared with controls, teachers reported less deviation from controls and the investigators observed levels of problems equal to those of controls.

Both high risk groups had an elevated prevalence of anxiety as well as stress and adjustment disorders. FHR-BP children displayed a significantly elevated prevalence of pervasive developmental disorders compared with controls, whereas the elevated prevalence in FHR-SZ children did not reach significance. Only FHR-SZ children had an elevated prevalence of ADHD and disruptive behavior disorders compared with controls. Thus, even though both high risk groups show elevated levels of unspecific psychopathology, there are also differences between their psychopathological profiles.

ADHD and disruptive behavior disorders in familial high risk children

We found significantly higher levels of ADHD and disruptive behavior disorders in FHR-SZ children compared with controls, which is in line with findings of impaired attention and disruptive behaviors in previous studies[8-10,12,46]. However, earlier studies have reported conflicting results on ADHD and disruptive behavior disorders in FHR-BP children[19,47]. In particular, Duffy et al[29] suggested that ADHD only precedes bipolar disorder in offspring of bipolar parents who do not respond to lithium treatment.

We did not find a higher prevalence of diagnoses of ADHD and disruptive behavior disorders in FHR-BP children at this

Table 3 Estimated means and percentage differences adjusted for child's gender between familial high risk groups on CBCL, TRF, TOF, STAI-CH and CGAS total and broad-band scores

	FHR-SZ		FHR-BP		PBC		Estimated differences in percentage (95% CI)		
Test (informant)	**N**	**Mean (95% CI)**	**N**	**Mean (95% CI)**	**N**	**Mean (95% CI)**	**FHR-SZ vs. PBC**	**FHR-BP vs. PBC**	**FHR-SZ vs. FHR-BP**
CBCL (caregiver)									
Total	192	27.2 (24.4-30.3)	111	23.4 (20.2-27.0)	191	17.0 (15.1-19.1)	59.9% (36.4-87.5)****	37.6% (14.2-65.8)***	16.2% (– 3.0 to 39.2)
Internalizing	194	6.6 (5.9-7.4)	110	6.6 (5.7-7.7)	191	4.9 (4.3-5.5)	35.4% (13.8-61.2)***	36.0% (11.2-66.3)**	−0.4% (– 17.9 to 20.9)
Externalizing	193	7.8 (6.8-8.8)	111	6.1 (5.1-7.3)	191	4.1 (3.5-4.8)	90.4% (56.1-132.3)****	50.9% (19.3-90.9)***	26.2% (1.5-56.9)*
TRF (teacher)									
Total	167	26.2 (22.7-30.2)	103	20.0 (16.5-24.2)	166	14.7 (12.5-17.2)	78.3% (43.9-120.9)****	36.2% (6.2-74.7)*	30.9% (3.2-66.1)*
Internalizing	168	5.7 (4.9-6.6)	103	5.5 (4.6-6.7)	167	3.7 (3.1-4.3)	56.0% (25.1-94.7)****	50.6% (17.1-93.6)**	3.6% (– 18.3 to 31.5)
Externalizing	168	6.5 (5.3-7.9)	103	4.5 (3.4-5.9)	167	3.0 (2.4-3.8)	113.3% (56.2-191.4)****	47.3% (1.9-112.8)*	44.9% (2.7-104.3)*
TOF (tester)									
Total	194	34.9 (30.7-39.7)	116	24.9 (20.9-29.8)	190	25.0 (21.8-28.7)	39.4% (15.6-68.1)***	−0.4% (–20.3 to 24.5)	39.9% (12.6-73.8)**
Internalizing	194	7.6 (6.5-8.9)	116	5.7 (4.6-7.1)	190	4.9 (4.1-5.9)	53.7% (21.8-93.9)***	15.3% (–12.7 to 52.1)	33.4% (2.5-73.6)*
Externalizing	194	13.2 (11.2-15.6)	116	8.0 (6.3-10.1)	190	9.0 (7.5-10.9)	46.4% (14.5-87.0)**	−12.0% (–34.9 to 19.0)	66.2% (24.3-122.4)***
STAI-CH (child)									
State-Anxiety	193	8.1 (7.5-8.8)	117	7.2 (6.5-7.9)	195	6.9 (6.4-7.5)	17.2% (5.2-30.6)**	3.2% (– 9.2 to 17.3)	13.6% (0.2-28.7)*
Trait-Anxiety	190	12.6 (11.5-13.7)	116	12.2 (10.9-13.6)	194	10.4 (9.5-11.4)	20.9% (6.9-36.6)**	17.4% (2.0-35.1)*	3.0% (– 10.3 to 18.2)
CGAS	199	68.2 (66.3-70.2)	118	73.7 (71.2-76.3)	197	77.9 (75.9-79.9)	−12.4% (–15.7 to −8.9)****	−5.4% (–9.4 to – 1.2)*	−7.4% (–11.5 to – 3.1)***

FHR-SZ – children with familial high risk for schizophrenia spectrum psychosis, FHR-BP – children with familial high risk for bipolar disorder, PBC – population-based controls, CI – confidence interval, CBCL – Child Behavior Checklist school-age version, TRF – Teacher's Report Form, TOF – Test Observation Form, STAI-CH – State-Trait Anxiety Inventory for Children, CGAS – Children's Global Assessment Scale

*$p < 0.05$, **$p < 0.01$,***$p < 0.001$,****$p < 0.0001$

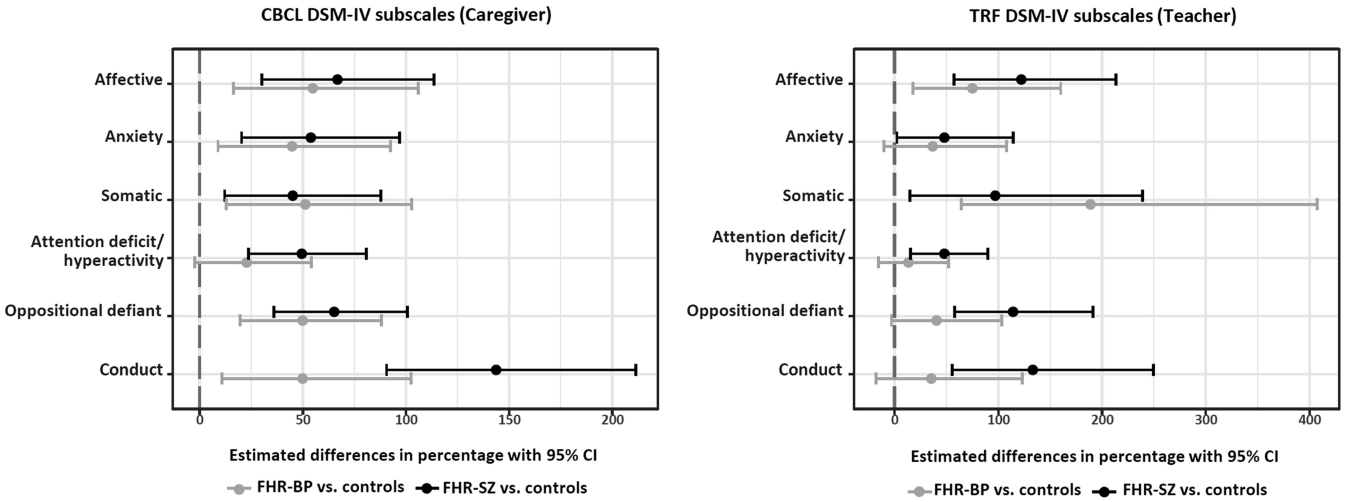

Figure 2 Percentage differences in mean scores of subscales of the Child Behavior Checklist (CBCL) and the Teacher's Report Form (TRF). The population-based control group is set as reference (the vertical dashed line). FHR-SZ – children with familial high risk for schizophrenia spectrum psychosis, FHR-BP – children with familial high risk for bipolar disorder

early age compared with controls. Further, among children with a diagnosis of ADHD, those from the FHR-BP group most often had the predominantly inattentive type, whereas children from the FHR-SZ and control groups most often had the combined or predominantly hyperactive-impulsive type. Likewise, in the dimensional measures, we only found some evidence of elevated symptoms of ADHD and disruptive behavior disorders in FHR-BP children, in the form of elevated scores on the Inattention subscale and the subscale of oppositional defiant disorder problems of the caregiver version of mADHD-RS. Detection of inattention in a classroom setting may be more challenging than the observation of hyperactivity and impulsivity, which may explain why the difference between the FHR-BP group and controls only showed a trend towards significance in teachers' ratings of inattention.

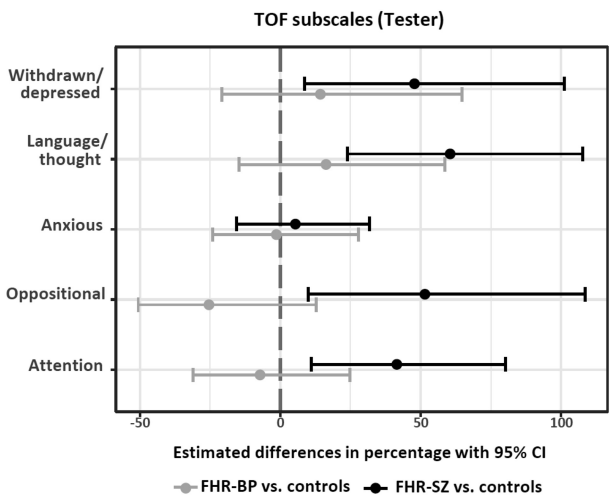

Figure 3 Percentage differences in mean scores of subscales of the Test Observation Form (TOF). The population-based control group is set as reference (the vertical dashed line). FHR-SZ – children with familial high risk for schizophrenia spectrum psychosis, FHR-BP – children with familial high risk for bipolar disorder

Strengths and limitations

An important strength of this study is the use of Danish national registers to recruit the families, which contributes to the high representativeness of this large nationwide cohort.

The narrow age range of the children is also a major strength of the study, since the prevalence and nature of psychopathological disorders and symptoms are highly age-dependent. The prevalence of psychopathology could be compared between the study groups with higher precision and power.

Psychopathology was evaluated both categorically and dimensionally with state-of-the-art assessment instruments through multiple informants in different settings. This provided a comprehensive understanding of the children's psychopathology in different contexts.

Another major strength of the study is the inclusion of FHR-SZ and FHR-BP children in the same study, which allowed to explore possible shared and different antecedents between these groups.

This study also has some limitations. The FHR-BP group consisted of only 120 children. Some of the non-significant differences between FHR-BP and controls may thus be due to an insufficient statistical power. However, the FHR-BP group scored lower than the FHR-SZ group on most psychopathology scales, which is more likely the reason why the latter group differed significantly from controls on more scales than did the former one.

Some studies have suggested that parental mood influences the parental reports on children's psychopathology, although results have been conflicting[48]. This could potentially explain

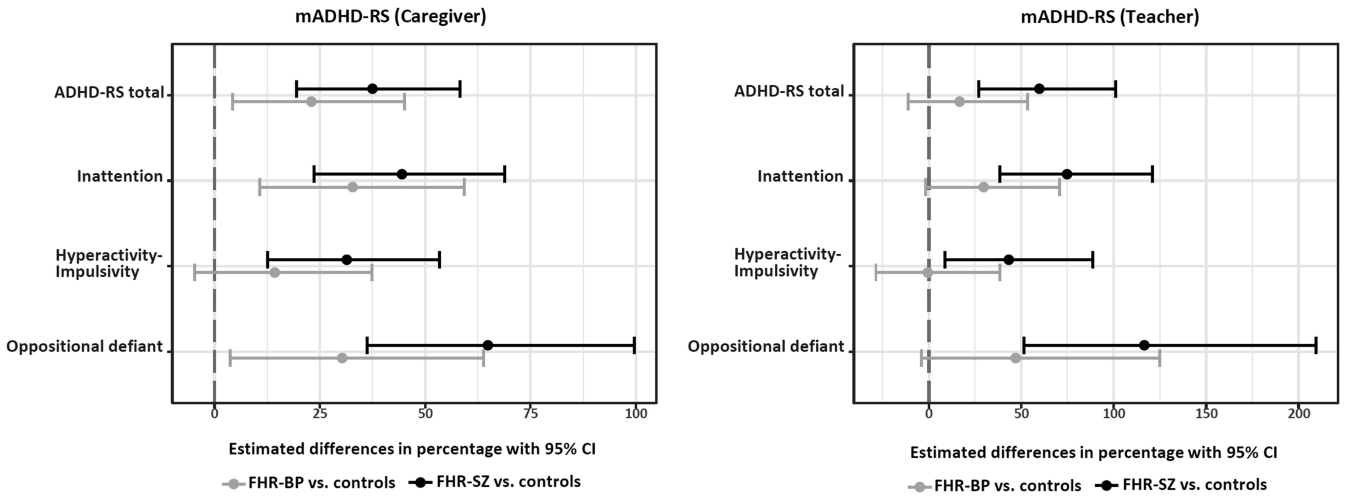

Figure 4 Percentage differences in mean scores of subscales of the modified version of the ADHD-Rating Scale (mADHD-RS). The population-based control group is set as reference (the vertical dashed line). ADHD-RS total – sum score of the Inattention and the Hyperactivity-Impulsivity subscales, FHR-SZ – children with familial high risk for schizophrenia spectrum psychosis, FHR-BP – children with familial high risk for bipolar disorder

why parents from the FHR-BP group reported more dimensional psychopathology than teachers and investigators.

As these results are from the first wave of assessments, we cannot determine whether the high rates of psychopathology found in these children are a transient phenomenon or rather a part of different trajectories towards more severe illnesses. We need to monitor the prevalence of psychopathological symptoms in the familial high risk groups over time, and explore if they may predict schizophrenia or bipolar disorder later in life. Also, follow-up studies are needed to identify resilience factors that can protect children with psychopathology from developing severe mental illness.

Implications

Children from the familial high risk groups displayed significantly more dimensional psychopathology and psychiatric disorders compared with controls. The finding of high levels of psychopathology at this early age in FHR-SZ and FHR-BP children could have implications for school performance, peer relations and other important developmental aspects. A preventive strategy could be to offer these children and their families special and enhanced attention and support from teachers and health care professionals. Also, our findings highlight the need to strengthen the collaboration between adult and child psychiatry in the treatment of these families.

Furthermore, longitudinal familial high risk studies are needed to identify which psychopathological symptoms predict conversion to severe mental disorders in FHR-SZ and FHR-BP children and which resilience factors help these children compensate and protect them from conversion. The next wave of assessment of this cohort at age 11 began in March 2017 and is called the Danish High Risk and Resilience Study - VIA 11.

Finally, our findings emphasize the need for clinical trials of primary interventions towards this vulnerable group of children to prevent their unspecific psychopathological symptoms from converting into severe mental disorders and to increase their daily level of functioning.

At this stage, we cannot determine whether the signs and symptoms of psychopathology found in these children at familial high risk represent transitory states that they will eventually grow out of or antecedents of more severe disorders. However, we can assert that some of these children have symptoms which impair their current level of functioning and call for interventions to support their healthy development.

ACKNOWLEDGEMENTS

This work was supported by the Mental Health Services of the Capital Region of Denmark, the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), Aarhus University, the Tryg Foundation and the Beatrice Surovell Haskell Fund for Child Mental Health Research of Copenhagen. The authors would like to express their gratitude to the dedicated families participating in the study; to M. Skjærbæk, A. Ranning, H. Jensen, M. Melau, C. Gregersen, H. Stadsgaard, K. Kold Zahle and M. Toft Henriksen for contributing to data collection; to C. Bøcker Pedersen and M. Giørtz Pedersen for retrieving the register extract; to M. Chaine and J. Ohland for help with data management; and to P.B. Mortensen, T. Werge, D. Hougaard and A. Børglum for collaboration in iPSYCH.

REFERENCES

1. Fusar-Poli P, McGorry PD, Kane JM. Improving outcomes of first-episode psychosis: an overview. World Psychiatry 2017;16:251-65.
2. Yung AR. Treatment of people at ultra-high risk for psychosis. World Psychiatry 2017;16:207-8.
3. Rapoport JL, Giedd JN, Gogtay N. Neurodevelopmental model of schizophrenia: update 2012. Mol Psychiatry 2012;17:1228-38.
4. Owen MJ, O'Donovan MC. Schizophrenia and the neurodevelopmental continuum: evidence from genomics. World Psychiatry 2017;16:227-35.

5. Laurens KR, Luo L, Matheson SL et al. Common or distinct pathways to psychosis? A systematic review of evidence from prospective studies for developmental risk factors and antecedents of the schizophrenia spectrum disorders and affective psychoses. BMC Psychiatry 2015;15:205.
6. Dean K, Stevens H, Mortensen PB et al. Full spectrum of psychiatric outcomes among offspring with parental history of mental disorder. Arch Gen Psychiatry 2010;67:822-9.
7. Rasic D, Hajek T, Alda M et al. Risk of mental illness in offspring of parents with schizophrenia, bipolar disorder, and major depressive disorder: a meta-analysis of family high-risk studies. Schizophr Bull 2014;40:28-38.
8. de la Serna E, Baeza I, Andrés S et al. Comparison between young siblings and offspring of subjects with schizophrenia: clinical and neuropsychological characteristics. Schizophr Res 2011;131:35-42.
9. Keshavan M, Montrose DM, Rajarethinam R et al. Psychopathology among offspring of parents with schizophrenia: relationship to premorbid impairments. Schizophr Res 2008;103:114-20.
10. Ross RG, Compagnon N. Diagnosis and treatment of psychiatric disorders in children with a schizophrenic parent. Schizophr Res 2001;50:121-9.
11. Hans SL, Auerbach JG, Styr B et al. Offspring of parents with schizophrenia: mental disorders during childhood and adolescence. Schizophr Bull 2004;30:303-15.
12. Sanchez-Gistau V, Romero S, Moreno D et al. Psychiatric disorders in child and adolescent offspring of patients with schizophrenia and bipolar disorder: a controlled study. Schizophr Res 2015;168:197-203.
13. Donatelli JA, Seidman LJ, Goldstein JM et al. Children of parents with affective and nonaffective psychoses: a longitudinal study of behavior problems. Am J Psychiatry 2010;167:1331-8.
14. Dworkin RH, Green SR, Small NE et al. Positive and negative symptoms and social competence in adolescents at risk for schizophrenia and affective disorder. Am J Psychiatry 1990;147:1234-6.
15. Niemi LT, Suvisaari JM, Tuulio-Henriksson A et al. Childhood developmental abnormalities in schizophrenia: evidence from high-risk studies. Schizophr Res 2003;60:239-58.
16. Birmaher B, Axelson D, Monk K et al. Lifetime psychiatric disorders in school-aged offspring of parents with bipolar disorder: the Pittsburgh Bipolar Offspring study. Arch Gen Psychiatry 2009;66:287-96.
17. Chang K, Steiner H, Ketter T. Studies of offspring of parents with bipolar disorder. Am J Med Genet 2003;123C:26-35.
18. Henin A, Biederman J, Mick E et al. Psychopathology in the offspring of parents with bipolar disorder: a controlled study. Biol Psychiatry 2005;58: 554-61.
19. Vandeleur C, Rothen S, Gholam-Rezaee M et al. Mental disorders in offspring of parents with bipolar and major depressive disorders. Bipolar Disord 2012;14:641-53.
20. Frías Á, Palma C, Farriols N et al. Characterizing offspring of bipolar parents: a review of the literature. Actas Esp Psiquiatr 2015;43:221-34.
21. Singh MK, DelBello MP, Stanford KE et al. Psychopathology in children of bipolar parents. J Affect Disord 2007;102:131-6.
22. Morón-Nozaleda MG, Díaz-Caneja CM, Rodríguez-Toscano E et al. A developmental approach to dimensional expression of psychopathology in child and adolescent offspring of parents with bipolar disorder. Eur Child Adolesc Psychiatry 2017;26:1165-75.
23. Mesman E, Birmaher BB, Goldstein BI et al. Categorical and dimensional psychopathology in Dutch and US offspring of parents with bipolar disorder: a preliminary cross-national comparison. J Affect Disord 2016;205:95-102.
24. Diler RS, Birmaher B, Axelson D et al. Dimensional psychopathology in offspring of parents with bipolar disorder. Bipolar Disord 2011;13:670-8.
25. Giles LL, DelBello MP, Stanford KE et al. Child behavior checklist profiles of children and adolescents with and at high risk for developing bipolar disorder. Child Psychiatry Hum Dev 2007;38:47-55.
26. Dienes KA, Chang KD, Blasey CM et al. Characterization of children of bipolar parents by parent report CBCL. J Psychiatr Res 2002;36:337-45.
27. Egeland JA, Shaw JA, Endicott J et al. Prospective study of prodromal features for bipolarity in well Amish children. J Am Acad Child Adolesc Psychiatry 2003;42:786-96.
28. Maoz H, Goldstein T, Axelson DA et al. Dimensional psychopathology in preschool offspring of parents with bipolar disorder. J Child Psychol Psychiatry 2014;55:144-53.
29. Duffy A, Horrocks J, Doucette S et al. The developmental trajectory of bipolar disorder. Br J Psychiatry 2014;204:122-8.
30. Goetz M, Sebela A, Mohaplova M et al. Psychiatric disorders and quality of life in the offspring of parents with bipolar disorder. J Child Adolesc Psychopharmacol 2017;27:483-93.
31. Maziade M, Gingras N, Rouleau N et al. Clinical diagnoses in young offspring from eastern Québec multigenerational families densely affected by schizophrenia or bipolar disorder. Acta Psychiatr Scand 2008;117:118-26.
32. Thorup AAE, Jepsen JR, Ellersgaard DV et al. The Danish High Risk and Resilience Study - VIA 7 - a cohort study of 520 7-year-old children born of parents diagnosed with either schizophrenia, bipolar disorder or neither of these two mental disorders. BMC Psychiatry 2015;15:233.
33. Pedersen CB. The Danish Civil Registration System. Scand J Publ Health 2011;39(Suppl. 7):22-5.
34. Mors O, Perto GP, Mortensen PB. The Danish Psychiatric Central Research Register. Scand J Publ Health 2011;39(Suppl. 7):54-7.
35. Kaufman J, Birmaher B, Brent D et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children - Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 1997;36:980-8.
36. Shaffer D. A Children's Global Assessment Scale (CGAS). Arch Gen Psychiatry 1983;40:1228.
37. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms & Profiles. Burlington: University of Vermont, Research Center for Children, Youth, & Families, 2001.
38. Makransky G, Bilenberg N. Psychometric properties of the parent and teacher ADHD Rating Scale (ADHD-RS): measurement invariance across gender, age, and informant. Assessment 2014;21:694-705.
39. DuPaul G, Power TJ, Anastopoulos A et al. ADHD Rating Scale-IV. New York: Guilford, 1998.
40. Barkley R, Gwenyth EH, Arthur LR. Defiant teens. A clinician's manual for assessment and family intervention. New York: Guilford, 1999.
41. McConaughy SH, Achenbach TM. Manual for the Test Observation Form for ages 2-18. Burlington: University of Vermont, Center for Children, Youth, & Families, 2004.
42. Spielberger CD, Edwards CD, Lushene R et al. State-Trait Anxiety Inventory for Children: sampler set, manual, test booklet, scoring key. Palo Alto: Mind Garden, 1973.
43. Zapf A, Castell S, Morawietz L et al. Measuring inter-rater reliability for nominal data - which coefficients and confidence intervals are appropriate? BMC Med Res Methodol 2016;16:93.
44. Duffy A, Horrocks J, Doucette S et al. Childhood anxiety: an early predictor of mood disorders in offspring of bipolar parents. J Affect Disord 2013;150: 363-9.
45. Duffy A, Doucette S, Lewitzka U et al. Findings from bipolar offspring studies: methodology matters. Early Interv Psychiatry 2011;5:181-91.
46. Keshavan MS, Sujata M, Mehra A et al. Psychosis proneness and ADHD in young relatives of schizophrenia patients. Schizophr Res 2003;59:85-92.
47. Duffy A. The nature of the association between childhood ADHD and the development of bipolar disorder: a review of prospective high-risk studies. Am J Psychiatry 2012;169:1247-55.
48. Maoz H, Goldstein T, Goldstein BI et al. The effects of parental mood on reports of their children's psychopathology. J Am Acad Child Adolesc Psychiatry 2014;53:1111-22.e5.

DOI:10.1002/wps.20527

Self and schizophrenia: current status and diagnostic implications

The notion of self-disorders in schizophrenia emerged in contemporary psychiatry at the beginning of this millennium[1]. It was considered an unorthodox novelty, and neither the DSM-IV nor the DSM-5 contains a reference to disordered self in the schizophrenia spectrum.

However, that notion was historically co-constitutive of the concept of schizophrenia. Bleuler[2] listed experiential "ego-disorders" among the fundamental symptoms of schizophrenia and reported patients complaining of being only "reflections of themselves", unable to "catch up with themselves" or having "lost their individual self". All classic texts on schizophrenia contain a reference to disordered self[1]. The concept of "disintegration", widely used in psychiatry and psychoanalysis, makes only sense in the presence of some notion of self that is at stake.

The DSM-III glossary of terms linked disturbance in the "sense of self" to schizophrenia, and the ICD-9 definition of schizophrenia referred to disturbance of fundamental features of personality (e.g., uniqueness and autonomy), whereas in the ICD-10 the term "personality" was removed. The disappearance of "disordered self" was perhaps linked to the simplification of fundamental symptoms into the so-called "four A's" (autism, ambivalence, association and affect disorders) and a difficulty with conceptualizing the notion of autism.

What kind of self is disordered in the schizophrenia spectrum conditions? It is useful to follow a distinction of contemporary philosophy of mind and phenomenology between the so-called "narrative self" and the "core self".

The narrative self refers to features which characterize and individualize a person and which easily lend themselves to linguistic self-description (e.g., "I have a tendency to act impulsively") and descriptions from the third-person perspective ("she is acting impulsively"). These features comprise biographical, characterological and cognitive characteristics and are heavily dependent on language and memory.

The notion of core self refers, instead, to the first-person perspective which is an intrinsic structural feature of all experience and which provides us with an immediate or pre-reflective sense of subjectivity and self-familiarity as an "I-me-myself". This can be extended to comprise a sense of temporal persistency, self-coincidence, substantiality-embodiment, and demarcation. All these features are never an object of ordinary experience, but provide a first-person structure for the narrative level of experiencing oneself as, for example, "impulsive" or "suspicious". However, these features are experientially accessible when we reflect upon the way in which our experience articulates itself.

We have previously proposed that the essential feature of schizophrenia spectrum disorders is a disturbance of the core self in its immediate relation to the world[3]. It is important to emphasize that we are not talking about a lack or a deficit (as in "too much or too little") but rather of an instability or *dis*-order[4]. This basic disturbance of self-world relation is the generative component in the Gestalt of autism[3], which "appears nowhere else in this particular fashion"[2] and which imbues schizophrenia with an air of un-understandability[5].

Empirical studies[1] from different groups and on different samples clearly show a selective hyper-aggregation of disorders of core self in schizophrenia and schizotypal disorder as opposed to bipolar disorder and other psychiatric disorders. Self-disorders typically begin in childhood or adolescence, are observed in populations at ultra-high-risk for psychosis, and predict subsequent schizophrenia spectrum outcome[1].

Two studies have demonstrated temporal persistence and similarity of patterns of self-disorders five years apart[6]. Self-disorders are unrelated to IQ[1], and preliminary data fail to show any substantial correlation with neurocognitive disorders. In sum, empirical research seems to corroborate Bleuler's idea that these phenomena are to be considered as trait features of the schizophrenia spectrum.

This structural instability of self-world relation is the background for the development of psychotic symptoms, which in their form contain an imprint of disordered selfhood[4,7]. For example, the characteristic auditory verbal hallucinations are often a progression from the state of anonymization and spatialization of thinking, where the patient's "I think" becomes transformed into "it thinks in me". The phenomenon of thought broadcasting is a flamboyant expression of the loss of sense of demarcation. And the characteristic double-bookkeeping involves a construction of a private world or alternative ontological framework[7,8].

The recognition of self-disorders entails important nosological consequences. Currently, we see a decrease in the diagnosis of disorganized schizophrenia, a very uncommon use of the schizotypal diagnosis and an increasing frequency in the use of the borderline personality disorder diagnosis. This latter diagnosis is over-inclusive and often applied to patients which would in the ICD-9 be diagnosed with a schizophrenia spectrum condition[9]. It seems to us that it is nearly impossible to conceptualize a core psychopathological difference between the notion of schizotypy and the contemporary clinical application of the DSM-5 diagnosis of borderline personality disorder[9].

This diagnostic confusion is multidetermined, but mostly due to a very tolerant use of the ninth borderline disorder criterion ("transient, stress-related paranoid ideation or severe dissociative symptoms") and the unclarity of the borderline disorder criteria of "identity disturbance" and "feelings of emptiness". "Feelings of emptiness" are undefined, and the identity disturbance criterion, although apparently referring to the narrative level of selfhood, is not sufficiently differentiated from disturbances of core self[10]. We find it crucial to sharpen the distinction between schizophrenia spectrum psychopathology (involving disturbances of both core and narrative self) and disorders of personality (which do not involve structural disturbances of the core self).

Contemporary classification is striving for simplicity and reliability, with much research being performed by for-the-purpose-trained lay interviewers. The disappointment with the slow progress of pathogenetic research encourages critical voices advocating abandonment of phenotypic categories altogether. However, the story of self-disorder research may inspire us to reconsider the phenotypic classification with a more refined psychopathological approach.

Josef Parnas[1,2], Maja Zandersen[1]
[1]Mental Health Centre Glostrup, Institute of Clinical Medicine, University Hospital of Copenhagen, Copenhagen, Denmark; [2]Center for Subjectivity Research, University of Copenhagen, Copenhagen, Denmark

1. Parnas J, Henriksen MG. Harv Rev Psychiatry 2014;22:251-65.
2. Bleuler E. Dementia praecox oder Gruppe der Schizophrenien. Leipzig: Deuticke, 1911.
3. Parnas J. World Psychiatry 2012;11:67-9.
4. Stephensen H, Parnas J. J Phenom Cogn Sci (in press).
5. Jaspers K. Allgemeine Psychopathologie, 1st ed. Berlin: Springer, 1913.
6. Nordgaard J, Nilsson LS, Sæbye D et al. Eur Arch Psychiatry Clin Neurosci (in press).
7. Parnas J, Henriksen MG. Conscious Cogn 2016;43:75-88.
8. Henriksen MG, Parnas J. Schizophr Bull 2013;40:542-7.
9. Zandersen M, Henriksen MG, Parnas J. J Pers Disord (in press).
10. Zandersen M, Parnas J. Schizophr Bull (in press).

DOI:10.1002/wps.20528

The schizophrenia spectrum anhedonia paradox

Anhedonia, defined as a diminished capacity to experience pleasure, has been considered a core symptom of schizophrenia since the earliest descriptions of the disorder. It is longitudinally stable and associated with a range of poor clinical outcomes[1]. Unfortunately, interventions targeting this symptom have produced minimal benefits, and no drug has received US Food and Drug Administration's approval for this indication.

Limited progress in effectively treating anhedonia results in part from a lack of conceptual clarity regarding the nature of the symptom. Evidence for anhedonia in schizophrenia has primarily come from data obtained via clinical interviews, which indicate that the majority of those diagnosed with that disorder are anhedonic. Clinicians have long assumed that such self-reports indicate that individuals with schizophrenia have a diminished capacity to experience positive emotion. However, laboratory-based studies provide evidence that contradicts this notion, indicating that schizophrenia patients self-report as much positive emotion as healthy controls in response to pleasant stimuli[2] and show intact neurophysiological responses in key reward circuits during receipt of rewarding outcomes[3].

It has been argued that this apparent discrepancy can be resolved if one examines the anchors and probes used in negative symptom interviews[4]. Upon careful inspection, it is clear that what interviewers are rating is the frequency of reward-seeking behavior, rather than the extent to which patients enjoy pleasurable activities when engaged in them. Based on this evidence, as well as on results from ecological momentary assessment studies, the field has gradually shifted away from the view that schizophrenia patients have a reduced hedonic capacity. Rather, schizophrenia appears to be associated with a behavioral deficit characterized by a reduction in the frequency of pleasurable activity[4].

The disconnect between behavior and hedonic capacity has been termed the "liking-wanting anhedonia paradox", and spurred research attempting to determine why apparently normal hedonic responses do not translate into motivated behaviors aimed at obtaining rewards in schizophrenia. Several conceptual models attempted to answer this question, proposing that impairments in various aspects of reward processing (e.g., reinforcement learning, value representation, effort-cost computation, reward anticipation), that rely on cortico-striatal circuitry, prevent fully intact hedonic responses from influencing decision-making processes needed to guide action selection and initiate motivated behavior[4]. These models have received significant empirical support and are beginning to influence the development of treatments targeting these underlying mechanisms.

However, there is a second "anhedonia paradox" that has emerged over recent years. We refer to this as the "schizophrenia spectrum anhedonia paradox". Specifically, there is growing evidence that, although patients with schizophrenia have intact hedonic capacity[4], individuals with schizotypy and youth in the prodromal phase of illness do not. People with schizotypy self-report less positive emotion in response to pleasant stimuli than healthy controls and show reduced neurophysiological response during the receipt of reward outcomes[5]. Youth at clinical high risk for psychosis also have diminished neurophysiological and self-reported responses to pleasant stimuli[6]. Since schizophrenia is a more severe form of psychopathology in nearly every conceivable way, this apparent discrepancy is paradoxical: why would the less severe forms of pathology show deficits in hedonic capacity, whereas the more severe form does not? Below we discuss some plausible explanations, hoping to promote future studies aimed at resolving this paradox.

A first possibility is that mood and anxiety symptoms produce diminished hedonic response in schizotypy and clinical high risk youth more than in schizophrenia. Consistent with this notion is evidence indicating that youth at clinical high risk for psychosis and those with schizotypy have higher rates of comorbid depression and anxiety than people with schizophrenia, and that greater severity of depression and anxiety is associated with reduced hedonic response in those individuals[6].

A second possibility is that antipsychotics have a normalizing effect on reward processing. Studies examining the neural response to rewarding stimuli in schizophrenia suggest that second generation antipsychotics are associated with intact response to reward outcomes in the ventral striatum[7]. Since individuals with schizotypy and youth at clinical high risk for psychosis are much less likely to be prescribed antipsychotics, the apparent paradox may reflect medication effects that become evident with more severe pathology.

Third, schizophrenia is associated with more severe cognitive impairment and poorer insight into clinical symptomatology than schizotypy or clinical high risk states. It is possible that impaired cognition and insight are paradoxically protective, causing schizophrenia patients to have less awareness of hedonic deficits that may actually exist. Those with schizotypy and clinical high risk youth may be better able to accurately report their hedonic state because of higher cognitive function and insight.

Fourth, environmental and stress effects may have a greater impact on youth at clinical high risk for psychosis and those with schizotypy. Schizophrenia is associated with impoverished quality of life, which for many patients reflects an environment and daily routine with restricted social, cognitive and affective demands. For individuals with schizotypy and clinical high risk youth, environments and daily routines are generally more complex and stressful. It is possible that this stress attenuates reward system responsivity. Supporting this, individuals with schizotypy seem to enjoy solitary activities, yet report activities with others as being taxing, stressful and unenjoyable[8]. Animal models and studies on humans support the notion of a "stress-induced anhedonia"[9]; however, this phenomenon has yet to be directly investigated in the schizophrenia spectrum.

The schizophrenia spectrum anhedonia paradox harkens back to the seminal writings of P. Meehl[10], who proposed that anhedonia is one of several polygenic potentiators that comprise the endophenotype for schizotaxia. Meehl distinguished between primary and secondary anhedonia. Primary anhedonia refers to one's hedonic capacity. This capacity is polygenically determined and dependent on neurotransmitter function and neural circuitry responsible for reward responsivity. Capacity varies on a continuous dimension, but may be taxonic at the extreme end. It seems that only a small proportion of schizophrenia patients would fall at the extreme end of the continuum, whereas substantially more individuals with schizotypy and clinical high risk youth would display primary anhedonia. Primary anhedonia may be a risk factor for the development of many forms of psychopathology, with schizophrenia being a less common outcome than others.

Meehl also proposed the existence of "secondary anhedonia", which is measured via verbal report (through clinical interview or questionnaire) of one's hedonic response. He proposed that such reports are influenced by "aversive drift", which reflects heightened trait negative affect that becomes increasingly more prominent as the course of illness progresses. The aversive drift construct has yet to receive significant empirical attention; however, there is evidence from laboratory and ecological momentary assessment studies that those on the schizophrenia spectrum display elevated negative emotion that is contextually invariant[2]. Thus, even in instances where the neural machinery for hedonic response is intact, co-activation of negative emotion may bleed over into potentially rewarding situations, lowering the overall net hedonic value of stimuli that would otherwise be rewarding.

Whether secondary anhedonia worsens with illness progression has yet to be determined. However, a greater frequency of secondary anhedonia in the chronic phase of schizophrenia than in earlier phases of illness could be another viable explanation for the schizophrenia spectrum anhedonia paradox.

Gregory P. Strauss[1], Alex S. Cohen[2]
[1]Department of Psychology, University of Georgia, Athens, GA, USA; [2]Department of Psychology, Louisiana State University, Baton Rouge, LA, USA

1. Herbener ES, Harrow M. J Abnorm Psychol 2002;111:237-48.
2. Cohen AS, Minor KS. Schizophr Bull 2010;36:143-50.
3. Radua J, Schmidt A, Borgwardt S et al. JAMA Psychiatry 2015;72:1243-51.
4. Kring AM, Barch DM. Eur Neuropsychopharmacol 2014;24:725-36.
5. Cohen AS, Callaway DA, Najolia GM et al. J Abnorm Psychol 2012;121:407-15.
6. Gruber JM, Strauss GP, Dombrecht L et al. Schizophr Res (in press).
7. Juckel G, Schlagenhauf F, Koslowski M et al. Psychopharmacology 2006;187:222-8.
8. Kwapil TR, Silvia PJ, Myin-Germeys I et al. J Res Pers 2009;43:103-6.
9. Pizzagalli DA. Annu Rev Clin Psychol 2014;10:393-423.
10. Meehl PE. J Abnorm Psychol 2001;110:188-93.

DOI:10.1002/wps.20529

Peer delivered services in mental health care in 2018: infancy or adolescence?

Peer support is now considered to be a central component of the behavioral health care system in countries such as the US, Canada, Australia and the UK. Professionals looking to improve their ability to promote recovery have strategies and training programs that include collaborating with peers in their services (e.g., Boston University's Recovery Promoting Competencies Toolkit).

In 2012, Davidson et al[1] characterized in this journal peer delivered services as being still in their infancy. They pointed out that, while there was a proliferation of peer support work-

ers in the mental health field, their roles and tasks were unclear and the existing research base mostly focused on feasibility studies, often with significant methodological problems. It is our contention that, six years on, the field of peer delivered services has matured significantly.

Recent work has aimed to achieve a common understanding of roles and possible quality indicators for peer services. For example, Chinman et al[2] are developing a peer specialist fidelity measure for two content areas: services provided by peer specialists and factors that support or hamper the performance of those services. Cronise et al[3] conducted a US national survey to identify the roles, tasks, settings, job training and compensation currently offered to those with the title of "peer specialist". They found that peers are no longer just part-time workers in community based settings. Data from 597 respondents revealed that more than 64% worked in full-time positions in a wide range of settings, including treatment and forensic organizations.

The incorporation of peers into the standard mental health workforce has seen the majority hired into roles regarded as unique to a person with personal mental health experience. However, others hold positions in which peers' personal experiences are not required, but considered additive. This includes, for example, rehabilitation workers such as case managers, employment specialists and job coaches.

Agreement about a common set of practitioner competencies (https://www.samhsa.gov/brss-tacs/recovery-support-tools/peers/core-competencies-peer-workers), complementing a set of agreed-upon national guidelines for peer support services in behavioral health in the US (https://inaops.org/national-standards), has begun to emerge since Davidson et al's article.

Peer support has also evolved in Europe, as evidenced by new training programs for peers, including, for example, those identified by the European Union's Compass Consortium[4] as a best practice (e.g., Peer2Peer in Spain), as well as university based programs, such as a two-year Associate Degree program in "Experts by Experience" at Hanze University in the Netherlands.

Functions or processes overlapping across the totality of these efforts suggest that there is growing agreement about some basic qualities unique to peer support (i.e., relationships based on shared lived experience/validation of experiential knowledge and a deliberate focus on enhancing strengths, hope and empowerment, among other qualities).

Recent systematic reviews[5] have confirmed that, while peers and clinicians typically performed fairly equally on traditional outcome measures (e.g., rehospitalization, relapse), peer outcomes were better in areas such as self-efficacy, hope, empowerment, engagement, and others more related to recovery processes.

Responding to earlier criticisms, research on peer support has advanced, with growing examples of well controlled studies. For example, Mahlke et al[6], in their randomized controlled trial (RCT) of 261 peers, clearly indicated tasks to be delivered, specified a standardized training for the peers, and selected peer support workers with similar experiences. The study found that one-to-one peer support plus treatment as usual was associated with significantly higher scores of self-efficacy at six-month follow-up compared to treatment as usual alone.

Manualized interventions created by peers themselves or led by peers have multiplied over the past six years. They have made better controlled studies possible, and thus provided better evidence for outcomes, than the less well defined service of one-to-one peer support. For example, in an RCT, Wellness Recovery Action Planning[7] has been found more effective in reducing psychiatric symptoms, enhancing participant hopefulness and improving quality of life in people with severe and persistent mental illness, as compared to usual care. A peer-led manualized intervention to combat self-stigma was investigated in an initial RCT with positive results[8]. Other programs such as the peer-developed Spanish program Education: Tool to Fight Stigma and Discrimination are now included in European Union's Compass Consortium Best Practices guide[4].

Peer-led health interventions such as self-management, as well as the use of peer navigators to link individuals with mental illnesses to health services, have shown evidence for their effectiveness[9]. A recent RCT of a program named Toward Recovery, Empowerment and Experiential Expertise – TREE, developed by peers in the Netherlands, found the intervention, added to care as usual, to be more effective than care as usual only, in terms of outcomes such as empowerment, mental health confidence and loneliness[10].

Taken together, the number of manualized peer developed/peer led interventions, and the expanding number being studied in RCTs, reflect the emerging sophistication of peer delivered services.

While implementation and methodological issues still exist, the greater clarity around what peer support actually is, the greater variety of available training programs, the current development of a fidelity scale, the suggestion of standardized competencies, better designed RCTs, and the emergence of manualized peer developed/led interventions, are exciting advances in the growth of peer delivered services over the past six years.

These advances justify characterizing peer support services as well beyond their infancy. Rather, they are an established, maturing area of development and study, with great promise for the future of services to promote recovery.

Marianne Farkas[1], Wilma Boevink[2]
[1]Center for Psychiatric Rehabilitation, Boston University, Boston, MA, USA; [2]Trimbos Institute, Utrecht, The Netherlands

M. Farkas acknowledges the support of the US National Institute on Disability, Independent Living and Rehabilitation Research; the Administration for Community Living; the Substance Abuse and Mental Health Services Administration; and the Department of Health and Human Services. The contents of this paper do not necessarily represent the policy of the above entities.

1. Davidson L, Bellamy C, Guy K et al. World Psychiatry 2012;11:123-8.
2. Chinman M, McCarthy S, Mitchell-Miland C et al. Psychiatr Rehabil J 2016;39:256-65.

3. Cronise R, Teixeira C, Rogers ES et al. Psychiatr Rehabil J 2016;39:211-21.
4. European Union Compass Consortium. Good practices in mental health and well being. 2014. https://ec.europa.eu/health/sites/health/files/mental_health/.
5. Bellamy C, Schmutte T, Davidson L. Ment Health Soc Inclusion 2017;21:161-7.
6. Mahlke C, Priebe S, Heumann K et al. Eur Psychiatry 2017;42:103-10.
7. Cook JA, Copeland ME, Jonikas JA et al. Schizophr Bull 2012;38:881-91.
8. Russinova Z, Rogers ES, Gagne C et al. Psychiatr Serv 2014;65:242-6.
9. Cabasa l, Camacho D, Velez Grau C et al. J Psychiatr Res 2016;84:80-9.
10. Boevink W, Kroon H, Van Vugt M et al. Psychosis 2016;8:287-300.

DOI:10.1002/wps.20530

Hoarding disorder has finally arrived, but many challenges lie ahead

In 2010, the DSM-5 Obsessive-Compulsive and Related Disorders Sub-Workgroup recommended the inclusion of hoarding disorder as a new mental disorder in the diagnostic system[1]. Following an expert survey[2], a field trial[3], and a period of public consultation, the new disorder was approved for inclusion in December 2012.

Unlike other proposed changes in DSM-5, the separation of hoarding disorder from obsessive-compulsive disorder (OCD) was met with wide support from both colleagues and patients, who largely felt that the OCD label did not accurately reflect their patients' and their own experiences, respectively.

The uncontroversial acceptance of hoarding disorder can be further ascribed to a number of factors, including the recognition that: a) most patients' symptoms cannot be easily attributable to other mental disorders (including OCD); b) there are a number of important differences between hoarding disorder and OCD with respect to phenomenology of the symptoms, onset and course of the disorder, and neural correlates, among others; c) patients are less likely to respond to evidence-based treatments for OCD[4]; d) hoarding is a prevalent problem affecting persons of both genders and across different cultures[5]; and e) the risk of pathologizing normal behaviour (i.e., normative collecting) is low. The planned inclusion of hoarding disorder in the ICD-11[6] is a welcome addition, which will result in a truly global recognition of this disabling condition.

Individuals with hoarding disorder experience persistent difficulties discarding or parting with possessions, regardless of their actual value. This is due to a perceived need to save the items and distress associated with discarding them. This results in the accumulation of possessions that congest and clutter active living areas and substantially compromise their intended use, causing clinically significant distress or impairment. These symptoms must not be attributable to another physical or mental disorder.

Most people with this disorder excessively acquire items that they do not need or for which no space is available, and typically experience distress if they are unable or are prevented from acquiring items (excessive acquisition specifier). A substantial proportion of sufferers lack insight into their difficulties and are reluctant to seek help for their problems (insight specifier). Other common features of the disorder (not required for diagnosis) include indecisiveness, perfectionism, avoidance, procrastination, difficulty with planning and organizing tasks, and distractibility. Some individuals live in various degrees of unsanitary conditions (*squalor*), that may be a logical consequence of severely cluttered spaces and/or related to planning and organizing difficulties. Persons with the disorder may experience conflicts with neighbours or landlords, and legal proceedings regarding housing evictions or loss of custody of children are not uncommon.

Hoarding disorder affects at least 1.5% of men and women[5]. Most patients usually come to the attention of services when they are in their 50s, but the symptoms may first emerge much earlier, during adolescence. Symptoms typically start interfering with the individual's everyday functioning by the mid-20s, and cause clinically significant impairment by the mid-30s[7]. A progressive worsening of symptoms is typically reported over each decade of life[7]. Once symptoms begin, the course of hoarding is often chronic, with few individuals reporting a waxing and waning course[7]. As expected from a newly recognized disorder, the causes of hoarding disorder are largely unknown, but twin studies suggest that both genetic and environmental risk factors are important[8]. Anecdotal links between material deprivation (e.g., childhood poverty) and hoarding have received no support in the literature.

The diagnosis is usually made on the basis of a direct interview to establish whether the person meets the diagnostic criteria. Because hoarding may not always be the initial reason for consultation, clinicians often need to ask direct questions such as "Do you find it difficult to discard or part with possessions?" or "Do you have a large number of possessions that congest and clutter the main rooms in your home?". A home visit is recommended for the assessment of clutter, impairment, and associated risks. If a home visit is not feasible, the clinician should try to gather additional information from reliable informants, such as a spouse or relative (with the patient's consent). This is particularly important for persons with limited insight, because they may underestimate the extent and consequences of their difficulties. The evaluation should include a thorough risk assessment. Attention should be paid to potential fire hazards, the risk of clutter avalanches, the presence of rodent or insect infestation, and unsanitary living conditions that pose a risk to health. In addition, it is important to establish whether other vulnerable persons (e.g., children, elderly people) live with the person who hoards.

Few treatment studies have specifically included individuals fulfilling DSM-5 criteria for hoarding disorder and, therefore, the evidence to guide treatment choice is incomplete. Currently, the intervention with the strongest evidence base for the disorder is a multicomponent psychological treatment

that is based on a cognitive behavioural model[9]. The intervention includes: office and in-home sessions; motivational interviewing methods to address ambivalence about therapy; education about hoarding; goal-setting; organizing, decision-making and problem-solving skills training; exposure to sorting, discarding, and not acquiring; and cognitive strategies to facilitate this work. This intervention has been evaluated in a few controlled clinical trials with promising results. However, the outcomes are modest and the long-term prognosis unclear[10].

While the official recognition of hoarding disorder as a *bona fide* mental disorder is a huge step in the right direction, numerous challenges lie ahead, some related to the disorder itself and others to the limited research into effective treatments and service development. Some patient-related challenges include that many sufferers have limited insight into their difficulties and they actively or passively resist intervention. Even patients with good insight are deeply ashamed and feel stigmatized, so may still not seek help for their difficulties.

Since the disorder was included in DSM-5, research has been slow. Current treatment options are very limited and only available in a handful of university clinics worldwide. The disorder is frequently underdiagnosed. When correctly diagnosed, colleagues have limited or no referral options. Regular OCD or anxiety disorder clinics are ill-equipped to handle intensive behavioural interventions requiring home visits over extended periods.

These challenges can only be met with substantial investments in research on key strategic areas: prevalence and cost of illness studies; improving detection and reducing stigma; treatment development; service development; and development of legislative frameworks to help reconcile the rights and needs of the patients (who need but may not want help) with those of dependents (e.g., children), neighbours, or landlords who may be adversely affected by the disorder.

David Mataix-Cols, Lorena Fernández de la Cruz
Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

1. Mataix-Cols D, Frost RO, Pertusa A et al. Depress Anxiety 2010;27:556-72.
2. Mataix-Cols D, Fernández de la Cruz L, Nakao T et al. Psychol Med 2011; 41:2475-84.
3. Mataix-Cols D, Billotti D, Fernández de la Cruz L et al. Psychol Med 2013; 43:837-47.
4. Bloch MH, Bartley CA, Zipperer L et al. Mol Psychiatry 2014;19:1025-30.
5. Nordsletten AE, Reichenberg A, Hatch SL et al. Br J Psychiatry 2013;203: 445-52.
6. Stein DJ, Kogan CS, Atmaca M et al. J Affect Disord 2016;190:663-74.
7. Tolin DF, Meunier SA, Frost RO et al. Depress Anxiety 2010;27:829-38.
8. Iervolino AC, Perroud N, Fullana MA et al. Am J Psychiatry 2009;166: 1156-61.
9. Frost RO, Hartl TL. Behav Res Ther 1996;34:341-50.
10. Thompson C, Fernández de la Cruz L, Mataix-Cols D et al. Asian J Psychiatry 2017;27:53-66.

DOI:10.1002/wps.20531

Scaling up psychological treatments for common mental disorders: a call to action

Empirically supported psychological treatments – spanning interpersonal, cognitive and behavioural therapies – are recommended as first-line interventions to address the significant burden of depression, anxiety and stress-related disorders worldwide. Nevertheless, they remain inaccessible for the wide majority of the world's population, both in low- and middle-income countries (LMICs), where less than 5% of people with major depressive disorder receive minimally adequate treatment[1], and in high-income countries (HICs), where the corresponding figure reaches only 20%[1].

This massive treatment gap for such effective treatments is unprecedented in medicine and, as the experience of HICs shows, is not simply a challenge which can be addressed by more mental health care providers. Here we summarize a range of potential strategies through which these treatments might be scaled up to achieve their full potential and reduce the global burden of common mental disorders.

Three major barriers prevent delivery of psychological treatments: the lack of skilled providers, limited access, and low demand for mental health care. Each is an obstacle in most countries, but all have viable, evidence-based solutions.

If we assume that a "skilled" provider is a health care professional who has been trained in one of the mental health disciplines (social work, psychology or psychiatry), then there is no chance of overcoming the first barrier. There are large gaps between the required and actual numbers of mental health professionals in all countries. Furthermore, the methods typically used to train these specialized persons are expensive, time-intensive and requiring of another, even more experienced specialist to conduct regular supervision for an extended period of time.

An effective strategy to address this barrier is "task sharing" or training non-specialist providers – i.e., individuals with no formal training or background in mental health care – to deliver brief, low-intensity psychological treatments. The concept of non-specialist providers originated from para-professional movements in the US and UK. They include nurse practitioners, community health workers, teachers and peers, and are selected because of their availability, low cost, access to and close ties with the population they serve[2]. Not only can non-specialist providers in LMICs be trained to deliver treatments (and as effectively as specialists in HICs[2,3]), but recent evaluations demonstrate that they can ensure high quality of therapy through peer-led supervision[4]. This, in turn, addresses the bottleneck of the need for supervision provided by mental health specialists.

Recent evidence also makes clear that utilizing a core set of common treatment "elements" (such as behavioural activation, exposure, problem solving and communication skills) can reduce the complexity of needing to learn diverse psychological treatment packages for specific clinical phenotypes (such as depressive, anxiety and stress-related disorders). For example, the COBRA trial in the UK demonstrated that non-specialist junior mental health workers with no previous professional training in mental health services successfully delivered a treatment package that focused on the core element of behavioural activation. Results showed equivalent effectiveness in reducing depressive symptom severity as specialists delivering longer courses of cognitive behavioural therapy[5]. Similarly, in India, lay counsellors trained over 3 months to deliver a culturally adapted version of behavioural activation attained improved remission rates and sustained outcomes in primary care attenders with moderately severe to severe depression[6].

The second barrier is limited access to psychological treatments. In most countries, psychological treatments are accessible only to a minority of individuals who can afford private treatment or who are supported by generous insurance programs. Furthermore, these provider-centered treatments are typically delivered face-to-face, in urban specialist facilities, and at a time that is most suitable for the provider. In contrast, evidence-based solutions involve the delivery of psychological treatments in settings and at times that are convenient to the patient (for example, at home and during the weekend). In addition, the use of telemedicine and other digital platforms can facilitate this flexibility as well as guided self-care. Delivery of a treatment through a digital platform may be as effective as in-person treatment, but preferred by the recipient and with better sustained outcomes[7]. Moreover, recent evaluations have demonstrated that therapists can be efficiently trained through digital platforms[8].

In all contexts, these feasible and cost-effective solutions may be particularly beneficial for individuals with limited financial, social or physical capacity to travel to health facilities, such as mothers with infants, individuals with physical disabilities, or people who are homebound for various reasons, including due the impact of mental disorders.

Third, there is a low demand for psychological interventions, particularly from lower social classes and ethnic minorities, and treatment retention of most psychological treatments is less than 50% in most patient populations. The solutions to these problems reflect lessons learned from a community engagement model used for psychosis[9]. There is growing evidence of the benefits of: engaging a "grassroots" perspective when developing and designing mental health services; avoiding biomedical labels and using patients' own explanatory models; targeting social determinants concurrently with psychological symptoms; and engaging the individuals' relationships and resources, including their partner and community at large. Furthermore, a common elements approach is also likely to be

more acceptable, as it is brief, focused, and entails mastering only a limited set of skills.

Despite the growing evidence base supporting these exciting innovations, access to psychological treatments remains an exception. One unique exemplar of scaling up these treatments is the UK's Improving Access to Psychological Treatments (IAPT). IAPT services treat more than 537,000 patients with depression and anxiety annually, train non-specialist providers and specialists with brief accredited courses, and assess the progress of almost all (98%) patients using a unique monitoring outcome system[10]. Their results show that stepped care models of delivery are clinically effective, facilitate short wait times to improve patient attendance, and ultimately increase collaboration between therapists and patients.

In order to integrate and optimize new models of delivery beyond a mental health specialist providing individual care, we must develop, implement and evaluate stepped care systems. As demonstrated by IAPT, this model of care would consist of two levels: an entry, low-intensity step (Step 1) for the majority of patients with mild to moderate symptoms; and a high-intensity step (Step 2) for the minority of patients suffering from severe symptoms and those who do not respond to the first step.

Step 1 would involve either guided self-care or non-specialist professionals performing a range of tasks such as screening, delivering brief evidence-based psychological treatments, and acting as case managers to link the patient, family physician and specialists from mental health or other disciplines. In Step 2, mental health specialists would treat the more severe spectrum of these disorders, monitor use and adherence to medication when appropriate, and ensure treatment quality by training and supervising non-specialist professionals.

This stepped care model emphasizes patient-centered approaches and collaboration with local communities. This includes receiving input on how treatment could be best delivered in order to reduce administrative barriers, and engaging patient advocates in planning and improving the navigation of existing systems. In addition, we can target relevant co-occurring risk factors through integrated health programmes such as parenting platforms, chronic disease interventions and community-based care. In doing so, we may also have the opportunity to reach marginalized groups who may not typically seek mental health care.

We call on the mental health community at large to embrace these evidence-based strategies into routine health care delivery platforms, as a cost-effective approach to reducing the astonishingly large treatment gap for common mental disorders worldwide.

Daisy R. Singla[1,2], Giuseppe Raviola[3,4], Vikram Patel[3,5]
[1]Department of Psychiatry, University of Toronto and Sinai Health System, Toronto, ON, Canada; [2]Lunenfeld-Tanenbaum Research Institute, Toronto, ON, Canada; [3]Department of Global Health & Social Medicine, Harvard Medical School, Boston, MA, USA; [4]Boston Children's Hospital, Boston, MA, USA; [5]Sangath, Alto Porvorim, Goa, India

The authors would like to thank R. Maunder for his contributions to this letter.

1. Thornicroft G, Chatterji S, Evans-Lacko S et al. Br J Psychiatry 2017;210:119-24.
2. Singla DR, Kohrt B, Murray LK et al. Annu Rev Clin Psychol 2017;13:149-81.
3. Cuijpers P, Karyotaki E, Reijnders M et al. World Psychiatry 2018;17:90-101.
4. Singla DR, Weobong B, Nadkarni A et al. Behav Res Ther 2014;60:53-9.
5. Richards DA, Ekers D, McMillan D, et al. Lancet 2016;388:871-80.
6. Patel V, Weobong B, Weiss HA et al. Lancet 2017;389:176-85.
7. Andrews G, Cuijpers P, Craske MG et al. PLoS One 2010;5:e13196.
8. Fairburn CG, Allen E, Bailey-Straebler S et al. J Med Internet Res 2017;19:e214.
9. Hunter A, Riger S. J Community Psychol 1986;14:55-71.
10. Clark DM, Canvin L, Green J et al. Lancet 2018;391:679-86.

DOI:10.1002/wps.20532

Progress in developing a classification of personality disorders for ICD-11

In appointing a Working Group charged with developing recommendations in the area of personality disorders (PDs) for the ICD-11, the World Health Organization (WHO) Department of Mental Health and Substance Abuse highlighted several problems with the classification of PDs in the ICD-10.

First, PDs appeared to be substantially underdiagnosed relative to their prevalence among individuals with other mental disorders. Second, of the ten specific PDs, only two (emotionally unstable personality disorder, borderline type and dissocial personality disorder) were recorded with any frequency in publicly available databases. Third, rates of co-occurrence were extremely high, with most individuals with severe disorders meeting the requirements for multiple PDs. Fourth, the typical description of PD as persistent across many years was inconsistent with available evidence about its lack of temporal stability.

The WHO, therefore, asked the Working Group to consider changes in the basic conceptualization of PDs and specifically to explore the utility and feasibility of a dimensional approach. At the same time, the WHO emphasized that any classification system of PDs for the ICD-11 must be usable and useful for health care workers in lower-resource settings who are not highly trained specialist mental health professionals[1].

The Working Group, under the leadership of P. Tyrer, took the WHO's requests very seriously in developing its proposal for ICD-11. PD was conceptualized in terms of a general dimension of severity, continuous with normal personality variation and

sub-threshold personality difficulty. After meeting the general requirements for a diagnosis of PD, an individual would be assigned a mild, moderate or severe PD diagnosis, based primarily on the extent of interpersonal dysfunction and the risk of harm to self or others. The ICD-10 specific PDs were abandoned entirely in favour of five broad trait domains grounded in the scientific literature on personality[2]: negative affectivity, disinhibition, detachment, dissociality and anankastia.

Descriptions of the Working Group proposal were subsequently published in specialty and more general scientific journals[3,4]. It should be noted that, although the essence of the ICD-11 proposal was conceptually compatible with what came to be the "alternative model" of PD diagnosis in the DSM-5, the Working Group recommended against adoption of that model for ICD-11 because it was seen as too complicated for implementation in most clinical settings around the world.

The WHO became aware of significant concerns among some members of the practice community and some PD researchers about various aspects of the proposal. This led to a meeting of the WHO with representatives from the European Society for the Study of Personality Disorders (ESSPD), the International Society for the Study of Personality Disorders (ISPPD), and the North American Society for the Study of Personality Disorders (NASSPD). A description of the concerns of members of the leadership of these organizations about the original Working Group proposal has recently been published[5], although these concerns were not universal[6]. Nevertheless, the WHO believed it was important to attempt to engage a process that would help to avoid further divisiveness and acrimony in this area.

The WHO thus convened a Task Group consisting of members appointed by ISSPD/ESSPD/NASSPD and members of the original Working Group, which was asked to develop recommendations for responding to the concerns. Through discussions over several months, it became clear that the ISSPD/ESSPD/NASSPD representatives were willing to accept a dimensional model of PDs, but felt that the one that had been proposed provided insufficient information about the nature of individual personality disturbance to support case conceptualization, treatment selection, and management.

The other major issue to be addressed was the diagnostic status of borderline PD. Some research suggests that borderline PD is not an independently valid category, but rather a heterogeneous marker for PD severity[7,8]. Other researchers view borderline PD as a valid and distinct clinical entity, and claim that 50 years of research support the validity of the category[9]. Many – though by no means all – clinicians appear to be aligned with the latter position. In the absence of more definitive data, there seemed to be little hope of accommodating these opposing views. However, the WHO took seriously the concerns being expressed that access to services for patients with borderline PD, which has increasingly been achieved in some countries based on arguments of treatment efficacy, might be seriously undermined.

In September 2017, the Task Group held a face-to-face meeting in Heidelberg, Germany, with the leadership and support of S.C. Herpertz, then ISSPD President. The purpose of the meeting was to develop specific proposals for modifications to the ICD-11 guidelines that would address the issues of concern. The main recommended changes were as follows:

- *Systematic incorporation of self functioning in the core diagnostic guidelines for PD.* PD is conceptualized as an enduring disturbance characterized by problems in functioning of aspects of the self (e.g., identity, self-worth, accuracy of self-view, self-direction) and/or interpersonal dysfunction.
- *A substantially richer and more clinically informative operationalization of PD severity.* The degree and pervasiveness of disturbances in functioning of aspects of the self; of interpersonal dysfunction across various contexts and relationships (e.g., romantic relationships, school/work, parent-child, family, friendships, peer contexts); of emotional, cognitive and behavioural manifestations of the personality dysfunction; as well as of associated distress or functional impairment should be considered in making a severity determination for individuals who meet the general diagnostic requirements for PD.
- *A substantially richer and more clinically informative operationalization of trait qualifiers.* Each should describe the core feature of the trait domain, followed by a description of the common manifestations of that domain in individuals with PD.
- *A complete description of PD includes the severity rating and the applicable trait domain qualifiers.* The WHO acknowledges that it will not be feasible to conduct such a complete evaluation in all settings.
- *Provision of an optional qualifier for "borderline pattern".* This qualifier may enhance clinical utility by facilitating the identification of individuals who may respond to certain psychotherapeutic treatments. Whether it will provide information that is non-redundant with the trait domain qualifiers is an empirical question.

A revision of the diagnostic guidelines for PDs based on the above recommendations has been approved by the ICD-11 Working Group and the ISSPD/ESSPD/NASSPD representatives. These guidelines are available for review and comment at http://gcp.network, and are now being used in field testing.

Geoffrey M. Reed
Department of Mental Health and Substance Abuse, World Health Organization, Geneva, Switzerland

The views expressed in this letter are those of the author and do not necessarily represent the official policies or positions of the WHO. Members of the ICD-11 PD Working Group included P. Tyrer (Chair), R. Blashfield, L.A. Clark (DSM liaison), M. Crawford, A. Farnam, A. Fossati, Y.-R. Kim, N. Koldobsky, D. Lecic-Tosevski, R. Mulder, D. Ndetei and M. Swales. Representatives of ISSPD/ESSPD/NASSPD included S.C. Herpertz, M. Bohus, S.K. Huprich and C. Sharp. The WHO acknowledges the major contributions of L.A. Clark and M.B. First to the revision of the diagnostic guidelines described above.

1. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. World Psychiatry 2011;10:86-92.
2. Krueger RF, Markon KE. Annu Rev Clin Psychol 2014;10:477-501.
3. Tyrer P, Crawford M, Mulder R et al. Personal Ment Health 2011;5:246-59.
4. Tyrer P, Reed GM, Crawford MJ. Lancet 2015;385:717-26.
5. Herpertz SC, Huprich SK, Bohus M et al. J Pers Disord 2017;31:577-89.
6. Hopwood SJ, Kotov R, Krueger RF et al. Personal Ment Health 2018;12:82-6.
7. Sharp C, Wright AGC, Fowler JC et al. J Abnorm Psychol 2015;124:387-98.
8. Williams TF, Scalco MD, Simms LJ. Psychol Med 2018;48:834-84.
9. Clarkin JF, Lenzenweger MF, Yeomans F et al. J Pers Disord 2007;21:474-99.

DOI:10.1002/wps.20533

Neurocognitive disorders in ICD-11: the debate and its outcome

In the ICD-11, the chapters "06. Mental, behavioural or neurodevelopmental disorders" and "08. Diseases of the nervous system" are going to include, respectively, the groupings of "Neurocognitive disorders" and "Disorders with neurocognitive impairment as a major feature". Concern over the "wrong" allocation of dementias in the diagnostic system had produced many critical reactions from mental health professionals, due to the anticipated adverse consequences for treatment and care. Here we summarize the background and outcome of these reactions.

In late 2016, the World Health Organization (WHO) moved the dementia categories – contrary to the "traditional" location of clinical manifestations in ICD-10 (F00-F03) – from chapter 06 to chapter 08 of the ICD-11 draft. This step, following a Neurology Topic Advisory Group proposal, generated written protest notes by about two dozens of national and international scientific associations, mainly from psychiatry, old age psychiatry, psychology and other mental health workforce. In early 2017, the WHO corrected the previous step in that the dementia categories were moved back to chapter 06.

What was the rationale of these moves? According to the ICD-11 Reference Guide, the guiding principles for "allocation of entities" are "to maintain the structural and functional integrity" of the classification and "to preserve consistency with previous versions". Classification should be changed only with a "strong rationale", and categories should be kept in their "legacy location" if they "could arguably be in two or more places".

Neurocognitive disorders such as Alzheimer dementia are being classified in ICD-10 according to the dagger-asterisk system, with the *clinical manifestation* in chapter F (F00*) and the *aetiology* in chapter G (G30†). In ICD-11, according to this "legacy location", Alzheimer dementia should continue to be classified both in chapter 06 ("disorders") for its manifestation and in chapter 08 ("diseases") for its aetiology, using the new post-coordination coding.

Despite increasing knowledge on aetiopathogenesis and biomarkers, dementias are generally still diagnosed clinically and classified according to their manifestation. The proposal to move them to chapter 08 may have been either misled by concept or misread by the WHO, although the ultimate aim of classifying disease entities is indeed to primarily build on aetiologies and dysfunctional body systems and not solely on clinical manifestations. Despite Griesinger's dictum "mental disease is brain disease"[1], and although involvement of brain dysfunction is increasingly recognized and important to consider, most "mental" disorders cannot be treated as "brain disorders" or diseases with *monocausal* brain pathology.

Arguments against the move of dementias to chapter 08 were referring to WHO managing issues (move contrary to the joint recommendation by Mental Health and Neurology Topic Advisory Groups), conceptual and methodological issues (lack of evidence for the move; the need for a biopsychosocial approach in integrated care), treatment and service issues (resulting limitation of access to care; importance of neuropsychological vs. biomedical measures in treatment and care), professional and interdisciplinary issues (cross-national variation in responsibility of specialties, but usually major role of psychiatrists in treatment and care; importance of keeping the balance among disciplines), economic issues (problems with reimbursement by insurance companies in several countries if dementia is withdrawn from chapter on mental and behavioural disorders), psychopathological issues (behavioural symptoms do not belong in the "neurology" section, while being a major burden for patients and carers and hence a significant focus for treatment), and classification analogies in ICD-11 (e.g., chapters on cardiovascular, infectious and endocrinological diseases).

As an outcome of the debate, the WHO has moved dementias back to mental disorders in chapter 06, analogously to ICD-10 and DSM-5. Chapter 08 covers in its neurocognitive section only "diseases", e.g. Alzheimer disease, which can be associated by post-coordination coding with "6E00 Dementia due to Alzheimer disease". Options for post-coordination coding have now also been implemented for "6D91 Mild neurocognitive disorder" (F06.7 in ICD-10), which can be associated with any of the diseases in chapter 08, or with diseases classified elsewhere, as a result of commentaries by the Japanese Society of Psychiatry and Neurology (JSPN), the German Association of Psychiatry and Psychotherapy, and the American Psychiatric Association.

Another proposal by JSPN was the introduction of specifiers for behavioural symptoms in the diagnosis of dementias, because of their high burden for patients and carers. This has been implemented by the WHO under "6E20 Behavioural or psychological disturbances in dementia".

In conclusion, we have witnessed successful outcomes from a worldwide interactive process with the WHO on classifying neurocognitive disorders taking into account clinical utility[2]. In keeping abreast of the ever developing state of the art, the ICD-11 will need ongoing adaptation, e.g., taking into account the progress in preclinical classification of Alzheimer dementia

and biomarker-based diagnosis[3].

Wolfgang Gaebel[1,2], Frank Jessen[3,4], Shigenobu Kanba[5,6]
[1]Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany; [2]WHO Collaborating Centre on Quality Assurance and Empowerment in Mental Health, Düsseldorf, Germany; [3]Department of Psychiatry, Medical Faculty, University of Cologne, Cologne, Germany; [4]German Center for Neurodegenerative Diseases, Bonn, Germany; [5]Department of Neuropsychiatry, Kyushu University, Fukuoka, Japan; [6]Japanese Society of Psychiatry and Neurology, Tokyo, Japan

1. Griesinger W. Die Pathologie und Therapie der psychischen Krankheiten für Ärzte und Studierende, Vol.1. Stuttgart: Krabbe, 1845.
2. Reed GM. Prof Psychol Res Pr 2010;41:457-64.
3. Dubois B, Hampel H, Feldman HH et al. Alzheimers Dement 2016;12:292-323.

DOI:10.1002/wps.20534

Digital interventions in severe mental health problems: lessons from the Actissist development and trial

Severe mental health problems are characterized by repeated relapse, yet timely access to treatment remains problematic[1]. Within current health care systems, the delivery of treatment by scheduled appointment can result in warning signs being missed or treated too late. Recognizing the need for innovative, timely and efficient solutions to improve the speed and quality of treatment delivery, digital strategies are being developed worldwide[2].

Grounded in the cognitive model of psychosis, and following an extensive period of co-design with patients and stakeholders, we developed Actissist[3], a theory-informed smartphone app targeting areas of distress in early psychosis. Actissist uses question and answer dialogues with a branched design to provide cognitive or behavioral-informed feedback to participants, based on the information they input into the app. The app also contains a menu of multi-media options (e.g., links to external sites, patient stories, relaxation sessions) designed to complement and support the feedback from the intervention domains.

In a proof-of-concept, single-blind, randomized controlled trial, 36 early psychosis patients were randomly allocated to receive either Actissist plus treatment as usual (N=24) or ClinTouch[4], a symptom monitoring app, plus treatment as usual (N=12) over 12 weeks, with blind assessor follow-up at 12 and 22 weeks[3]. Participants were recruited over 7 months from several early intervention for psychosis services in the North West of England.

Nearly two thirds (38/59; 64.4%) of referred people participated in the study. We found that Actissist was feasible (75% participants used it at least once a day over the 12-week intervention period; 97% participants remained in the trial until the end), acceptable (90% participants declared they would recommend Actissist to others in a similar position), and safe (no serious adverse events related to the study). The treatment effects at 12 weeks favoured the Actissist group, with a Cohen's D standardized effect size of −0.85 (95% CI: −1.44 to −0.25) for the total score on the Positive and Negative Syndrome Scale, and of −0.65 (95% CI: −1.28 to −0.02) for the total score on the Calgary Depression Scale for Schizophrenia.

The next stage of Actissist is being tested in a powered randomized controlled trial (RCT). However, there are at present several clear challenges to both the conduction of standard RCTs in this area and the implementation of digital health interventions in ordinary practice.

In standard RCTs, the intervention is fixed at the onset of the trial and is not permitted to evolve during the trial. For many drugs under investigation or complex interventions, this is reasonable. However, this is problematic for digital health interventions due to the pace of change in technology. Fixing the intervention at trial onset can render the technology outdated or even obsolete by the time the trial results are available. Adaptive interventions, which are designed to systematically and efficiently optimize behavioural interventions, might be one possible solution to this problem[5].

Furthermore, the success of digital health interventions is not merely determined by patient uptake; it will ultimately be determined by patients *and* staff, both of whom are key end-users. We have found that mental health professionals and patients often express concerns about data security, safety and risk information being robustly handled[6]. However, given reassurances from reputable and trusted organizations, patients recognize the value of digital health interventions in enhancing their connection with services, and perceive digital approaches as not only destigmatizing but also a relevant way of receiving health care. Perhaps most importantly, patients view these interventions as empowering, affording them meaningful choice and the opportunity to take active control of their health care.

Staff attitudes, however, are a potentially major barrier to digital health care implementation[6]. In our work, staff often expressed the opinion that resources would be better spent on professionals' training than on technology development. Integrating a steady stream of data into patients' records was sometimes perceived as overwhelming, adding to already stretched workloads and professional responsibilities. Without considering issues around implementation during the early stages of the development and delivery of digital health interventions, it is unlikely that these approaches will be disseminated beyond research studies and into the service setting.

Moreover, a clear set of strategies regarding closer involvement of patients in the development of digital innovations as well as engagement of stakeholders with digitally-enabled services is lacking. More research is needed worldwide to under-

stand patients' and stakeholders' perspectives on digital health systems, to maximize implementation. We achieved this in Actissist[3] by holding quarterly meetings with an expert reference group comprising patient representatives and other stakeholders, who were actively involved in all aspects of trial design and app development. We also integrated extensive qualitative work with patients and other stakeholders from before the trial commenced right through to trial exit interviews post follow-up.

Finally, from a global perspective, there is a need to address the exclusion of low-income individuals who cannot access the technology necessary to run digital health tools. Evidence-based digital systems should be a health care cost covered by routine processes, rather than billed to patients. The digital divide also relates to staff using digital systems in the health care context. In our qualitative work, staff often described concerns about their own ability to use technology as well as lack of confidence in the ability of health services to successfully implement a coherent and fully integrated digital system, highlighting the need for all individuals using mental health services and those delivering services to be fully trained and supported[6].

One final consideration is the lack of theory-driven work underpinning apps being developed across the health setting. It is through theoretical development and innovation that we advance our discipline.

Each of the challenges set out above will need significant programmes of research, considering not only methods of evaluating digital health interventions, but also drawing on implementation science principles. Taken together, these challenges define a prioritized research agenda for digital health interventions for mental health. The promise shown in this field will only be turned into significant progress through multi-disciplinary working.

Sandra Bucci[1], Shon Lewis[1], John Ainsworth[2], Gillian Haddock[1], Matthew Machin[2], Katherine Berry[1], Natalie Berry[1], Dawn Edge[1], Richard Emsley[3]

[1]Division of Psychology and Mental Health, School of Health Sciences, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK; [2]Division of Informatics Imaging and Data Sciences, University of Manchester and Health eResearch Centre, Farr Institute for Health Informatics Research, University of Manchester, Manchester, UK; [3]Division of Population Health, Health Services Research and Primary Care, University of Manchester, Manchester, UK

1. Murdoch C, Kendall T. Implementing the five year forward view for mental health. Redditch: NHS England, 2016.
2. Bhugra D. Tasman A, Pathare S et al. Lancet Psychiatry 2017;4:775-818.
3. Bucci S, Barrowclough C, Ainsworth J et al. Schizophr Bull (in press).
4. Palmier-Claus J, Ainsworth J, Machin M et al. BMC Psychiatry 2012;12:172.
5. Collins LM, Nahum-Shani I, Almirall D. Clin Trials 2014;11:426-34.
6. Berry N, Bucci S, Lobban F. JMIR Mental Health 2017:4:e52.

DOI:10.1002/wps.20535

Rethinking progress and challenges of mental health care in China

The rapid socio-economic development and extensive public health reform in China has led to considerable changes in the mental health service system, as previously described[1,2]. However, an update on the recent progress and challenges is now warranted.

Due to various reasons, China has faced major deficits in mental health resources in the past decades. For example, in 2004 there were only 16,103 licensed psychiatrists and psychiatric registrars (1.24 per 100,000 population), 24,793 psychiatric nurses (1.91 per 100,000), and 557 psychiatric hospitals with 129,314 psychiatric beds (9.95 per 100,000) nationwide[1]. Through strengthening the mental health service and education systems nationally, by 2015, there were 27,733 psychiatrists and psychiatric registrars (2.02 per 100,000 population), 57,591 psychiatric nurses (4.19 per 100,000) and 2,936 mental health services with approximately 433,000 psychiatric beds (31.5 per 100,000)[3]. In contrast, based on the World Health Organization (WHO)'s Mental Health Atlas[4], the proportion of psychiatrists in 2014 was 0.3 per 100,000 in India, 0.87 per 100,000 in Thailand, and 20.1 per 100,000 in Japan.

Although the number of mental health professionals has increased in China, there remains a comparative shortage in human resources. Furthermore, these resources are mostly located in urban psychiatric hospitals, making services far less accessible for at least half of China's 1.39 billion people living in rural areas. Moreover, the lack of qualified community mental health professionals, which applies to many urban areas even today, remains a major barrier.

To effectively manage millions of community-dwelling patients with severe psychiatric disorders, a national community-based model named "The management and treatment program for severe mental illness with subsidy from the central government" or the "686 Program" was initiated in 2004. We were involved in the development and training components of this program, which integrates the resources of hospital services, community case management, neighborhood committees and the police to provide comprehensive monitoring, treatment, rehabilitation and prevention services. The program prioritized patients with psychiatric disorders and relatively high risk of violent behaviours, namely those with schizophrenia, schizoaffective disorder, paranoid psychosis, bipolar disorder, and epilepsy and mental retardation associated with mental disorders.

Since 2004, the central and local governments have so far invested CNY 2.24 billion (US$ 325 million) in this program. By 2015, a total of 5.4 million patients with severe mental illness (of which around three quarters with schizophrenia) have been registered at 2,774 districts/counties in 31 provinces, municipalities and autonomous regions. Of the registered patients, 88.7% received regular services and follow-up monitoring[3]. Despite these large figures, the treatment coverage is

relatively limited, considering that there are an estimated 173 million Chinese people in the community who suffer from psychiatric disorders[5]. Much improvement in the scale of appropriate services is therefore needed in order to reduce the treatment gap.

Comprehensive epidemiological data are important in informing policy and service developments to address the treatment gap in China. Initiated by the Ministry of Health with the support of the WHO, two early large-scale psychiatric surveys had been conducted in 1982 and 1993. Between July 2013 and March 2015, the first national mental health survey, involving 28,140 respondents, was conducted in 31 provinces, municipalities and autonomous regions of China[6]. The preliminary findings were announced by the National Health and Family Planning Commission of China on April 7, 2017. These included the prevalence of mood disorders, depression and anxiety disorders, being 4.06%, 3.59% and 4.98%, respectively[3].

However, epidemiological data on special populations, such as immigrant workers, children, adolescents and older adults, are still lacking in China. Further, critical information on the health burden and impact of psychiatric disorders such as illness severity, duration and degree of disability, and associated physical comorbidities, remains largely unavailable.

China's population is rapidly aging, mostly due to the increased life expectancy as well as the one-child family policy that was instituted for 35 years. In 2000, only 7% of Chinese population was over the age of 65 years, but the figure is expected to reach 23% by 2050[7]. Due to the one-child family policy, the proportion of "empty nest family" in China had grown to 25% of all elderly households in 2003, with a projected increase to 90% by 2030[8]. The change in family structure may have significant impact on the access to social care and financial independence, as well as on the mental health of the elderly.

Currently, the availability of psychogeriatric services is deficient, and general mental health services and even treatment guidelines or intervention models for older people are poorly developed in China. The burden of care may therefore ultimately fall on family caregivers. As consequence, many family members may experience psychological problems and poor quality of life, as well as limited employment opportunities.

The protection of the rights of psychiatric patients in China remains an important concern[9]. Of note, the National Mental Health Law finally came into effect on May 1, 2013. It provides the legal framework to uphold the rights of psychiatric patients to receive dignified and appropriate treatment. According to one study, the prevalence of physical restraint in Chinese psychiatric patients decreased from 30.7% to 22.4% following the implementation of the legislation[10]. Psychiatric patients and/or their families have the right to apply for an independent medical assessment by a third party if there is a dispute about an involuntary admission[9]. However, in practice, qualified independent third parties are not easily accessible and are only available in major cities.

In summary, due to its rapid economic growth and fast-changing social structure, China still faces enormous mental health challenges. Although China's mental health legislation is a critical part of mental health reform, effective implementation of high quality services will require sustained investment in community mental health care. Decreasing the treatment gap and promoting community integration require sufficient workforce as well as innovative service models. As such, the government is investing in doubling the number of psychiatrists by 2020 and promoting digital approaches in mental health care.

Yu-Tao Xiang[1,2], Chee H. Ng[3], Xin Yu[4], Gang Wang[2]
[1]Faculty of Health Sciences, University of Macau, Macao SAR, China; [2]National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China; [3]Department of Psychiatry, University of Melbourne, Melbourne, Australia; [4]Peking University Institute of Mental Health, Beijing, China

The first and last author contributed equally to this work.

1. Liu J, Ma H, He YL et al. World Psychiatry 2011;10:210-6.
2. Xiang YT, Yu X, Sartorius N et al. Lancet 2012;380:1715-6.
3. National Health and Family Planning Commission of China. News conference, April 7, 2017. www.nhfpc.gov.cn.
4. World Health Organization. Mental health atlas – 2014 country profiles. Geneva: World Health Organization, 2015.
5. Phillips MR, Zhang J, Shi Q et al. Lancet 2009;373:2041-53.
6. Liu Z, Huang Y, Lv P et al. Soc Psychiatry Psychiatr Epidemiol 2016;51:1547-57.
7. Chu L-W, Chi I. J Am Med Dir Assoc 2008;9:237-43.
8. Xie LQ, Zhang JP, Peng F et al. Arch Gerontol Geriatr 2010;50:24-9.
9. Xiang YT, Yu X, Ungvari GS et al. Lancet 2012;379:780-2.
10. An FR, Sha S, Zhang QE et al. Psychiatry Res 2016;241:154-8.

DOI:10.1002/wps.20500

Psychotic experiences as an independent risk factor for angina pectoris in 48 low- and middle-income countries

People with schizophrenia are known to have a life expectancy reduced by 10-20 years compared to the general population[1], which is largely attributable to their increased risk for cardiovascular diseases (CVDs)[2].

There is some evidence that people with psychotic experiences (PEs) who do not reach the clinical threshold for a psychosis diagnosis are also at higher risk of premature mortality (by 5 years)[3], which likewise may be explained by a higher likelihood of CVDs[4,5]. Furthermore, similar to schizophrenia, there is increasing evidence that PEs are associated with adverse health behaviors, diabetes and mental health problems, which may all increase CVD risk[5].

PEs are common in the general population (their lifetime prevalence is estimated to be 7.2%[6]), with an onset earlier than many other known risk factors for CVDs. Thus, if epidemiological data show that PEs are associated with clinically meaningful elevated odds of CVDs, and this association is present also in younger adults, being independent from other known vulnerability factors, then PEs may constitute an important risk factor for CVDs, a leading cause of death worldwide. The aims of the current study were to examine the association between PEs and angina pectoris, and to explore the factors influencing this association.

We analyzed cross-sectional, community-based data from 48 low- and middle-income countries that participated in the World Health Surveys[7]. Multistage clustered sampling was used to select individuals aged ≥18 years. Angina pectoris was operationalized as a previous medical diagnosis of that condition and/or a positive result on the World Health Organization (WHO) Rose Angina Questionnaire. Participants were asked questions on psychotic symptoms (delusional mood, delusions of reference and persecution, delusions of control, hallucinations) which came from the WHO Composite International Diagnostic Interview (CIDI) 3.0. They were considered to have PEs if they endorsed at least one psychotic symptom.

The potential mediators in our study included past 7-day heavy episodic drinking, current smoking, low physical activity (<150 minutes of moderate-to-vigorous physical activity per week), inadequate fruit and vegetable consumption (<2 fruit or <3 vegetable servings/day), diabetes (self-reported), obesity (defined as a body mass index ≥30 kg/m^2 based on self-reported weight and height), depression (lifetime or past 12-month DSM-IV diagnosis based on CIDI), antipsychotic use (past 2 weeks), and past 30-day perceived stress, anxiety and sleep problems. Socio-demographic variables included gender, age, education, wealth quintiles, and living environment (urban/rural).

Statistical analyses were performed with Stata 14.1. Multivariable logistic regression analysis was conducted to assess the association between PEs (exposure) and angina pectoris (outcome). Analyses using the overall and country-, age- and gender-stratified samples were conducted. We carried out mediation analysis with the khb (Karlson Holm Breen) command in Stata. This method decomposes the total effect of a variable into direct and indirect effects, and allows calculation of the mediated percentage.

Each potential mediator was included in the model individually. All regression models were adjusted for age, gender, education, wealth, living environment and country. The sample weighting and complex study design were taken into account. The level of statistical significance was set at $p<0.05$.

The mean age of the final sample (N=224,842) was 38.3± 16.0 years, and 49.3% were males. Overall, after adjustment for age, gender, education, wealth, living environment and country, the presence of PEs was associated with a 2.29 (95% CI: 2.13-2.47) times higher odds for angina pectoris. Similar associations were observed among males (OR=2.45, 95% CI: 2.16-2.78) and females (OR=2.20, 95% CI: 2.00-2.42). The association was strongest in the youngest age group: OR=2.59, 95% CI: 2.36-2.85 for 18-44 years; OR=2.06, 95% CI: 1.80-2.37 for 45-64 years; OR=1.62, 95% CI: 1.31-2.01 for ≥65 years. The OR was >1 in all countries, with a significant association being observed in all but five countries (OR range: 1.24 for Ukraine to 10.85 for Myanmar).

Mediation analysis showed that depression, anxiety, perceived stress and sleep problems explained 20.6%, 11.4%, 10.1% and 9.5% of the PE-angina pectoris association, respectively (collectively 33.0%). Diabetes (1.7%) and alcohol consumption (0.8%) explained a small proportion of the association. Smoking, physical activity, fruit/vegetable consumption, obesity and antipsychotic use were not significant mediators. The association between PEs and angina pectoris continued to be significant even after adjustment for all eleven potential mediators (OR=1.72, 95% CI: 1.52-1.93).

The value of PEs as a clinically useful predictor depends on their independent associations with CVD when adjusting for known CVD risk factors. We found that PEs are indeed associated with greater odds of angina pectoris even when accounting for a broad range of psychiatric, demographic, behavioral and lifestyle risk factors.

Even more intriguing, we found that this effect was greatest among younger and middle-aged adults, when the effect of risk factors for CVD is expected to be attenuated due to shorter lifetime exposure. This, therefore, provides compelling epidemiological evidence that screening for PEs in young adulthood may provide independent predictive information regarding lifetime CVD risk, which can potentially facilitate clinical prevention of CVDs, pending future translational research studies.

In particular, prospective studies of clinical samples of adults at risk for CVDs can be used to determine whether screening for PEs adds unique predictive value to existing CVD risk screens. If so, PE screening can be a valuable contribution to CVD risk detection, particularly given the early onset of PEs relative to CVDs.

Ai Koyanagi[1,2], Hans Oh[3], Brendon Stubbs[4-6], Nicola Veronese[7,8], Davy Vancampfort[9,10], Josep Maria Haro[1,2], Jordan E. DeVylder[11]

[1]Research and Development Unit, Parc Sanitari Sant Joan de Déu, Universitat de Barcelona, Barcelona, Spain; [2]Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, Madrid, Spain; [3]School of Social Work, University of Southern California, Los Angeles, CA, USA; [4]Physiotherapy Department, South London and Maudsley NHS Foundation Trust, London, UK; [5]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [6]Faculty of Health, Social Care and Education, Anglia Ruskin University, Chelmsford, UK; [7]Neuroscience Institute, National Research Council, Padua, Italy; [8]Department of Geriatric Care, OrthoGeriatrics and Rehabilitation, E.O. Galliera Hospital, Genoa, Italy; [9]KU Leuven Department of Rehabilitation Sciences, Leuven, Belgium; [10]University Psychiatric Center KU Leuven, Leuven-Kortenberg, Belgium; [11]School of Social Service, Fordham University, New York, NY, USA

1. Liu NH, Daumit GL, Dua T et al. World Psychiatry 2017;16:30-40.
2. Correll CU, Solmi M, Veronese N et al. World Psychiatry 2017;16:163-80.
3. Sharifi V, Eaton WW, Wu LT et al. Br J Psychiatry 2015;207:30-6.
4. Oh H, DeVylder J. World Psychiatry 2015;14:101-2.
5. Moreno C, Nuevo R, Chatterji S et al. World Psychiatry 2013;12:251-7.

6. Linscott RJ, van Os J. Psychol Med 2013;43:1133-49.
7. Ustun TB, Chatterji S, Mechbal A et al. In: Murray CJL, Evans DB (eds). Health systems performance assessment: debates, methods and empiricism. Geneva: World Health Organization, 2003:797-808.

DOI:10.1002/wps.20536

Feasibility of a guided self-help intervention to reduce psychological distress in South Sudanese refugee women in Uganda

Implementing evidence-based psychological interventions in low-resource refugee settings is challenging, because of the need for an extensive workforce of trainers, supervisors and facilitators[1,2]. Self-Help Plus (SH+) was developed by the World Health Organization (WHO) as a tool potentially applicable in those settings[3].

SH+ is a guided self-help intervention consisting of five audio-recorded sessions and an illustrated self-help manual[3]. It can be provided to large groups (20 to 30 participants) and facilitated by lay helpers with minimal training. It aims to reduce psychological distress in people with a range of common mental disorders and subthreshold symptoms. It is based on acceptance and commitment therapy, a third wave cognitive behavioral therapy focused on enhancing psychological flexibility[4].

We adapted SH+ for South Sudanese refugees and conducted a feasibility cluster randomized controlled trial of the intervention in Rhino Camp, a refugee settlement area in northern Uganda[5]. Our focus in this study was on women, since prior intervention adaptation and piloting had shown the need for separate evaluation efforts with men and women. We randomly allocated one village to SH+ and one to enhanced usual care. Within each village, we randomly selected households and screened one Juba Arabic-speaking consenting woman (age $\geq$18 years) until 25 eligible women were identified per village.

We screened for moderate psychological distress using the Kessler 6 (K6) (primary outcome, cut-off $\geq$5)[6]. We assessed exclusion criteria (imminent risk of suicide; observable signs of severe mental disorder; severe cognitive impairment) using structured questionnaires. With eligible and consenting women, we assessed secondary outcomes: disability (WHO Disability Assessment Schedule 2.0, WHODAS 2.0); self-defined psychosocial concerns (Psychological Outcome Profiles instrument, PSYCHLOPS); depression symptoms (Patient Health Questionnaire, PHQ-9); post-traumatic stress disorder (PTSD) symptoms (PTSD Checklist Civilian, PCL-6); hazardous alcohol use (two survey questions); feelings of anger (shortened explosive anger index); inter-ethnic relations (three survey questions); subjective wellbeing (WHO Wellbeing Index, WHO-5); psychological flexibility (Acceptance and Action Questionnaire, AAQ-II). We also assessed attendance, health service use, cost indicators, and exposure to potentially traumatic events.

The SH+ workshops were facilitated by four Juba Arabic-speaking Ugandan women from the settlement area without prior mental health training. Facilitators were trained by international experts (KC, FB) over a four-day period and supervised weekly by a Ugandan social worker. Enhanced usual care consisted of one psychoeducation session focused on psychological distress delivered by a trained community health worker, which included information on where to access existing mental health services delivered by the implementing organization, the Peter C. Alderman Foundation.

Assessors were blinded to allocation of villages to study condition, and conducted interviews one week pre- and post-intervention. All participants provided written or verbal informed consent. Ethical procedures were approved by the WHO Ethics Review Committee, the MildMay Uganda Research Ethics Commission, and the Uganda National Council for Science and Technology.

We screened 50 women, all of whom were eligible and consented. Their mean age was 29.5$\pm$8.5 years and 68% of them were married. Half of participants were managing households; 60% had no schooling or completed primary school.

Fidelity checks (clinical supervisor attending 10% of sessions) showed that all sections of the audio were delivered correctly at each session. Weekly supervision was provided to SH+ facilitators and covered reporting of any adverse events, requests for additional help from participants or problems in running the course. Few problems were reported and supervision was brief. Attendance was good (90% of women attended each session).

We found that our research protocol was feasible. Randomization resulted in balanced groups at baseline despite the small sample. We did not find differences between groups at baseline on socio-demographic characteristics. There were larger mean post-intervention differences for the SH+ condition on all outcome measures. These were statistically significant for the K6 ($p<0.05$) and the WHO-5 ($p<0.001$). Blinding was maintained: assessors guessed correctly which participants were part of which study condition at chance level (50% of cases). Similarly, contamination did not appear to be a major concern: none of the participants in the control condition had seen the SH+ self-help book, attended workshops, nor heard about SH+. Only two women receiving enhanced usual care were lost to follow-up – an attrition of 4%.

In conclusion, we found that the SH+ intervention and research protocols were feasible in Uganda among South Sudanese refugees, with promising results related to randomization, fidelity, adherence, contamination, blinding, and sensitivity to change. If efficacy is confirmed in a forthcoming larger fully-powered trial, SH+ will represent a promising and

potentially scalable evidence-based mental health intervention for addressing psychological distress in refugees and other populations affected by adversities at a time of great need.

Wietse A. Tol[1,2], Jura Augustinavicius[2], Kenneth Carswell[3], Marx R. Leku[1], Alex Adaku[1], Felicity L. Brown[4,5], Claudia García-Moreno[3], Peter Ventevogel[6], Ross G. White[7], Cary S. Kogan[8], Richard Bryant[9], Mark van Ommeren[3]
[1]Peter C. Alderman Foundation, Uganda; [2]Johns Hopkins University, Baltimore, MD, USA; [3]World Health Organization, Geneva, Switzerland; [4]War Child Holland, Amsterdam, The Netherlands; [5]Harvard School of Public Health, Boston, MA, USA; [6]United Nations High Commissioner for Refugees, Geneva, Switzerland; [7]University of Liverpool, Liverpool, UK; [8]University of Ottawa, Ottawa, Canada; [9]University of New South Wales, Sydney, Australia

1. Silove D, Ventevogel P, Rees S. World Psychiatry 2017;16:130-9.
2. Tol WA, Barbui C, Bisson J et al. PLoS Med 2014;11:e1001769.
3. Epping-Jordan JE, Harris R, Brown FL et al. World Psychiatry 2016;15:295-6.
4. Hayes SC, Luoma JB, Bond FW et al. Behav Res Ther 2006;44:1-25.
5. Adaku A, Okello J, Lowry B et al. Confl Health 2016;10:18.
6. Prochaska JJ, Sung HY, Max W et al. Int J Methods Psychiatr Res 2012;21: 88-97.

DOI:10.1002/wps.20537

Correction

It has been brought to our attention that on p. 79 of the paper "Income inequality and depression: a systematic review and meta-analysis of the association and a scoping review of mechanisms", by Patel et al, published in the February 2018 issue of *World Psychiatry,* the ethnicity of the participants in the study by Fernández-Niño et al (2014) was incorrectly reported as "Hispanic", whereas it was mixed, also including people of indigenous descent.

Mainstreaming psychiatry: implementing the WPA Action Plan 2017-2020

The vision of the WPA is a world in which people live in conditions that promote mental health and have access to mental health treatment and care that meet appropriate professional and ethical standards, integrate public health principles and respect human rights. The Action Plan sets out a strategy for expanding the contribution of psychiatry to improved mental health for people across the globe[1,2]. It proposes working with partners to reach people who face adversity and disadvantage.

The WPA's Action Plan has two areas of strategic focus. Its institutional work, supported by the Executive Committee and dedicated colleagues, aims to strengthen the capacity of psychiatry through accelerating the adoption worldwide of quality approaches and advances in scientific, educational, service development and advocacy activities. Its development work includes the signature initiatives for the triennium. These focus on mainstreaming the contributions of psychiatry to meet the mental health needs of children, young women and young men living with adversity or in the midst of emergencies[3].

The first step for the development work was WPA's pivotal role in shaping the citiesRISE platform[4] and its successful launch in 2017. The platform is a collective action program seeking to promote the mental health of children and young people in adverse settings. The work is being implemented in six cities, called "nodes". The six nodes provide a foundation for broader uptake in new locations.

The WPA works through the platform with multiple partners including the World Bank, International Medical Corps (IMC), King's College London, Harvard University, Grand Challenges Canada, local and national civic and government organizations, and others. Working alongside the partnership, the World Health Organization (WHO) has a key technical role in sharing knowledge and practice. WPA and citiesRISE emphasize the prevention and early treatment of mental illness and the promotion of mental health, as well as partnerships with innovators in the development and use of digital technologies in mental health.

Major international donors have committed to support the development of the citiesRISE platform and the implementation of several projects in the WPA Action Plan, beginning with its signature initiative: the Alliance for Mental Health Responses to Emergencies, Conflicts and Adversity (the Alliance).

The second step in the development work was convening a WPA workshop in Madrid in March 2018, hosted and supported by the Juan José López-Ibor Foundation, in collaboration with citiesRISE. The Foundation has had a catalytic role in the development of the Alliance initiative. This role will continue as the work becomes operational globally.

During the workshop we defined that WPA's Alliance initiative will operate through the citiesRISE platform as a central program of the WPA Action Plan 2017-2020. It will have an initial reporting period of two years. There is also an invitation from citiesRISE to make a six-year plan to work together on the platform. The work will start in Bogotá and Chennai, two of the citiesRISE nodes. It is expected to extend to two other cities in the citiesRISE network – in Africa and the USA – during the triennium. Support will come from citiesRISE, WPA and the Foundation.

The Alliance initiative will begin with building a coalition of mental health professionals in Bogotá and Chennai. It will build on previous work with WPA[5] and existing experiences[6] to develop and implement a program of capacity building that will support and sustain the readiness of psychiatrists and other mental health professionals to respond to conflict, emergency and adversity. The work in Bogotá will be directed towards the needs of people affected by displacement from their homes, as well as uncertainty, insecurity, poverty and the physical impacts of emergencies and conflicts. The work in Chennai is expected to focus on young women and young men living in slums and facing similar conditions. In each case the local collective action groups – including young people, experts and researchers – will work with the WPA and global research partners to refine the choice, implementation and evaluation of the initiatives.

Large numbers of people are exposed to extreme stressors. Supporting the bonds between people within communities can promote resilient responses and mitigate the psychological impact of emergencies[7]. However, tackling community recovery needs mental health and psychosocial support interventions[8]. The challenge now is to evaluate and refine programs and good practice in mental health promotion in emergencies. It is equally important to add a graded response to the needs of people with mental ill health. Psychiatrists have vital roles as part of the response, as advocates, facilitators, trainers and clinicians.

The Alliance initiative aims to support psychiatrists to: work with partners to perform their roles in emergency responses, and by analogy in other situations of adversity such as slums, with a special focus on human rights and cultural competencies; train their peers and other clinicians and community-based workers in their own countries and regions; and support the development of new and existing community-based services in innovative and community-directed ways. The work will align with international protocols on mental health and psychosocial support in humanitarian emergencies[9].

We anticipate working with partners including IMC and WHO to adapt resources, guidelines and protocols that will become a lasting source of support for the initiative in these regions and eventually elsewhere. The initiative will include support for the development of new and existing community-based services. It will aim to demonstrate and disseminate best practice in the role of psychiatrists

supporting appropriate responses to conflicts, emergencies and adversity.

The Alliance initiative is the first example of work to be implemented through the WPA-citiesRISE platform. In 2018-2019 we will work towards incorporating the other Action Plan initiatives that focus on the mental health needs of people in adversity. These initiatives include suicide prevention, with a focus on the needs of young women and young men in low income and emergency settings; support for human rights and quality care in institutional and other mental health care settings; and further development of capacity building projects with community mental health and primary care providers, in partnership with other organizations. Continuing the work to describe and disseminate examples of best practice in working between practitioners, service users and carers[10] will remain a focus.

The work we are commencing is a critical priority for a world in which threats to the mental health and well-being of people in adverse situations remain high. Many organizations have worked for a long time to tackle global challenges in mental health that also concern WPA. Working together and choosing the best way to contribute to these efforts, we are beginning to leverage new resources to serve our collective goals. We welcome the growing involvement of WPA components and our other partners in the expansion of this effort over the next six months and the years to come.

Helen Herrman
President, World Psychiatric Association

1. Herrman H. World Psychiatry 2017;15:190-1.
2. World Psychiatric Association. Action Plan, 2017-20. www.wpanet.org.
3. Herrman H. World Psychiatry 2016;15:190-1.
4. citiesRISE. http://cities-rise.org/
5. Maj M. World Psychiatry 2009;8:129-30.
6. International Medical Corps. https://internationalmedicalcorps.org.
7. Greenberg N, Wessely S. World Psychiatry 2017;16:249-50.
8. Ventevogel P, van Ommeren M, Schilperoord M et al. Bull World Health Organ 2015;93:666.
9. Inter-Agency Standing Committee. IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. http://www.who.int/mental_ health.
10. Wallcraft J, Amering M, Freidin J et al. World Psychiatry 2011;10:229-36.

DOI:10.1002/wps.20545

WPA Position Statement on Banning the Participation of Psychiatrists in the Interrogation of Detainees

Though torture is illegal, as stipulated by a number of well-known conventions and treaties, and thus subjected to international prosecution, psychiatrists have been sometimes involved in situations connected to ill-treatment and torture which are also ethically unacceptable on any grounds.

The purpose of this Position Statement is to provide ethical guidelines for practice, in which psychiatrists are explicitly forbidden, and must refrain, from participating in any procedure linked to the interrogation of a detainee. An exception is the specific case of assessing the liability, when the person is being or has been submitted to ill-treatment or torture, and such events and possible consequences have to be documented.

The Statement is the outcome of a two-year consultation process among key stake holders and members of WPA, including a roundtable in Cape Town[1,2] (November 2016), and has been approved by the WPA General Assembly in October 2017.

- The Madrid Declaration establishes the ethical standards for psychiatric practice. Article 2 of the section on "specific situations" says: "Psychiatrists should not take part in any process of mental or physical torture, even when authorities attempt to force their involvement in such acts".
- The WPA reiterates its position that psychiatrists should not participate in, or otherwise assist or facilitate, the commission of torture of any person under any circumstance. Psychiatrists who become aware that torture has occurred, is occurring, or being planned must report it promptly to a person or persons in a position to take corrective action.
- Every person in military or civilian detention is entitled to appropriate medical care. Denial of adequate health care to a detainee may be considered as ill-treatment or torture, when this is intentionally done by state agents according to one of the purposes stated in the United Nations Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment[3].
- Psychiatrists working in detention facilities under any kind of contract, either private or public, have a duty to act for the benefit of detainees and not to do harm. Therefore, they should not participate or assist in any way, whether directly or indirectly, overtly or covertly, in the interrogation of any person deprived of liberty on behalf of military, civilian security agencies or law enforcement authorities, nor participate in any other professional intervention that would be considered coercive and against the benefit of the detainee in that context.
- "Interrogation" refers to the attempt to elicit from a person deprived of liberty information that is not intended for the therapeutic benefit of the person. It refers to a deliberate attempt to elicit information from a person deprived of liberty for the purposes of incriminating the detainee, identifying or incriminating other persons, or otherwise obtaining information that might be of value to those who control the detainee. It may involve the creation of environments intended to undermine the self-identity of the detainee or break his or her autonomy, self-determination or will, including but not limited to humiliation, debasement or punishment. It does not include interviews or other interactions with a person deprived of

liberty that have been appropriately authorized by a court or by counsel for the detainee or a medical interview that is conducted as part of a therapeutic or forensic process under demand or proper informed consent of the person deprived of liberty.
- Requesting, releasing or causing transfer of medical records or clinical data or allowing access to clinical files for interrogation purposes is a violation of professional ethics.
- Participation includes, but is not limited to, intervention in the environment where the prisoner is held, advising on ways to confuse or debilitate the person to act against his or her will, doing psychological or medical examinations to certify the health of prisoners or detainees for interrogation, being present in the interrogation room, suggesting strategies, asking or suggesting questions, or advising authorities on the use of specific techniques of interrogation with particular detainees.
- Psychiatrists may provide training to military or civilian investigative or law enforcement personnel on the adequate care of detained persons, proper cognitive interview techniques, recognizing and responding to persons with mental illnesses, the possible adverse medical and psychological effects of techniques and conditions of interrogation, and other areas within their professional expertise that will not result in harm to the physical or psychological health or well-being of the person.

Many regimes around the world put pressure on the medical profession. If the reader feels this is his/her case or wishes to contact the WPA Section on Psychological Consequences of Torture and Persecution, the relevant e-mail address is pauperez@arrakis.es.

Pau Pérez-Sales[1,2], Joost Jan den Otter[1], Lilla Hardi[1,3], Thomas Wenzel[1,4], Gabriel Diaconu[1,5], Graciela Cors[1,6], Marianne Kastrup[1,7]

[1]WPA Section on Psychological Consequences of Torture and Persecution; [2]SiR(a) Centre, Madrid, Spain; [3]Cordelia Foundation, Budapest, Hungary; [4]Medical University of Vienna, Vienna, Austria; [5]Medlife Memorial Hospital, Bucharest, Romania; [6]Latif Clinics, Shfar'am, Israel; [7]Copenhagen, Denmark

1. Milles S. Torture J 2017;27:86-7.
2. Soldz S. Torture J 2017;27:88-9.
3. United Nations. Convention against torture and other cruel, inhuman or degrading treatment or punishment. New York: United Nations, 1984.

DOI:10.1002/wps.20539

The WPA website: rich in content, excellent in performance

The WPA website (www.wpanet.org) has been thoroughly re-designed with state-of-the-art features and has gone live since October 2015. The website is based on a responsive design, which means that the dimensions of its pages now get automatically modified so as to make them properly fit the screens of various devices such as smart phones and tablets.

Media gallery on the website showcases latest photos, videos and audios. Photo gallery shows photos of international conferences and other events. The section on videos and audios contains videos of the speeches of WPA President and other officials as well as various educational videos. The homepage showcases announcements and updates on World Congresses and other WPA conferences in the near future. Latest updates regarding WPA are also given.

The homepage prominently displays the latest news from WPA Member Societies, Scientific Sections, Zonal Representatives and Affiliated Associations, along with the WPA Action Plan[1], the latest issue of the WPA official journal *World Psychiatry* and of the WPA Newsletter, along with all the past issues. The e-learning section, available for all registered users, includes various educational videos and other materials of clinically relevance.

A specially designed login for WPA Executive Committee members has been introduced in the website, where details of agendas of various meetings are available. A group talk feature is available, where all WPA Executive Committee members can join together for discussions.

The educational section of the website contains details about educational resources, essentials of the WPA international guidelines for diagnostic assessment, along with a public educational gallery which includes several articles on common mental disorders. The website also provides information on how to join a WPA Scientific Section, details on the various Sections, and a list of office bearers. The website is linked with social media such as Twitter, Facebook and the WPA YouTube channel.

World Psychiatry, the WPA official journal, is frequently visited on our website. The new impact factor of the journal is 26.561, consolidating its position as no. 1 among psychiatric journals worldwide. The journal is also now no. 1 in the overall Social Sciences Citation Index (SSCI) category. Back issues from 2002, along with translations in several languages, are available for free download.

Several recent WPA documents are available on the website, which include WPA position statements on good psychiatric practice, substance use disorders, safeguarding children, roles and responsibilities of the psychiatrist in the 21st century, mental health and well-being of psychiatrists, cultural competency in mental health care, rights of children and adults with intellectual disability, mental health in the workplace, e-mental health, and homelessness and mental health.

The site's relevance and the popularity of its contents are documented by the fact that it has constantly remained on the high Google page rank of 6. This is an algorithm used by Google that measures how many links point to a website or page, and more importantly the quality or importance of the sites that provide those links.

The performance report of our website for the year 2017 bears testimony to its increasing influence. There has been

a total of 121,776 visitors, with 87,866 new and 15,204 returning. Users visited the site from 205 different countries and from 8,003 different cities around the world. Maximum users (33%) were in the 25-34 years age group.

In the future, the launch of WPA mobile apps for Android and iOS versions may become necessary. We are planning to finalize the designs of the apps and their scope. There is also provision for live streaming using the WPA YouTube channel with a link on our website. The live streaming of the various meetings and congresses can thus reach much larger audiences. Further plans involve more aggressive search engine optimization (SEO), to increase the site publicity and reach, which will make the website come on top on Google page results.

As the Editor of the website, the WPA Secretary General works with the guidance and support of the WPA Executive Committee, Board, Council and all other WPA components[2-4]. Our collaborative work will undoubtedly make the WPA website an instrument to usher in progressive changes in psychiatry and mental health.

Roy Abraham Kallivayalil
WPA Secretary General

1. Herrman H. World Psychiatry 2017;16:329-30.
2. Kallivayalil RA. World Psychiatry 2015;14:374-5.
3. Kallivayalil RA. World Psychiatry 2017;16:114.
4. Kallivayalil RA. World Psychiatry 2017;16:330-1.

DOI:10.1002/wps.20540

The contribution of the WPA to the production of the ICD-11 chapter on mental, behavioural or neurodevelopmental disorders

The WPA has been actively supporting the World Health Organization (WHO) in the production of the chapter on mental, behavioural or neurodevelopmental disorders of the 11th edition of the International Classification of Diseases and Related Health Problems (ICD-11).

WPA Member Societies have participated in the WPA/WHO Global Survey of Psychiatrists' Attitudes Towards Mental Disorders Classification, whose results have significantly influenced the process of development of the chapter. The survey involved 4,887 psychiatrists in 44 countries, being the largest and most broadly international survey ever conducted of psychiatrists' attitudes towards the classification of mental disorders. Through the survey, psychiatrists provided strong endorsement of a focus on clinical utility, which was indeed the main objective of the development of the new diagnostic system. Over two thirds of the participants also indicated that they would prefer a system of flexible guidance allowing for cultural variation and clinical judgment, as opposed to a system based on strict operational criteria, a preference which has been actually reflected in the structure of the ICD-11.

Several WPA Member Societies and experts have been involved in ICD-11 field trials. These included the so-called formative field studies (aimed to guide decisions about the basic structure and content of the classification by exploring clinicians' conceptualization of the interrelationships among categories of mental disorders); the Internet-based field studies, implemented through the Global Clinical Practice Network (which used vignette methodologies to examine clinical decision-making in relationship to the proposed ICD-11 diagnostic categories and guidelines); and the clinic-based (or ecological implementation) field studies (conducted to assess the reliability and clinical utility of the diagnostic guidelines with real patients).

Several WPA officers and experts have served as chairpersons or members of ICD-11 Working Groups. The chairpersons have included W. Gaebel (Working Group on Psychotic Disorders), M. Maj (Working Group on Mood and Anxiety Disorders), P. Tyrer (Working Group on Personality Disorders), L. Salvador-Carulla (Working Group on Intellectual Disabilities), O. Gureje (Working Group on Somatic Distress and Dissociative Disorders) and D. Stein (Working Group on Obsessive-Compulsive and Related Disorders). Prof. M. Maj has represented the WPA in the ICD-11 International Advisory Board.

A number of lectures, symposia and workshops on various issues related to the ICD-11 development have been held at the 16th and 17th World Congresses of Psychiatry, taking place respectively in Madrid and Berlin. The workshops, in particular, represented a unique opportunity for psychiatrists from many countries to directly familiarize with the various sections of the diagnostic system and exercise in the application of the ICD-11 clinical descriptions and diagnostic guidelines.

World Psychiatry has been one of the main channels through which the international psychiatric community has been informed about the ICD-11 development. In particular, the journal has hosted some of the main papers summarizing the philosophy of the entire process and the structure of the diagnostic system, as well as many articles dealing with specific sections of the classification, as well as individual papers or forums on general topics of classification in psychiatry. Several examples can be found in recent issues of the journal[1-15]. All these contributions are freely downloadable from the WPA website (www.wpanet.org).

Mario Luciano
WHO Collaborating Center for Research and Training in Mental Health, University of Campania "L. Vanvitelli", Naples, Italy

1. Reed GM, First MB, Medina-Mora ME et al. World Psychiatry 2016;15:112-3.
2. Reed GM, Drescher J, Krueger RB et al. World Psychiatry 2016;15:205-21.
3. Gureje O, Reed GM. World Psychiatry 2016;15: 291-2.
4. Maciejewski PK, Maercker A, Boelen PA et al. World Psychiatry 2016;15:266-75.

5. Mann K, Fauth-Bühler M, Higuchi S et al. World Psychiatry 2016;15:297-8.
6. Leibenluft E. World Psychiatry 2017;16:100-1.
7. Silove D, Manicavasagar V, Pini S. World Psychiatry 2016;15:113-5.
8. Bucci P. World Psychiatry 2017;16:115-6.
9. Steardo L Jr. World Psychiatry 2017;16:331-2.
10. Jablensky A. World Psychiatry 2016;15:26-31.
11. Kendler KS. World Psychiatry 2016;15:5-12.
12. Maj M. World Psychiatry 2016;15:193-4.
13. Frances A. World Psychiatry 2016;15:32-3.
14. Tandon R. World Psychiatry 2016;15:133-4.
15. van Os J, Reininghaus U. World Psychiatry 2016; 15:118-24.

DOI:10.1002/wps.20541

Acknowledgement
This publication has been partially supported by an unrestricted educational grant from Otsuka Pharmaceutical Italy S.r.l., which is hereby gratefully acknowledged.