World Psychiatry

OFFICIAL JOURNAL OF THE WORLD PSYCHIATRIC ASSOCIATION (WPA)

Volume 15, Number 3 *October 2016*

WILEY Blackwell

NEW IMPACT FACTOR: 20.205

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 135, spanning 118 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 69 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Research Reports containing unpublished data are welcome for submission to the journal.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).

Narrowing the gap between ICD/DSM and RDoC constructs: possible steps and caveats

One century ago K. Jaspers, discussing the status of research on the neural correlates of mental disorders, stated that "we only know the end links in the chain of causation from soma to psyche and vice versa, and from both these terminal points we endeavour to advance"[1].

There has certainly been some progress from Jaspers' time. We know today some more links in that chain of causation, on one side and on the other. But the gap does remain and we – clinicians and psychopathologists on one side and neuroscientists on the other – do have to advance from our respective terminal points in order to narrow it.

Indeed, the recent debate following the publication of the DSM-5 and the launch of the Research Domain Criteria (RDoC) project has made it clear that we have today on the one hand a problem with several ICD/DSM constructs, in that they are too distant from the level of inquiry of neurosciences, and on the other a problem with at least part of RDoC constructs, in that they are somewhat distant from the level of "the actual clinical phenomena that bring patients to the clinic"[2]. This gap seems to be particularly sensible in the area of psychoses: according to B. Cuthbert, "one hears informal comments at conferences that psychosis is a 'black box' in RDoC"[2].

How can we, clinicians and psychopathologists, from our side of the chain of causation Jaspers was referring to, contribute to narrow this gap?

Actually, after an initial phase in which the RDoC project was presented (or perceived) as being in competition with current diagnostic systems, a dialogue has started between the US National Institute of Mental Health (NIMH) and both the World Health Organization (WHO) and the American Psychiatric Association (APA) (see also Sanislow[3] in this issue of the journal). Particularly productive in this respect have been the WHO/NIMH meeting held in Madrid in February 2014 entitled "Future directions in research diagnostic criteria for mental and behavioural disorders" and the APA/NIMH symposium that took place in Atlanta in May 2016 entitled "DSM-5 and RDoC: moving towards a common agenda for understanding mental disorders".

Having been the chairperson in the Madrid meeting and the discussant in the Atlanta symposium, I would like to share here with the readers of *World Psychiatry*: a) a list of the possible steps, emerging from those meetings, that we clinicians and psychopathologists can implement in order to contribute to narrow the gap between ICD/DSM and RDoC constructs, and b) a list of caveats that we will have to take into account, because not all of the assumptions which are at the basis of the RDoC project can be fully endorsed by us at the current state of development of our discipline.

A first possible step that we clinicians and psychopathologists can implement in order to narrow the above-mentioned gap is a redefinition and dissection of some complex symptoms and signs. Indeed, while the characterization of psychiatric syndromes has been repeatedly refined in the past four decades, that of psychiatric symptoms and signs has remained more or less the same, with the result that several symptoms, especially composite and heterogeneous ones (e.g., delusions, hallucinations, anhedonia), are characterized in the DSM-5 glossary in a way that is somewhat outdated and too far from the level of neuroscientific inquiry.

A second possible step is the identification of experiential intermediate phenotypes which can be added to those, mostly behavioural, that are included in the RDoC framework. Primary psychotic experiences – of which aberrant salience, in part corresponding to Jaspers' delusional atmosphere, represents a good example – may indeed be a more reasonable and meaningful target for neuroscientific inquiry than, say, delusional ideas. Of course, these primary psychotic experiences will have to be characterized in a way that is clear and reliable, and a good example of how this can be done is given by the instrument called Examination of Anomalous Self-Experience (EASE)[4], developed by J. Parnas and some other European psychopathologists.

A third possible step is a refinement of currently identified dimensions of mental disorder and the delineation of crosswalks between some of those dimensions and some RDoC constructs. This is the aim of the effort that P. Wang and D. Clarke are conducting within the APA[5], which has been described at the above-mentioned Atlanta symposium.

A fourth possible step is a more precise and detailed characterization of broader dimensional groupings or spectra, such as internalizing and externalizing, but also neuroticism. The research conducted by R. Krueger and his group on these spectra and their neurobiological correlates, also illustrated at the Atlanta symposium, and their recent effort to explore possible interrelationships between those spectra and some RDoC constructs[6], represent a good example of how this strategy can be productively pursued.

A fifth possible step is the refinement of recent attempts to delineate several stages in the development of some mental disorders, again particularly in the area of psychoses[7]. Some of those stages, especially early ones, may be closer to the level of neuroscientific inquiry than the full-blown syndromes described in the ICD and DSM.

A sixth possible step is an in-depth exploration of the dynamics within networks of symptoms. Recent research in this area, in fact, has suggested that reciprocal interactions may exist between the symptoms of a mental disorder, so that an adverse event may trigger one or more specific symptoms, which may in turn activate other symptoms, that may then modulate the expression of the former[8]. These dynamics can be relevant to neuroscientific inquiry.

This is, of course, a very tentative list, that will need to be extensively refined and to which further elements can be added.

Let's come now to the conceptual caveats, all emerging from the recent rich literature in the field of philosophy of psychiatry. Incidentally, according to T. Kuhn[9], the "recourse to philosophy and to debate over fundamentals" is one of the symptoms of the "transition from normal to extraordinary science" which characterizes paradigm shifts within a given discipline. There is no doubt that this symptom is quite evident in the current phase of development of our discipline.

The first caveat is that there are several different levels of observation and explanation of abnormal mental phenomena, and there is no *a priori* reason why one of these levels should be regarded as more fundamental than the others[10]. Of course, all abnormal experiences and behaviours are likely to be implemented through neural circuits, but this does not mean that the level of neural circuits will necessarily be the most useful and efficient at which to observe and explain those experiences and behaviours. What is the most useful and efficient level will depend upon the purpose for which the observation and explanation is required. If our purpose is to develop new psychotropic drugs, the level of neural circuits is likely to be the most efficient, but if our purpose is to develop new psychotherapies or psychosocial interventions, other levels of observation and explanation may be more useful and efficient.

A second, related but more radical, caveat is that, although all abnormal mental phenomena are likely to be implemented through neural circuits, this does not mean that those circuits will have to be necessarily themselves "faulty", and in need to be "fixed". Some forms of mental dysfunction may involve maladaptive operating rules acquired by learning[11], which may be subtended by a reconfiguration of neural activity that is not in itself "dysfunctional", although different from ordinary patterns. In other terms, the level of the dysfunction may be higher than that of neural circuits, and intervening on those circuits may not be an adequate way to act on that dysfunction.

The third caveat was already voiced by K. Jaspers one century ago: "the method of living mosaic – i.e., the idea that disease entities are mosaic-like structures composed from a variety of individual and identical pieces – turns psychopathological investigation and diagnosis into something mechanical and petrifies discovery"[1]. In other terms, whether it is really possible to decompose currently identified mental disorders into "pieces" (variables or dimensions), which recur exactly with the same characteristics and presumably with the same neurobiological correlates in all those disorders, remains to be proved. A given symptom may instead have a different meaning and different underlying pathogenetic processes depending on the overall psychopathological context within which it emerges.

A fourth caveat is that, while the problematic issue of test-retest and inter-rater reliability of psychopathological measures has been repeatedly emphasized in the past few years, the possibly even more problematic issue of test-retest and inter-laboratory reliability of neurobiological measures in psychiatry has not been equally emphasized. This is a problem to be taken into account, if the aim is to develop measures to be used in ordinary clinical practice.

On the basis of all this, what tentative conclusions can be drawn, concerning both clinical practice and progress of knowledge?

As to clinical practice, it will be empirical evidence to show whether, or to which extent, the characterization of individual patients in terms of behavioural and neurobiological variables can augment (or, as originally proposed in the RDoC project, even replace) our current psychopathological characterization in the pursuit of what remains our main objective, i.e., the efficient prediction of outcomes, in particular response to treatments. One thing are promises and statements; another thing are facts and data. What we badly need today is the latter: an empirical evidence which is solid, convincing, widely replicated and clinically relevant. The history of biological psychiatry has not started yesterday. We have seen so many biological findings which have been just forgotten after some years, without being either confirmed or disproved, and which have never been even considered for application in ordinary practice.

As to progress of knowledge, clinicians and psychopathologists on one side and neuroscientists on the other are still "exploring from opposite directions", as Jaspers was pointing out, a continent which is largely unknown[1]. Neither team can state that its effort is *a priori* more fundamental, worthwhile or scientific than the other, or claim present or future complete control over the territory. The best service we can do to our profession and our patients is to work actively to advance from our respective terminal points in the chain of causation, in a spirit of collaboration and mutual respect.

Mario Maj
Department of Psychiatry, University of Naples SUN, Naples, Italy

1. Jaspers K. Allgemeine Psychopathologie. Berlin: Springer, 1913.
2. Cuthbert BN. World Psychiatry 2014;13:28-35.
3. Sanislow C. World Psychiatry 2016;15:222-3.
4. Parnas J, Møller P, Kircher T et al. Psychopathology 2005;38:236-58.
5. Clarke DE, Kuhl EA. World Psychiatry 2014;13:314-6.
6. Krueger RF, DeYoung CG. Psychophysiology 2016;53:351-4.
7. McGorry P, Keshavan M, Goldstone S. World Psychiatry 2014;13:2011-23.
8. Borsboom D, Cramer AO. Annu Rev Clin Psychol 2013;9:91-121.
9. Kuhn TS. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.
10. Berenbaum H. J Abnorm Psychol 2013;122:894-901.
11. Bolton D. What is mental disorder? Oxford: Oxford University Press, 2008.

DOI:10.1002/wps.20370

Traditional marijuana, high-potency cannabis and synthetic cannabinoids: increasing risk for psychosis

Robin M. Murray, Harriet Quigley, Diego Quattrone, Amir Englund, Marta Di Forti

Institute of Psychiatry, Psychology and Neuroscience, King's College, De Crespigny Park, London, SE5 8AF, UK

Epidemiological evidence demonstrates that cannabis use is associated with an increased risk of psychotic outcomes, and confirms a dose-response relationship between the level of use and the risk of later psychosis. High-potency cannabis and synthetic cannabinoids carry the greatest risk. Experimental administration of tetrahydrocannabinol, the active ingredient of cannabis, induces transient psychosis in normal subjects, but this effect can be ameliorated by co-administration of cannabidiol. This latter is a constituent of traditional hashish, but is largely absent from modern high-potency forms of cannabis. Argument continues over the extent to which genetic predisposition is correlated to, or interacts with, cannabis use, and what proportion of psychosis could be prevented by minimizing heavy use. As yet, there is not convincing evidence that cannabis use increases risk of other psychiatric disorders, but there are no such doubts concerning its detrimental effect on cognitive function. All of the negative aspects are magnified if use starts in early adolescence. Irrespective of whether use of cannabis is decriminalized or legalized, the evidence that it is a component cause of psychosis is now sufficient for public health messages outlining the risk, especially of regular use of high-potency cannabis and synthetic cannabinoids.

Key words: Cannabis, psychosis, marijuana, synthetic cannabinoids, cognitive function, brain structure, genetic predisposition, early adolescence

(World Psychiatry 2016;15:195–204)

The use of cannabis has been illegal in most countries since the 1930s, but this has not deterred use[1]. Currently, cannabis is used by around 180 million people globally[2]. The tensions produced by this unsatisfactory situation have resulted in much attention being paid to the legal status of cannabis.

Possession of the drug in small quantities has been decriminalized officially in countries such as Portugal and the Netherlands, and unofficially in many more. In 2013, Uruguay became the first nation to legalize the sale, cultivation and distribution of cannabis[3]. Four US states have also legalized recreational use, and another twenty-five US states as well as Canada permit so-called "medicinal marijuana". While Uruguay has strict rules concerning access, laws vary state by state in the US, with policy being increasingly driven by entrepreneurs in search of profit, and law makers in search of taxes.

Given the above, it seems likely that consumption of cannabis will increase rather than decrease. This makes it imperative to understand the possible adverse consequences of use, even if they only affect a minority of users. In this paper we start by reviewing cannabinoids and the endocannabinoid system. We then focus on cannabis use and risk of psychiatric disorder, particularly psychosis, before touching on the effects on cognition and brain structure.

CANNABINOIDS AND THE ENDOCANNABINOID SYSTEM

Cannabis contains over one hundred cannabinoids[4], the most important of which are tetrahydrocannabinol (THC) and cannabidiol (CBD). These are produced in tiny crystal formations around the flowering tops. Recreational cannabis has been traditionally available as herb (marijuana, grass, weed) or resin (hashish, hash). In some countries such as the US it is smoked by itself, while in much of Europe it is smoked with tobacco. When smoked or inhaled, effects come on after a few minutes and last 2-3 hours; if eaten it can take 2 hours for the effects to be felt and they can last up to 8 hours.

Cannabinoids exert their effects primarily by interacting with the endocannabinoid system, which comprises endogenous ligands, their receptors, and the enzymes that synthesize and degrade them[5].

There are two specific receptors: cannabinoid receptor type-1 (CB1) and cannabinoid receptor type-2 (CB2). The CB1 receptor is widespread throughout the brain, with high concentrations in the neocortex, basal ganglia and hippocampus[6]. CB1 receptors are located pre-synaptically on the terminals of GABAergic and glutamatergic neurons, where they act homeostatically to counteract the over- or under-activity of these systems by modulating pre-synaptic neurotransmitter release[7]. The CB2 receptor, initially thought to be confined to immune cells and peripheral tissues[8], has recently also been found in the cerebellum and brain stem.

The best known endogenous cannabinoid receptor ligands are N-arachidonoylethanolamide (anandamide, AEA) and 2-arachidonoylglycerol (2-AG). These are biosynthesized post-synaptically in an activity-dependent manner before being cleared by a reuptake mechanism and enzymatic hydrolysis.

THC is responsible for the euphoria and feelings of increased sociability and insightfulness, "the high" that users enjoy. It is a partial agonist at the CB1 receptor[9]. As the endocannabinoid system normally operates "on-demand" in an activity-dependent manner[10], exogenous THC appears to overwhelm the endogenous system[11-15], with resulting lower levels, for example, of AEA[16].

Administering THC to healthy volunteers impairs learning, attention and memory in a dose-response manner[17-22]. Such impairment is likely why drivers, under the influence of cannabis, are at double risk of traffic accidents[23]. Experimental studies have also shown that a sufficiently high dose of intravenous THC can induce short-lived psychotic symptoms, including paranoia and hallucinations[19,24,25]. It also increases paranoid thoughts in a virtual reality setting[26].

CBD lacks significant affinity for the CB1 receptor[27,28], but it is able to displace THC at low nanomolar concentrations[29]. It may act antagonistically against CB1 agonists via a non-orthosteric binding site[30]. It appears to block or ameliorate many of the effects of THC. For example, the co-administration of CBD significantly reduces THC-induced tachycardia[31], the anxiogenic effects of THC[32], and the detrimental effects of THC on perception[33,34] and memory[35].

THE CHANGING NATURE OF RECREATIONAL CANNABINOIDS

The proportion of THC in the commonly used herbal cannabis (marijuana) and resin (hashish) was 3% or less in the 1960s, but subsequently it began to rise. Growers cross-bred plants to increase potency. Then, they found that preventing pollination increased THC, as in this situation the female plant converts its energy into producing more cannabinoids rather than seeds[36]. This type of cannabis is referred to as *sinsemilla*, which means "without seed" in Spanish, but is sometimes colloquially termed "skunk", because of its strong smell. Plants bred to produce a high concentration of THC cannot simultaneously produce a lot of CBD, so the product contains only traces of the latter[37].

By the early years of the 21th century, the average proportion of THC had risen to 16 and 20% in England and Holland respectively, and *sinsemilla* had taken over much of the traditional market from resin[37,38]. Similarly, Australia saw a shift towards high-potency cannabis, with mean THC around 15%[39], while in the US potency reached an average of 12% by 2014[40].

In the US states where recreational cannabis or "medicinal marijuana" have been legalized, an increasingly wide variety of products are becoming available, including oils and "edibles" such as biscuits, chocolates and cakes. Novel ways of extracting THC from the plant have produced resin oil with up to 80% THC content, while other innovations delivering high THC concentrations include "vaping" and "wax dabbing".

J.W. Huffman spent over 25 years seeking to synthetize cannabinoids for therapeutic use[41]. However, in the late 2000s, some of his compounds started to be used as "legal highs", often termed "Spice". Subsequently, the use of such synthetic cannabinoids increased dramatically, often taken sprayed on herbal mixtures. While THC is a partial agonist with weak affinity for the CB1 receptor, synthetic cannabinoids are full agonists and generally have higher affinity. Not surprisingly, they pose a greater health risk compared to plant cannabis[42-44]. A survey of 80,000 drug users showed that those who used synthetic cannabinoids were thirty times more likely to end up in an emergency unit than users of traditional cannabis[45]. Acute physical reactions include nausea and vomiting, breathlessness, hypertension, tachycardia, chest pain, and occasionally acute renal failure.

Over 200 synthetic cannabinoids have been reported available on the Internet[46]. As each has a slightly different molecular structure, they can have unpredictable side effects. Furthermore, they cannot be detected by routine drug tests, making them particularly attractive to those in prison and in the army.

PSYCHOSIS

Concern that use of cannabis might induce psychosis is not new. For example, in 1896, the Scottish psychiatrist T. Clouston visited the Cairo asylum and noted that 40 out of 253 people in the hospital had insanity attributed to the use of hashish[47]. However, by the 1960s, this view was commonly ridiculed as "reefer madness", with the implication that it was those who believed that cannabis could induce psychosis who were mad, rather than those who consumed the drug.

In the first prospective study to explore whether cannabis played a causal role in psychosis, Andréasson et al[48] traced 45,750 young men who had been asked about their drug use when they were conscripted into the Swedish army. Those who had used cannabis more than fifty times were six times more likely to develop schizophrenia over the next fifteen years than those who had never used it. Surprisingly, the findings were mostly ignored. Even *The Lancet*, which had published Andréasson et al's paper in 1987, carried an editorial in 1995 stating the prevailing view that "the smoking of cannabis, even long term, is not harmful to health"[49].

However, there has now been a raft of longitudinal prospective studies[50,51]. Nine out of twelve found that cannabis use was associated with a significantly increased risk of psychotic symptoms or psychotic illness; the remaining three showed a trend in the same direction[52-64] (Table 1). Marconi et al[65] performed a meta-analysis and showed that the more extensive the cannabis use the greater the risk for psychosis in all of the studies included. There was an odds ratio of almost four for risk of psychosis-related outcomes among the heaviest users compared to the non-users.

Is use of higher potency types of cannabis more risky than traditional forms? Di Forti et al[66] examined 410 patients with their first episode of psychotic disorder and 390 healthy controls. People using high-potency cannabis on a daily basis were five times more likely than non-users to suffer from a psychotic disorder. Use of hashish was not related to an increased risk of psychosis, possibly due to its lower THC content combined with the presence of CBD[66-68].

Table 1 Longitudinal studies concerning the role of cannabis as a risk factor for psychosis

Study	Country	Design	No. participants	Follow-up (years)	OR (95% CI) (adjusted risk)
Tien & Anthony[52]	US	Population based	4,494	1	2.4 (1.2-7.1)
Zammit et al[53]	Sweden	Conscript cohort	50,053	27	3.1 (1.7-5.5)
Manrique-Garcia et al[54]				35	1.8 (1.3-2.3)
van Os et al[55]	The Netherlands	Population based	4,045	3	2.8 (1.2-6.5)
Weiser et al[56]	Israel	Population based	9,724	4-15	2.0 (1.3-3.1)
Fergusson et al[57]	New Zealand	Birth cohort	1,265	3	1.8 (1.2-2.6)
Arseneault et al[58]	New Zealand	Birth cohort	1,034	15	4.5 (1.1-18.2)
Ferdinand et al[59]	The Netherlands	Population based	1,580	14	2.8 (1.79-4.43)
Henquet et al[60]	Germany	Population based	2,437	4	1.7 (1.1-1.5)
Wiles et al[61]	UK	Population based	8,580	1.5	1.5 (0.55-3.94)
Rössler et al[62]	Switzerland	Community survey	591	30	1.8 (0.96-3.2)
Gage et al[63]	UK	Birth cohort	1,756	2	1.1 (0.76-1.65)
Rognli et al[64]	Sweden	Cohort of discharged prisoners	6,217	5	2.6 (1.40-5.0)

Similarly, in a Dutch survey of 2,000 cannabis users, those who preferred cannabis with the highest CBD content had experienced fewer psychotic-like experiences[69]. Morgan and Curran[70], who tested hair for cannabinoids, showed that users with both detectable THC and CBD had fewer psychotic symptoms than those with only THC. Finally, in an experimental study of 48 healthy volunteers, treatment with oral CBD before administration of intravenous THC significantly reduced the occurrence of psychotic symptoms[35].

Reports have begun to emerge of cases of psychosis following the use of types of cannabis with much higher THC content, for example "wax dabs"[71]. Psychiatric symptoms are also increasingly being reported consequent upon use of synthetic cannabinoids[72]. Papanti et al[73] carried out a systematic review and reported that agitation, anxiety, paranoia and psychosis can result; these reactions are sometimes referred to as "spiceophrenia". Mounting evidence suggests that more chronic psychotic disorders can occur in persistent users of synthetic cannabinoids[74].

The existence of a cannabis psychosis distinct from schizophrenia is dubious. It is true that sudden high consumption can induce a state of acute intoxication which usually rapidly resolves. This is not uncommon with consumption of edibles, where it is more difficult to titrate one's ingestion than with smoked cannabis. Use of plant or synthetic cannabinoids for a relatively short time may induce an acute psychosis from which people recover over a period of days or weeks. But the longer use continues, the more the clinical picture merges into that of schizophrenia-like psychosis[54,64].

Nevertheless, there are differences between people with a cannabis-associated psychosis and non-using psychotic patients. Cannabis-using patients tend to have a significantly earlier onset than psychosis patients who never used cannabis[75]. One study showed a dose-response association, with daily users of high-potency cannabis experiencing their first episode of psychosis, on average, 6 years younger than never users[68].

Cannabis-using psychotic patients also tend to have higher IQ and better neurocognition than non-using psychotic patients[76,77]. They also have higher premorbid IQ and better premorbid social function[78] and are less likely to show neurological soft signs[79]. The likely explanation is that many non-drug-using schizophrenic patients have some neurodevelopmental impairment and consequent poor premorbid cognition and social function. In contrast, those who have used cannabis are often initially clever and sociable; introduced to cannabis by their friends, they are sufficiently socially adept to be able to conceal their habit from their parents.

CRITICISMS OF THE CAUSAL HYPOTHESIS

Most European and Australasian experts are now convinced that cannabis is one of a number of contributory causes of schizophrenia. However, three sceptical articles have recently appeared from North America[80-82]. We will now review the main criticisms.

One suggestion has been that those who use cannabis may be psychologically more vulnerable than those who do not. However, the Dunedin study from New Zealand controlled for psychotic symptoms at age 11, and still found a link between cannabis use and later psychotic symptoms[58].

Might some people be taking cannabis in an attempt to self-medicate symptoms of psychosis or its precursors? There is little evidence for this. A second New Zealand study, this time from Christchurch, showed that once minor psychotic symptoms developed, people tended to smoke less[83]. Further-

more, when psychotic patients are asked why they use cannabis, they report the same hedonic reasons as the rest of the population, i.e., for enjoyment[84]. Indeed, even though many know that they will develop paranoid ideas, the immediate "high" outweighs this.

A common suggestion has been that those cannabis users who go psychotic have also been using other drugs. However, a number of studies have addressed this question and not found the effect sufficient to negate the impact of cannabis[58], even when use of tobacco was accounted for[66,67].

Another argument states that cannabis use became more common in the latter part of the 20th century without an obvious change in the incidence of schizophrenia. In fact, there is little reliable information on temporal trends in the incidence of schizophrenia, so it is difficult to know whether this is true or not. To our knowledge, the only competent study spanning several decades and using the same research criteria for schizophrenia reported that the incidence doubled between 1965 and 1999, and that the proportion of schizophrenic patients using cannabis increased disproportionally compared with other psychiatric patients[85].

GENETIC PREDISPOSITION OR GENE X ENVIRONMENT INTERACTION?

A popular explanation for the association between cannabis use and psychosis is shared genetic vulnerability[80,81,86]. Cannabis-using psychotic patients not uncommonly have other relative(s) who are psychotic[87]. However, often the other psychotic member(s) of the family are also using cannabis.

One can now examine the relationship between predisposition to psychosis, as measured by the polygenic risk score for schizophrenia, and cannabis use. Power et al[88] examined the effect of the polygenic risk score on cannabis use in a large sample of Australians. The score was responsible for only a very small proportion of cannabis consumption. In a similar manner, Gage et al[89] suggested that those who used high-potency cannabis might be especially genetically predisposed to psychosis. However, Di Forti et al[90], who examined the polygenic risk score for schizophrenia in users of low- and high-potency cannabis, found no evidence to support this view.

A more likely possibility is that some individuals are more vulnerable to the psychotogenic effects of cannabis than others. No published study has yet examined a possible interaction between the polygenic risk score for schizophrenia and cannabis use in causing psychosis. However, schizophrenia patients with large, rare deletions are less likely to have comorbid cannabis abuse over their lifetime than those without such copy number variants[91]. This provides support for a threshold model of risk, with those carrying a copy number variant needing fewer adverse environmental exposures to become frankly psychotic.

Other work has examined candidate genes involved in the dopamine system. Caspi et al[92] suggested that variation in the catechol-O-methyltransferase (COMT) gene might moderate lability to cannabis-induced psychosis, but attempted replications have been inconsistent. Most recently, an experimental study[93] found no effect of this COMT polymorphism on THC-induced psychotic symptoms, but those with the val/val genotype had a greater decrement in working memory.

Two case-control studies have reported that a variant of AKT1 increases risk of psychotic illness among cannabis users, and a third has shown that those who carry this variant show a greater psychotogenic response to smoked cannabis[94-96]. Another report indicates that a variant in the D2 receptor gene may also increase psychosis risk, and that the risk is even greater in carriers of both this variant and the above-mentioned AKT1 polymorphism[97].

WHAT IS THE MECHANISM OF ACTION?

Bianconi et al[84] showed that cannabis-using psychotic patients appeared to be more sensitive to both the positive and negative effects of the drug. Similar findings have been reported from individuals at high clinical risk of developing psychosis[98]. D'Souza et al[24] showed that people with schizophrenia had a stronger reaction to the psychotogenic and cognitive effects of intravenous THC compared to healthy controls.

In animal studies, administration of THC reliably leads to increased dopamine release, but human studies have been more equivocal. One positron emission tomography (PET) study reported an increase in striatal dopamine release, but another found no significant effect. A re-analysis combining data from the two studies reported a small but significant increase in THC-induced dopamine release[99].

Several PET studies have shown that cannabis users, like other drug abusers, have a low capacity to synthetize and release striatal dopamine. However, Volkow et al[100] reported that, unlike other drug abusers, cannabis users show no alteration in striatal D2/D3 receptors. Furthermore, following an amphetamine challenge, psychotic patients who use cannabis, despite the absence of marked elevation in dopamine release, present a greater exacerbation of their symptoms compared to patients who never used it. These findings might be explained by cannabis use inducing post-synaptic dopamine supersensitivity[101], as was found by Ginovart et al[102] in their study of animals given chronic THC. This hypothesis is strengthened by the genetic evidence, reviewed above, that variation in post-synaptic genes may predispose to cannabis-associated psychosis.

OUTCOME AND TREATMENT

A recent meta-analysis showed that psychotic patients who continued cannabis use had higher relapse rates, longer hospital admissions, and more severe positive symptoms than either former users who discontinued cannabis or never-users[103].

Unfortunately, persuading cannabis users to stop is not easy. A variety of therapies, especially cognitive behavior therapy and motivational interviewing, have been tried, but so far without great success. Given tokens for cannabis-free urine tests is currently under trial. The only pharmacological treatment that has had any success is clozapine: a double-blind trial showed it to have a useful effect in diminishing craving for cannabis[104].

OTHER PSYCHIATRIC DISORDERS

Cannabis dependence

Withdrawal symptoms are usually relatively minor, because cannabis remains in the body for several weeks. However, anxiety and craving, irritability, insomnia, appetite disturbance, dysphoria and depression can develop.

Almost 10% of users will become dependent[105,106], and some claim that the rate goes as high as 17% if use starts in adolescence[107]. Certainly, cannabis dependence is an increasingly common cause of help seeking in Australia, UK, continental Europe and North America[23,108]. An Internet survey[109] reported that high-potency cannabis use was associated with an especially increased likelihood of dependence.

Depression and anxiety disorders

Cross-sectional studies report a high prevalence of depression and anxiety disorders in cannabis users[110-113], but the direction of effect remains unclear[112,114-116].

The Swedish conscript cohort showed no evidence of increased risk of depression in cannabis users[117], and systematic reviews have provided only weak evidence that cannabis use increases the risk of affective outcomes[118,119]. However, one such review concluded that cannabis use was associated with a modestly increased risk for depression, with heavy use accounting for a slightly stronger risk[120].

On the other hand, a prospective study of a large US cohort found that cannabis use was associated with increased odds of alcohol, nicotine and other drug use, but not of mood or anxiety disorders[121].

Post-traumatic stress disorder

People with post-traumatic stress disorder (PTSD) are especially likely to use cannabis[122-124], but again the nature of this relationship is uncertain. Some studies show that traumatic experiences and subsequent PTSD increase the risk of drug abuse[125,126].

Cannabis has become popular among US military veterans suffering from PTSD, and several US states have approved its medicinal use for such symptoms. However, as yet there is no evidence concerning the safety or efficacy of this practice.

Attention-deficit/hyperactivity disorder

There is a high prevalence of attention-deficit/hyperactivity disorder (ADHD) in adults seeking treatment for cannabis use disorders[127]. Prospective studies show that cannabis use increases risk of adult ADHD[128], while childhood hyperactivity/impulsivity predicts early substance use[129].

It remains controversial whether medicinal use of cannabis reduces the use of stimulant medication. A small placebo controlled trial on adults with ADHD is underway[130].

Summary

The evidence that cannabis use increases the risk of depression, anxiety disorders, PTSD or ADHD is much less convincing than that for psychosis. Indeed, it remains possible, but not proven, that cannabis may be helpful for people with PTSD and ADHD.

EFFECTS ON BRAIN AND COGNITION

There are many reports that cannabis use can alter brain structure. However, many of the studies are small, the control groups are inadequate, and most have not fully controlled for the effects of alcohol consumption (heavy cannabis users also tend to be heavy alcohol users)[131].

Two recent large studies found no main effect of cannabis on brain structure[132,133]. However, the former study[132] stands out in that the investigators found an interaction with the polygenic risk score for schizophrenia, such that individuals with a high (but not low) polygenic risk score who used cannabis did show decreased cortical thickness. Thus, people with a vulnerability to schizophrenia may also be more vulnerable to the adverse effects of cannabis on the brain.

Potency has not generally been taken into consideration in imaging studies. However, Yucel et al[134] found that those using high-potency cannabis showed hippocampal volume decrements, while those who had used preparations containing CBD did not. Similarly, in another study, cannabis users with hair samples higher in CBD were found to show less decrement in the volume of the right hippocampus than users with less CBD[135]. A further magnetic resonance imaging study found that use of high-potency cannabis was associated with disturbed white matter connections in the corpus callosum, an effect which was absent in hashish users[136].

Cannabis users perform worse on executive function, attention, verbal ability and memory tasks than non-users[137,138]. Follow-up of the Dunedin cohort showed a decline in IQ scores of six points between ages 13 and 38 among those who had been repeatedly diagnosed with cannabis use disorder[139]. However, other shorter studies have failed to replicate this finding[140,141]. Recently, in a study following up 5,115 young men and women for 25 years, past exposure to marijuana was

associated with worse verbal memory, but did not appear to affect executive function or processing speed[142].

As recently summarized by Hall and Lynskey[143], "case-control studies have generally found poorer verbal learning, memory, and attention in those who regularly use marijuana than in controls; the size of these differences usually has been related to the duration and frequency of marijuana use". Some studies suggest that cognition can recover fully when use stops[144], while others indicate that only partial recovery is possible[142].

Once again, CBD may ameliorate the negative impact of THC. A naturalistic study with 134 users found that participants using cannabis higher in CBD displayed no cognitive impairment[145]. The same group explored memory functioning in 120 users: participants whose hair tested positive for CBD and THC displayed significantly better performance than those with only THC[146].

ARE ADOLESCENTS ESPECIALLY VULNERABLE?

Some brain imaging studies have found greater brain changes in those who started heavy cannabis use in adolescence as opposed to adult life, including decreased volume in several cortical and subcortical regions, together with evidence of white matter disruption and abnormal brain activation responses to cognitive tasks[138]. These reports await confirmation.

Pope et al[147] found that the initiation of cannabis use before age 17 was associated with lower verbal IQ scores in long-term heavy cannabis users. There was also greater IQ decline in those Dunedin cohort members who started use in adolescence[148], but social decline was not so associated with age of onset[149].

Silins et al[150] reviewed 2,500 young people in Australasia and found that daily cannabis use before age 17 was associated with "clear reductions" in the likelihood of completing high school and obtaining a university degree. Similarly, a 1-year follow-up of 1,155 adolescents found that weekly cannabis use was related to poorer performance in maths and English tests[151].

In the original report from the Dunedin cohort concerning psychosis, those who started to use cannabis at age 18 or later showed only a small, non-significant increase in the risk of schizophrenia-like psychosis by age 26, but the risk increased fourfold among those starting at age 15 or earlier[58].

A possible explanation for the above reports is that the brain is still developing in those who start cannabis in their teens. Exposing the juvenile brain to the drug might permanently impair the endocannabinoid system, and impact adversely on brain and neurotransmitter function[138].

THERAPEUTIC USE OF CANNABIS AND ITS COMPONENTS

The problems associated with the recreational use of cannabis should not blind us to the possibility that some of its constituents may have useful therapeutic effects, as for example with opiates.

A German clinical trial[152] found that CBD had antipsychotic actions equivalent to a standard antipsychotic, amisulpride, in patients with schizophrenia. Furthermore, in a study of psychotic patients only partially responding to antipsychotics, the addition of CBD rather than placebo led to a significant improvement in the score on a psychosis scale[153].

Cannabinoid receptors modulate pain perception, so not surprisingly there are reports of therapeutic use of exogenous cannabinoids in human pain. A beneficial effect of smoked THC on the pain of HIV-associated neuropathy has been reported[154], and inhaled cannabis was found to provide short-term relief from chronic neuropathic pain, with a number-needed-to-treat of 5.6[155].

Several cannabinoid drugs are already available. For example, THC has long been used as an antiemetic. THC or a combination product of THC and CBD, marketed in some countries as an oromucosal spray (nabiximols), can be a useful option for pain or painful spasms in patients with multiple sclerosis[156,157]. CBD may be effective in the treatment of some patients with epilepsy[158-161], but the data are insufficient to provide definitive evidence[162].

CONCLUSIONS

As there is no good animal model of psychosis, it is difficult to conclusively prove any environmental cause. Thus, it is unclear what changes an exogenous cannabinoid would need to induce in an animal in order to provide definitive proof that cannabis can cause psychosis. Given the lack of an equivalent of painting tobacco tar on mice to demonstrate its carcinogenicity, is it sensible to wait for absolute proof that exogenous cannabinoids are a component cause of psychosis?

Gage et al[163], who exhaustively scrutinized the epidemiological literature for possible confounding, bias, misclassification, reverse causation and other explanations for the association, concluded that "epidemiologic studies provide strong enough evidence to warrant a public health message that cannabis use can increase the risk of psychotic disorders".

Of course, it is important not to overstate our knowledge in any public health campaign. For example, there is still uncertainty over the extent to which cannabis use can induce psychosis in the absence of genetic vulnerability. There remains argument over the proportion of psychosis that could be prevented if nobody used cannabis; estimates range from 8 to 24%[66]. The effects of cannabis on the brain also remain to be clarified. Moreover, we need to take care that public education does not get confused with the highly charged debate for and against decriminalization or legalization[164].

On the other hand, changes in legislation in several countries provide "natural" experiments concerning the effects of population exposure to cannabis. Will legalization result in an increase in consumption? Early reports are contradictory[165,166].

Will liberalization of laws lead to use of more potent forms of cannabis, or will it popularize safer varieties? Will educational campaigns focusing on the risks of regular use of high-potency cannabis or synthetic cannabinoids be effective? Will diminution of legal constraints on adult use result in greater use by those in their early teens who seem most susceptible to adverse effects? Will the mental health and addiction services be able to cope? It is important that researchers take the opportunity to monitor changes in the legal status of cannabis use and their effects on mental health.

REFERENCES

1. Room R. Cannabis policy: moving beyond stalemate. Oxford: Oxford University Press, 2010.
2. United Nations Office on Drugs and Crime. World drug report 2015. New York: United Nations, 2015.
3. Room R. Legalizing a market for cannabis for pleasure: Colorado, Washington, Uruguay and beyond. Addiction 2014;109:345-51.
4. Hanus LO. Pharmacological and therapeutic secrets of plant and brain (endo)cannabinoids. Med Res Rev 2009;29:213-71.
5. Mechoulam R, Parker LA. The endocannabinoid system and the brain. Annu Rev Psychol 2013;64:21-47.
6. Howlett AC, Bidaut-Russell M, Devane WA et al. The cannabinoid receptor: biochemical, anatomical and behavioral characterization. Trends Neurosci 1990;13:420-3.
7. Katona I, Freund TF. Multiple functions of endocannabinoid signaling in the brain. Annu Rev Neurosci 2012;35:529-58.
8. Piomelli D. The molecular logic of endocannabinoid signalling. Nat Rev Neurosci 2003;4:873-84.
9. Pertwee RG. The diverse CB1 and CB2 receptor pharmacology of three plant cannabinoids: delta9-tetrahydrocannabinol, cannabidiol and delta9-tetrahydrocannabivarin. Br J Pharmacol 2008;153:199-215.
10. Lutz B. On-demand activation of the endocannabinoid system in the control of neuronal excitability and epileptiform seizures. Biochem Pharmacol 2004;68:1691-8.
11. Bocker KB, Gerritsen J, Hunault CC et al. Cannabis with high delta9-THC contents affects perception and visual selective attention acutely: an event-related potential study. Pharmacol Biochem Behav 2010;96:67-74.
12. D'Souza DC, Fridberg DJ, Skosnik PD et al. Dose-related modulation of event-related potentials to novel and target stimuli by intravenous Delta(9)-THC in humans. Neuropsychopharmacology 2012;37:1632-46.
13. Ilan AB, Gevins A, Coleman M et al. Neurophysiological and subjective profile of marijuana with varying concentrations of cannabinoids. Behav Pharmacol 2005;16:487-96.
14. Murray RM, Morrison PD, Henquet C et al. Cannabis, the mind and society: the hash realities. Nat Rev Neurosci 2007;8:885-95.
15. Morrison PD, Nottage J, Stone JM et al. Disruption of frontal theta coherence by Delta9-tetrahydrocannabinol is associated with positive psychotic symptoms. Neuropsychopharmacology 2011;36:827-36.
16. Morgan CJ, Page E, Schaefer C et al. Cerebrospinal fluid anandamide levels, cannabis use and psychotic-like symptoms. Br J Psychiatry 2013;202:381-2.
17. Miller LL, Cornett TL. Marijuana: dose effects on pulse rate, subjective estimates of intoxication, free recall and recognition memory. Pharmacol Biochem Behav 1978;9:573-7.
18. Tinklenberg J, Melges F, Hollister L et al. Marijuana and immediate memory. Nature 1970;226:1171-2.
19. Morrison PD, Zois V, McKeown DA et al. The acute effects of synthetic intravenous Delta9-tetrahydrocannabinol on psychosis, mood and cognitive functioning. Psychol Med 2009;39:1607-16.
20. Hart CL, Ilan AB, Gevins A et al. Neurophysiological and cognitive effects of smoked marijuana in frequent users. Pharmacol Biochem Behav 2010; 96:333-41.
21. Nordstrom BR, Hart CL. Assessing cognitive functioning in cannabis users: cannabis use history an important consideration. Neuropsychopharmacology 2006;31:2798-9.
22. Schoeler T, Bhattacharyya S. The effect of cannabis use on memory function: an update. Subst Abuse Rehabil 2013;4:11-27.
23. Hall W. What has research over the past two decades revealed about the adverse health effects of recreational cannabis use? Addiction 2015;110: 19-35.
24. D'Souza DC, Abi-Saab WM, Madonick S et al. Delta-9-tetrahydrocannabinol effects in schizophrenia: implications for cognition, psychosis, and addiction. Biol Psychiatry 2005;57:594-608.
25. D'Souza DC, Ranganathan M, Braley G et al. Blunted psychotomimetic and amnestic effects of delta-9-tetrahydrocannabinol in frequent users of cannabis. Neuropsychopharmacology 2008;33:2505-16.
26. Freeman D, Dunn G, Murray RM et al. How cannabis causes paranoia: using the intravenous administration of 9-tetrahydrocannabinol (THC) to identify key cognitive mechanisms leading to paranoia. Schizophr Bull 2015;41:391-9.
27. Bisogno T, Hanus L, De Petrocellis L et al. Molecular targets for cannabidiol and its synthetic analogues: effect on vanilloid VR1 receptors and on the cellular uptake and enzymatic hydrolysis of anandamide. Br J Pharmacol 2001;134:845-52.
28. Thomas A, Ross RA, Saha B et al. 6"-Azidohex-2"-yne-cannabidiol: a potential neutral, competitive cannabinoid CB1 receptor antagonist. Eur J Pharmacol 2004;487:213-21.
29. Thomas A, Baillie GL, Phillips AM et al. Cannabidiol displays unexpectedly high potency as an antagonist of CB1 and CB2 receptor agonists in vitro. Br J Pharmacol 2007;150:613-23.
30. McPartland JM, Duncan M, Di Marzo V et al. Are cannabidiol and Delta(9)-tetrahydrocannabivarin negative modulators of the endocannabinoid system? A systematic review. Br J Pharmacol 2015;172:737-53.
31. Karniol IG, Shirakawa I, Kasinski N et al. Cannabidiol interferes with the effects of delta9-tetrahydrocannabinol in man. Eur J Pharmacol 1974;28: 172-7.
32. Zuardi AW, Shirakawa I, Finkelfarb E et al. Action of cannabidiol on the anxiety and other effects produced by delta 9-THC in normal subjects. Psychopharmacology 1982;76:245-50.
33. Hindocha C, Freeman TP, Schafer G et al. Acute effects of delta-9-tetrahydrocannabinol, cannabidiol and their combination on facial emotion recognition: a randomised, double-blind, placebo-controlled study in cannabis users. Eur Neuropsychopharmacol 2015;25:325-34.
34. Leweke FM, Schneider U, Radwan M et al. Different effects of nabilone and cannabidiol on binocular depth inversion in man. Pharmacol Biochem Behav 2000;66:175-81.
35. Englund A, Morrison PD, Nottage J et al. Cannabidiol inhibits THC-elicited paranoid symptoms and hippocampal-dependent memory impairment. J Psychopharmacol 2013;27:19-27.
36. Potter DJ. A review of the cultivation and processing of cannabis (Cannabis sativa L.) for production of prescription medicines in the UK. Drug Test Anal 2014;6:31-8.
37. Hardwick S, King LA. Home Office cannabis potency study 2008. St. Albans: Home Office Scientific Development Branch, 2008.
38. Pijlman FT, Rigter SM, Hoek J et al. Strong increase in total delta-THC in cannabis preparations sold in Dutch coffee shops. Addict Biol 2005;10: 171-80.
39. Swift W, Wong A, Li KM et al. Analysis of cannabis seizures in NSW, Australia: cannabis potency and cannabinoid profile. PLoS One 2013;8: e70052.
40. ElSohly MA, Mehmedic Z, Foster S et al. Changes in cannabis potency over the last 2 decades (1995-2014): analysis of current data in the United States. Biol Psychiatry 2016;79:613-9.
41. Seely KA, Lapoint J, Moran JH et al. Spice drugs are more than harmless herbal blends: a review of the pharmacology and toxicology of synthetic cannabinoids. Prog Neuropsychopharmacol Biol Psychiatry 2012;39:234-43.
42. Spaderna M, Addy PH, D'Souza DC. Spicing things up: synthetic cannabinoids. Psychopharmacology 2013;228:525-40.
43. European Monitoring Center for Drugs and Drug Addiction. European drug report 2016. Lisbon: European Monitoring Center for Drugs and Drug Addiction, 2016.
44. Tait RJ, Caldicott D, Mountain D et al. A systematic review of adverse events arising from the use of synthetic cannabinoids and their associated treatment. Clin Toxicol 2016;54:1-13.
45. Winstock A, Lynskey M, Borschmann R, et al. Risk of emergency medical treatment following consumption of cannabis or synthetic cannabinoids in a large global sample. J Psychopharmacol 2015;29:698-703.
46. Schifano F, Orsolini L, Duccio Papanti G et al. Novel psychoactive substances of interest for psychiatry. World Psychiatry 2015;14:15-26.

47. Clouston TS. The Cairo Asylum – Dr. Warnock on hasheesh insanity. Br J Psychiatry 1896;42:790-5.
48. Andréasson S, Engström A, Allebeck P et al. Cannabis and schizophrenia. A longitudinal study of Swedish conscripts. Lancet 1987;330:1483-6.
49. Anonymous. Deglamorising cannabis. Lancet 1995;346:1241.
50. Gage SH, Hickman M, Zammit S. Association between cannabis and psychosis: epidemiologic evidence. Biol Psychiatry 2016;79:549-56.
51. Murray RM, Di Forti M. Cannabis and psychosis: what degree of proof do we require? Biol Psychiatry 2016;79:514-5.
52. Tien AY, Anthony JC. Epidemiological analysis of alcohol and drug use as risk factors for psychotic experiences. J Nerv Ment Dis 1990;178:473-80.
53. Zammit S, Allebeck P, Andréasson S et al. Self reported cannabis use as a risk factor for schizophrenia in Swedish conscripts of 1969: historical cohort study. BMJ 2002;325:1199.
54. Manrique-Garcia E, Zammit S, Dalman C et al. Cannabis, schizophrenia and other non-affective psychoses: 35 years of follow-up of a population-based cohort. Psychol Med 2012;42:1321-8.
55. van Os J, Bak M, Hanssen M et al. Cannabis use and psychosis: a longitudinal population-based study. Am J Epidemiol 2002;156:319-27.
56. Weiser M, Knobler HY, Noy S et al. Clinical characteristics of adolescents later hospitalized for schizophrenia. Am J Med Genet 2002;114:949-55.
57. Fergusson DM, Horwood LJ, Swain-Campbell NR. Cannabis dependence and psychotic symptoms in young people. Psychol Med 2003;33:15-21.
58. Arseneault L, Cannon M, Poulton R et al. Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study. BMJ 2002;325:1212-3.
59. Ferdinand RF, Sondeijker F, van der Ende J et al. Cannabis use predicts future psychotic symptoms, and vice versa. Addiction 2005;100:612-8.
60. Henquet C, Krabbendam L, Spauwen J et al. Prospective cohort study of cannabis use, predisposition for psychosis, and psychotic symptoms in young people. BMJ 2005;330:11.
61. Wiles NJ, Zammit S, Bebbington P et al. Self-reported psychotic symptoms in the general population: results from the longitudinal study of the British National Psychiatric Morbidity Survey. Br J Psychiatry 2006;188:519-26.
62. Rössler W, Hengartner MP, Angst J et al. Linking substance use with symptoms of subclinical psychosis in a community cohort over 30 years. Addiction 2012;107:1174-84.
63. Gage SH, Hickman M, Heron J et al. Associations of cannabis and cigarette use with psychotic experiences at age 18: findings from the Avon Longitudinal Study of Parents and Children. Psychol Med 2014;44:3435-44.
64. Rognli EB, Berge J, Håkansson A et al. Long-term risk factors for substance-induced and primary psychosis after release from prison. A longitudinal study of substance users. Schizophr Res 2015;168:185-90.
65. Marconi A, Di Forti M, Lewis CM et al. Meta-analysis of the association between the level of cannabis use and risk of psychosis. Schizophr Bull 2016;42:1262-9.
66. Di Forti M, Marconi A, Carra E et al. Proportion of patients in south London with first-episode psychosis attributable to use of high potency cannabis: a case-control study. Lancet Psychiatry 2015;2:233-8.
67. Di Forti M, Morgan C, Dazzan P et al. High-potency cannabis and the risk of psychosis. Br J Psychiatry 2009;195:488-91.
68. Di Forti M, Sallis H, Allegri F et al. Daily use, especially of high-potency cannabis, drives the earlier onset of psychosis in cannabis users. Schizophr Bull 2014;40:1509-17.
69. Schubart CD, Sommer IE, van Gastel WA et al. Cannabis with high cannabidiol content is associated with fewer psychotic experiences. Schizophr Res 2011;130:216-21.
70. Morgan CJ, Curran HV. Effects of cannabidiol on schizophrenia-like symptoms in people who use cannabis. Br J Psychiatry 2008;192:306-7.
71. Pierre JM, Gandal M, Son M. Cannabis-induced psychosis associated with high potency "wax dabs". Schizophr Res 2016;172:211-2.
72. Castaneto MS, Gorelick DA, Desrosiers NA et al. Synthetic cannabinoids: epidemiology, pharmacodynamics, and clinical implications. Drug Alcohol Depend 2014;144:12-41.
73. Papanti D, Schifano F, Botteon G et al. "Spiceophrenia": a systematic overview of "spice"-related psychopathological issues and a case report. Hum Psychopharmacol 2013;28:379-89.
74. Fattore L. Synthetic cannabinoids – further evidence supporting the relationship between cannabinoids and psychosis. Biol Psychiatry 2016;79:539-48.
75. Large M, Sharma S, Compton MT et al. Cannabis use and earlier onset of psychosis: a systematic meta-analysis. Arch Gen Psychiatry 2011;68:555-61.
76. Loberg EM, Helle S, Nygard M et al. The cannabis pathway to non-affective psychosis may reflect less neurobiological vulnerability. Front Psychiatry 2014;5:159.
77. Arnold C, Allott K, Farhall J et al. Neurocognitive and social cognitive predictors of cannabis use in first-episode psychosis. Schizophr Res 2015;168:231-7.
78. Ferraro L, Russo M, O'Connor J et al. Cannabis users have higher premorbid IQ than other patients with first onset psychosis. Schizophr Res 2013;150:129-35.
79. Ruiz-Veguilla M, Callado LF, Ferrin M. Neurological soft signs in patients with psychosis and cannabis abuse: a systematic review and meta-analysis of paradox. Curr Pharm Des 2012;18:5156-64.
80. Ksir C, Hart CL. Cannabis and psychosis: a critical overview of the relationship. Curr Psychiatry Rep 2016;18:12.
81. Hill M. Perspective: Be clear about the real risks. Nature 2015;525:S14.
82. Haney M, Evins AE. Does cannabis cause, exacerbate or ameliorate psychiatric disorders? An oversimplified debate discussed. Neuropsychopharmacology 2016;41:393-401.
83. Fergusson DM, Boden JM, Horwood LJ. Psychosocial sequelae of cannabis use and implications for policy: findings from the Christchurch Health and Development Study. Soc Psychiatry Psychiatr Epidemiol 2015;50:1317-26.
84. Bianconi F, Bonomo M, Marconi A et al. Differences in cannabis-related experiences between patients with a first episode of psychosis and controls. Psychol Med 2016;46:995-1003.
85. Boydell J, van Os J, Caspi A et al. Trends in cannabis use prior to first presentation with schizophrenia, in South-East London between 1965 and 1999. Psychol Med 2006;36:1441-6.
86. Ksir C, Hart CL. Correlation still does not imply causation. Lancet Psychiatry 2016;3:401.
87. McGuire PK, Jones P, Harvey I et al. Morbid risk of schizophrenia for relatives of patients with cannabis-associated psychosis. Schizophr Res 1995;15:277-81.
88. Power RA, Verweij KJ, Zuhair M et al. Genetic predisposition to schizophrenia associated with increased use of cannabis. Mol Psychiatry 2014;19:1201-4.
89. Gage SH, Munafò MR, MacLeod J et al. Cannabis and psychosis. Lancet Psychiatry 2015;2:380.
90. Di Forti M, Vassos E, Lynskey M et al. Cannabis and psychosis - Authors' reply. Lancet Psychiatry 2015;2:382.
91. Martin AK, Robinson G, Reutens D et al. Cannabis abuse and age at onset in schizophrenia patients with large, rare copy number variants. Schizophr Res 2014;155:21-5.
92. Caspi A, Moffitt TE, Cannon M et al. Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-O-methyltransferase gene: longitudinal evidence of a gene x environment interaction. Biol Psychiatry 2005;57:1117-27.
93. Tunbridge EM, Dunn G, Murray RM et al. Genetic moderation of the effects of cannabis: catechol-O-methyltransferase (COMT) affects the impact of Delta9-tetrahydrocannabinol (THC) on working memory performance but not on the occurrence of psychotic experiences. J Psychopharmacol 2015;29:1146-51.
94. van Winkel R, van Beveren NJ, Simons C. AKT1 moderation of cannabis-induced cognitive alterations in psychotic disorder. Neuropsychopharmacology 2011;36:2529-37.
95. Di Forti M, Iyegbe C, Sallis H et al. Confirmation that the AKT1 (rs2494732) genotype influences the risk of psychosis in cannabis users. Biol Psychiatry 2012;72:811-6.
96. Morgan CJ, Freeman TP, Powell J et al. AKT1 genotype moderates the acute psychotomimetic effects of naturalistically smoked cannabis in young cannabis smokers. Transl Psychiatry 2016;6:e738.
97. Colizzi M, Iyegbe C, Powell J et al. Interaction between functional genetic variation of DRD2 and cannabis use on risk of psychosis. Schizophr Bull 2015;41:1171-82.
98. Gill KE, Poe L, Azimov N et al. Reasons for cannabis use among youths at ultra high risk for psychosis. Early Interv Psychiatry 2015;9:207-10.
99. Bossong MG, Mehta MA, van Berckel BN et al. Further human evidence for striatal dopamine release induced by administration of 9-tetrahydrocannabinol (THC): selectivity to limbic striatum. Psychopharmacology 2015;232:2723-9.

100. Volkow ND, Wang GJ, Telang F et al. Decreased dopamine brain reactivity in marijuana abusers is associated with negative emotionality and addiction severity. Proc Natl Acad Sci USA 2014;111:E3149-56.
101. Murray RM, Mehta M, Di Forti M. Different dopaminergic abnormalities underlie cannabis dependence and cannabis-induced psychosis. Biol Psychiatry 2014;75:430-1.
102. Ginovart N, Tournier BB, Moulin-Sallanon M, et al. Chronic Delta(9)-tetrahydrocannabinol exposure induces a sensitization of dopamine D(2)/(3) receptors in the mesoaccumbens and nigrostriatal systems. Neuropsychopharmacology 2012;37:2355-67.
103. Schoeler T, Monk A, Sami MB et al. Continued versus discontinued cannabis use in patients with psychosis: a systematic review and meta-analysis. Lancet Psychiatry 2016;3:215-25.
104. Brunette MF, Dawson R, O'Keefe CD et al. A randomized trial of clozapine vs. other antipsychotics for cannabis use disorder in patients with schizophrenia. J Dual Diagn 2011;7:50-63.
105. Budney AJ, Roffman R, Stephens RS et al. Marijuana dependence and its treatment. Addict Sci Clin Pract 2007;4:4-16.
106. Lopez-Quintero C, Perez de los Cobos J, Hasin DS et al. Probability and predictors of transition from first use to dependence on nicotine, alcohol, cannabis, and cocaine: results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Drug Alcohol Depend 2011;115:120-30.
107. Anthony J. The epidemiology of cannabis dependence. In: Roffman RA, Stephens RS (eds). Cannabis dependence: its nature, consequences and treatment. Cambridge: Cambridge University Press, 2006:58-105.
108. Health & Social Care Information Centre. Statistics on drug misuse. England 2014. London: Health & Social Care Information Centre, 2014.
109. Freeman T, Winstock A. Examining the profile of high-potency cannabis and its association with severity of cannabis dependence. Psychol Med 2015;45:3181-9.
110. Toftdahl NG, Nordentoft M, Hjorthøj C. Prevalence of substance use disorders in psychiatric patients: a nationwide Danish population-based study. Soc Psychiatry Psychiatr Epidemiol 2016;51:129-40.
111. Degenhardt L, Hall W, Lynskey M. The relationship between cannabis use, depression and anxiety among Australian adults: findings from the National Survey of Mental Health and Well-Being. Soc Psychiatry Psychiatr Epidemiol 2001;36:219-27.
112. Feingold D, Weiser M, Rehm J et al. The association between cannabis use and anxiety disorders: results from a population-based representative sample. Eur Neuropsychopharmacol 2016;26:493-505.
113. Kedzior KK, Laeber LT. A positive association between anxiety disorders and cannabis use or cannabis use disorders in the general population – a meta-analysis of 31 studies. BMC Psychiatry 2014;14:1-22.
114. Chen C-Y, Wagner AF, Anthony CJ. Marijuana use and the risk of major depressive episode. Soc Psychiatry Psychiatr Epidemiol 2002;37:199-206.
115. Feingold D, Weiser M, Rehm J et al. The association between cannabis use and mood disorders: a longitudinal study. J Affect Disord 2015;172:211-8.
116. Crippa JA, Zuardi AW, Martín-Santos R et al. Cannabis and anxiety: a critical review of the evidence. Hum Psychopharmacol 2009;24:515-23.
117. Manrique-Garcia E, Zammit S, Dalman C et al. Cannabis use and depression: a longitudinal study of a national cohort of Swedish conscripts. BMC Psychiatry 2012;12:112.
118. Degenhardt L, Hall W, Lynskey M. Exploring the association between cannabis use and depression. Addiction 2003;98:1493-504.
119. Moore THM, Zammit S, Lingford-Hughes A et al. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. Lancet 2007;370:319-28.
120. Lev-Ran S, Roerecke M, Le Foll B et al. The association between cannabis use and depression: a systematic review and meta-analysis of longitudinal studies. Psychol Med 2014;44:797-810.
121. Blanco C, Hasin DS, Wall MM et al. Cannabis use and risk of psychiatric disorders: prospective evidence from a US national longitudinal study. JAMA Psychiatry 2016;73:388-95.
122. Cougle JR, Bonn-Miller MO, Vujanovic AA et al. Posttraumatic stress disorder and cannabis use in a nationally representative sample. Psychol Addict Behav 2011;25:554-8.
123. Kevorkian S, Bonn-Miller MO, Belendiuk K et al. Associations among trauma, posttraumatic stress disorder, cannabis use, and cannabis use disorder in a nationally representative epidemiologic sample. Psychol Addict Behav 2015;29:633-8.
124. Kilpatrick DG, Acierno R, Saunders B et al. Risk factors for adolescent substance abuse and dependence: data from a national sample. J Consult Clin Psychol 2000;68:19-30.
125. Chilcoat HD, Breslau N. Posttraumatic stress disorder and drug disorders: testing causal pathways. Arch Gen Psychiatry 1998;55:913-7.
126. Vlahov D, Galea S, Resnick H et al. Increased use of cigarettes, alcohol, and marijuana among Manhattan, New York, residents after the September 11th terrorist attacks. Am J Epidemiol 2002;155:988-96.
127. Notzon DP, Pavlicova M, Glass A et al. ADHD is highly prevalent in patients seeking treatment for cannabis use disorders. J Atten Disord (in press).
128. Fergusson DM, Boden JM. Cannabis use and adult ADHD symptoms. Drug Alcohol Depend 2008;95:90-6.
129. Chang Z, Lichtenstein P, Larsson H. The effects of childhood ADHD symptoms on early-onset substance use: a Swedish twin study. J Abnorm Child Psychol 2012;40:425-35.
130. Experimental Medicine in ADHD – Cannabinoids (EMA-C). https://clinicaltrials.gov/.
131. Weiland BJ, Thayer RE, Depue BE et al. Daily marijuana use is not associated with brain morphometric measures in adolescents or adults. J Neurosci 2015;35:1505-12.
132. French L, Gray C, Leonard G et al. Early cannabis use, polygenic risk score for schizophrenia and brain maturation in adolescence. JAMA Psychiatry 2015;72:1002-11.
133. Pagliaccio D, Barch DM, Bogdan R et al. Shared predisposition in the association between cannabis use and subcortical brain structure. JAMA Psychiatry 2015;72:994-1001.
134. Yucel M, Lorenzetti V, Suo C et al. Hippocampal harms, protection and recovery following regular cannabis use. Transl Psychiatry 2016;6:e710.
135. Demirakca T, Sartorius A, Ende G et al. Diminished gray matter in the hippocampus of cannabis users: possible protective effects of cannabidiol. Drug Alcohol Depend 2011;114:242-5.
136. Rigucci S, Marques TR, Di Forti M et al. Effect of high-potency cannabis on corpus callosum microstructure. Psychol Med 2016;46:841-54.
137. Schoeler T, Kambeitz J, Behlke I et al. The effects of cannabis on memory function in users with and without a psychotic disorder: findings from a combined meta-analysis. Psychol Med 2016;46:177-88.
138. Volkow ND, Swanson JM, Evins AE et al. Effects of cannabis use on human behavior, including cognition, motivation, and psychosis: a review. JAMA Psychiatry 2016;73:292-7.
139. Meier MH, Caspi A, Ambler A, et al. Persistent cannabis users show neuropsychological decline from childhood to midlife. Proc Natl Acad Sci USA 2012;109:E2657-64.
140. Mokrysz C, Landy R, Gage SH et al. Are IQ and educational outcomes in teenagers related to their cannabis use? A prospective cohort study. J Psychopharmacol 2016;30:159-68.
141. Jackson NJ, Isen JD, Khoddam R et al. Impact of adolescent marijuana use on intelligence: results from two longitudinal twin studies. Proc Natl Acad Sci USA 2016;113:E500-8.
142. Auer R, Vittinghoff E, Yaffe K et al. association between lifetime marijuana use and cognitive function in middle age: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. JAMA Intern Med 2016;176:352-61.
143. Hall W, Lynskey M. Long-term marijuana use and cognitive impairment in middle age. JAMA Intern Med 2016;176:362-3.
144. Schreiner AM, Dunn ME. Residual effects of cannabis use on neurocognitive performance after prolonged abstinence: a meta-analysis. Exp Clin Psychopharmacol 2012;20:420-9.
145. Morgan CJ, Schafer G, Freeman TP et al. Impact of cannabidiol on the acute memory and psychotomimetic effects of smoked cannabis: naturalistic study. Br J Psychiatry 2010;197:285-90.
146. Morgan CJ, Gardener C, Schafer G et al. Sub-chronic impact of cannabinoids in street cannabis on cognition, psychotic-like symptoms and psychological well-being. Psychol Med 2012;42:391-400.
147. Pope HG Jr, Gruber AJ, Hudson JI et al. Early-onset cannabis use and cognitive deficits: what is the nature of the association? Drug Alcohol Depend 2003;69:303-10.
148. Meier MH, Hill ML, Small PJ et al. Associations of adolescent cannabis use with academic performance and mental health: a longitudinal study of upper middle class youth. Drug Alcohol Depend 2015;156:207-12.
149. Cerdá M, Moffitt TE, Meier MH et al. Persistent cannabis dependence and alcohol dependence represent risks for midlife economic and social problems: a longitudinal cohort study. Clin Psychol Sci (in press).

150. Silins E, Horwood LJ, Patton GC et al. Young adult sequelae of adolescent cannabis use: an integrative analysis. Lancet Psychiatry 2014;1:286-93.
151. Stiby AI, Hickman M, Munafo MR et al. Adolescent cannabis and tobacco use and educational outcomes at age 16: birth cohort study. Addiction 2015;110:658-68.
152. Leweke FM, Piomelli D, Pahlisch F et al. Cannabidiol enhances anandamide signaling and alleviates psychotic symptoms of schizophrenia. Transl Psychiatry 2012;2:e94.
153. McGuire P. A double-blind, randomised, placebo-controlled, parallel group trial of cannabidiol as adjunctive therapy in the first line treatment of schizophrenia or related psychotic disorder. Presented at the 5th Schizophrenia International Research Society Conference, Florence, April 2016.
154. Abrams DI, Jay C, Shade S et al. Cannabis in painful HIV-associated sensory neuropathy. A randomized placebo-controlled trial. Neurology 2007; 68:515-21.
155. Andreae MH, Carter GM, Shaparin N et al. Inhaled cannabis for chronic neuropathic pain: a meta-analysis of individual patient data. J Pain 2015; 16:1221-32.
156. Patti F, Messina S, Solaro C et al. Efficacy and safety of cannabinoid oromucosal spray for multiple sclerosis spasticity. J Neurol Neurosurg Psychiatry 2016;87:944-51.
157. Koppel BS, Brust JCM, Fife T et al. Systematic review: Efficacy and safety of medical marijuana in selected neurologic disorders: Report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology 2014;82:1556-63.
158. Brust JC, Ng SK, Hauser AW et al. Marijuana use and the risk of new onset seizures. Transactions of the American Clinical and Climatological Association 1992;103:176-81.
159. Gloss D, Vickrey B. Cannabinoids for epilepsy. Cochrane Database Syst Rev 2012;6.
160. Porter BE, Jacobson C. Report of a parent survey of cannabidiol-enriched cannabis use in pediatric treatment-resistant epilepsy. Epilepsy Behav 2013;29:574-7.
161. Press CA, Knupp KG, Chapman KE. Parental reporting of response to oral cannabis extracts for treatment of refractory epilepsy. Epilepsy Behav 2015;45:49-52.
162. Friedman D, Devinsky O. Cannabinoids in the treatment of epilepsy. N Engl J Med 2015;373:1048-58.
163. Gage SH, Hickman M, Zammit S. Association between cannabis and psychosis: epidemiologic evidence. Biol Psychiatry 2016;79:549-56.
164. Large M. The need for health warnings about cannabis and psychosis. Lancet Psychiatry 2016;3:188-9.
165. Hasin DS, Wall M, Keyes KM et al. Medical marijuana laws and adolescent marijuana use in the USA from 1991 to 2014: results from annual, repeated cross-sectional surveys. Lancet Psychiatry 2015;2:601-8.
166. Shi Y, Lenzi M, An R. Cannabis liberalization and adolescent cannabis use: a cross-national study in 38 countries. PLoS One 2015;10:e0143562.

DOI:10.1002/wps.20341

Disorders related to sexuality and gender identity in the ICD-11: revising the ICD-10 classification based on current scientific evidence, best clinical practices, and human rights considerations

Geoffrey M. Reed[1,2], Jack Drescher[3], Richard B. Krueger[4], Elham Atalla[5], Susan D. Cochran[6], Michael B. First[4], Peggy T. Cohen-Kettenis[7], Iván Arango-de Montis[8], Sharon J. Parish[9], Sara Cottler[10], Peer Briken[11], Shekhar Saxena[1]

[1]Department of Mental Health and Substance Abuse, World Health Organization, Geneva, Switzerland; [2]School of Psychology, Universidad Nacional Autónoma de México, Mexico City, Mexico; [3]Department of Psychiatry, New York Medical College, New York, NY, USA; [4]Department of Psychiatry, Columbia University, College of Physicians and Surgeons, New York State Psychiatric Institute and New York Presbyterian Hospital, New York, NY, USA; [5]Primary Care and Public Health Directorate, Ministry of Health, Manama, Bahrain; [6]Fielding School of Public Health, University of California, Los Angeles, CA, USA; [7]Department of Medical Psychology, VU University Medical Centre, and Center of Expertise on Gender Dysphoria, Amsterdam, The Netherlands; [8]Instituto Nacional de Psiquiatria Ramón de la Fuente Muñiz, Mexico City, Mexico; [9]Departments of Medicine and Psychiatry, Weill Cornell Medical College and New York Presbyterian Hospital/Westchester Division, White Plains, NY, USA; [10]Department of Reproductive Health and Research, World Health Organization, Geneva, Switzerland; [11]Institute for Sex Research and Forensic Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

In the World Health Organization's forthcoming eleventh revision of the International Classification of Diseases and Related Health Problems (ICD-11), substantial changes have been proposed to the ICD-10 classification of mental and behavioural disorders related to sexuality and gender identity. These concern the following ICD-10 disorder groupings: F52 Sexual dysfunctions, not caused by organic disorder or disease; F64 Gender identity disorders; F65 Disorders of sexual preference; and F66 Psychological and behavioural disorders associated with sexual development and orientation. Changes have been proposed based on advances in research and clinical practice, and major shifts in social attitudes and in relevant policies, laws, and human rights standards. This paper describes the main recommended changes, the rationale and evidence considered, and important differences from the DSM-5. An integrated classification of sexual dysfunctions has been proposed for a new chapter on Conditions Related to Sexual Health, overcoming the mind/body separation that is inherent in ICD-10. Gender identity disorders in ICD-10 have been reconceptualized as Gender incongruence, and also proposed to be moved to the new chapter on sexual health. The proposed classification of Paraphilic disorders distinguishes between conditions that are relevant to public health and clinical psychopathology and those that merely reflect private behaviour. ICD-10 categories related to sexual orientation have been recommended for deletion from the ICD-11.

Key words: International Classification of Diseases, ICD-11, sexual health, sexual dysfunctions, transgender, gender dysphoria, gender incongruence, paraphilic disorders, sexual orientation, DSM-5

(World Psychiatry 2016;15:205–221)

The World Health Organization (WHO) is in the process of developing the eleventh revision of the International Classification of Diseases and Related Health Problems (ICD-11). The ICD-11 is expected to be approved by the World Health Assembly in May 2018. The ICD-10 was approved in 1990, making the current period between revisions the longest in the history of the ICD.

In 2007, the WHO Department of Mental Health and Substance Abuse appointed the International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders, to provide policy guidance and consultation throughout the development of the ICD-11 classification of mental and behavioural disorders[1]. As the revision process advanced, a series of Working Groups in different disorder content areas were also appointed to review available evidence and develop recommendations regarding needed revisions in specific diagnostic groupings[2].

From early in the revision process, it was clear that there were a series of complex and potentially controversial issues associated with the ICD-10 categories related to sexuality and gender identity, including the following disorder groupings: F52 Sexual dysfunctions, not caused by organic disorder or disease; F64 Gender identity disorders; F65 Disorders of sexual preference; and F66 Psychological and behavioural disorders associated with sexual development and orientation. During the more than 25 years since the approval of ICD-10, there have been substantial advances in research relevant to these categories, as well as major changes in social attitudes and in relevant policies, laws, and human rights standards.

Due to the complexity of this context and the need to take a broad perspective in order to develop scientifically and clinically sound recommendations that would facilitate access to health services, the WHO Departments of Mental Health and Substance Abuse and of Reproductive Health and Research have worked together to propose revisions in these areas. The two WHO departments appointed a joint Working Group on Sexual Disorders and Sexual Health to assist in the development of specific recommendations.

The first task of the Working Group was to review available scientific evidence as well as relevant information on health policies and health professionals' experience with the ICD-10 diagnostic categories identified above. These issues were examined within various settings, including primary care and specialist health care settings, as well as social service and forensic contexts. Also considered were human rights issues pertinent to diagnostic classification in each of the areas under the Working Group's purview. The Working Group was also asked to review what were then proposals for the American

Psychiatric Association's DSM-5[3], and to consider the clinical utility of those proposals and their suitability for global implementation in various settings. Finally, the Working Group was asked to prepare specific proposals, including the placement and organization of categories, and to draft diagnostic guidelines for the ICD-11 recommended diagnostic categories, in line with the overall ICD revision requirements[2].

The following sections describe the main recommended changes for the above-mentioned four areas in the ICD-11 as compared to ICD-10. The ICD-10 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders[4], the version intended for use by specialist mental health professionals, is used as the frame of reference for this comparison. The rationale for changes, the evidence considered, and specific comments on differences from DSM-5 are also provided.

PROPOSED CHANGES TO F52 SEXUAL DYSFUNCTIONS, NOT CAUSED BY ORGANIC DISORDER OR DISEASE

The ICD-10 classification of Sexual dysfunctions (F52) is based on a Cartesian separation of "organic" and "non-organic" conditions. Sexual dysfunctions considered "non-organic" are classified in the ICD-10 chapter on Mental and Behavioural Disorders, and most "organic" sexual dysfunctions are classified in the chapter on Diseases of the Genitourinary System. However, substantial evidence has accumulated since ICD-10's publication indicating that the origin and maintenance of sexual dysfunctions frequently involves the interaction of physical and psychological factors[5]. The ICD-10 classification of sexual dysfunctions is therefore not consistent with current, more integrative clinical approaches in sexual health[6-9].

The Working Group on Sexual Disorders and Sexual Health has proposed an integrated classification of sexual dysfunctions for ICD-11 (see Table 1) that is more closely informed by current evidence and best practices, to be included in a new ICD-11 chapter on Conditions Related to Sexual Health[10]. The proposed integrated classification encompasses the sexual dysfunctions listed in the ICD-10 chapter on Mental and Behavioural Disorders and many of those currently found in the chapter on Diseases of the Genitourinary System[11].

In the proposed diagnostic guidelines for ICD-11, sexual response is described as a complex interaction of psychological, interpersonal, social, cultural, physiological and gender-influenced processes. Any of these factors may contribute to the development of sexual dysfunctions[8], which are described as syndromes that comprise the various ways in which people may have difficulty experiencing personally satisfying, non-coercive sexual activities.

The proposed ICD-11 diagnostic guidelines organize Sexual dysfunctions into four main groups: Sexual desire and arousal dysfunctions; Orgasmic dysfunctions; Ejaculatory dysfunctions; and Other specified sexual dysfunctions. In addition, a separate grouping of Sexual pain disorders has been proposed. Where possible, categories in the proposed classification of sexual dysfunctions apply to both men and women, emphasizing commonalities in sexual response[12,13] (e.g., Hypoactive sexual desire dysfunction, Orgasmic dysfunction), without ignoring established sex differences in the nature of these experiences[14]. Men and women exhibit similar central nervous system pathways of activation and deactivation and similar neurotransmitter activity related to sexual desire. Dynamic alterations of sexual response are similarly modulated and reinforced by behaviour, experience and neuroplasticity. Separate sexual dysfunctions categories for men and women are provided where sex differences are related to distinct clinical presentations (e.g., Female sexual arousal dysfunction in women as compared to Erectile dysfunction in men).

The proposed guidelines indicate that, in order to be considered a sexual dysfunction, the problem or difficulty should generally: a) have been persistent or recurrent over a period of at least several months; b) occur frequently, although it may fluctuate in severity; and c) be associated with clinically significant distress. However, in cases where there is an immediate acute cause of the sexual dysfunction (e.g., a radical prostatectomy or injury to the spinal cord in the case of Erectile dysfunction; breast cancer and its treatment in Female sexual arousal dysfunction), it may be appropriate to assign the diagnosis even though the duration requirement has not been met, in order to initiate treatment.

The proposed diagnostic guidelines make clear that there is no normative standard for sexual activity. "Satisfactory" sexual functioning is defined as being satisfying to the individual, i.e. the person is able to participate in sexual activity and in a sexual relationship as desired. If the individual is satisfied with his/her pattern of sexual experience and activity, even if it is different from what may be satisfying to other people or what is considered normative in a given culture or subculture, a sexual dysfunction should not be diagnosed. Unrealistic expectations on the part of a partner, a discrepancy in sexual desire between partners, or inadequate sexual stimulation are not valid bases for a diagnosis of sexual dysfunction.

The proposed ICD-11 classification uses a system of harmonized qualifiers that may be applied across categories to identify the important clinical characteristics of the sexual dysfunctions. A *temporal qualifier* indicates whether the sexual dysfunction is *lifelong*, i.e. the person has always experienced the dysfunction from the time of initiation of relevant sexual activity, or *acquired*, i.e. the onset of the sexual dysfunction has followed a period of time during which the person did not experience it. A *situational qualifier* is used to indicate whether the dysfunction is *generalized*, i.e. the desired response is absent or diminished in all circumstances, including masturbation, or *situational*, i.e. the desired response is absent or diminished in some circumstances but not in others (e.g., with some partners or in response to some stimuli).

An innovative feature of the proposed ICD-11 classification of Sexual dysfunctions and Sexual pain disorders, and an

Table 1 Classification of Sexual dysfunctions in ICD-11 (proposed), ICD-10 and DSM-5

Proposed ICD-11	ICD-10	DSM-5	Comments
***Chapter*:** Conditions Related to Sexual Health ***Grouping*:** Sexual dysfunctions	***Chapter*:** Mental and Behavioural Disorders ***Grouping*:** Behavioural syndromes associated with physiological disturbances and physical factors ***Subgrouping*:** Sexual dysfunction, not caused by organic disorder or disease ***Chapter*:** Diseases of the Genitourinary System ***Grouping*:** Diseases of male genital organs ***Subgrouping*:** Other disorders of penis ***Grouping*:** Noninflammatory disorders of female genital tract ***Subgrouping*:** Pain and other conditions associated with female genital organs and menstrual cycle	***Grouping*:** Sexual dysfunctions	• In ICD-11, Sexual dysfunctions have been included in a new chapter called Conditions Related to Sexual Health. • ICD-11 Sexual dysfunctions proposals represent an integrated classification, including conditions listed in Mental and Behavioural Disorders chapter in ICD-10 and many of those currently found in Diseases of the Genitourinary System. • In ICD-11, there are four main groupings of sexual dysfunctions: Sexual desire and arousal dysfunctions; Orgasmic dysfunctions; Ejaculatory dysfunctions; and Other specified sexual dysfunctions. There is another separate grouping of Sexual pain disorders. • DSM-5 classification of Sexual dysfunctions excludes those caused by a nonsexual medical disorder, by the effects of a substance or medication, or by a medical condition. ICD-11 classification allows for a diagnosis of Sexual dysfunction when it represents an independent focus of treatment; contributory factors may be coded using etiological qualifiers.
***Category*:** Hypoactive sexual desire dysfunction	***Category*:** Lack or loss of sexual desire	***Category***: Female sexual interest/arousal disorder; Male hypoactive sexual desire disorder	• In ICD-11, Hypoactive sexual desire dysfunction can be applied to both men and women; In DSM-5, Female sexual interest/arousal disorder is separated from Male hypoactive sexual desire disorder.
Category*: Recommended for deletion*	***Category***: Sexual aversion	***Category***: Not included	• In ICD-11, the ICD-10 category Sexual aversion would be classified under Sexual pain-penetration disorder or under Specific phobia, depending on specific nature of symptoms. • In DSM-5, that category would similarly be classified as Genital-pelvic pain/penetration disorder or under Specific phobia.
Category: Female sexual arousal dysfunction	***Category***: Failure of genital response; Lack of sexual enjoyment	***Category***: Female sexual interest/arousal disorder	• In ICD-11, separate categories are provided for men and women to replace ICD-10 Failure of genital response, because of anatomical and physiological differences that underlie distinct clinical presentations. • In ICD-11, the psychological component of arousal involved in ICD-10 Lack of sexual enjoyment is also subsumed in women under Female sexual arousal dysfunction.
Category: Erectile dysfunction	***Category***: Failure of genital response; Impotence of organic origin	***Category***: Erectile disorder	• In ICD-11, separate categories are provided for men and women to replace ICD-10 Failure of genital response, because of anatomical and physiological differences that underlie distinct clinical presentations. • ICD-11 includes "organic" Erectile dysfunctions.
Category: Orgasmic dysfunction	***Category***: Orgasmic dysfunction	***Category***: Female orgasmic disorder	• In ICD-11, Orgasmic dysfunction can be applied to both men and women. • In ICD-11, there is a distinction between subjective experience of orgasm in men and ejaculation.

Table 1 Classification of Sexual dysfunctions in ICD-11 (proposed), ICD-10 and DSM-5 *(continued)*

Proposed ICD-11	ICD-10	DSM-5	Comments
Category: Early ejaculation	***Category***: Premature ejaculation	***Category***: Premature (early) ejaculation	• Terminology in ICD-11 changed from Premature ejaculation to Early ejaculation.
Category: Delayed ejaculation	***Category***: Orgasmic dysfunction	***Category***: Delayed ejaculation	• DSM-5 does not distinguish between subjective experience of orgasm and ejaculation in men.
Category: Other specified sexual dysfunction	***Category***: Other sexual dysfunction, not caused by organic disorder or disease; Other specified disorders of penis; Other specified conditions associated with female genital organs and menstrual cycle	***Category***: Other specified sexual dysfunction	• DSM-5 classification of Sexual dysfunctions excludes those caused by a nonsexual medical disorder, by the effects of a substance or medication, or by a medical condition. ICD-11 classification allows for a diagnosis of Sexual dysfunction when it represents an independent focus of treatment; contributory factors may be coded using etiological qualifiers.
Category: Unspecified sexual dysfunction	***Category***: Unspecified sexual dysfunction, not caused by organic disorder or disease; Disorder of penis, unspecified; Unspecified condition associated with female genital organs and menstrual cycle	***Category***: Unspecified sexual dysfunction	• DSM-5 classification of Sexual dysfunctions excludes those caused by a nonsexual medical disorder, by the effects of a substance or medication, or by a medical condition. ICD-11 classification allows for a diagnosis of Sexual dysfunction when it represents an independent focus of treatment; contributory factors may be coded using etiological qualifiers.
Category: Sexual pain-penetration disorder (in separate grouping of Sexual pain disorders)	***Category***: Nonorganic vaginismus; Vaginismus (organic)	***Category***: Genito-pelvic pain/penetration disorder	• In ICD-11, Sexual pain penetration disorder includes Vaginismus and excludes Dyspareunia and Vulvodynia, which are classified in the Genitourinary chapter. • In DSM-5, Genito-pelvic pain/penetration disorder groups includes Dyspareunia and Vulvodynia if it occurs during penetration attempts or vaginal intercourse.

important one for a system that does not attempt to divide "organic" and "non-organic" dysfunctions, is a system of *etiological qualifiers* that may be applied to these categories. These qualifiers are not mutually exclusive, and as many may be applied as are considered to be relevant and contributory in a particular case. Proposed qualifiers include the following:

- *Associated with disorder or disease classified elsewhere, injury or surgical treatment* (e.g., diabetes mellitus, depressive disorders, hypothyroidism, multiple sclerosis, female genital mutilation, radical prostatectomy)[15-19];
- *Associated with a medication or substance* (e.g., selective serotonin reuptake inhibitors, histamine-2 receptor antagonists, alcohol, opiates, amphetamines)[20,21];
- *Associated with lack of knowledge* (e.g., about the individual's own body, sexual functioning, and sexual response)[22];
- *Associated with psychological or behavioural factors* (e.g., negative attitudes toward sexual activity, adverse past sexual experiences, poor sleep hygiene, overwork)[23,24];
- *Associated with relationship factors* (e.g., relationship conflict, lack of romantic attachment)[25,26];
- *Associated with cultural factors* (e.g., culturally-based inhibitions about the expression of sexual pleasure, the belief that loss of semen can lead to weakness, disease or death)[27,28].

Other changes that have been proposed include the elimination of the ICD-10 category F52.7 Excessive sexual drive from the classification of Sexual dysfunctions. The ICD-10 category F52.0 Loss or lack of sexual desire is more specifically categorized in ICD-11 as Hypoactive sexual desire dysfunction in women and men, Female sexual arousal dysfunction in women, or Erectile dysfunction in men. The ICD-10 category F52.10 Sexual aversion is classified in ICD-11 under Sexual pain-penetration disorder or under the grouping of Anxiety and fear-related disorders if it is used to describe a phobic response. The ICD-10 category F52.11 Lack of sexual enjoyment, which the ICD-10 indicates is more common in women, is captured primarily in the ICD-11 under Female sexual arousal dysfunction. Other possible reasons for lack of sexual enjoyment, including hypohedonic orgasm and painful orgasm[29], would be classified under Other specified sexual dysfunctions. The ICD-10 category F52.2 Failure of genital response is separated into two categories: Female sexual arousal dysfunction in women, and Erectile dysfunction in men.

Comparison with DSM-5

The proposed classification of sexual dysfunctions in ICD-11 is different from the DSM-5 in its attempt to integrate

dysfunctions that may have a range of etiological or contributory dimensions. The DSM-5 acknowledges that an array of factors may be relevant to etiology and treatment and may contribute to sexual dysfunctions; these include partner, relationship, individual vulnerability, cultural, religious, and medical factors. At the same time, the DSM-5 indicates that, if a sexual dysfunction is caused by a nonsexual medical disorder, the effects of a substance or medication, or a medical condition, a diagnosis of Sexual dysfunction would not be assigned. This is logical given the DSM-5's purpose as a classification of mental and behavioural disorders (even though it differs from the approach that DSM-5 has taken to Sleep-wake disorders and Neurocognitive disorders). Because ICD-11 is a classification of all health conditions, it provides the possibility for greater integration. The proposed ICD-11 classification allows for assigning a Sexual dysfunction diagnosis in situations in which this is an independent focus of treatment, regardless of presumed etiology. The presence of a variety of contributory factors may be recorded using the etiological qualifiers.

The DSM-5 has combined dysfunctions of sexual desire and sexual arousal in women in the category Female sexual interest/arousal disorder[30], which has proved to be quite controversial[31-35]. In contrast, the proposed ICD-11 category Hypoactive sexual desire dysfunction can be applied to both men and women, while Female sexual arousal dysfunction is classified separately. The separation of desire and arousal in women into distinct dysfunctions is supported by several lines of evidence, including genetic evidence from twin studies[36], studies of specific single nucleotide polymorphisms and the use of serotonergic antidepressant medications[37,38], and neuroimaging studies[39]. There is also evidence that Hypoactive desire disorder in women and men respond to similar treatments[40], and that these are different from treatments that are effective for Female sexual arousal disorder[41-43]. Although there is significant comorbidity between desire and arousal dysfunction, the overlap of these conditions does not mean that they are one and the same; research suggests that management should be targeted toward their distinct features[44].

The proposed classification of sexual pain in ICD-11 provides the possibility of identifying specific types of pain syndromes without excluding those in which another medical condition is considered to be contributory. The DSM-5 category Genito-pelvic pain/penetration disorder includes vaginismus, dyspareunia and vulvodynia not completely attributable to other medical conditions. A similar category of Sexual pain-penetration disorder has been proposed for ICD-11, but it does not include dyspareunia and vulvodynia, which have been retained as separate categories in the ICD-11 genitourinary chapter. These syndromes are characterized by different etiologies, occur in different populations, and have distinct treatment approaches[45-47].

Finally, the DSM-IV-TR category Male orgasmic disorder has been replaced in DSM-5 by Delayed ejaculation. This decision seems to have been largely based on a Medline search that indicated infrequent usage of terminology including orgasm as opposed to terminology specifying ejaculation for male disorders[48]. Another rationale for DSM-5 to modify the term was the small number of cases of male orgasmic disorder seen in clinical practice[49]. However, this was not only a modification of terminology but rather the lumping of two separate phenomena into a single category. The proposed ICD-11 classification of Sexual dysfunctions emphasizes the subjective experience of orgasm and separates it from the ejaculatory phenomenon, consistent with available research[50].

PROPOSED CHANGES TO F64 GENDER IDENTITY DISORDERS

Over the past several years, a range of civil society organizations as well as the governments of several Member States and the European Union Parliament have urged the WHO to remove categories related to transgender identity from its classification of mental disorders in the ICD-11[51-53].

One impetus for this advocacy has been an objection to the stigmatization that accompanies the designation of any condition as a mental disorder in many cultures and countries. The WHO Department of Mental Health and Substance Abuse is committed to a variety of efforts to reduce the stigmatization of mental disorders[54]. However, the stigmatization of mental disorders *per se* would not be considered a sufficient reason to eliminate or move a mental disorder category. The conditions listed in the ICD Mental and Behavioural Disorders chapter are intended to assist in the identification of people who need mental health services and in the selection of appropriate treatments[1], in fulfillment of WHO's public health objectives.

Nevertheless, there is substantial evidence that the current nexus of stigmatization of transgender people and of mental disorders has contributed to a doubly burdensome situation for this population, which raises legitimate questions about the extent to which the conceptualization of transgender identity as a mental disorder supports WHO's constitutional objective of "the attainment by all peoples of the highest possible level of health"[55]. Stigma associated with the intersection of transgender status and mental disorders appears to have contributed to precarious legal status, human rights violations, and barriers to appropriate health care in this population[56-58].

The WHO's 2015 report on *Sexual health, human rights, and the law*[58] indicates that, in spite of recent progress, there are still very few non-discriminatory, appropriate health services available and accessible to transgender people. Health professionals often do not have the necessary competence to provide services to this population, due to a lack of appropriate professional training and relevant health system standards[59-61]. Limited access to accurate information and appropriate health services can contribute to a variety of negative behavioural and mental health outcomes among transgender people, including increased HIV-related risk behaviour, anxiety, depression, substance abuse, and suicide[62-65]. Additionally,

Table 2 Classification of conditions related to gender identity in ICD-11 (proposed), ICD-10 and DSM-5

Proposed ICD-11	ICD-10	DSM-5	Comments[71,72]
Chapter: Conditions Related to Sexual Health ***Grouping***: Gender incongruence	***Chapter***: Mental and Behavioural Disorders ***Grouping***: Disorders of adult personality and behaviour ***Subgrouping***: Gender identity disorders	***Grouping***: Gender dysphoria	• ICD-11 does not classify Gender incongruence as a mental and behavioural disorder; Gender dysphoria is listed as a mental disorder in DSM-5. • ICD-11's primary focus is experience of incongruence between experienced gender and assigned sex; DSM-5 emphasizes distress related to gender identity through name of category and criteria.
Category: Gender incongruence of adolescence and adulthood	***Category***: Transsexualism	***Category***: Gender dysphoria in adolescents and adults	• ICD-11 contains four broad essential features and two are required for diagnosis; DSM-5 contains six criteria and two are required for diagnosis. • In ICD-11, distress and functional impairment are described as common associated features, particularly in disapproving social environments, but are not required; DSM-5 requires clinically significant distress or impairment for diagnosis. • ICD-11 requires a duration of several months; DSM-5 requires six months.
Recommended for deletion	***Category***: Dual-role transvestism	*Not included*	• Recommended for deletion from ICD-11 due to lack of public health or clinical relevance (not in DSM-5).
Category: Gender incongruence of childhood	***Category***: Gender identity disorder of childhood	***Category***: Gender dysphoria in children	• ICD-11 contains three essential features, all of which are required for diagnosis; DSM-5 contains eight diagnostic criteria, six of which must be present. • In ICD-11, distress and functional impairment are described as common associated features, particularly in disapproving social environments, but are not required; DSM-5 requires clinically significant distress or impairment for diagnosis. • ICD-11 requires a duration of two years, suggesting that the diagnosis cannot be made before approximately age 5; DSM-5 requires six months and does not set a lower age limit.
Recommended for deletion	***Category***: Other gender identity disorders	***Category***: Other specified gender dysphoria	• Recommended for deletion in ICD-11 to prevent misuse for clinical presentations involving only gender variance.
Recommended for deletion	***Category***: Gender identity disorder, unspecified	***Category***: Unspecified gender dysphoria	• Recommended for deletion in ICD-11 to prevent misuse for clinical presentations involving only gender variance.

many transgender people self-administer hormones of dubious quality obtained through illicit markets or online without medical supervision[66,67], with potentially serious health consequences[68-70]. For example, in a recent study of 250 transgender people in Mexico City, nearly three-quarters of participants had used hormones, and nearly half of these had begun using them without medical supervision[71].

In spite of WHO's concerted advocacy for mental health parity[54], a primary mental disorder diagnosis can exacerbate problems for transgender people in accessing health services, particularly those that are not considered to be mental health services. Even in countries that recognize the need for transgender-related health services and where professionals with relevant expertise are relatively available, private and public insurers often specifically exclude coverage for these services[58]. Classification as a mental disorder has also contributed to the perception that transgender people must be treated by psychiatric specialists, further restricting access to services that could reasonably be provided at other levels of care.

In most countries, the provision of health services requires the diagnosis of a health condition that is specifically related to those services. If no diagnosis were available to identify transgender people who were seeking related health services, these services would likely become even less available than they are now[72,73]. Thus, the Working Group on Sexual Disorders and Sexual Health has recommended retaining gender incongruence diagnoses in the ICD-11 to preserve access to health services, but moving these categories out of the ICD-11 chapter on Mental and Behavioural Disorders (see Table 2). After consideration of a variety of placement options[72], these

categories have been provisionally included in the proposed new ICD-11 chapter on Conditions Related to Sexual Health.

The Working Group has recommended reconceptualizing the ICD-10 category F64.0 Transsexualism as Gender incongruence of adolescence and adulthood[72] and the ICD-10 category F64.2 Gender identity disorder of childhood as Gender incongruence of childhood[73]. The proposed diagnostic requirements for Gender incongruence of adolescence and adulthood include the continuous presence for at least several months of at least two of the following features: a) a strong dislike or discomfort with primary or secondary sex characteristics due to their incongruity with the experienced gender; b) a strong desire to be rid of some or all of one's primary or secondary sex characteristics (or, in adolescence, anticipated secondary sex characteristics); c) a strong desire to have the primary or secondary characteristics of the experienced gender; and d) a strong desire to be treated (to live and be accepted as) a person of the experienced gender. As in the ICD-10, the diagnosis of Gender incongruence of adolescence and adulthood cannot be assigned before the onset of puberty. The duration requirement is reduced from two years in ICD-10 to several months in ICD-11.

The ICD-11 abandons ICD-10 terms such as "opposite sex" and "anatomic sex" in defining the condition, using more contemporary and less binary terms such as "experienced gender" and "assigned sex". Unlike ICD-10, the proposed ICD-11 diagnostic guidelines do not implicitly presume that all individuals seek or desire full transition services to the "opposite" gender. The proposed guidelines also explicitly pay attention to the anticipated development of secondary sex characteristics in young adolescents who have not yet reached the last physical stages of puberty, an issue that is not addressed in ICD-10.

The proposed ICD-11 diagnostic requirements for Gender incongruence of childhood are considerably stricter than those of ICD-10, in order to avoid as much as possible the diagnosis of children who are merely gender variant. All three of the following essential features must be present: a) a strong desire to be, or an insistence that the child is, of a different gender; b) a strong dislike of the child's own sexual anatomy or anticipated secondary sex characteristics, or a strong desire to have the sexual anatomy or anticipated secondary sex characteristics of the desired gender; and c) make believe or fantasy play, toys, games, or activities and playmates that are typical of the experienced gender rather than the assigned sex. The third essential feature is not meaningful without the other two being present; in their absence it is merely a description of gender variant behaviour. These characteristics must have been present for at least two years in a prepubertal child, effectively meaning that the diagnosis cannot be assigned prior to the age of approximately 5 years. The ICD-10 does not mention a specific duration requirement or a minimum age at which it is appropriate to assign the diagnosis.

The proposed diagnostic guidelines for both Gender incongruence of adolescence and adulthood and Gender incongruence of childhood indicate explicitly that gender variant behaviour and preferences alone are not sufficient for making a diagnosis; some form of experienced anatomic incongruence is also necessary. Importantly, the diagnostic guidelines for both categories indicate that gender incongruence may be associated with clinically significant distress or impairment in social, occupational, or other important areas of functioning, particularly in disapproving social environments and where protective laws and policies are absent, but that neither distress nor functional impairment is a diagnostic requirement.

The area of transgender health is characterized by calls for change in health system responses[58,74,75], by rapid change in social attitudes in some countries, and by controversy. As a part of this work, the Working Group on Sexual Disorders and Sexual Health received proposals and opinions from a wide range of civil societies, professional organizations, and other interested parties[72,73]. The most controversial issue has been the question of whether the childhood diagnostic category should be retained[73]. The main argument advanced against retaining the category is that stigmatization associated with being diagnosed with *any* health condition – not just a mental disorder diagnosis – is potentially harmful to children who will in any case not be receiving medical interventions before puberty[76]. A more substantive critique is that, if it is the case that the problems of extremely gender-variant children arise primarily from hostile social reactions and victimization, assigning a diagnosis to the child amounts to blaming the victim[77]. This latter concern suggests a need for further research as well as a broader social conversation. The Working Group has recommended retaining the category based on the rationale that it will preserve access to treatment for this vulnerable and already stigmatized group. Treatment most often consists of specialized supportive mental health services as well as family and social (e.g., school) interventions[73], while treatments aimed at suppressing gender-variant behaviours in children are increasingly viewed as unethical.

The diagnosis also serves to alert health professionals that a transgender identity in childhood often does not develop seamlessly into an adult transgender identity. Available research instead indicates that the majority of children diagnosed with DSM-IV Gender identity disorder of childhood, which was not as strict in its requirements as those proposed for ICD-11, grow up to be cisgender (non-transgender) adults with a homosexual orientation[78-80]. In spite of the claims of some clinicians to be able to distinguish between children whose transgender identity is likely to persist into adolescence and adulthood and those likely to be gay or lesbian, there is considerable overlap between these groups in all predictors examined[80], and no valid method of making a prediction at an individual level has been published in the scientific literature. Therefore, while medical interventions are not currently recommended for prepubertal gender incongruent children, psychosocial interventions need to be undertaken with caution and based on considerable expertise so as not to limit later choices[59,81,82]. The inclusion of the category in the ICD-11 is intended to provide better opportunities for much-needed education of health professionals, the development of stand-

ards and pathways of care to help guide clinicians and family members, including adequate informed consent procedures, and future research efforts.

Finally, the ICD-10 category F64.1 Dual-role transvestism – occasionally dressing in clothing typical of another gender in order to "enjoy the temporary experience of membership of the opposite sex, but without any desire for a more permanent sex change"[4] or accompanying sexual arousal – has been recommended for deletion from the ICD-11, due to its lack of public health or clinical relevance.

Comparison with DSM-5

The most important difference between the proposals for ICD-11 and the DSM-5 is that the latter has retained the categories related to gender identity as a part of its classification of mental disorders. Both childhood and adult forms of Gender identity disorder in DSM-IV have been renamed in DSM-5 as Gender dysphoria, defined by "marked incongruence between one's experienced/expressed gender and assigned gender of at least 6 months' duration" and "clinically significant distress or impairment in social, school, or other important areas of functioning"[3]. Both the name of the DSM-5 condition – dysphoria – and the diagnostic criteria, therefore, emphasize distress and dysfunction as integral aspects of the condition. They are also the central rationale for classifying these conditions as mental disorders; without distress or dysfunction, gender dysphoria would not fulfill the requirements of DSM-5's own definition of a mental disorder.

In contrast, the proposal for ICD-11 is to include child and adult Gender incongruence categories in another chapter that explicitly integrates medical and psychological perspectives, Conditions Related to Sexual Health. The proposed ICD-11 diagnostic guidelines indicate that distress and dysfunction, although not necessary for a diagnosis of Gender incongruence, may occur in disapproving social environments and that individuals with gender incongruence are at increased risk for psychological distress, psychiatric symptoms, social isolation, school drop-out, loss of employment, homelessness, disrupted interpersonal relationships, physical injuries, social rejection, stigmatization, victimization, and violence. At the same time, particularly in countries with progressive laws and policies, young transgender people living in supportive environments still seek health services, even in the absence of distress or impairment. The ICD-11 approach provides for this.

A challenge to DSM-5 conceptualization of Gender dysphoria is, therefore, the question of whether distress and dysfunction related to the social consequences of gender variance (e.g., stigmatization, violence) can be distinguished from distress related to transgender identity itself[83,84]. A recent study of 250 transgender adults receiving services at the only publicly funded clinic in Mexico City providing comprehensive services for transgender people[71] found that distress and dysfunction associated with emerging transgender identity were very common, but not universal. However, more than three-quarters of participants reported having experienced social rejection and nearly two-thirds had experienced violence related to their gender identity during childhood or adolescence. Distress and dysfunction were more strongly predicted by experiences of social rejection and violence than by features related to gender incongruence. These data provide further support for ICD-11's conceptualization and the removal of gender incongruence from the classification of mental disorders.

Finally, there are several technical differences between the proposals for ICD-11 and DSM-5 in relation to these categories. The most substantive is that the DSM-5 diagnosis of Gender dysphoria of childhood requires a duration of only six months, in contrast to two years in the ICD-11 proposal, and does not specify a lower age limit at which the diagnosis can be applied.

PROPOSED CHANGES TO F65 DISORDERS OF SEXUAL PREFERENCE

From WHO's perspective, there is an important distinction between conditions that are relevant to public health and indicate the need for health services versus those that are simply descriptions of private behaviour without appreciable public health impact and for which treatment is neither indicated nor sought. This distinction is based on the ICD's central function as a global public health tool that provides the framework for international public health surveillance and health reporting. It is also related to the increasing use of the ICD over the past several decades by WHO Member States to structure clinical care and define eligibility for subsidized health services[1]. The regulation of private behaviour without health consequences to the individual or to others may be considered in different societies to be a matter for criminal law, religious proscription, or public morality, but is not a legitimate focus of public health or of health classification.

This requirement is particularly pertinent to the classification of atypical sexual preferences commonly referred to as paraphilias. The Working Group on Sexual Disorders and Sexual Health noted that the diagnostic guidelines provided for ICD-10's classification of Disorders of sexual preference often merely describe the sexual behaviour involved. For example, the ICD-10 diagnostic guidelines define F65.1 Fetishistic transvestism as "the wearing of clothes of the opposite sex principally to obtain sexual excitement"[4], without requiring any sort of distress or dysfunction and without reference to the public health or clinical relevance of this behaviour. This is at odds with ICD-10's general guidance for what constitutes a mental disorder and contradicts ICD-10's own statement that "social deviance or conflict alone, without personal dysfunction, should not be included in mental disorder"[4]. According to this principle, specific patterns of sexual arousal that are merely relatively unusual[85,86], but are not associated with distress,

dysfunction or harm to the individual or to others[87,88], are not mental disorders. Labeling them as such does not contribute meaningfully to public health surveillance or to the design of health services, and may create harm to individuals so labeled[89]. Thus, a major consideration for the recommended revisions for ICD-11 in this area was whether an atypical sexual arousal pattern represented a condition of public health significance and clinical importance.

The Working Group recommended that Disorders of sexual preference be renamed as Paraphilic disorders to reflect the terminology used in the current scientific literature and in clinical practice[90]. The Group proposed that the paraphilic disorders included in ICD-11 consist primarily of patterns of atypical sexual arousal that focus on non-consenting others, as these conditions could be considered to have public health implications (see Table 3). The core proposed diagnostic requirements for a Paraphilic disorder in ICD-11 are: a) a sustained, focused and intense pattern of sexual arousal – as manifested by persistent sexual thoughts, fantasies, urges, or behaviours – that involves others whose age or status renders them unwilling or unable to consent (e.g., pre-pubertal children, an unsuspecting individual being viewed through a window, an animal); and b) that the individual has acted on these thoughts, fantasies or urges or is markedly distressed by them. There is no requirement in the proposed ICD-11 diagnostic guidelines that the relevant arousal pattern be exclusive or preferential.

This conceptualization has resulted in the recommendation to retain three ICD-10 categories in this section, each labeled specifically as a disorder rather than simply naming or describing the behaviour involved. These include Exhibitionistic disorder, Voyeuristic disorder, and Pedophilic disorder. In addition, two new named categories have been proposed: Coercive sexual sadism disorder and Frotteuristic disorder.

Coercive sexual sadism disorder is defined by a sustained, focused and intense pattern of sexual arousal that involves the infliction of physical or psychological suffering on a non-consenting person. This arousal pattern has been found to be prevalent among sex offenders treated in forensic institutions[92-96] and among individuals who have committed sexually motivated homicides[97]. The new proposed nomenclature of Coercive sexual sadism disorder was selected to clearly distinguish this disorder from consensual sadomasochistic behaviours that do not involve substantial harm or risk.

Frotteuristic disorder is defined by a sustained, focused and intense pattern of sexual arousal that involves touching or rubbing against a non-consenting person in public places. Frotteurism has been found to be among the most common of paraphilic disorders[98-102] and is a significant problem in some countries[103]. It was also included in DSM-IV and has been retained in DSM-5.

In addition, the category Other paraphilic disorder involving non-consenting individuals is proposed for use when the other diagnostic requirements for a paraphilic disorder are met but the specific pattern of sexual arousal does not fit into any of the available named categories and is not sufficiently common or well researched to be included as a named category (e.g., arousal patterns involving corpses or animals).

Based on the concerns described above, the Working Group proposed that three named ICD-10 categories – F65.0 Fetishism, F65.1 Fetishistic transvestism, and F65.5 Sadomasochism – be removed from the classification. Indeed, several countries (Denmark, Sweden, Norway and Finland) have already removed these categories from their national lists of accepted ICD-10 diagnoses, in response to similar concerns[104]. Instead, the proposed additional category Other paraphilic disorder involving solitary behaviour or consenting individuals may be used when the pattern of sexual arousal does not focus on non-consenting individuals but is associated with marked distress or significant risk of injury or death (e.g., asphyxophilia, or achieving sexual arousal by restriction of breathing).

One additional requirement in the proposed diagnostic guidelines is that, when a diagnosis of Other paraphilic disorder involving solitary behaviour or consenting individuals is assigned based on distress, the distress should not be entirely attributable to rejection or feared rejection of the arousal pattern by others (e.g., a partner, family, society). In these cases, codes related to counselling interventions from the ICD-11 chapter on Factors Influencing Health Status and Contact with Health Services may be considered. These are non-disease categories that indicate reasons for clinical encounters and include Counselling related to sexual knowledge and sexual attitude, Counselling related to sexual behaviour and sexual relationships of the patient, and Counselling related to sexual behaviour and sexual relationship of the couple. These categories recognize the need for health services, including mental health services, that may be legitimately provided in the absence of diagnosable mental disorders[11].

The proposed diagnostic guidelines make clear that the mere occurrence or a history of specific sexual behaviours is insufficient to establish a diagnosis of a Paraphilic disorder. Rather, these sexual behaviours must reflect a sustained, focused, and intense pattern of paraphilic sexual arousal. When this is not the case, other causes of the sexual behaviour need to be considered. For example, many sexual crimes involving non-consenting individuals reflect actions or behaviours that may be transient or occur impulsively or opportunistically rather than reflecting either a persistent pattern of sexual arousal or any underlying mental disorder. However, sexual behaviours involving non-consenting individuals may also occur in the context of some mental and behavioural disorders, such as manic episodes or dementia, or in the context of substance intoxication. These do not satisfy the definitional requirements of a Paraphilic disorder.

The Working Group on Sexual Disorders and Sexual Health has also recommended that the proposed ICD-11 grouping of Paraphilic disorders be retained within the chapter on Mental and Behavioural Disorders rather than being moved to the proposed new chapter on Conditions Related to Sexual Health, for two main reasons. First, the assessment and treatment of Paraphilic disorders, which often takes place in forensic con-

Table 3 Classification of Paraphilic disorders in ICD-11 (proposed), ICD-10 and DSM-5

Proposed ICD-11	ICD-10	DSM-5	Comments[90]
***Chapter*:** Mental and Behavioural Disorders ***Grouping*:** Paraphilic disorders	***Chapter*:** Mental and Behavioural Disorders ***Grouping*:** Disorders of adult personality and behaviour ***Subgrouping*:** Disorders of sexual preference	***Grouping*:** Paraphilic disorders	• ICD-11 name changed to be consistent with current scientific literature and clinical practice; brings it in line with DSM-5. • ICD-11 distinguishes between conditions that are relevant to public health and clinical psychopathology on the one hand and private behaviours that are not a legitimate focus of health classification on the other. • Requirements for named Paraphilic disorders in ICD-11 are: a) a sustained, focused and intense pattern of sexual arousal that involves others whose age or status renders them unwilling or unable to consent; and b) that the individual has acted on the arousal patterns or is markedly distressed by it.
***Category*:** Exhibitionistic disorder	***Category*:** Exhibitionism	***Category***: Exhibitionistic disorder	• DSM-5 diagnosis may be assigned based on functional impairment, though without specification of how impairment is to be evaluated or based on whose perspective. ICD-11 guidelines require either action or distress; not including functional impairment is consistent with overall guidance for ICD-11 Mental and Behavioural Disorders.
Category: Voyeuristic disorder	***Category***: Voyeurism	***Category***: Voyeuristic disorder	• DSM-5 diagnosis may be assigned based on functional impairment, though without specification of how impairment is to be evaluated or based on whose perspective. ICD-11 guidelines require either action or distress; not including functional impairment is consistent with overall guidance for ICD-11 Mental and Behavioural Disorders.
Category: Pedophilic disorder	***Category***: Paedophilic disorder	***Category***: Pedophilic disorder	• DSM-5 diagnosis may be assigned based on functional impairment, though without specification of how impairment is to be evaluated or based on whose perspective. ICD-11 guidelines require either action or distress; not including functional impairment is consistent with overall guidance for ICD-11 Mental and Behavioural Disorders. • In DSM-5, diagnosis may be assigned based on the presence of "interpersonal difficulty" due to the arousal pattern, in the absence of action, distress, or functional impairment. • DSM-5 includes a variety of specifiers, which have been criticized for lack of consistency and questionable validity[91].
Category: Coercive sexual sadism disorder	*Not included*	*Not included*	• Defined by sustained, focused and intense pattern of sexual arousal that involves the infliction of physical or psychological suffering on a non-consenting person. • Not equivalent to DSM-5 Sexual sadism disorder or ICD-10 Sadomasochism, which do not distinguish between arousal patterns involving consenting and non-consenting others.
Category: Frotteuristic disorder	*Not included*	***Category***: Frotteuristic disorder	• DSM-5 diagnosis may be assigned based on functional impairment, though without specification of how impairment is to be evaluated or based on whose perspective. ICD-11 guidelines require either action or distress; not including functional impairment is consistent with overall guidance for ICD-11 Mental and Behavioural Disorders.
Recommended for deletion	***Category***: Sadomasochism	***Category***: Sexual masochism disorder	• If consensual behaviour is involved, may be classified as in ICD-11 as Other paraphilic disorder involving solitary behaviour or consenting individuals, if accompanied by marked distress that is not entirely attributable to rejection or feared rejection of the arousal pattern by others (e.g., a partner, family, society) or by significant risk of injury or death. • If arousal pattern focuses on the infliction of suffering on non-consenting individuals, may be classified in ICD-11 as Coercive sexual sadism disorder.
Not included	*Combined with Sexual masochism*	***Category***: Sexual sadism disorder	• In ICD-11, may be classified as Other paraphilic disorder involving solitary behaviour or consenting individuals, if accompanied by marked distress that is not entirely attributable to rejection or feared rejection of the arousal pattern by others (e.g., a partner, family, society) or by significant risk of injury or death.

Table 3 Classification of Paraphilic disorders in ICD-11 (proposed), ICD-10 and DSM-5 *(continued)*

Proposed ICD-11	ICD-10	DSM-5	Comments[90]
Recommended for deletion	***Category***: Fetishism	***Category***: Fetishistic disorder	• In ICD-11, may be classified as Other paraphilic disorder involving solitary behaviour or consenting individuals, if accompanied by marked distress that is not entirely attributable to rejection or feared rejection of the arousal pattern by others (e.g., a partner, family, society) or by significant risk of injury or death.
Recommended for deletion	***Category***: Fetishistic transvestism	***Category***: Transvestic disorder	• In ICD-11, may be classified as Other paraphilic disorder involving solitary behaviour or consenting individuals, if accompanied by marked distress that is not entirely attributable to rejection or feared rejection of the arousal pattern by others (e.g., a partner, family, society) or by significant risk of injury or death.
Recommended for deletion	***Category***: Multiple disorders of sexual preference	*Not included*	• This ICD-10 category was not considered to be clinically informative. Multiple paraphilic disorder diagnoses may be assigned in both ICD-11 and DSM-5.
Category: Other paraphilic disorder involving non-consenting individuals	*Not included*	*Not included*	• May be used when the diagnostic requirements for a Paraphilic disorder are met but the specific pattern of sexual arousal does not fit into available named categories (e.g., arousal patterns involving corpses or animals).
Category**:** Other paraphilic disorder involving solitary behaviour or consenting individuals	*Not included*	*Not included*	• May be used when the pattern of sexual arousal does not focus on non-consenting individuals but is associated with marked distress or significant risk of injury or death.
Recommended for deletion	***Category***: Other disorders of sexual preference	***Category***: Other specified paraphilic disorder	• Replaced in ICD-11 by above two "Other paraphilic disorder" categories, which specify whether arousal pattern involves: a) non-consenting individuals; or b) consenting individuals or solitary behaviour.
Recommended for deletion	***Category***: Disorder of sexual preference, unspecified	***Category***: Unspecified paraphilic disorder	• Recommended for deletion in ICD-11 to prevent misuse for clinical presentations involving only relatively unusual patterns of sexual arousal that are not associated with distress, dysfunction, or harm to the individual or to others.

texts, requires specialized mental health expertise. Evidence-based treatments for Paraphilic disorders are almost entirely psychological and psychiatric in nature and require substantial mental health expertise to administer. When adjunctive somatic treatments are used (e.g., anti-androgen drugs), they are controversial and legally and clinically complex and must be administered within a psychiatric framework.

Second, a substantial portion of the assessment and treatment of Paraphilic disorders relates to the civil commitment, mitigation, and treatment of specific classes of sex offenders. This is a complex and controversial legal area that must be considered in defining how Paraphilic disorders should be classified. In many countries – including the US, Germany, the UK, Canada, and other countries whose legal systems are based on the British or German systems – there are laws that allow for the civil commitment and preventive detention of certain sexual offenders who are sometimes termed sexually violent predators. These laws permit involuntary commitment of such individuals to psychiatric facilities after they have completed mandatory prison sentences, to allow for continued treatment and minimization of risk to the community where these offenders are to be released.

In countries where the constitutionality of such laws has been challenged, they have been upheld[105]. However, crucial to the finding of constitutionality has been the determination by relevant courts that a risk of dangerousness by itself is not sufficient grounds for civil commitment under such statutes. Rather, the constitutional requirement specifically rests on a finding of the presence of a mental disorder as the basis for civil commitment because it "narrows the class of persons eligible for confinement to those who are unable to control their dangerousness"[106].

Although there are continuing controversies about the application of these laws in many countries[107,108], the Working Group on Sexual Disorders and Sexual Health did not consider that moving Paraphilic disorders out of the Mental and Behavioural Disorders chapter would be an appropriate or helpful way to address these concerns.

Comparison with DSM-5

The changes proposed for Paraphilic disorders in ICD-11 represent a major departure from ICD-10, which was developed during the late 1980s. In many ways, these changes align the ICD-11 more closely with the DSM-5. At the same time, there are substantive differences between the two systems. Sexual masochism disorder, Fetishistic disorder, and Transvestic disorder are included as named mental disorders in DSM-

5, while in ICD-11 these phenomena can be diagnosed under Other paraphilic disorder involving solitary behaviour or consenting individuals only if they are associated with significant distress or significant risk of injury or death.

The duration requirement proposed for Paraphilic disorders in ICD-11 is more flexible than the six-month requirement in DSM-5, which does not appear to have specific empirical support[109]. The ICD-11 guidelines require a clinical judgment that the arousal pattern is sustained, focused, and intense, making clear that a single instance of behaviour or criminal act does not meet this requirement. Functional impairment is included relatively automatically in diagnostic criteria for DSM-5, but has not been included as a part of the proposed ICD-11 diagnostic guidelines for Paraphilic disorders, in keeping with the general principle for ICD-11 Mental and Behavioural Disorders that impairment should only be used when necessary to distinguish a disorder from normality[1].

PROPOSED CHANGES TO F66 PSYCHOLOGICAL AND BEHAVIOURAL DISORDERS ASSOCIATED WITH SEXUAL DEVELOPMENT AND ORIENTATION

The ICD-10 explicitly states that "sexual orientation by itself is not to be considered a disorder"[4]. Nevertheless, the ICD-10 grouping of Psychological and behavioural disorders associated with sexual development and orientation suggests that there do exist mental disorders uniquely linked to sexual orientation. These categories include F66.0 Sexual maturation disorder, F66.1 Egodystonic sexual orientation, and F66.2 Sexual relationship disorder (see Table 4).

The Working Group on Sexual Disorders and Sexual Health emphasized that, although the ICD-10 F66 categories mention gender identity in their definitions, historically they emerged from concerns related to sexual orientation[89]. Over the last half century, international classification systems of mental disorders, including the ICD and the DSM, but also various national and regional classifications, have gradually removed diagnostic categories that defined homosexuality *per se* as a mental disorder. This reflects emerging human rights standards[56,110], the recognition that homosexual behaviour is a widely prevalent aspect of human behaviour[111], and the lack of empirical evidence to support pathologization and medicalization of variations in sexual orientation expression[112,113].

As noted earlier, the ICD-10 also indicates that "social deviance or conflict alone, without personal dysfunction, should not be included in mental disorder"[4]. The Working Group viewed this exclusion as essential to the consideration of diagnostic categories linked to sexual orientation[89]. Given that expression of same-sex orientation continues to be heavily stigmatized in parts of the world[56,110], psychological and behavioural symptoms seen in non-heterosexual individuals may be products of persistently hostile social responses rather than expressions of inherent psychopathology. This perspective is supported by robust empirical evidence from international studies[114-116]. Violence, stigma, exclusion and discrimination linked to same-sex orientations is a worldwide phenomenon and has been documented as especially vicious, often showing a high degree of brutality[117]. In some countries, criminal law is still applied to consensual same-sex sexual activity, though international, regional and national human rights bodies have explicitly called for States to end this practice[56]. Thus, the Working Group concluded that, if a disease label is to be attached to a social condition, it is essential that the condition have demonstrable public health and clinical utility, for example by identifying a legitimate mental health need.

The core diagnostic features of F66.0 Sexual maturation disorder in the ICD-10 are: a) uncertainty about one's gender identity or sexual orientation and b) distress *about the uncertainty* rather than about the particular gender identity or sexual orientation. Research has repeatedly demonstrated that same-sex sexual orientation emerges over time[118], with the process typically beginning in late childhood or early adolescence. Often there is a substantial level of anti-gay stigma in the individual's social environment that creates stress for the individual. As distress arising from stigma cannot be considered as indicative of a mental disorder under the ICD-10 social conflict exclusion, the Working Group considered that this category conflates normative developmental patterns observed in gay, lesbian, bisexual, and transgender people with psychopathological processes.

The concept of egodystonic homosexuality (F66.1 Egodystonic sexual orientation in ICD-10) first entered mental disorders classifications in DSM-III, as part of a negotiation related to removing homosexuality *per se* from that diagnostic system[119]. The compromise was that, while homosexuality itself might not be a disorder, homosexuality could still provide the basis for a psychiatric diagnosis, but only if the individual was distressed about it. This construction was dropped from American Psychiatric Association's classification in 1987[113]. In what appears to have been a parallel process in the subsequent revisions leading to ICD-10, the concept of Egodystonic sexual orientation was incorporated in the ICD-10, approved in 1990, when the ICD-9 diagnostic category for homosexuality *per se* was removed. According to the ICD-10, it is theoretically possible to apply this category to individuals with a heterosexual orientation who wish it were otherwise, but is hard to see this as anything other than an attempt to deflect criticism regarding the purpose of the category[120].

Lesbian, gay, and bisexual individuals often report higher levels of distress than their heterosexual counterparts in international surveys, but this has been linked strongly to experiences of social rejection and stigmatization[114-116]. Because distress related to social adversity cannot be considered as indicative of a mental disorder, any more than can distress related to other socially stigmatized conditions such as poverty or physical illness, the Working Group considered the existence of this distress as lacking in evidentiary value.

F66.2 Sexual relationship disorder in ICD-10 describes a situation in which the individual's sexual orientation (or gender

Table 4 Classification of disorders related to sexual orientation in ICD-11 (proposed), ICD-10 and DSM-5

Proposed ICD-11	ICD-10	DSM-5	Comments[89]
Recommended for deletion	***Chapter*:** Mental and Behavioural Disorders ***Grouping*:** Disorders of adult personality and behaviour ***Subgrouping*:** Psychological and behavioural disorders associated with sexual development and orientation	*Not included*	• All categories in this ICD-10 grouping have been recommended for deletion. • These categories or their equivalents are not included in DSM-5, and were not included in DSM-IV. • No scientific interest in these conditions since ICD-10 was published. • No evidence-based treatments. • Working Group determined that these categories confound responses to adverse social circumstances, normal developmental patterns, and psychopathology. • If requirements for depression, anxiety, or another disorder are met, that diagnosis should be used. These diagnoses do not depend on thematic content of associated concerns. • Otherwise, Counselling related to sexuality codes from ICD-11 chapter on Factors Influencing Health Status and Contact with Health Services are more appropriate.
Recommended for deletion	***Category*:** Sexual maturation disorder	*Not included*	• ICD-10 defines category based on uncertainty about gender identity or sexual orientation, which causes anxiety or depression.
Recommended for deletion	***Category***: Egodystonic sexual orientation	*Not included*	• According to ICD-10, should be used when the gender identity or sexual preference is not in doubt, but the individual wishes it were different because of associated psychological and behavioural disorders.
Recommended for deletion	***Category***: Sexual relationship disorder	*Not included*	• According to ICD-10, should be used when the gender identity or sexual preference abnormality is responsible for difficulties in forming or maintaining a relationship with a sexual partner. • Difficulties in intimate relationships are common, occur for many reasons, and are dyadic. Working Group concluded that there was no justification for category based on the co-occurrence of an issue related to sexual orientation or gender identity with a relationship problem.
Recommended for deletion	***Category***: Other psychosexual development disorder	*Not included*	• This is a residual category for the ICD-10 grouping, which is recommended for deletion in ICD-11.
Recommended for deletion	***Category***: Psychosexual development disorder, unspecified	*Not included*	• This is a residual category for the ICD-10 grouping, which is recommended for deletion in ICD-11.
Recommended for deletion	***Qualifiers*:** *(May be applied to all categories in grouping)* • *Heterosexual* • *Homosexual* • *Bisexual* • *Other, including prepubertal*	*Not included*	• These categories specify sexual orientation of individual receiving any of the above ICD-10 diagnoses, which are recommended for deletion.

identity) has created a disturbance in a primary sexual relationship. Difficulties in intimate relationships are common, occur for many reasons, and are, by their nature, dyadic. The Working Group concluded that there was no justification for creating a mental disorder category specifically based on the co-occurrence of an issue related to sexual orientation or gender identity with a relationship problem.

The Working Group's review concluded that gay, lesbian, and bisexual people receive mental health services for the same reasons that heterosexual people do, and also could find no evidence that concerns about sexual orientation that accompany other mental disorders such as depression or anxiety require different methods of treatment[121]. Further, there are no evidence-based practices related to the F66 categories, and therapeutic attempts to change sexual orientation are considered to be outside the scope of ethical practice[122]. There is also a risk that misattributing symptoms of other mental disorders to conflicts about sexual orientation may interfere with appropriate treatment selection[89].

Moreover, the F66 categories have attracted no scientific interest since the ICD-10 was published. The Working Group conducted a search of Medline, Web of Science, and PsycINFO, and failed to find a single reference to Sexual maturation disorder or Sexual relationship disorder. The last peer-reviewed, indexed reference to "egodystonic homosexuality" was published more than two decades ago. The F66 categories do not

contribute meaningfully to public health surveillance, are not routinely reported by any country, and are not used in WHO's calculation of disease burden. At the same time, they selectively target individuals with same-sex orientation or gender nonconformity, with no apparent justification. Individuals with needs for information or who experience distress specifically related to sexual orientation that is not diagnosable as another disorder (e.g., Adjustment disorder) can still receive services through the use of codes related to counselling interventions from the ICD-11 chapter on Factors Influencing Health Status and Contact with Health Services described earlier in this paper.

The Working Group has therefore proposed the elimination of the entire grouping of F66 disorders from the ICD-11.

Comparison with DSM-5

The proposed changes for ICD-11 in this area bring it in line with DSM-5. No equivalent to any of the ICD-10 F66 categories is included in DSM-5 or was included in DSM-IV.

CONCLUSIONS

In the more than quarter century since the approval of the ICD-10, there have been substantial gains in scientific, clinical, social, and human rights understandings relevant to diagnostic categories related to sexuality and gender identity. These different streams of evidence have been considered in the development of a set of proposals for ICD-11 that departs markedly from the descriptions of categories related to sexuality and gender identity in the ICD-10. The inclusion of mental and behavioural disorders alongside all other diagnostic entities in health care is a central feature of the ICD, and has uniquely positioned the current revision effort to contemplate a broader and more integrative set of classification options with respect to these categories.

The ICD-10 classification of Sexual dysfunctions was substantially outdated in its view of psychological and physical causes of sexual dysfunction as separable and separate, making it inconsistent with current evidence regarding the etiology and treatment of these conditions. For the ICD-11, an innovative, integrated system has been proposed, including a set of qualifiers to indicate the range of factors that the clinician considers to be contributory. It must be emphasized that the WHO does not consider the ICD-11 chapters to constitute scope of practice boundaries between medical specialties, but intends and expects that psychiatrists and other mental health professionals with appropriate training will continue to engage in the treatment of these common and costly conditions and that the reformulated classification of these conditions will encourage broader availability of treatment.

The role of psychiatry in many countries is likely to evolve in substantive ways with respect to the evaluation and treatment of Gender incongruence, proposed to replace Gender identity disorders in the ICD-10. The best health care services for transgender people are by definition multidisciplinary[59]. But psychiatrists in some countries have been unfortunately positioned as gatekeepers to enforce elaborate and burdensome requirements in order to access these services[83], ostensibly in order to verify that transgender people are certain about their decision to seek health services to make their bodies align with their experienced identity. However, in the recent Mexican study described above[71], the average delay between reported awareness of transgender identity and initiation of hormones – by far the most common treatment received – was found to be more than 12 years, and nearly half of participants had initiated hormones without medical supervision, exposing themselves to serious health risks. While these figures are not broadly generalizable, they are likely more reflective of the situation in most of the world than those reported in available studies from the US or Western Europe, given that more that 80% of the global population lives in low- and middle-income countries. Psychiatrists and other mental health professionals have a major role to play in improving the health status of this often mistreated population[58,74,75].

With respect to the classification of Paraphilic disorders, the Working Group on Sexual Disorders and Sexual Health has attempted to grapple with thorny issues related to how best to distinguish between conditions that are relevant to public health and clinical psychopathology on the one hand and private behaviours that are not a legitimate focus of health classification on the other. At the same time, proposals in this area affirm the status of persistent and intense sexual arousal patterns focusing on individuals who do not or cannot consent as psychiatric in their nature and management[90]. In contrast, the Working Group concluded that there are no legitimate public health or clinical objectives served by mental disorder categories uniquely linked to sexual orientation[89].

In summary, the Working Group on Sexual Disorders and Sexual Health has proposed changes in the classification of these conditions that it considers to be: a) more reflective of current scientific evidence and best practices; b) more responsive to the needs, experience, and human rights of vulnerable populations; and c) more supportive of the provision of accessible and high-quality health care services. Proposed diagnostic guidelines for the disorders described in this paper will be made available for review and comment by members of WHO's Global Clinical Practice Network (http://gcp.network)[123], and subsequently for public review prior to finalization of the ICD-11. We hope that this paper will serve to encourage further scientific and professional discussion.

ACKNOWLEDGEMENTS

Most of the authors of this paper were members of or consultants to the WHO ICD-11 Working Group on Sexual Disorders and Sexual Health. G.M. Reed, S. Cottler and S. Saxena are members of the WHO Secretariat. S.J. Parish co-chaired a consultation meeting on the classification of sexual dysfunctions sponsored by the International Society for the Study of Women's Sexual Health and the World Association of Sexual Health. The authors are grateful to the other members of the Working Group, to other members of the WHO Secretar-

iat, and to the participants at the consultation meeting on the classification of sexual dysfunctions who contributed to discussion of the issues addressed in this paper. P. Briken's work on this paper was supported by a grant from the Federal Centre for Health Education, Germany. Unless specifically stated, the views expressed in this paper are those of the authors and do not represent the official policies or positions of WHO.

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. 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.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington: American Psychiatric Association, 2013.
4. World Health Organization. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1992.
5. Perelman MA. The sexual tipping point: a mind/body model for sexual medicine. J Sex Med 2009;6:629:32.
6. Lewis RW, Fugl-Meyer KS, Corona G et al. Definitions/epidemiology/risk factors for sexual dysfunction. J Sex Med 2010;7:1598-607.
7. McCabe MP, Sharlip ID, Atalla E et al. Definitions of sexual dysfunctions in women and men: a consensus statement from the Fourth International Consultation on Sexual Medicine 2015. J Sex Med 2016;13:135-43.
8. McCabe MP, Sharlip ID, Lewis R et al. Risk factors for sexual dysfunction among women and men: a consensus statement from the Fourth International Consultation on Sexual Medicine 2015. J Sex Med 2016;13:153-67.
9. Carvalho J, Nobre P. Biopsychosocial determinants of men's sexual desire: testing an integrative model. J Sex Med 2011;8:754-63.
10. Chou D, Cottler S, Khosla R et al. Sexual health in the International Classification of Diseases (ICD): implications for measurement and beyond. Reprod Health Matters 2015;23:185-92.
11. World Health Organization. International statistical classification of diseases and related health problems, 10th revision (ICD-10). Geneva: World Health Organization, 1992.
12. Pfaus JG. Pathways of sexual desire. J Sex Med 2009;6:1506-33.
13. Kingsberg SA, Clayton AH, Pfaus JG. The female sexual response: current models, neurobiological underpinnings and agents currently approved or under investigation for the treatment of Hypoactive Sexual Desire Disorder. CNS Drugs 2015;29:915-33.
14. Basson R. The female sexual response: a different model. J Sex Marital Ther 2000;26:51-65.
15. Hackett G, Krychman M, Baldwin D et al. Coronary heart disease, diabetes, and sexuality in men. J Sex Med 2016;13:887-904.
16. Atlantis E, Sullivan T. Bidirectional association between depression and sexual dysfunction: a systematic review and meta-analysis. J Sex Med 2012;9:1497-507.
17. El Miedany Y, El Gaafary M, El Aroussy N et al. Sexual dysfunction in rheumatoid arthritis patients: arthritis and beyond. Clin Rheumatol 2012; 31:601-6.
18. Mohammadi K, Rahnama P, Mohseni SM et al. Determinants of sexual dysfunction in women with multiple sclerosis. BMC Neurol 2013;13:83.
19. Catania L, Abdulcadir O, Puppo V et al. Pleasure and orgasm in women with female genital mutilation/cutting (FGM/C). J Sex Med 2007;4:1666-78.
20. Johnson SD, Phelps D, Cottler LB. The association of sexual dysfunction and substance use among a community epidemiological sample. Arch Sex Behav 2004;33:55-63.
21. Clayton AH, Croft HA, Handiwala L. Antidepressants and sexual dysfunction: mechanisms and clinical implications. Postgrad Med 2014;126:91-9.
22. Nobre PJ, Pinto-Gouveia J. Dysfunctional sexual beliefs as vulnerability factors to sexual dysfunction. J Sex Res 2006;43:68-75.
23. Dunn KM, Croft PR, Hackett GI. Association of sexual problems with social, psychological, and physical problems in men and women: a cross sectional populations survey. J Epidemiol Community Health 1999;53:144-8.
24. Öberg K, Fugl-Meyer KS, Fugl-Meyer AR. On sexual well being in sexually abused Swedish women: epidemiological aspects. Sex Relation Ther 2002;17:329-41.
25. Brotto LA, Atallah S, Johnson-Agbakwu C et al. Psychological and interpersonal dimensions of sexual function and dysfunction. J Sex Med 2016; 13:538-71.
26. Laumann EO, Nicolosi A, Glasser DB et al. Sexual problems among women and men aged 40-80 y: prevalence and correlates identified in the Global Study of Sexual Attitudes and Behaviors. Int J Impot Res 2005;17: 39-57.
27. Atallah S, Johnson-Agbakwu C, Rosenbaum T et al. Ethical and sociocultural aspects of sexual function and dysfunction in both sexes. J Sex Med 2016;13:591-606.
28. Oniz A, Keskinoglu P, Bezircioglu I. The prevalence and causes of sexual problems among premenopausal Turkish women. J Sex Med 2007;4: 1575-81.
29. Laan E, Rellini AH, Barnes T. Standard operating procedures for female orgasmic disorder: consensus of the International Society for Sexual Medicine. J Sex Med 2013;10:74-82.
30. Brotto LA. The DSM diagnostic criteria for hypoactive sexual desire disorder in women. Arch Sex Behav 2010;39:221-39.
31. Parish SJ, Hahn SR. Hypoactive sexual desire disorder: a review of epidemiology, biopsychology, diagnosis, and treatment. Sex Med Rev 2016;4: 103-20.
32. DeRogatis LR, Clayton AH, Rosen RC et al. Should sexual desire and arousal disorders in women be merged? Arch Sex Behav 2011;40:217-9.
33. Sungur MZ, Gündüz A. A comparison of DSM-IV-TR and DSM-5 definitions for sexual dysfunctions: critiques and challenges. J Sex Med 2014; 11:364-73.
34. Sarin S, Amsel RM, Binik YM. Disentangling desire and arousal: a classificatory conundrum. Arch Sex Behav 2013;42:1079-100.
35. Balon R, Clayton AH. Female sexual interest/arousal disorder: a diagnosis out of thin air. Arch Sex Behav 2014;43:1227-9.
36. Burri A, Greven C, Leupi M et al. A multivariate twin study of female sexual dysfunction. J Sex Med 2012;9:2671-81.
37. Bishop JR, Moline J, Ellingrod VL et al. Serotonin 2A-1438 G/A and G-protein Beta3 subunit C825T polymorphisms in patients with depression and SSRI-associated sexual side-effects. Neuropsychopharmacology 2006;31:2281-8.
38. Bishop JR, Ellingrod VL, Akroush M et al. The association of serotonin transporter genotypes and selective serotonin reuptake inhibitor (SSRI)-associated sexual side effects: possible relationship to oral contraceptives. Hum Psychopharmacol 2009;24:207-15.
39. Arnow BA, Millheiser L, Garrett A et al. Women with hypoactive sexual desire disorder compared to normal females: a functional magnetic resonance imaging study. Neuroscience 2009;158:484-502.
40. Pyke RE, Clayton AH. Psychological treatment trials for hypoactive sexual desire disorder: a sexual medicine critique and perspective. J Sex Med 2015;12:2451-8.
41. Berman JR, Berman LA, Toler SM et al. Safety and efficacy of sildenafil citrate for the treatment of female sexual arousal disorder: a double-blind, placebo controlled study. J Urology 2003;170:2333-8.
42. Brotto LA, Basson R, Luria M. A mindfulness-based group psychoeducational intervention targeting sexual arousal disorder in women. J Sex Med 2008;5:1646-59.
43. Brotto LA, Chivers ML, Millman RD et al. Mindfulness-based sex therapy improves genital-subjective arousal concordance in women with sexual desire/arousal difficulties. Arch Sex Behav (in press).
44. Maserejian NN, Shifren J, Parish SJ et al. Sexual arousal and lubrication problems in women with clinically diagnosed hypoactive sexual desire disorder: preliminary findings from the hypoactive sexual desire disorder registry for women. J Sex Marital Ther 2012;38:41-62.
45. Binik YM. Should dyspareunia be retained as a sexual dysfunction in DSM-V? A painful classification. Arch Sex Behav 2005;34:11-21.
46. Pukall CF, Goldstein AT, Bergeron S et al. Vulvodynia: definition, prevalence, impact, and pathophysiological factors. J Sex Med 2016;13:291-304.
47. Bornstein J, Goldstein AT, Stockdale CK et al. 2015 ISSVD, ISSWSH, and IPPS consensus terminology and classification of persistent vulvar pain and vulvodynia. J Sex Med 2016;127:745-51.

48. Segraves RT. Considerations for a better definition of male orgasmic disorder in DSM V. J Sex Med 2010;7:690-9.
49. Waldinger M. Male ejaculation and orgasm disorders. In: Balon R, Segraves R (eds). Handbook of sexual dysfunction. Boca Raton: Taylor and Francis, 2005:215-48.
50. Wylie K, Ralph D, Levin RJ et al. Comments on "Considerations for a better definition of male orgasmic disorder in DSM V". J Sex Med 2010;7: 695-9.
51. Global Action for Trans* Equality. It's time for reform. Trans health issues in the International Classification of Diseases. Report on the GATE Experts Meeting. The Hague, November 2011.
52. European Commission. Trans and intersex people: discrimination on the grounds of sex, gender identity and gender expression. Luxembourg: European Union, 2012.
53. European Parliament. Resolution of 28 September 2011 on human rights, sexual orientation and gender identity at the United Nations. Strasbourg: European Parliament, 2011.
54. World Health Organization. Mental health action plan 2013-2020. Geneva: World Health Organization, 2013.
55. World Health Organization. Basic documents, 48th ed. Geneva: World Health Organization, 2014.
56. United Nations High Commissioner for Human Rights. Discriminatory laws and practices and acts of violence against individuals based on their sexual orientation and gender identity. New York: United Nations General Assembly, 2011.
57. Council of Europe. Discrimination on grounds of sexual orientation and gender identity in Europe, 2nd ed. Strasbourg: Council of Europe Publishing, 2011.
58. World Health Organization. Sexual health, human rights and the law. Geneva: World Health Organization, 2015.
59. World Professional Association for Transgender Health. Standards of care for the health of transsexual, transgender and gender non-conforming people, version 7. World Professional Association for Transgender Health, 2011.
60. United Nations Development Programme. Discussion paper: Transgender health and human rights. New York: United Nations Development Programme, 2013.
61. Sood N. Transgender people's access to sexual health and rights: a study of law and policy in 12 Asian countries. Kuala Lumpur: Asian-Pacific Resource and Research Centre for Women, 2009.
62. Nuttbrock L, Hwahng S, Bockting W et al. Psychiatric impact of gender-related abuse across the life course of male-to-female transgender persons. J Sex Res 2010;47:12-23.
63. Grossman AH, D'Augelli AR. Transgender youth: invisible and vulnerable. J Homosex 2006;51:111-28.
64. Sugano E, Nemoto T, Operario D. The impact of exposure to transphobia on HIV risk behavior in a sample of transgendered women of color in San Francisco. AIDS Behav 2006;10:217-25.
65. Grossman AH, D'Augelli AR, Salter NP. Male-to-female transgender youth: gender expression milestones, gender atypicality, victimization, and parents' responses. J GLBT Fam Stud 2006;2:71-92.
66. Rotondi NK, Bauer GR, Scanlon K et al. Nonprescribed hormone use and self-performed surgeries: "do-it-yourself" transitions in transgender communities in Ontario, Canada. Am J Public Health 2013;103:1830-6.
67. Sanchez NF, Sanchez JP, Danoff A. Health care utilization, barriers to care, and hormone usage among male-to-female transgender persons in New York City. Am J Public Health 2009;99:713-9.
68. Asscheman H, Giltay EJ, Megens JA et al. A long-term follow-up study of mortality in transsexuals receiving treatment with cross-sex hormones. Eur J Endocrinol 2011;164:635-42.
69. Hembree WC, Cohen-Kettenis P, Delemarre-van de Waal HA et al. Endocrine treatment of transsexual persons: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2009;94:3132-54.
70. Mueller A, Gooren L. Hormone-related tumors in transsexuals receiving treatment with cross-sex hormones. Eur J Endocrinol 2008;159:197-202.
71. Robles R, Fresán A, Vega-Ramírez H et al. Removing transgender identity from the classification of mental disorders: a Mexican field study for ICD-11. Lancet Psychiatry 2016;3:850-9.
72. Drescher J, Cohen-Kettenis P, Winter S. Minding the body: situating gender identity diagnoses in the ICD-11. Int Rev Psychiatry 2012;24:568-77.
73. Drescher J, Cohen-Kettenis PT, Reed GM. Gender incongruence of childhood in the ICD-11: controversies, proposal, and rationale. Lancet Psychiatry 2016;3:297-304.
74. Lo S, Horton R. Transgender health: an opportunity for global health equity. Lancet 2016;388:316-8.
75. Reisner SL, Poteat T. Keatley J et al. Global health burden and needs of transgender populations: a review. Lancet 2016;388:412-36.
76. Winter S, Settle E, Wylie K et al. Synergies in health and human rights: a call to action to improve transgender health. Lancet 2016;388:318-21.
77. Cabral M, Suess A, Ehrt J et al. Removal of gender incongruence of childhood diagnostic category: a human rights perspective. Lancet Psychiatry 2016;3:405-6.
78. Drescher J, Byne W. Gender dysphoric/gender variant (GD/GV) children and adolescents: summarizing what we know and what we have yet to learn. J Homosex 2012;59:501-10.
79. Drescher J, Byne W. Treating transgender children and adolescents: an interdisciplinary discussion. New York: Routledge, 2013.
80. Steensma TD, McGuire JK, Kreukels BP et al. Factors associated with desistence and persistence of childhood gender dysphoria: a quantitative follow-up study. J Am Acad Child Adolesc Psychiatry 2013;52:582-90.
81. Steensma TD, Biemond R, de Boer F et al. Desisting and persisting gender dysphoria after childhood: a qualitative follow-up study. Clin Child Psychol Psychiatry 2011;16:499-516.
82. Byne W, Bradley SJ, Coleman E, et al. Report of the APA Task Force on Treatment of Gender Identity Disorder. Arch Sex Behav 2012;41:759-96.
83. Bouman WP, Bauer GR, Richards C et al. World Professional Association for Transgender Health consensus statement on considerations of the role of distress (Criterion D) in the DSM diagnosis of gender identity disorder. Int J Transgend 2010;12:100-6.
84. Ehrbar RD, Witty MC, Ehrbar HG et al. Clinician judgment in the diagnosis of gender identity disorder in children. J Sex Marital Ther 2008;34:385-412.
85. Joyal CC. How anomalous are paraphilic interests? Arch Sex Behav 2014; 43:1241-3.
86. Joyal CC. Defining "normophilic" and "paraphilic" sexual fantasies in a population-based sample: on the importance of considering subgroups. Sex Med 2015;3:321-30.
87. Richters J, de Visser RO, Rissel CE et al. Demographic and psychosocial features of participants in bondage and discipline, "sadomasochism" or dominance and submission (BDSM): data from a national survey. J Sex Med 2008;5:1660-8.
88. Reiersøl O, Skeid S. The ICD diagnoses of fetishism and sadomasochism. J Homosex 2006;50:243-62.
89. Cochran SD, Drescher J, Kismödi E et al. Proposed declassification of disease categories related to sexual orientation in the International Statistical Classification of Diseases and Related Health Problems (ICD-11). Bull World Health Organ 2014;92:672-9.
90. Krueger RB, Reed GM, First MB et al. Paraphilic disorders in the International Classification of Disease and Related Health Problems, Eleventh Revision (ICD-11). Arch Sex Behav (in press).
91. Briken P, Fedoroff JP, Bradford JW. Why can't pedophilic disorder remit? Arch Sex Behav 2014;43:1237-9.
92. Becker JV, Stinson J, Tromp S et al. Characteristics of individuals petitioned for civil commitment. Int J Offend Ther 2003;47:185-95.
93. Berner W, Berger P, Hill A. Sexual sadism. Int J Offend Ther 2003;47:383-95.
94. Briken P, Bourget D, Dufour M. Sexual sadism in sexual offenders and sexually motivated homicide. Psychiatr Clin North Am 2014;37:215-30.
95. Elwood RW, Doren DM, Thornton D. Diagnostic and risk profiles of men detained under Wisconsin's sexually violent person law. Int J Offend Ther 2010;54:187-96.
96. Packard, RL, Levenson JL. Revisiting the reliability of diagnostic decisions in sex offender civil commitment. Sexual Offender Treatment 2006;1:1-15.
97. Krueger RB. The DSM diagnostic criteria for sexual sadism. Arch Sex Behav 2010;39:325-45.
98. Abel GG, Becker JV, Mittelman M et al. Self-reported sex crimes of nonincarcerated paraphiliacs. J Interpers Violence 1987;2:3-25.
99. Ahlers CJ, Schaefer GA, Mundt IA et al. How unusual are the contents of paraphilias? Paraphilia-associated sexual arousal patterns in a community-based sample of men. J Sex Med 2011;8:1362-70.
100. Bradford JMW, Boulet J, Pawlak A. The paraphilias: a multiplicity of deviant behaviors. Can J Psychiatry 1992;37:104-8.
101. Långström N. The DSM diagnostic criteria for exhibitionism, voyeurism, and frotteurism. Arch Sex Behav 2010;39:317-24.

102. Templeman TL, Stinnett RD. Patterns of sexual arousal and history in a "normal" sample of young men. Arch Sex Behav 1991;20,137-50.
103. Johnson RS, Ostermeyer B, Sikes KA et al. Prevalence and treatment of frotteurism in the community: a systematic review. J Am Acad Psychiatry Law 2014;42:478-83.
104. Nordic Centre for Classifications in Health Care. Removed ICD-10 codes in categories F64 and F65 in the Nordic Countries. Helsinki: Nordic Centre for Classifications in Health Care, 2015.
105. First MB, Halon RL. Use of DSM paraphilia diagnoses in sexually violent predator commitment cases. J Am Acad Psychiatry Law 2008;36:443-54.
106. US Supreme Court. Kansas v. Hendricks, 521 U.S. 346 (1997).
107. Janus E. Sexually violent predator laws: psychiatry in service to a morally dubious enterprise. Lancet 2004;364:50-1.
108. Zonana H. The civil commitment of sex offenders. Science 1997;278:1248-9.
109. First MB. DSM-5 and paraphilic disorders. J Am Acad Psychiatry Law 2014;42:191-201.
110. O'Flaherty M, Fisher J. Sexual orientation, gender identity and international human rights law: contextualising the Yogyakarta Principles. Human Rights Law Rev 2008;8:207-48.
111. Caceres CF, Konda K, Segura ER et al. Epidemiology of male same-sex behaviour and associated sexual health indicators in low- and middle-income countries: 2003-2007 estimates. Sex Transm Infect 2008;84(Suppl. 1):i49-56.
112. Hooker E. Reflections of a 40-year exploration: a scientific view on homosexuality. Am Psychol 1993;48:450-3.
113. Drescher J. Queer diagnoses: parallels and contrasts in the history of homosexuality, gender variance, and the Diagnostic and Statistical Manual. Arch Sex Behav 2010;39:427-60.
114. King M, McKeown E, Warner J et al. Mental health and quality of life of gay men and lesbians in England and Wales: controlled, cross-sectional study. Br J Psychiatry 2003;183:552-8.
115. Mays VM, Cochran SD. Mental health correlates of perceived discrimination among lesbian, gay, and bisexual adults in the United States. Am J Public Health 2001;91:1869-76.
116. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003;129:674-97.
117. Organization for Security and Cooperation in Europe/Office for Democratic Institutions and Human Rights. Hate crimes in the OSCE region – incidents and responses. Annual report for 2006. Warsaw: Organization for Security and Cooperation in Europe/Office for Democratic Institutions and Human Rights, 2007.
118. Calzo JP, Antonucci TC, Mays VM et al. Retrospective recall of sexual orientation identity development among gay, lesbian, and bisexual adults. Dev Psychol 2011;47:1658-73.
119. Spitzer RL. The diagnostic status of homosexuality in DSM-III: a reformulation of the issues. Am J Psychiatry 1981;138:210-5.
120. van Drimmelen-Krabbe JJ, Ustun TB, Thompson DH et al. Homosexuality in the International Classification of Diseases: a clarification. JAMA 1994; 272:1660.
121. American Psychological Association. Guidelines for psychological practice with lesbian, gay, and bisexual clients. Am Psychol 2012;67:10-42.
122. Pan American Health Organization. "Cures" for an illness that does not exist. Washington, DC: Pan American Health Organization, 2012.
123. 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.

DOI:10.1002/wps.20354

Updating the Research Domain Criteria

Two and a half years ago, *World Psychiatry* published a Forum about the US National Institute of Mental Health (NIMH) Research Domain Criteria (RDoC) initiative[1]. In it, there was spirited commentary with divergent views of diagnosis and critical examination of RDoC. Some criticisms were based on an incomplete understanding of RDoC[2], but the discussion captured the challenge of shifting the paradigm for psychiatric nosology. Here I provide an update on RDoC, while articulating some fundamental issues.

The RDoC idea was introduced in the 2008 NIMH Strategic Plan to "develop, for research purposes, new ways of classifying mental disorders based on dimensions of observable behavior and neurobiological measures", because advances in integrative neuroscience were not being realized in patient care. Decades since Robins and Guze[3] proposed validity criteria that included laboratory tests, and years since Wakefield[4] articulated the "harmful dysfunction" definition of mental disorder, there was still no valid laboratory test – biological or otherwise – linking any psychiatric diagnosis to an internal mechanism. Syndrome-based diagnoses (DSM/ICD) did not map to disrupted neural mechanisms[5], and there was a steep decline in the development of new therapeutic agents[6]. It was argued that clinical syndromes were too distal from genetic mechanisms to research connections, and suggested that an intermediate phenotype approach would be more viable[7]. Regardless, funding for psychopathology research in the US remained harnessed to the DSM. RDoC was introduced to allow researchers to pursue funding for translational research without the limitation of DSM independent variables.

RDoC differs from syndrome-based research by incorporating a dimensional approach. Rather than beginning with diagnoses based on clinical description and then trying to connect them to mechanisms, RDoC research begins with dysfunctional mechanisms and works toward clinical symptoms. The process is to identify a mechanism used for functional behavior, and to link its improper function to clinical problems. The aim is to inform nosology, including future DSM and ICD revisions, to ultimately help those suffering clinical problems by clarifying homogeneous treatment targets. RDoC is not a clinical manual or a replacement for the DSM or the ICD. It is an evolving structure, intended to facilitate translational research.

The RDoC matrix is organized with Domains and Constructs in the rows, and Units of Analysis in the columns (see www.nimh.nih.gov/research-priorities/rdoc). It was developed in consultation with the scientific community, including NIMH workshops where experts in each of the respective domains refined construct definitions. Constructs required validity for a functional unit of behavior and a link to clinical problems. Connection to a neural circuit was emphasized to fill the gap between neuroscience and psychopathology research, and does not necessitate a reductionist philosophy; indeed, observations of different systems provide different forms of important evidence[8].

There are at present five RDoC Domains: Negative Valence, Positive Valence, Cognitive Systems, Social Processes, Arousal and Regulatory Systems. Within each domain are a number of constructs. Evidence for a new domain, Motor Systems, is under review with an NIMH-sponsored workshop, and there are plans for annual consideration to accommodate new findings indicating revisions. Across the columns of the matrix are Units of Analysis, which include Genes, Molecules, Cells, Circuits, Physiology, Behavior, and Self-Reports (self-reports include patient-reported symptoms). Elements in each cell hold findings that correspond to respective constructs, within a particular unit of analysis.

A unique column, Paradigms, is for behavioral tasks designed for valid and reliable assessment of a specific mechanism or circuit (e.g., an "n-back" task for working memory). Such tasks are routine in experimental research but, for RDoC, these approaches need to be more fully developed to meet acceptable psychometric standards, including sensitivity and specificity. A National Advisory Mental Health Council Workgroup was convened and evaluated the state of the research in this area and formulated recommendations, reported at its September 2016 meeting.

Since the *World Psychiatry* Forum, RDoC impact has accelerated. RDoC Workgroup members now have over 30 papers, published or in press, detailing the rationale, description, and development of RDoC (see www.nimh.nih.gov/research-priorities/rdoc). Many of these are in collaboration with non-NIH affiliated clinical scientists. Seminal papers published by the RDoC Workgroup have collectively been cited over 2,000 times. In addition to *World Psychiatry*, several journals have devoted special sections to RDoC (e.g., *International Journal of Psychophysiology, JAMA Psychiatry, Journal of Abnormal Psychology, Neuropsychiatric Genetics, and Psychophysiology*), with others in the offing. Presently, there are 38 active projects funded by NIMH RDoC Requests for Applications (RFAs), three active program announcements. A search with NIH Reporter (https://projectreporter.nih.gov/reporter.cfm) reveals the term "RDoC" in 273 currently funded grants. RDoC figures prominently in the recent NIMH Strategic Plan (www.nimh.nih.gov/about/strategic-planning-reports).

Regular RDoC Internal Workgroup meetings continue to solicit and incorporate feedback from the field, while guiding ongoing development. In 2014, the RDoC Unit was created within the Office of the NIMH Director. The Unit has been instrumental for codifying a number of RDoC-related efforts, including improvements to the on-line version of the matrix, which now links constructs to definitions, as well as common elements across Units of Analysis. In this form, the matrix can be used to facilitate research design and identify areas where more knowledge is needed. It can also serve as a teaching tool.

Future possibilities include linking matrix elements to the US National Library of Medicine. The RDoC Database repository (RDoC-db) collates subject-level data, creating a common-use data set for future data mining (https://data-archive.nimh.nih.gov/rdocdb). Other web-based tools include a discussion forum, and an ongoing series of webinars, archived on the website. Planning is underway for another major update of the RDoC website.

RDoC does not mandate a unitary approach to translational research. Rather, it provides a scaffold to organize findings, and on which a nomological net may be constructed. Theories of development, environmental influences, and psychopathology are needed to spell out the connections between constructs. A research exemplar is the Bipolar Schizophrenia Network on Intermediate Phenotypes (B-SNIP), which recently reported emergent biotypes that overlapped in different degrees across psychosis spectrum patients, evidencing systematically varying levels of cognitive control as well as differences in grey matter[9]. With RDoC, these biotypes offer new possibilities for independent variables in future studies[10].

As a research tool, RDoC has the luxury to evolve incrementally, with regular updates. Its web-based format facilitates faster-paced change and allows open access. In contrast, any modification to a DSM or ICD diagnosis can have immediate consequences for patients (e.g., treatment decisions, reimbursement for services, disability accommodations). For meaningful progress, ongoing collaboration with stakeholders is valued (e.g., professional and health organizations, regulatory agencies, and patient advocacy groups). Combined efforts will encourage new ways to think about diagnosis for clinicians and researchers alike.

Charles A. Sanislow
Department of Psychology and Program in Neuroscience and Behavior, Wesleyan University, Middletown, CT, USA

C.A. Sanislow is a member of the NIMH RDoC Internal Workgroup. The opinions expressed in this piece are those of the author, and not necessarily those of the NIMH, NIH or the US government. Current members of the NIMH RDoC Internal Workgroup are: B. Cuthbert (Head), S. Morris (Acting Head), W. Carpenter, M. Chai, R. Garcia, M. Garvey, D. Greenstein, A. Kadam, K. McLinden, J. Pacheco, D. Pine, M. Rudorfer, C. Sanislow, J. Simmons, U. Vaidyanathan and C. Zarate. D. Barch and M.B. First serve as external consultants.

1. Cuthbert BN. World Psychiatry 2014;13:28-35.
2. Cuthbert BN. World Psychiatry 2014;13:196-7.
3. Robins E, Guze SB. Am J Psychiatry 1970;126:983-7.
4. Wakefield J. Psychol Rev 1992;99:232-47.
5. Sanislow CA, Pine DS, Quinn KJ et al. J Abnorm Psychol 2010;119:631-9.
6. Pankevich DE, Altevogt BM, Dunlop J et al. Neuron 2014;84:546-53.
7. Meyer-Lindenberg A, Weinberger DR. Nat Rev Neurosci 2006;7:818-27.
8. Kozak MJ, Cuthbert BN. Psychophysiology 2016;53:286-97.
9. Clementz BA, Sweeney JA, Hamm JP et al. Am J Psychiatry 2016;173:373-84.
10. Insel TR, Cuthbert BN. Science 2015;348:499-500.

DOI:10.1002/wps.20374

Adopting a continuous improvement model for future DSM revisions

The approach used for making changes to both the DSM and ICD has been up to now to revise the manuals in their entirety at fixed (albeit variable) intervals. Typically, these diagnostic revision efforts have been multi-year affairs involving the appointment of committees of experts tasked with making changes with the goal of improving the validity, reliability, and clinical utility of the diagnostic systems[1,2].

While this approach has the advantage of facilitating standardized communication among users of the classifications by ensuring the uniformity and stability of the diagnostic definitions in the time interval during which that edition of the manual is in effect, it prevents the incorporation of new scientific knowledge into the manual as it emerges, a limitation that has become especially problematic given the extended intervals between revisions that have characterized the most recent editions of the DSM and ICD (19 years and 23 years, respectively).

Advances in digital publishing that allow instantaneous dissemination of changes at minimal cost have paved the way towards the American Psychiatric Association (APA) adopting a continuous improvement model for the DSM, in which revisions are pegged to specific scientific advances. Thus, rather than waiting until the next wholesale revision to implement a clinically useful change (such as incorporating a solidly validated biomarker into the definition of a disorder), such a change could be put into effect as soon as it has been determined that it is diagnostically advantageous to do so. Moreover, implementing a continuous data-driven approach has the added advantage of discouraging changes that are not well supported by empirical evidence. As described by Kendler in his accounting of the history of the DSM-5 Scientific Review Committee, there is an inherent trend towards making changes built into the DSM revision process: "for workgroup members, it is a natural source of pride to 'make a difference', to 'put their mark' on the document" (3).

A new DSM web portal (www.dsm5.org) has been set up by the APA to field proposals for changes on a continuous basis. Submissions will be web-based, with proposers required to provide supportive information in a structured format, including the reasons for the change, the magnitude of change, data documenting improvements in validity across a range of validators, evidence of reliability and clinical utility, and a consideration of current or potential deleterious consequences associated with the proposed change. It is anticipated that most submissions will come from interested persons (e.g., psychiatric researchers, individual clinicians) or organizations

(e.g., psychiatric subspecialty groups, advocacy organizations, APA components) that are external to the APA committee overseeing the DSM revision process. This is in marked contrast to the prior DSM revision efforts, in which the proposals were drafted by workgroup members, who were also responsible for providing supportive literature reviews and conducting data re-analyses.

The revision process will be overseen by a Steering Committee (analogous to the DSM Task Force) whose members have expertise in psychiatric nosology, psychiatric research, clinical psychiatry, and the DSM. Five standing Review Committees (analogous to the DSM Workgroups), which cover broad domains of psychiatric diagnosis, will work with the Steering Committee to review the proposals and draft the revisions. For example, a single review committee will cover so-called "serious mental illnesses", which include schizophrenic spectrum and other psychotic disorders, bipolar disorders, and neurodegenerative disorders. Final review of the proposals (which will be posted for public comment) in terms of whether criteria for approval have been met will be conducted by the Steering Committee and, if so, will be referred to the APA Board of Trustees for official approval. Once approved, each change will be publicized by the APA and digital versions of the manual will be updated to reflect the change.

Three types of proposals that require substantial empirical support have been identified, each with explicit criteria regarding the type of evidence that is expected to be submitted. Type 1 proposals involve changes to an existing diagnostic criteria set. Submitted evidence should document that the change would markedly improve the validity, reliability, or clinical utility of a criteria set, or that it would substantially reduce identified deleterious consequences associated with a criteria set. Type 2 proposals involve the addition of a new diagnostic category, subtype or specifier, and supporting evidence should document that the new category: a) meets the criteria for a mental disorder provided in the DSM-5[1]; b) has strong evidence of validity; c) can be applied reliably; d) has substantial clinical value (e.g., it identifies a group of patients now not receiving appropriate clinical attention); e) avoids substantial overlap with existing diagnoses; and f) has a positive benefit/harm ratio (e.g., low risk of harm due to social or forensic consequences). Type 3 proposals entail deleting an existing category or subtype/specifier, and require evidence that the proposed item to be deleted has only weak validity, minimal utility (e.g., it is rarely used in clinical practice or research) or is better conceptualized as a subtype of an existing diagnosis.

Proposals involving corrections and clarifications of existing criteria that do not require empirical support will be considered on an expedited basis by a subcommittee of the Steering Committee. Included are instances of lack of clarity or ambiguity in the meaning of the wording of criteria or text; inconsistencies or contradictions within text or criteria (e.g., diagnostic criteria in conflict with descriptive text), and errors of omission or inadvertent inclusion (e.g., omission of a disorder in the "not better explained by" list in an exclusion criterion).

In conclusion, it is hoped that the implementation of an empirically rigorous continuous improvement process for the DSM will facilitate the inclusion of scientific advances in a timelier manner than was possible using the current revision process, which should ultimately result in a more valid and clinically useful diagnostic classification.

Michael B. First
Columbia University Department of Psychiatry, New York State Psychiatric Institute, New York, NY, USA

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington: American Psychiatric Association, 2013.
2. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. World Psychiatry 2011;10:86-93.
3. Kendler KS. Psychol Med 2013;43:1793-800.

DOI:10.1002/wps.20342

Prescribing according to diagnosis: how psychiatry is different

Most countries in the modern world have a formal system of medicines regulation. Medicinal drugs, or more specifically their formulation and dosage, are licensed or "labelled" for certain specific indications in which their use has been shown to be broadly safe and beneficial. Prescribers can, as a consequence, have confidence in the assumption that a drug licensed for a particular precisely defined ailment is likely to be effective in that condition. Outside this regulatory framework is a body of literature which may relate to other, unlicensed or "off-label" uses. Sometimes the extra-regulatory evidence base amounts to no more than a handful of case reports, but often there are published and successful randomized controlled trials. The licensed uses of a medicine are thus merely those for which formal approval has been sought and obtained by the manufacturer. Other beneficial uses are certainly not precluded by a drug's label.

In most areas of medicine, prescribing is very closely aligned with drug labeling. This is because drugs usually have a known mode of action which can be understood in the context of what is known about the condition: penicillin is bactericidal to specific bacteria; insulin replaces what the body fails to produce; antihypertensives reduce cardiac output or vascular resistance. Off-label prescribing is therefore fairly rare in physical medicine, because prescribers invariably make a firm diagnosis and prescribe a drug precisely indicated for that condition.

Psychiatry is different. The biochemical or pathological basis of most mental health conditions is, at best, poorly understood. Our knowledge of the precise actions of effective drugs is likewise

incomplete. Diagnostic criteria shift and sway like in no other area of medicine and named mental health conditions appear and disappear with a disturbing frequency. Alongside these factors is the troubling awareness of the frequent overlap of symptoms across formal diagnostic entities. It seems trite to point out, for example, that people with depression sometimes hear voices, or that people with bipolar disorder may suffer from anxiety.

When I first worked in psychiatry in the 1980s, I was struck by the extent to which drugs were prescribed outside their license. The most common prescription I saw was for amitriptyline combined with chlorpromazine; the latter prescribed for anxiety and insomnia. Treatment of choice for alcohol withdrawal was thioridazine, a drug previously only known to me as an antipsychotic. This was also a time when antipsychotics were prescribed to almost everyone with behavioral disturbances associated with dementia.

Over the years since then, my experience and observations of prescribing decision-making have been fairly uniform. Rarely is a firm diagnosis arrived at before prescribing and, perhaps more tellingly, rarely do prescribers know the precise licensed indications for every drug they prescribe. So, antipsychotics tend to be prescribed for psychotic and manic symptoms, and antidepressants tend to be prescribed for symptoms of low mood. Lithium is prescribed for mood fluctuations. Benzodiazepines are prescribed for pretty much everything. Never, in over 25 years in psychiatry, have I observed prescribers make any noticeable attempt to match a diagnosis with the licensed indication of a drug. I, for my part, would not now think to point out to a prescriber that, say, valproate has no label for prophylaxis in bipolar disorder.

These informal observations are backed-up by published data. In the US, off-label prescribing of antipsychotics amounted to 74% of all prescriptions in 1995 and 60% (nine million prescriptions) in 2008[1]. In the UK, 63% of risperidone prescribing in primary care was for unlicensed conditions in the period from 2007 to 2011[2]. Even in secondary care, where prescribing is often more tightly controlled by local protocols, off-label prescribing is common – in my own unit we recorded off-label prescribing rates of 11% for risperidone long-acting injection[3], 23% for oral aripiprazole[4] and 33% for depot paliperidone[5]. In each case these figures represent off-label prescribing in consecutive patients during the first few months or years after the introduction of the drug or formulation. Prescribers' willingness to prescribe outside a license is thus not solely a consequence of positive experiences of prescribing for specific off-label indications – it starts straightaway.

This is not necessarily poor practice: some of the most effective medicines in a particular condition do not have a license for that condition. Fluoxetine may be the most effective selective serotonin reuptake inhibitor in generalized anxiety disorder (GAD)[6], but is not labeled for this condition. Sertraline is recommended as first line treatment for GAD by the National Institute for Health and Care Excellence (NICE) in the UK, but it too has no license for this indication. Quetiapine is licensed in most countries for schizophrenia, schizoaffective disorder, mania, bipolar depression and unipolar depression, but it is also effective in other conditions such as GAD[7]; its range of actions is such that diagnosis becomes almost irrelevant.

The example of quetiapine encapsulates the issue in question. We may call it an antipsychotic but it is much more than that, because it has multiple pharmacological actions and numerous active metabolites[8]. The nonsense of current nomenclature – which is closely aligned with labeling – is well recognized, and other systems have been suggested. For example, neuroscience-based nomenclature classifies drugs according to their pharmacological profile rather than their primary or initial therapeutic indication[9].

The classification of mental health conditions gives us a false sense of order and a crude system for prescription reimbursement. It has little or no relevance to psychotropic drug action and as a consequence an accurate diagnosis is not required for optimal prescribing.

David Taylor
South London and Maudsley NHS Foundation Trust, Pharmacy Department, Maudsley Hospital, Denmark Hill, London, UK

1. Alexander GC, Gallagher SA, Mascola A et al. Pharmacoepidemiol Drug Saf 2011;20:177-84.
2. Marston L, Nazareth I, Petersen I et al. BMJ Open 2014;4:e006135.
3. Taylor DM, Fischetti C, Sparshatt A et al. J Clin Psychiatry 2009;70:196-200.
4. Taylor D, Atkinson J, Fischetti C et al. Acta Psychiatr Scand 2007;116:461-6.
5. Attard A, Olofinjana O, Cornelius V et al. Acta Psychiatr Scand 2014;130: 46-51.
6. Baldwin D, Woods R, Lawson R et al. BMJ 2011;342:d1199.
7. Kreys TJ, Phan SV. Pharmacotherapy 2015;35:175-88.
8. Fisher DS, Handley SA, Flanagan RJ et al. Ther Drug Monit 2012;34:415-21.
9. Zohar J, Stahl S, Möller HJ et al. Eur Neuropsychopharmacol 2015;25:2318-25.

DOI:10.1002/wps.20343

Rising suicide rates: an under-recognized role for the Internet?

According to recent data from the Centers for Disease Control and Prevention, the age-adjusted suicide rate increased by 24% in the US from 1999 to 2014, after a period of consistent decline from 1986 to 1999[1]. The increase, whose pace rose after 2006, occurred in males, females, and the entire 10 to 74 age range, but was more pronounced among females aged 10 to 14 and males aged 45 to 64. Several explanations have been proposed in the wide coverage that the data received: the black-box warning issued by the Food and Drug Administration in 2004 linking antidepressants and suicidality; the advancing age of puberty; inadequate health insurance coverage; stressors related to the Great Recession; the rising rate of

divorce; and the increased use of heroin and opiate drugs. Another possible contributor has received much less attention: the Internet, which dominated the period covered by the data, may have played a facilitating role.

The Internet may have become the first stop for many individuals contemplating suicide. There, pro-suicide sites account for over 11% of suicide-related search results according to one study[2]. Whether they support the freedom to commit suicide for everyone or focus on "death with dignity" for the terminally ill, pro-suicide sites often relay information on how to successfully commit suicide. They do so by "perfecting" well-known means that employ easily procurable ingredients, or introducing obscure new methods that, before the age of the search engine, would have required highly specialized knowledge and access (e.g., suicide by inhaling helium or ether, extracting and drinking nicotine, drinking excessive quantities of water). Beyond the informational aspect, the suicide-positive attitude of some sites can appeal to lonely sufferers desperate to commune with like-minded individuals. One effect may be to transform suicidal planning from a solitary choice widely seen as pathological into a shared experience, normalized by a supportive community that encourages the behavior. Extreme examples include online suicide pacts, group suicides and "live" suicides, where death is streamed spectacle-like to an abetting audience.

Why might an at-risk individual faced with such online content be more likely to pass to the act? The answer may partially lie in the interplay between impulsivity as a well-known trait in self-harm and impulsivity as a defining feature of the online experience.

Several studies have linked impulsivity with self-harm. One, conducted in 215 subjects with bipolar disorder, showed that impulsivity significantly increased the risk of suicidal acts[3]. Behavioral and neuroimaging data lend support to the suicidality-impulsivity link. Suicidal ideation has been associated with impairment on behavioral challenges that capture impulsive decision-making, such as the Iowa gambling and go/no-go tasks[4]. In neuroimaging studies, self-harm has been linked to diminished reward signal in areas that have been associated with impulse control[5].

Impulsivity, then, appears to be an important feature of suicide. It is also a defining feature of the online experience. Impulsive and disinhibited online behavior was described early in the Internet era. And, when disordered Internet use was considered as a possible new psychopathology, it was partly conceptualized as an impulse control disorder, with the definitions for "problematic Internet use" borrowing from diagnostic criteria for pathological gambling disorder, then categorized as an impulse control disorder in the DSM-IV[6].

A look at how other impulse-driven behaviors have been affected by the Internet may help explain how online content can influence a suicidal individual. Pathological gambling and compulsive buying have long been conceptualized as impulse control disorders, and data suggest Internet-mediated exacerbations of these behaviors.

Early calls for controlling online gambling stemmed from concerns about an online gambler's ability to go undetected and uninterrupted for long periods of time and the lack of the usual "fail safes" that can protect vulnerable users offline. Subsequent online gambling prevalence data suggest that these worries were justified. While many studies are limited by their online design and reliance on self-selected samples, one used a weighted approach to analyzing data from an international sample of 12,521 gamblers[7]. Among Internet gamblers, only 39.9% were not problem or at-risk gamblers compared to 82.1% of non-Internet gamblers.

Similarly, it was initially thought that the Internet would diminish compulsive buying by facilitating comparisons, protecting against in-store marketing, and obviating the need to go to a physical store thereby saving time. Instead, preliminary research seems to suggest a negative effect. For example, in a study of 314 customers of an online retail store, 17.7% met criteria for compulsive buying behavior[8]. This prevalence rate is considerably higher than those seen in studies that were conducted before the arrival of online retail or that did not focus exclusively on it.

The Internet may similarly make self-harm more impulsive, automatic and difficult to resist by diminishing the "obstacles" faced offline, teaching about methods, and providing an encouraging environment. Formal data on the influence of the Internet on suicidal behavior, including suicidal impulsivity, mostly consist of case reports. Still, some studies that compared Internet search trends with population-level data have yielded troublesome results. For example, in a Japanese study, searches for "hydrogen sulfide", "hydrogen sulfide suicide" and "suicide hydrogen sulfide" correlated with the incidence of suicide 11 months following the search, and "suicide by jumping" correlated with suicide six months later[9].

Such data make formal research into the potential pro-suicide effects of new technologies an important public health matter. But research into the Internet as a *prevention* tool cannot be ignored. While adequate control of pro-suicide, impulse-promoting content may be somewhat unrealistic in the difficult-to-govern online world, more can be done to make education, support and prevention sites more accessible, appealing and effective. Meanwhile, educating mental health professionals about the need to incorporate suicide-related online behavior in the assessment of at-risk patients would seem like a well-justified strategy. "Have you been googling suicide lately?" has become a requisite question in the thorough suicide evaluation.

Elias Aboujaoude
OCD Clinic, Stanford University School of Medicine, Stanford, CA, USA

1. Curtin SC, Warner M, Hedegaard H. Increase in suicide in the United States, 1999-2014. Hyattsville: National Center for Health Statistics, 2016.
2. Recupero PR, Harms SE, Noble JM. J Clin Psychiatry 2008;69:878-88.
3. Jiménez E, Arias B, Mitjans M et al. Acta Psychiatr Scand 2016;133:266-76.
4. Westheide J, Quednow BB, Kuhn KU et al. Eur Arch Psychiatry Clin Neurosci 2008;258:414-21.

5. Dombrovski AY, Szanko K, Clark L et al. JAMA Psychiatry 2013;70:1.
6. Aboujaoude E, Koran LM, Gamel N et al. CNS Spectr 2006;11:750-5.
7. Wood RT, Williams RJ. New Media and Society 2011;13:1123-41.
8. Kukar-Kinney M, Ridgway NM, Monroe KB. J Retailing 2009;85:298-307.
9. Hagihara A, Miyazaki S, Abe T. Eur Arch Psychiatry Clin Neurosci 2012;262: 39-46.

DOI:10.1002/wps.20344

Abandoning personalization to get to precision in the pharmacotherapy of depression

Roy H. Perlis

Center for Experimental Drugs and Diagnostics, Department of Psychiatry and Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Effectiveness studies and analyses of naturalistic cohorts demonstrate that many patients with major depressive disorder do not experience symptomatic remission with antidepressant treatments. In an effort to better match patients with effective treatments, numerous investigations of predictors or moderators of treatment response have been reported over the past five decades, including clinical features as well as biological measures. However, none of these have entered routine clinical practice; instead, clinicians typically personalize treatment on the basis of patient preferences as well as their own. Here, we review the reasons why it has been challenging to identify and deploy treatment-specific predictors of response, and suggest strategies that may be required to achieve true precision in the pharmacotherapy of depression. We emphasize the need for changes in how depression care is delivered, measured, and used to inform future practice.

Key words: Antidepressants, major depression, precision medicine, risk stratification, personalized medicine, biomarkers, treatment matching

(World Psychiatry 2016;15:228–235)

After decades of effort to identify predictors of antidepressant treatment response, including more than 100 publications reporting genetic predictors, the approach to treating major depressive disorder remains one of trial and error. Initial management strategies vary widely across providers and health systems[1]. Next-step treatment is marked by even greater variation[2]. A recent survey of psychopharmacologists, for example, revealed roughly equal split between within- and across-antidepressant class switch following non-response to initial treatment[3]. This trial-and-error approach clearly matters to patients: a survey of Danish patients found that they would pay up to $280 to avoid a single medication change[4].

At the same time, pharmacogenomics has already made some clinical inroads in antidepressant prescribing. Among the more than 100 medication labels approved by the US Food and Drug Administration (FDA) that include information on genetic variation, at least 10 pertain to antidepressant pharmacotherapies or medications commonly used to augment antidepressants[5]. Multiple marketed assays are intended to guide antidepressant treatment; while none have yet pursued FDA approval, such diagnostic tests are available commercially from the laboratories that developed them. And clinical guidelines for the use of pharmacogenomic testing are available from US and international agencies[6]. Still, very few patients receive such testing, and its utility remains unclear, in part because of a relative lack of randomized controlled studies indicating benefit.

In this paper, we focus on the scientific challenges that have contributed to the persistence of artisanal prescribing of antidepressants even in the face of growing enthusiasm for the concept originally described as personalization, then stratification, and most recently precision medicine[7]. We also review the obstacles to translation of pharmacogenomic tools to common clinical practice. Finally, we address strategies that could be helpful in ensuring that the next decade does bring significant progress towards achieving true precision in the pharmacotherapy of depression.

WHAT ARE THE CHALLENGES IN PERSONALIZING ANTIDEPRESSANT TREATMENT?

Personalization is not precision

In oncology, the concept of matching treatments to patients to achieve and maintain remission is well established: there are particular tumor profiles that respond differentially to particular interventions. For major depressive disorder, while remission certainly remains a key goal, other considerations are also important: in addition to safety, clinicians may consider key symptoms to target and key adverse effects to avoid.

To this end, psychiatrists (and primary care physicians) already personalize treatment, albeit in a more artisanal and less scientific fashion than oncologists. A systematic approach to this process has been described by Preskorn[8]. Essentially, some medications are excluded on the basis of safety: for example, medications like bupropion that lower seizure threshold might be avoided in individuals at high risk for seizures. Others are avoided on the basis of adverse effects: in an obese patient, medications that commonly increase appetite, such as mirtazapine, would be excluded from initial consideration. Among the remaining options, some clinicians simply pick their favorite; others follow guidelines approved by their employer or payers, perhaps based on which medications are available at lowest cost; and others provide an individual patient with a few choices and discuss adverse effect profiles for each. The difficulty here is that, while most clinicians likely follow some variant of these approaches, there is no agreed-upon or evidence-based framework for such practices.

The evidence base for next-step interventions is even more modest. A particular challenge is the emphasis on randomized controlled trials, that tends to favor more recent industry-supported studies. Consider the case of augmentation: the strongest evidence base supports certain second-generation antipsychotics, simply because older strategies (for example, bupropion, buspirone or pramipexole) involve off-label use of medications long since generic. So, even on the basis of evidence-based personalization, the clinician cannot be strongly informed by treatment guidelines that tend to simply count large-scale positive trials.

In summary, clinicians already personalize, but in a haphazard and inconsistent way. Unfortunately, the very resistance to more systematic treatment approaches, like algorithms and guidelines, on the basis of the need to personalize actually hinders efforts at personalization: there is no agreed-upon standard on which to improve. Faced with an algorithm, many clinicians insist on the need to tailor treatment depending on particular clinical features, even in the absence of strong evidence that such features are truly predictive. The missing ingredient here may be humility: most clinicians likely rate themselves as above average in terms of ability to identify efficacious treatments, but clearly some are not. Ironically, one of the advantages of biologically-based treatment selection would be the ability to introduce more systematic approaches while preventing narcissistic injury to clinicians.

Treatment-specific effects are modest

Beyond a general resistance to external guidance on prescribing is the larger problem that treatment-specific differences in efficacy appear likely to be quite modest. While antidepressants are more effective than placebo, the magnitude of this difference is generally small, at least in the outpatient context[9]. This does not mean that prediction cannot be useful, just that some such prediction is actually pertaining to placebo response and thus by definition not treatment-specific. As discussed below, such non-specific predictors may still be useful in stratifying treatment intensity, if not specific treatment choice.

Data needed to compare active treatments are lacking

The regulation of medications in the US does not require active comparator studies: there is no obligation (or even expectation) that a new drug be superior to an existing one. So, not surprisingly, such studies are rarely done, and when they are, they are likely to be engineered to yield results which are misleading at best, with an active comparator group included only for "assay sensitivity" which may not even be analyzed in comparison with the study medication.

In the rare cases where straightforward comparator studies are done, they have tended to be a poor investment for the sponsor: treatment differences are likely very small on average, and the substantial placebo response places a floor effect on the performance of comparator drugs (unless comparators are actually worse than placebo, a phenomenon rarely encountered in psychopharmacology[10]).

Moreover, where large comparator studies are done, the data are typically held by the industry sponsor. Until recently, large pharmaceutical companies have been reluctant to share DNA or genotypic data in conjunction with treatment response, even where they did agree to sharing for association studies of disease. Presumably, the risk of finding a predictor of non-response, or perhaps the perceived need to involve the FDA in reporting genomic data pertaining to marketed drugs, has outweighed the scientific interest in such work.

Power is accordingly poor to find real effects

Combining a small effect with a small sample size represents a recipe for an underpowered study – one where the risk for both false positive and false negative findings is high[11]. Worse, because of the problem of "winner's curse", even when true effects are detected, they are likely to be overestimated – thus the pattern all too familiar in psychiatric pharmacogenomics in which initial exciting findings subsequently prove either to be false positives or of less importance than anticipated[12,13].

Standard statistical approaches to finding predictors of differential response between two or more interventions rely on a test for a treatment-by-predictor interaction, which is substantially less powerful than tests for main effect. Notably, such a test has greatest power when a predictor has opposite effects in two groups: for example, biomarker A is associated with greater-than-average response to fluoxetine, but worse-than-average response to bupropion. Biologically, this scenario seems implausible: more likely, biomarker A is associated with greater-than-average response to fluoxetine but no difference with bupropion. In this scenario, a test for interaction is even less powerful.

GETTING TO PRECISION

Begin using what we know rather than seeking a silver bullet

Efforts at personalization may have suffered from their ambition, with an unwillingness to employ more basic or mundane socio-demographic predictors in pursuit of a single powerful biomarker. In reality, multiple studies suggest that readily available patient-level features may at least help to set prior probability of response.

Phenomenology

Among the earliest putative predictors were the depressive subtypes, melancholic and atypical depression. An extensive literature explored these sets of symptoms in terms of phenomenology and associated peripheral markers. This literature illustrates some of the challenges in identifying response predictors.

Melancholic depression in general is highly correlated with depression severity, such that, while it is associated with poorer outcomes in general, these outcomes may better be explained by considering total severity. This underscores the importance of ensuring that putative predictors represent the easiest or most straightforward means of measuring a phenomenon. The value of total severity in this regard is further discussed below.

Atypical depression has been difficult to establish as a strong predictor of outcome because of problems in distinguishing individual symptoms from a true subtype. Empirical evidence suggested that, while reversed neurovegetative signs – hypersomnia rather than insomnia, or hyperphagia rather than loss of appetite – are common, they do not necessarily represent a distinct subgroup. That is, many patients may experience one or the other. Worse, as symptoms may fluctuate over time and across episodes, the determination of whether a patient meets criteria for this subtype likely depends upon when in the episode course the patient is assessed.

More recently, Fava et al[14] suggested additional depressive subtypes on the basis of questionnaires included in the baseline assessment for clinical trial participants – in particular, emphasizing the notion of hostile (irritable) and anxious depression. Both of these strongly predict poorer outcomes across multiple studies[15]. However, in addition to correlation with each other, they are also correlated with total severity, and like the other clinical subtypes may fluctuate within an episode.

Anxious depression in particular received some support from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, where it predicted poorer treatment response[16]. A subsequent replication effort in the Genome-based Therapeutic Drugs for Depression (GENDEP) study, however, did not provide further support[17]. This non-replication may suggest the importance of considering reference populations when attempting to derive predictors.

One of the most robust recent predictors of outcome was described by Uher et al[17] using results of factor analysis in lieu of the traditional depressive subtypes. They found that an interest-activity symptom dimension at baseline – which captured poor interest, decreased activity, indecisiveness, and anhedonia – was strongly associated with poor outcome both in GENDEP and in the larger STAR*D study. This association persisted despite control for overall severity and type of antidepressant.

As one of the best-validated predictors of outcome other than total severity, it would seem that the interest-activity factor could represent a good starting point for stratification. That it has not been so applied relates in part to the unwillingness of most clinical practices to employ systematic assessment of symptoms, notwithstanding the imposition of the Patient Health Questionnaire (PHQ-9) in primary care settings. This obstacle is discussed further below.

Notably, efforts to identify predictors of differential treatment response (often described as moderators of response[18]) also date back to the dawn of structured psychotherapies. These investigations often focus on specific scales quantifying the target of a particular kind of intervention. For example, the Coping Self-Efficacy Scale was a predictor of response to cognitive-behavioral therapy, delivered either by telephone or in person[19].

Another strategy attempts to integrate socio-demographic and clinical features to predict treatment resistance in major depressive disorder. From among a larger panel of variables, symptoms predictive of treatment resistance included insomnia and decreased energy, along with elements of history such as trauma exposure, post-traumatic stress disorder, and even mild psychotic-like symptoms. In an independent validation cohort also drawn from the STAR*D study, but from different sites, specificity for treatment resistant depression exceeded 0.91, although sensitivity was lower at 0.26[20]. This study also produced a risk visualization tool (http://trdrisk.mghcedd.org), intended to promote development of similar efforts integrating clinical and genomic data.

Employing any of these simple predictors would in no way preempt the use of biological predictors as they are identified. Indeed, even a simple baseline model would be a valuable basis for comparison with newer models – a starting point to be improved on by adding biological or other predictors. In this context, frameworks such as net reclassification improvement[21] may be more useful for understanding how the addition of a new marker improves prediction, compared to standard metrics such as area under receiver operating characteristic curve[22].

Genetic and genomic predictors

Among the potential biological predictors of outcome, cytochrome P450 (CYP450) variation has been understood to influence blood levels of multiple drugs for two decades or more. Unlike most genetic associations, the functional implications of the key variations have been described – that is, particular alleles are known to increase or decrease enzyme activity in a predictable way[23].

The central challenge to the use of CYP450 testing for antidepressant prescribing stems from the lack of a clear relationship between blood levels and either efficacy or adverse effects. At the extremes, some relationship is intuitive: individuals with undetectable blood levels will not respond to true drug effects (although they may still respond to placebo); individuals with supra-therapeutic blood levels should be more likely to experience adverse effects. However, for most antidepressants, even a simple dose-response relationship has been difficult to establish.

Given the clearer relationship of efficacy (and toxicity) of tricyclics to blood levels, it is unsurprising that this is the class of antidepressants with the strongest evidence that CYP450 testing is likely to inform dosing. Unfortunately, despite the substantial efforts expended to develop and promote guidelines for CYP450-informed dosing[7], this class has largely been superseded by other antidepressants on the basis of equivalent efficacy and wider therapeutic index (i.e., greater

margin of safety). So, the intervention where precision medicine in depression treatment may be most feasible is now also the one least clinically useful. The term in decision analysis for this scenario is a dominating choice: in most if not all circumstances, the cost-effectiveness of CYP450-guided tricyclic treatment will be less than that of simply prescribing a generic non-tricyclic.

For selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs), the impact of CYP450 variation is not fully understood. Most SSRIs and SNRIs are substrates for one of the common CYP450 enzyme systems, so it is possible to make predictions about changes in blood levels. What those levels mean, though, is not so clear: with the possible exception of modest data regarding fluoxetine[24,25] and venlafaxine[26], higher doses within the therapeutic range have generally not been shown to be more efficacious than lower doses. The evidence of poor tolerability at higher doses is rarely studied directly, particularly as it relates to CYP450 status: one study suggested that non-wild-type metabolizers of the CYP450 2C19 substrate citalopram experienced poorer tolerability with this treatment[27].

Further, even in circumstances where drug blood levels are important, CYP450 variation is only one contributor to such levels. Numerous environmental factors, including diet and other medications (as well as other, unmeasured genetic variation), may be important. One illustration of these effects was a study of venlafaxine-treated patients that examined the plasma ratio of venlafaxine to its metabolite desvenlafaxine in order to define individuals who were "functionally" poor metabolizers. Overall, 27% of individuals appeared to be poor metabolizers, even though only 4% were CYP450 2D6 poor metabolizers genotypically[28].

As venlafaxine is the pro-drug for desvenlafaxine, individuals who are poor metabolizers at CYP450 2D6 might be hypothesized to be less likely to respond to treatment (as they will have very low effective levels of the active drug). Indeed, in the four venlafaxine studies, poor metabolizers were less likely to achieve remission than wild-type metabolizers[29].

Other biomarkers

Efforts to identify predictors of differential antidepressant treatment response based on blood or other peripheral measures date back to the dawn of psychopharmacology. The dexamethasone suppression test (initially considered to diagnose depression, and later employed to guide treatment) presents a useful cautionary tale of a diagnostic tool deployed in psychiatry without sufficient consideration of its utility, or even what exactly it predicted[30,31].

An example of a prototypical predictor might be C-reactive protein (CRP), a marker of inflammation associated with cardiovascular disease. In the GENDEP study, a notable treatment-by-predictor interaction – exactly the sort that could potentially be informative for treatment selection – was identified with CRP. Specifically, symptomatic improvement was greater with escitalopram treatment among individuals with CRP levels lower than 1 mg/L, while it was greater with nortriptyline treatment among individuals with CRP levels higher than 1 mg/L. Still, given the poorer safety profile of tricyclic antidepressants, the modest difference in efficacy (three points on the Montgomery-Åsberg Depression Rating Scale) may not be sufficient to justify preferential use of nortriptyline even in the latter patient subset. While some frameworks for defining clinical significance exist – see, for example, the calculator at depressiontools.org[32] – the necessary effect size for utility of a given predictor depends critically on its context.

Numerous other minimally invasive markers are under active investigation for response prediction. Functional neuroimaging has been perhaps the most studied, with intriguing but not definitive results – not surprising given the relatively small cohorts studied. Similarly, quantitative electroencephalography has been applied to predict either overall treatment outcome or differential response. In a representative small study, a measure of frontal recordings at baseline and week 1 was associated with speed and probability of response to escitalopram over 13 weeks[33], consistent with a prior pilot study using fluoxetine[34]. The pilot study, importantly, included a placebo arm where no such an association was identified. Still, as noted earlier, the absence of any comparison drug makes the specificity of this effect unclear. One other notable aspect of these studies is the inclusion of a post-baseline (week 1) time point in the biomarker: prediction of outcome based on short-term treatment exposure, while not a standard strategy in psychopharmacology, may be easier than relying solely on baseline measures.

Educate clinicians and patients

In addition to patient education, preliminary experience with genomic testing suggests the necessity and value of clinician education[35], in terms of how results are presented to patients and families. These tests typically yield probabilistic results, very different from the dichotomous outcome yielded by many other tests in medicine, though common in other areas such as cancer, where estimates of survival are the coin of the realm. In one pilot pharmacogenomic study of antidepressant response, only 1/4 of consented patients were able to indicate an understanding of such testing[36].

A particular concern in psychopharmacology is the misinterpretation of CYP450 results as contraindicating a medication or class of medications. In light of the relative paucity of good therapeutic options, particularly for patients who do not remit with first-line treatments, ruling out a medication unnecessarily can be highly consequential. In reality, non-wild-type metabolizers simply require more cautious and informed titration: those who are poor metabolizers require lower doses of substrate drugs, while those who are ultrarapid metabolizers may require doses exceeding the FDA labeling, though still with careful titration. While simply avoiding

substrate drugs is a basic heuristic that may be reasonable when selecting initial treatments, such a heuristic can actually be detrimental as the range of reasonable options narrows. To this end, the tendency to present CYP450 results with color-coding – listing substrates in red, or with a stop sign, for example – may be unhelpful.

For both medications and diagnostic tools, clinician education can be mandated by the FDA within the approval process as part of the risk evaluation and mitigation strategy[37]. Similar education may be required for some interventions aimed at personalization of antidepressant prescribing, if only to limit the consequences of misinterpretation of test results.

Aim for stratification, not treatment-matching

Even where we cannot identify medication-specific predictors, distinguishing high- or low-risk groups may still be extremely useful. Three examples include greater depression severity, the interest-energy factor identified by Uher et al[17], and the treatment resistance risk score described earlier: to date, each of these appears to be a predictor of poorer outcome in general, rather than a feature that identifies an optimal treatment. So, while presence of greater risk may not help with selection of an individual treatment – venlafaxine versus fluoxetine, for example – it may instead indicate that a particular patient requires more intensive treatment in general. Individuals at high risk for treatment resistance could be triaged to more aggressive interventions – combination treatment, or incorporation of cognitive-behavioral therapy – or even more aggressive assessments, like specialist consultation or application of more intensive diagnostic tools.

Our approach to initial non-response needs to change

Protocols and randomization

Ironically, moving towards more truly personalized medicine may require moving away from traditional means of personalization by enrolling patients in protocol-driven treatment, much as is the case with cancer chemotherapies at academic medical centers. While clinicians maintain the importance of artisanal personalization, we are aware of no empirical data to indicate that such strategies improve upon uniform or standardized treatment selection (much less random selection among a small number of similar options). As much as it pains the expert psychopharmacologist to recognize this point, in general the clinician is at equipoise among multiple next-step strategies. Recent survey results reinforce this point[3]. But if this inconvenient reality were acknowledged and disclosed ("There are several reasonable next steps, we're going to let the computer select the first one to try"), it is possible that different strategies could be investigated.

Systematic measurement of outcomes

A related problem remains clinicians' reluctance to incorporate systematic measurement of outcomes – any outcomes – into their practices. The reasons for this resistance are manifold: the measures can be time consuming, they are rarely well integrated into clinical workflow, and they fail to capture the breadth of depressive symptomatology clinicians feel they need. While less often acknowledged, such measurement is likely to also create a bias to action: that is, identifying symptoms creates more requirement to act on these symptoms, or potentially liability for not acting on them.

Many health systems have elected to invest in the PHQ-9, a depression screening tool with limited utility for outcome measurement (a role it was never designed to fill). More recently, led in part by movements aiming for more patient-centered care followed by financial support from the Patient Centered Outcomes Research Institute, enthusiasm has grown for patient-reported outcomes – particularly measures of functional status and quality of life.

It seems reasonable to measure the improvement yielded by psychiatric interventions in a systematic way. To the clinicians who argue that the PHQ-9 captures only a limited amount of the benefit they provide, a reasonable response is to agree, and ask what better measures can be employed. Whatever psychosocial or pharmacological interventions do for depression, it should be possible to measure it. Either less intrusive and better integrated measures need to be found, or more resources need to be provided for clinicians to incorporate such measures. Notwithstanding the massive hyperbole currently attached to ambulatory monitoring devices, cell-phone-based survey tools may help to fill this void[38] – provided better platforms can be developed to safely and efficiently integrate these data for use by clinicians.

Use of electronic medical records and other large data sets

Yet another opportunity to improve precision in antidepressant treatment comes from the increasing availability of large clinical data sets, i.e., electronic medical records, with or without linkages to biobanks. These data sets provide a rich trove of clinical detail, typically far exceeding what is available from the health claims data sets employed for pharmacogivilance studies and health services research[39]. Compared to standard clinical trials, the patients and outcomes are likely to be more generalizable, as the biases inherent in patient recruitment are avoided. When biological materials – DNA or plasma, for example – are available, these resources also allow highly efficient *in silico* biomarker studies.

We have previously demonstrated the utility of electronic medical records for defining antidepressant treatment outcomes[40], and applying these metrics to characterize clinical[41] and genetic[42] predictors of non-response. A less well appreciated benefit of such large cohorts is the ability to study relatively rare but serious adverse effects, such as lithium-associated renal failure[43]. These designs

also facilitate investigation of quantitative drug effects, such as antidepressant-associated weight gain[44] or QT-interval prolongation[45].

Still, some important caveats apply to approaches using electronic medical records or national health registries. First, treatment assignment is not randomized, so the risk for confounding – particularly confounding by indication, in which the indication for a particular treatment confounds the result – is substantial (for an illustration of the impact of such confounding, see the study by Gallagher et al[41], in which treatment with non-steroidal anti-inflammatory drugs was associated with poorer antidepressant treatment outcome until the indication – e.g., pain– was controlled for). Statistical methods can help to control bias, but the risk for confounding cannot be entirely eliminated. Second, clinical care typically includes less precise measures of outcome as well as other relevant clinical covariates. In some cases proxy measures may suffice (hospitalization; treatment changes), but traditional clinical trial outcomes such as remission and response are more challenging to characterize. Indeed, one observation from studies based on electronic medical records[40] (and consistent with some mood disorder cohort studies[46]) is the extent to which episodic definitions of depression likely underestimate chronicity and persistence of residual symptoms relative to clinical cohorts.

Randomized trials of precision medicine will be needed... or will they?

Despite the utility of alternative approaches, randomized controlled trials remain the gold standard for investigating new interventions, pharmacological or otherwise. Even for pharmacotherapy, there has been continued innovation in the design and conduct of such studies. But for diagnostic tests, the optimal design of randomized trials remains subject to debate. For example, if subjects are to be randomized to assay-guided treatment compared to treatment as usual, how constrained or algorithmic should the treatment as usual be? Should clinicians stay unblinded, or should they receive a "dummy" or uninformative report? If the latter, is it ethical to delay reporting results (or even to report misleading results), and will clinicians be able to distinguish an uninformative from a placebo report? Design of a treatment-as-usual arm is particularly challenging as the inclination is to reduce heterogeneity by making this intervention more structured and algorithmic. However, as we have noted, standard of care is far from algorithmic at present, so this sort of comparator is artificial and itself likely to improve outcomes[47].

A further, practical problem is deciding who should pay for these studies. If the tools are developed by a for-profit entity, it is reasonable to require that the entity fund such studies. However, this barrier may be prohibitive for smaller companies or less costly tests. The shifting regulatory structure in the US, in which the FDA has allowed marketing of laboratory-developed tests without premarket review, but has indicated that it intends to increase oversight of this pathway[48], is likely to increase the pressure to conduct such randomized trials, if not the available resources.

To date, there is one small randomized trial investigating a pharmacogenetic assay for antidepressant prescribing, relying on a panel of CYP450 variants (2D6, 2C19 and 1A2) as well as some pharmacodynamic common variants. Among 51 outpatients with major depressive disorder, followed for eight weeks, the magnitude of improvement was numerically but not statistically significantly different between the treatment-as-usual (19%) and the assay-guided treatment (31%) groups ($p=0.3$). One of two unblinded cohort studies using the same assay did identify significantly greater improvement in the assay-guided treatment group[49]. In the absence of blinding or consideration of the impact of individual predictors, estimates of the benefit associated with specific variants await randomized trials.

In the meantime, electronic medical records or claims data may help to understand the potential impact of putative predictors of response. One approach uses cost-effectiveness analysis to examine the effect of a predictor, based on other assumptions about treatment costs and outcomes. In an illustration of this approach, we previously developed a model based on STAR*D data[50] considering a predictor of differential SSRI response. Under some assumptions, even a moderate difference between treatments was not cost-effective simply because using an alternate antidepressant was a dominating (better) choice. On the other hand, for a low-cost test, when the likelihood of an informative test is high, even relatively modest effect sizes could be cost-effective.

A major limitation in all such models is the need to make numerous assumptions about costs, probabilities and utilities. Their value is primarily in clarifying the circumstances where precision medicine may be most likely to be beneficial in antidepressant prescribing, as a means of designing future interventions.

Another perspective on cost-effectiveness comes from investigation of insurance claims databases in which some patients have already received pharmacogenomic testing. For the assay with the negative randomized trial described earlier, this investigation found that individuals receiving medications indicated to be "less desirable", based on an algorithm incorporating multiple variants, incurred greater past-year health care costs[51]. Whether such high-cost individuals represent an optimal population for deploying precision medicine is an intriguing, but as yet untested, hypothesis.

A direct but unrandomized assessment of cost-effectiveness comes from another study of health claims data that compared a cohort of 111 individuals who had received a commercial test combining CYP450 and pharmacodynamic variants with a propensity-score matched cohort who did not receive testing[52]. While not a substitute for randomization, this method allows some control of confounding by matching an unexposed (untested) group as closely as possible to the tested group. That study found, after matching and

adjustment, that outpatient treatment costs were 9.5% lower among tested patients. It also identified improvement in medication adherence among the tested group. Still, like other reports of pharmacogenomic testing outcomes, the absence of analysis by individual variant precludes an understanding of the elements of the assay most important for prediction.

CONCLUSIONS

Personalized medicine is already a reality in the treatment of depression, but precision medicine is not – that is, while clinicians routinely attempt to match treatments to patients, these strategies are neither systematic nor empirically supported. Making the transition to precision medicine will, first, require a commitment to the systematic practice of medicine: following algorithms or guidelines, and measuring patient outcomes to guide decision making. If physicians trained to rely on the art of medicine cannot make this transition, it is likely that nurse-clinicians and pharmacists will make it for them. Second, it will require a willingness to begin to study and deploy risk stratification tools that may not be perfect, but rather better than the current standard of care. A further benefit of these two steps will be an acceleration of the ability to develop and investigate new personalization strategies, because it will become more straightforward to identify biomarkers and study them in large clinical cohorts.

Evidence from effectiveness studies and clinical cohorts indicate that many patients remain poorly served by existing antidepressant treatments. Aiming for more precise treatment matching may help to ensure optimal outcomes even while the field strives for better treatment options.

REFERENCES

1. Johnson CF, Dougall NJ, Williams B et al. Patient factors associated with SSRI dose for depression treatment in general practice: a primary care cross sectional study. BMC Family Pract 2014;15:210.
2. von Wolff A, Meister R, Harter M et al. Treatment patterns in inpatient depression care. Int J Methods Psychiatr Res 2016;25:55-67.
3. Goldberg JF, Freeman MP, Balon R et al. The American Society of Clinical Psychopharmacology Survey of Psychopharmacologists' Practice Patterns for the Treatment of Mood Disorders. Depress Anxiety 2015;32:605-13.
4. Herbild L, Bech M, Gyrd-Hansen D. Estimating the Danish populations' preferences for pharmacogenetic testing using a discrete choice experiment. The case of treating depression. Value Health 2009;12:560-7.
5. National Research Council of the National Academies. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington: National Academies Press, 2011.
6. US Food and Drug Administration. Table of pharmacogenomic biomarkers in drug labeling 2014. www.fda.gov.
7. Hicks JK, Swen JJ, Thorn CF et al. Clinical Pharmacogenetics Implementation Consortium guideline for CYP2D6 and CYP2C19 genotypes and dosing of tricyclic antidepressants. Clin Pharmacol Ther 2013;93:402-8.
8. Preskorn SH. Outpatient management of depression: a guide for the primary-care practitioner, 2nd ed. Caddo: Professional Communications, 1999.
9. Fountoulakis KN, Möller HJ. Antidepressant drugs and the response in the placebo group: the real problem lies in our understanding of the issue. J Psychopharmacol 2012;26:744-50.
10. Pande AC, Crockatt JG, Janney CA et al. Gabapentin in bipolar disorder: a placebo-controlled trial of adjunctive therapy. Gabapentin Bipolar Disorder Study Group. Bipolar Disord 2000;2:249-55.
11. Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.
12. Lee CS, Cheng AT. Variant GADL1 and response to lithium in bipolar I disorder. N Engl J Med 2014;370:1859-60.
13. Chen CH, Lee CS, Lee MT et al. Variant GADL1 and response to lithium therapy in bipolar I disorder. N Engl J Med 2014;370:119-28.
14. Fava M, Uebelacker LA, Alpert JE et al. Major depressive subtypes and treatment response. Biol Psychiatry 1997;42:568-76.
15. Perlis RH, Uher R, Ostacher M et al. Association between bipolar spectrum features and treatment outcomes in outpatients with major depressive disorder. Arch Gen Psychiatry 2011;68:351-60.
16. Fava M, Rush AJ, Alpert JE et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. Am J Psychiatry 2008;165:342-51.
17. Uher R, Perlis RH, Henigsberg N et al. Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. Psychol Med 2012;42:967-80.
18. Kraemer HC, Stice E, Kazdin A et al. How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. Am J Psychiatry 2001;158:848-56.
19. Stiles-Shields C, Corden ME, Kwasny MJ et al. Predictors of outcome for telephone and face-to-face administered cognitive behavioral therapy for depression. Psychol Med 2015;45:3205-15.
20. Perlis RH. A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. Biol Psychiatry 2013;74:7-14.
21. Pencina MJ, D'Agostino RB Sr, Demler OV. Novel metrics for evaluating improvement in discrimination: net reclassification and integrated discrimination improvement for normal variables and nested models. Stat Med 2012;31:101-13.
22. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928-35.
23. De Gregori M, Allegri M, De Gregori S et al. How and why to screen for CYP2D6 interindividual variability in patients under pharmacological treatments. Curr Drug Metab 2010;11:276-82.
24. Fava M, Alpert J, Nierenberg A et al. Double-blind study of high-dose fluoxetine versus lithium or desipramine augmentation of fluoxetine in partial responders and nonresponders to fluoxetine. J Clin Psychopharmacol 2002;22:379-87.
25. Fava M, Rosenbaum JF, McGrath PJ et al. Lithium and tricyclic augmentation of fluoxetine treatment for resistant major depression: a double-blind, controlled study. Am J Psychiatry 1994;151:1372-4.
26. Berney P. Dose-response relationship of recent antidepressants in the short-term treatment of depression. Dialogues Clin Neurosci 2005;7:249-62.
27. Mrazek DA, Biernacka JM, O'Kane DJ et al. CYP2C19 variation and citalopram response. Pharmacogenet Genomics 2011;21:1-9.
28. Preskorn SH, Kane CP, Lobello K et al. Cytochrome P450 2D6 phenoconversion is common in patients being treated for depression: implications for personalized medicine. J Clin Psychiatry 2013;74:614-21.
29. Lobello KW, Preskorn SH, Guico-Pabia CJ et al. Cytochrome P450 2D6 phenotype predicts antidepressant efficacy of venlafaxine: a secondary analysis of 4 studies in major depressive disorder. J Clin Psychiatry 2010;71:1482-7.
30. Baldessarini RJ, Arana GW. Does the dexamethasone suppression test have clinical utility in psychiatry? J Clin Psychiatry 1985;46:25-9.
31. Perlis RH. Translating biomarkers to clinical practice. Mol Psychiatry 2011;16:1076-87.
32. Uher R, Tansey KE, Malki K et al. Biomarkers predicting treatment outcome in depression: what is clinically significant? Pharmacogenomics 2012;13:233-40.
33. Cook IA, Hunter AM, Gilmer WS et al. Quantitative electroencephalogram biomarkers for predicting likelihood and speed of achieving sustained remission in major depression: a report from the biomarkers for rapid identification of treatment effectiveness in major depression (BRITE-MD) trial. J Clin Psychiatry 2013;74:51-6.
34. Hunter AM, Cook IA, Greenwald SD et al. The antidepressant treatment response index and treatment outcomes in a placebo-controlled trial of fluoxetine. J Clin Neurophysiol 2011;28:478-82.
35. Roberts JS, Chen CA, Uhlmann WR et al. Effectiveness of a condensed protocol for disclosing APOE genotype and providing risk education for Alzheimer disease. Genet Med 2012;14:742-8.

36. Rose D, Russo J, Wykes T. Taking part in a pharmacogenetic clinical trial: assessment of trial participants understanding of information disclosed during the informed consent process. BMC Med Ethics 2013;14:34.
37. US Food and Drug Administration. Risk evaluation and mitigation strategies 2015. www.fda.gov.
38. Marcano Belisario JS, Jamsek J, Huckvale K et al. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015;7:MR000042.
39. Kohane IS. Using electronic health records to drive discovery in disease genomics. Nat Rev Genet 2011;12:417-28.
40. Perlis RH, Iosifescu DV, Castro VM et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. Psychol Med 2012;42: 41-50.
41. Gallagher PJ, Castro V, Fava M et al. Antidepressant response in patients with major depression exposed to NSAIDs: a pharmacovigilance study. Am J Psychiatry 2012;169:1065-72.
42. O'Dushlaine C, Ripke S, Ruderfer DM et al. Rare copy number variation in treatment-resistant major depressive disorder. Biol Psychiatry 2014;76:536-41.
43. Castro VM, Roberson AM, McCoy T et al. Stratifying risk for renal insufficiency among lithium-treated patients: an electronic health record study. Neuropsychopharmacology 2016; 41:1138-43.
44. Blumenthal SR, Castro VM, Clements CC et al. An electronic health records study of long-term weight gain following antidepressant use. JAMA Psychiatry 2014;71:889-96.
45. Castro VM, Clements CC, Murphy SN et al. QT interval and antidepressant use: a cross sectional study of electronic health records. BMJ 2013;346:f288.
46. Perlis RH, Dennehy EB, Miklowitz DJ et al. Retrospective age at onset of bipolar disorder and outcome during two-year follow-up: results from the STEP-BD study. Bipolar Disord 2009; 11:391-400.
47. Guo T, Xiang YT, Xiao L et al. Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. Am J Psychiatry 2015;172:1004-13.
48. US Food and Drug Administration. Laboratory developed tests 2015. www.fda.gov.
49. Hall-Flavin DK, Winner JG, Allen JD et al. Using a pharmacogenomic algorithm to guide the treatment of depression. Transl Psychiatry 2012;2:e172.
50. Perlis RH, Patrick A, Smoller JW et al. When is pharmacogenetic testing for antidepressant response ready for the clinic? A cost-effectiveness analysis based on data from the STAR*D study. Neuropsychopharmacology 2009; 34:2227-36.
51. Winner J, Allen JD, Altar CA et al. Psychiatric pharmacogenomics predicts health resource utilization of outpatients with anxiety and depression. Transl Psychiatry 2013;3:e242.
52. Fagerness J, Fonseca E, Hess GP et al. Pharmacogenetic-guided psychiatric intervention associated with increased adherence and cost savings. Am J Manag Care 2014;20:e146-56.

DOI:10.1002/wps.20345

Taking the depressed "person" into account before moving into personalized or precision medicine

Clinicians and patients suffering from major depression are confronted with the gap between guidelines produced by so-called evidence-based medicine and prescription patterns emerging from experience-based medicine, as well as with the gap between artisanal prescribing and the siren song of personalization, stratification and precision medicine.

Perlis' elegant paper describes the many challenges in abandoning personalization to get to precision in the pharmacotherapy of depression[1]. Currently, physicians indeed practice some form of personalization by taking into account the patient's symptom profile as well as the safety and tolerability of the different antidepressants, although taking the symptom profile into account is not or poorly empirically supported. Actually, the choice of a specific antidepressant is mainly based on the presence of a specific symptom (52%) or the wish to avoid a specific side effect (49%), and the specific symptoms considered are mainly anxiety (20%), insomnia (18%) and fatigue (14%)[2].

Before antidepressant treatment can start moving from artisanal prescribing to precision medicine, several issues should be addressed. Taking more into account the "anima" (the individual, the real person, as well as the illness) and not only the "persona" (mask, character imposed by our diagnostic and assessment tools) seems to be mandatory.

That randomized clinical trials (RCTs) really represent the gold standard is questionable: "never before have the inadequacies of RCTs been so apparent to so many; yet, equally, never before have those in position of authority – from regulators, to policy-makers, to doctors – relied so extensively on RCTs' evidence"[3]. Efficacy estimates are usually based upon RCTs, but only about 10 to 20% of daily practice patients "fit into" exclusion and inclusion criteria.

Furthermore, efficacy estimates taken from RCTs heavily depend on study design: response rates of 52% and 34% for antidepressant and placebo, respectively, in two-arm studies, 58% and 45% in three-arm studies (two antidepressant arms, one placebo arm) and 65% in studies comparing two antidepressants; these differences can only be explained by the changing probability of receiving placebo: 50%, 33% and 0%, respectively[4].

The role of patients' expectations was also shown by a trial comparing sertraline, hypericum and placebo, which found no effect of assigned treatment on clinical improvement, but a significant effect of patient's guess on which treatment he/she was assigned to: patients who guessed taking sertraline or hypericum had significantly higher response rates (56% and 68%, respectively) than patients who guessed taking placebo (24%)[5].

Many patients have ambivalent attitudes towards antidepressants that significantly influence outcome: patients with a rather negative, neutral or rather positive attitude towards taking antidepressants at baseline were found to have placebo response rates of 34%, 36% and 56%, respectively, and antidepressant response rates of 51%, 56% and 69%, respectively[6].

Socio-demographic characteristics are seldom taken into account, but variables such as living with other persons (versus living alone) or being unemployed (versus employed) dramatically influence the outcome of antidepressant treatment (OR: 2.81 and 0.27, respectively)[7]. There is also an ongoing debate on whether taking into account the patient's preference for pharmacotherapy or psychotherapy influences outcome. All these aspects should be considered before we try to improve our treatments for depression, be it by personalization, stratification or precision medicine.

In addition, the "persona" of the diagnostic criteria and of the assessment tools only partially represents the "anima" of the patient and of the depressive illness. A major depressive episode cannot be fully understood either by nine DSM or ten ICD criteria, or by ten Montgomery-Åsberg Depression Rating Scale (MADRS), seventeen Hamilton Depression Rating Scale (HAMD) or thirty Inventory of Depressive Symptomatology (IDS) items.

One important limitation of the DSM criteria for major depressive episode is the massive heterogeneity they cause: almost endless combinations of criteria are possible. Indeed, when you need five out of nine criteria and, moreover, most of these criteria are compound (e.g., psychomotor retardation or agitation), two patients with major depressive episode can have no symptom in common. This of course hampers "personalized" treatment as well as clinical and etiopathogenetic research.

When assessing change during treatment, the standard rating scales face the same problems. Moreover, the HAMD covers a lot of associated anxiety and neurovegetative symptoms, while the IDS has a 16-item version closely reflecting the DSM criteria and a 30-item version adding commonly associated symptoms (anxiety, irritability) and items relevant to depression "subtypes". DSM depression symptoms (included in the IDS 16-item version) do not seem to be of higher clinical relevance than non-DSM symptoms (additionally represented in the IDS 30-item version) with respect to either their centrality (connectedness of each symptom with all other symptoms) or their relation to psychosocial functioning or life stress[8].

Furthermore, the criteria/signs/symptoms we use for diagnosis and assessment do not reflect the patient's concerns. Who is the judge? It has been documented that physicians differ significantly from patients in what they consider important for "being cured for depression". For physicians, the top five items are negative feelings, feeling down, little interest or pleasure, disrupted social life, and feeling tired, while for patients the top five items are "to what extent is life meaningful", "how much do you enjoy life", "how satisfied you are with

yourself", "how able you are to concentrate", and "negative feelings"[9]. Patients do attach more importance to restoration of positive mood and cognitive functioning than to decrease of negative mood. However, standard rating scales do not assess positive mood.

We feel that, before we can move from artisanal prescription patterns into precision medicine, patients' characteristics, beliefs and attitudes should be better taken into account, and diagnostic and assessment tools should be revised.

Koen Demyttenaere
University Psychiatric Center, KU Leuven, Leuven, Belgium

1. Perlis RH. World Psychiatry 2016;15:228-35.
2. Zimmerman M, Posternak M, Friedman M et al. Am J Psychiatry 2014;161:1285-9.
3. McGoey L. Hist Hum Sci 2010;23:58-78.
4. Sinyor M, Levitt AJ, Cheung AH et al. J Clin Psychiatry 2010;71:270-9.
5. Chen JA, Papakostas GI, Youn SJ et al. J Clin Psychiatry 2011;72:1669-76.
6. Demyttenaere K, Reines EH, Lönn SL et al. Prim Care Companion CNS Disord 2011;13(4).
7. Demyttenaere K, Verhaeghen A, Dantchev N et al. Prim Care Companion J Clin Psychiatry 2009;11:307-31.
8. Fried EI, Epskamp S, Nesse RM et al. J Affect Disord 2016;189:314-20.
9. Demyttenaere K, Donneau AF, Albert A et al. J Affect Disord 2015;174:390-6.

DOI:10.1002/wps.20362

Right patient, right treatment, right time: biosignatures and precision medicine in depression

In contrast to diagnostic changes in the rest of medicine, mental disorders are still considered as behavioral, implying that an exclusive focus on symptoms would yield a precise diagnosis[1]. Thus, even though depression is characterized by biological heterogeneity and variable symptom presentation, diagnosis and treatment recommendations are traditionally given without reference to individual variability in genes, brain structure, function and/or psychological factors. Rather, clinical and health characteristics (e.g., age, weight, medical comorbidities, depression severity) serve as the sole method for treatment selection, despite limited consistency of these characteristics to yield strong associations to treatment response. As a result, treatment selection remains a trial and error process, and only one third of patients achieve remission with the first medication prescribed, with even lower rates of sustained remission in the longer term[2,3].

Much of the previous research in depression treatment has focused on predictor variables – that is, characteristics of individuals that are associated with treatment response (or non-response), independent of treatment. More recently, increased research has focused on moderator and mediator variables. Moderators are pre-treatment variables that predict differential response to different treatments; mediators are variables whose change during the course of treatment predicts eventual treatment outcomes. Clearly, our prior focus on predictor variables has yielded inconsistent and inadequate findings, and, even with the recent attention to moderator and mediator characteristics, we have yet to determine which patient will respond to which treatment. What is needed, instead, is a comprehensive panel of variables encompassing both clinical characteristics and biological factors that can lead us to identify the right treatment for the right patient.

A comprehensive approach for targeted drug treatment and prevention is precision medicine, which takes into account the complex interplay between individual variability in clinical phenotypes, genes and brain function[4]. Treatments for cancer and chronic heart disease have developed these models and, as a result, we have reduced morbidity and mortality through the development of targeted therapies for these diseases. Yet, mental illness often lags far behind. Recent focus of the US National Institute of Mental Health on the Research Domain Criteria (RDoC) and research in genetics, proteomics and brain imaging suggest that biological measures (or biomarkers) may help us to understand the heterogeneity within the symptoms of depression and other mental illnesses[5]. Identification of biomarkers of preclinical depression or of response to drug treatment will be crucial in the development of precision medicine, being propelled by recent technological advances in large-scale biologic databases (such as the human genome and connectome projects), powerful methods for characterizing patients (such as proteomics, metabolomics, genomics, diverse cellular assays), methods for detecting patterns of brain activity and structure, and effective computational tools for analyzing extremely large datasets.

Biomarkers are measurable characteristics of an organism that correspond to a particular physiological state. Biomarkers include compounds isolated from the blood, urine or other fluids as well as clinical, behavioral and neurocognitive parameters that are used to indicate the presence or severity of a particular disease state. Moderator biomarkers specify for whom or under what conditions the treatment works, and consequently help to clarify the best choice of inclusion and exclusion criteria or the best choice of patient stratification. Mediator biomarkers identify possible mechanisms through which a treatment might achieve its effect, and changes along with response to a particular intervention. Information gained from diagnostic or progression biomarkers should aid to tailor treatments for effective personalized medicine.

The development of biomarker predictors of antidepressant response languished after multiple candidates, most notably the dexamethasone suppression

test, proved to have inadequate prognostic clinical utility[6]. The recent emergence of low-cost pharmacogenomic techniques has sparked new interests in combinatorial use of allelic variations in drug transporters or metabolic genes as biomarkers that might predict drug response[7]. An initial generation of research identified a number of candidate genes with apparent validity as predictors of treatment efficacy and treatment-related side effects. These candidates include genes implicated in serotonergic function, the ABC family of xenobiotic transporters located in the blood-brain barrier, and the cytochrome P450 detoxification enzymes. However, to date, there are no effective biological methods to objectively assess depression endophenotypes, severity, or treatment response[8]. Previous efforts to achieve better treatment outcomes in psychiatry have led to the introduction of pharmacogenomics based decision-support tools[7], to help identify which patients are more or less likely to have a favorable outcome with specific pharmacotherapies, based on single nucleotide polymorphisms (SNPs) and gene variants in transporters and metabolizing enzymes.

Genome-wide association studies have revealed that common genetic variations are unlikely to explain sufficient variance in treatment response to guide selection of treatment for individual patients. Rare gene variants have greater explanatory power than common variants, but such individual markers would likely apply to relatively few patients. Thus, if neither common nor rare gene variants are likely to have widespread predictive value as "stand alone" predictors of treatment response in typical clinical trials, a new strategy is needed, one that integrates several types of clinical and neurobiological markers to guide clinical decision making for depressive disorders.

Since it is unlikely that a single biological alteration will have a one-to-one mapping with a DSM-defined or RDoC-specified mental phenomenon, a viable alternative to the single-biomarker approach is the development of biosignatures that aim to profile a diverse array of peripheral/serum growth factors, cytokines, hormones and metabolic markers. Additionally, integration with neurological, cognitive and psychological assessments will provide coverage of multiple abnormalities that contribute to the heterogeneity of depressive disorders. Such a biosignature will not only improve our ability to identify specific subtypes of depressive disorders, but will also assist with the selection of treatments that are likely to be more clinically useful[9,10].

Based on this, some of the most promising variables to evaluate include: comprehensive clinical phenotype; magnetic resonance imaging using measures of cortical structure; diffusion tensor imaging to assess cortical white matter tract integrity; functional magnetic resonance imaging assessing brain activation patterns to both emotional conflict and reward-dependent learning tasks; quantitative electroencephalography (EEG) to assess cortical and subcortical brain activation patterns; cortical evoked EEG potentials; behavioral neuropsychological tasks to assess reaction time and motor processing speed; DNA, mRNA, and plasma, urine and saliva protein and metabolomics samples, collected at baseline and throughout the study; socio-economic, demographic and life habits parameters.

Using this comprehensive approach, however, requires a large number of participants to be characterized in order to define subgroups in relation to treatment response. It also requires the use of effective computational tools to make integration of the wealth of knowledge generated from the diverse platforms possible. Herein lays our greatest challenge: developing large cohorts of depressed patients that will lead us to the discovery of not only new, meticulously-defined subtypes of depression, but also identification of precise treatments for each individual patient. If we are successful, this will propel the treatment of depression to equal the effectiveness of treatments for cancer and chronic heart disease.

Madhukar H. Trivedi
Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA

1. Kapur S, Phillips AG, Insel TR. Mol Psychiatry 2012;17:1174-9.
2. Rush AJ, Trivedi MH, Wisniewski SR et al. Am J Psychiatry 2006;163:1905-17.
3. Trivedi MH, Rush AJ, Wisniewski SR et al. Am J Psychiatry 2006;163:28-40.
4. Roychowdhury S, Chinnaiyan AM. Annu Rev Genomics Hum Genet 2014;15:395-415.
5. Biswal BB, Mennes M, Zuo XN et al. Proc Natl Acad Sci USA 2010;107:4734-9.
6. Rush AJ, Giles DE, Schlesser MA et al. J Clin Psychiatry 1996;57:470-84.
7. Biernacka JM, Sangkuhl K, Jenkins G et al. Transl Psychiatry 2015;5:e553.
8. Smith DF. Front Psychiatry 2013;4:57.
9. Trivedi MH. Biol Psychiatry 2013;74:2-4.
10. Trivedi MH, McGrath PJ, Fava M et al. J Psychiatr Res 2016;78:11-23.

DOI:10.1002/wps.20371

Person-centered measurement-based care for depression

It is evident that the same treatment will not work for all people with depression and that a major development is required to ameliorate the outcomes of depression in routine care. A symptom dimension of interest-activity robustly predicts treatment resistance[1], a blood test for inflammation may help select an antidepressant that works better for a given individual[2], and regular rating of symptom severity improves depression outcomes[3]. Yet, none of these simple measures that could improve treatment of depression are taken up in practice. On the other hand, some clinicians are using commercial pharmacogenetic tests in the absence of evidence that such tests could predict treatment outcomes[4,5]. R. Perlis eloquently describes how human motivations drive the paradoxes of contemporary health care[6]. Perhaps even more seriously, he argues that clinicians' insistence on artisanal prescribing hinders the accrual of data that is required

to meaningfully enhance the treatment of depression.

There may be a consensus that a serotonin reuptake inhibiting antidepressant is the first treatment to try in most individuals with the diagnosis of major depressive disorder, but we know that fewer than half of patients benefit sufficiently, that many experience side effects that are not matched by benefits, and that there is little evidence on what treatments should be attempted next. Many have lamented how it is possible that we still do not have personalized treatment given the amount of work that has been done. The number of articles published on this topic may be misleading. The reason why second and third line treatment for depression is still artisanal is that there is far too little data to personalize treatment choice.

The largest study of depression treatment completed to date – the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study – has failed to personalize the second and third line treatment choices for depression because it was too small. By the time STAR*D participants progressed to the third step, the numbers of patients allocated to specific treatments were too low to allow meaningful analysis of predictors. Genetic data were available for only half STAR*D participants, further compromising the power to find biomarkers that could facilitate the choice between second and third line treatments. Genetic case-control association studies of schizophrenia, depression and other disorders have taught us that sample sizes of many thousands are needed to leverage genomic information and enable meaningful predictions. For treatment predictions, these sample sizes have to be multiplied by the number of alternative treatments that need to be tested.

With today's technology, it is possible to create, combine and exploit datasets of hundreds of thousands for common disorders like depression. The way to do it may need to work with human motivation so that the process and not just the outcomes are meaningful for patients and for clinicians. The first step will be to motivate the collection of diagnostic information and regular outcome ratings in routine clinical practice. Person-centered care with active engagement of patients in clinical decisions offers a framework for achieving such routine information collection[7].

People living with depression come with their values and preferences and want to be actively involved in discussions about their care. Patients will complete regular outcome measures if they know that these meaningfully contribute to their care. Investigators of the Canadian Depression Research and Intervention Network have piloted a person-centered measurement-based care model where patients are given the option to complete regular measures on Internet-enabled devices and request feedback that serves to enhance their participation in collaborative decision making with their clinicians. Clinicians are able to access the information and also contribute diagnosis and rating scales. Based on the information provided by clinicians and patients, a feedback is generated that selects relevant recommendations from current best practice guidelines. In this model, patients are motivated to contribute data that serve both clinical and research purposes because they see the impact of the information on their care. They in turn motivate their clinicians to participate in the information gathering and feedback process. Patients are also asked for consent to use their data for clinical research and link their data with health care databases. The platform is improving outcomes of depression in real time, allows efficient evaluation of services, and at the same time contributes to the accrual of data that will eventually help personalize treatment for depression.

In a large database, it will be possible to look up individuals who resemble a given patient on a number of factors and recommend treatments that worked for that patient. Where two or more treatments are at equipoise, they can be compared using the efficient randomized registry design embedded in routine health care[8]. The results of such large pragmatic comparisons will gradually allow exploring further steps in treatment selection or testing novel treatments.

The vision outlined above has only been partially piloted. The early experience leads us to believe that the treatment for depression has to be person-centered and measurement-based before it can be meaningfully personalized.

Rudolf Uher
Dalhousie University Department of Psychiatry, Halifax, NS, Canada

The author is supported by the Canada Research Chairs Program.

1. Uher R, Perlis RH, Henigsberg N et al. Psychol Med 2012;42:967-80.
2. Uher R, Tansey KE, Dew T et al. Am J Psychiatry 2014;171:1278-86.
3. Guo T, Xiang Y-T, Xiao L et al. Am J Psychiatry 2015;172:1004-13.
4. Peters EJ, Slager SL, Kraft JB et al. PLoS One 2008;3:e1872.
5. GENDEP investigators, MARS investigators, STAR*D investigators. Am J Psychiatry 2013; 170:207-17.
6. Perlis RH. World Psychiatry 2016;15:228-35.
7. Dixon LB, Holoshitz Y, Nossel I. World Psychiatry 2016;15:13-20.
8. Lauer MS, D'Agostino RB Sr. N Engl J Med 2013;369:1579-81.

DOI:10.1002/wps.20363

Carving depression at its joints?

Personalization of treatments has long been an aspiration for medicine and has recently evolved into a sophisticated practice for the treatment of some diseases. Although in psychiatry treatment decisions are usually based on the individual patient and his/her needs, there is a lack of information about how the benefits and harms of individual pharmacological agents (and indeed treatments in other modalities) differ from patient to patient and very limited data on which to base the choice between treatment options for individual patients. The thoughtful paper by R. Perlis[1] addresses the challenges in

personalization of antidepressant treatment and highlights various important scientific questions thereof.

Perlis suggests that available phenomenological patient-level features may be of more help than generally acknowledged for establishing probability of response. Whilst many might have sympathy with this view, history is replete with debates about the therapeutic utility of various subdivisions of depression, perhaps most notably the prolonged dispute between the Newcastle categorical[2] and the Maudsley dimensional[3] approach. Such arguments remain inconclusive.

However, other recent work has focused on the cases of depression which have undiagnosed bipolar disorder and highlighted this as an area potentially important for personalizing treatment. The seminal paper by Angst et al[4] showed that broad diagnostic criteria (in comparison with DSM-IV-TR criteria) identified a large number of additional patients with major depressive episodes who were likely manifesting depression as part of a bipolar disorder. These authors suggest that additionally considering family history, illness course and clinical status, as well as diagnostic criteria, may provide useful information for physicians when assessing evidence of bipolarity in patients with major depressive episodes. Many such patients (with major depressive episodes as part of a bipolar disorder) will be treated with but not respond to antidepressants. This has led to the notion being promulgated that all antidepressants should be, as a regulatory requirement, tested in bipolar major depressive episodes as well as in unipolar depression[5].

Perlis also reviews biological approaches, but the question remains: are the currently available putative biomarkers of antidepressant response really more robust and consistent tools compared to "artisanal" practice? For example, the correlation between plasma drug levels and clinical response is weak and not only are drug plasma levels poorly associated with doses of drugs, but there is also a significant dissociation between brain and plasma kinetics, as demonstrated by positron emission tomography (PET) receptor occupancy studies[6]. Many factors, other than plasma levels, moderate drug action in the central nervous system. These factors will affect the predictive ability of pharmacogenomic biomarkers that are directly linked to pharmacokinetic variables, for example those which are genetic determinants of drug metabolism, and limit their potential contribution in increasing precision of pharmacotherapies.

The development of high precision pharmacotherapies is typically driven by the combination of three factors: a) treatments are potentially highly efficacious if the right treatment is given to the right person; b) treatments are very expensive; c) treatments may be associated with serious adverse effects. The need for careful pre-selection of a specific treatment for the right patients becomes highest in those diseases in which it is most important to direct expensive investments to the patients identified with highest-benefit and lowest-risk potential. For mood disorders as a whole, there is arguably less of a compelling need for this kind of "precision" treatments: pharmacotherapies for depression are relatively affordable, compared to those for autoimmune or neoplastic diseases, and very serious adverse effects are rare. Thus, clinicians may end up trusting more their own "artisanal" judgment based on experience than not very informative evidence-based medicine inspired treatment protocols and guidelines.

The integration of multimodal biomarker approaches may potentially increase precision, but at the moment their cost and complexity is high and the utility of this approach unproven. New biomarker approaches (transcriptomic, proteomic, genomic and telomeric) may potentially change this[7]. However, it will be important to establish how much higher remission rates can be achieved with such multimodal biomarkers informing personalized treatment before advocating this approach. Even if this could not be translated into clinical practice because of cost and complexity, proof-of-concept studies would answer crucial clinical research questions that have remained unresolved despite the overall progress in neuroscience.

Where does all of this leave us? It may be worthwhile pausing to reflect on how progress was achieved recently in other fields of medicine. Although we often feel that our problems are unique to psychiatry, the confounding effects of heterogeneity are not confined to mood disorders and have been addressed in other fields of study by focussing on the most reliable diagnoses which are most tractable for research[8]. This has produced major advances in the understanding of Alzheimer's disease, which have underpinned ongoing therapeutic research. This approach has also been shown to be practical in familial studies of lithium response[9]. Extending this approach further to mood disorders might mean focusing, for instance, on bipolar I disorder with a strong familial component. We could apply this "narrow" approach to therapeutic research in this area and combine it with the multimodal biomarker notions outlined above. Any relationship between biomarkers and therapeutic responses could then be further verified in larger clinical populations.

What of "broad" approaches, i.e. studying a multiplicity of factors in large groups of heterogeneous patients? This will undoubtedly continue and may be made potentially more fruitful by recent developments. The recent revision of DSM introduced new ways of splicing major depression, including the delineation of various facets of the clinical symptomatology. An interesting example of the potential advantages of this development is a recent study on major depression with mixed features[10].

Future progress will likely come from the application of both "narrow" and "broad" approaches, focused on valid and well-characterized patient samples, trying, to quote Socrates, to "divide things again by classes, where the natural joints are".

Allan H. Young, Alessandro Colasanti
Centre for Affective Disorders, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neurosciences, King's College London, London, UK

1. Perlis RH. World Psychiatry 2016;15:228-35.

2. Roth M. Acta Psychiatr Scand 1981;63(Suppl. 290):42-51.
3. Kendell RE. Scott Med J 1978;23:61-3.
4. Angst J, Azorin JM, Bowden CL et al. Arch Gen Psychiatry 2011;68:791-8.
5. Young AH. Aust N Z J Psychiatry 2012;46:293-4.
6. Suhara T. Arch Gen Psychiatry 2003;60:386-91.
7. Gururajan A, Clarke G, Dinan TG et al. Neurosci Biobehav Rev 2016;64:101-33.
8. Guerreiro R, Hardy J. Neurotherapeutics 2014; 11:732-7.
9. Grof P, Duffy A, Alda M et al. Acta Psychiatr Scand 2009;120:378-85.
10. Suppes T, Silva R, Cucchiaro J et al. Am J Psychiatry 2016;173:400-7.

DOI:10.1002/wps.20367

Pragmatic treatment options for depression and anxiety disorders are needed

Major depressive disorder (MDD) is a common disorder with a lifetime risk around 35%[1]. It is a significant cause of mortality and the second leading cause of years lived with disability worldwide[2]. At many levels, the similarities between MDD and the anxiety disorders are much stronger than the differences. For instance, these disorders share genetic, temperamental and environmental risk factors, frequently co-occur, and cognitive behavior therapy (CBT) and antidepressant drugs are recommended as principal treatments for both[3]. It will therefore be very difficult to locate pathologies and treatments that are specific to subtypes of depression without considering their relationship with anxiety. As a result, we broaden our comment on Perlis' proposal[4] to reduce personalized (or non-evidence based) medicine in favor of precise (evidence-based) medicine to include both depression and anxiety.

Despite the availability of evidence-based treatments and the energetic "Reduce Mental Illness" campaigns that have been carried out in many countries, the years lived with disability due to these disorders have not declined over the past two decades[2]. This burden persists, at least in part, because only 39% of adults who met criteria for a depressive and/or anxiety disorder in the past year sought help for their mental health problems[5]. Additionally, the burden associated with depression and anxiety continues, as Perlis suggests, because of the type and dose of treatment that patients receive from their mental health professionals.

Evidence-based guidelines for the treatment of depression and anxiety recommend a stepped approach to care where patients are prescribed treatments in order of their intensity, effect and cost[6]. Specifically, CBT is the recommended first-line psychological treatment for mild to moderate depression and the anxiety disorders, and in combination with pharmacotherapies for severe and complex cases[7,8]. Although these guidelines were developed to translate advances in medical research into clinical practice, their dissemination does not necessarily improve clinical outcomes. Indeed, of those people who met criteria for depression and/or an anxiety disorder in the past year and sought help for their mental health problems, 67% were offered an evidence-based treatment and only 41% received a minimally adequate dose of treatment[5].

Even when people receive a minimally adequate dose of treatment, a significant proportion continue to experience distress and impairment. This is because the recommended psychological and pharmacological interventions range in number needed to treat (NNT) from ~2 to 16, depending on the comparator[9]. Perlis argues that the NNT of antidepressants could be improved if: a) specific drugs were linked to specific disorder subtypes; b) diagnostic tools sensitive to these mechanisms of action were disseminated; and c) clinicians implemented these tools in practice. To this end, he states that "numerous investigations of predictors... of treatment response have been reported over the past five decades"[4]. Yet, to our knowledge, none of these predictors have been associated with Level 1 evidence that a certain treatment for MDD is superior for people who have certain characteristics. Level 1 evidence would require both independently replicated superiority (RCTs) with representative groups of patients in which the treatment is shown to remedy the core pathology in the target group *and* that there are no methodologically sound rebuttals.

We thank Perlis for mentioning a five-decade time frame. Five decades ago, we, in a study to identify core pathologies, divided a cohort of inpatients with a primary diagnosis of depression into those with "endogenous" or "neurotic" depression on the basis of clinical presentation, family history and longitudinal data. At intake, discrimination between endogenous and neurotic depression was strong, but the data were equally compatible with Grinker et al's five-division classification[10], which weakened the findings. At the 15-year follow-up, the overall outcomes of people identified with endogenous or neurotic depression were equally poor, with two relatively minor differences: endogenous depression showed more frequent but shorter hospital admissions, and the outcome in neurotic depression was more strongly related to the level of neuroticism at the index admission[11].

Identifying disorder subtypes for targeted treatment, as Perlis says, is likely to be time consuming and difficult. Indeed, seven types of first line drug therapies for MDD – selective serotonin reuptake inhibitors (SSRIs), noradrenergic and specific serotonergic antidepressants (NaSSAs), norepinephrine-dopamine reuptake inhibitors (NDRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), noradrenaline reuptake inhibitors (NARIs), melatonin agonists, serotonin modulators – that notionally target different pathologies are

listed in a recent clinical practice guideline[7]. If the relative power of each is weak, then superiority trials will need to be very large, costly and time consuming once it has been decided that the patients with the core pathologies can be reliably identified. The feasibility of such trials is questionable.

The possibility of identifying mechanisms of change that are associated with subtypes of depression and anxiety is attractive. In reality, it is likely that this will not occur for many years. We are mindful that the burden of these disorders is large and immediate. We therefore turn to practical ways of averting this burden that are available now. Studies that show that primary care patients prefer psychotherapy over antidepressant treatment[12], that psychotherapy alone has comparable long-term outcomes to combined treatment[13], and that the number needed to harm (NNH) associated with antidepressants can be considerable[14], support the extant recommendations for psychological interventions as first-line treatments.

Yet CBT suffers from four deficits in comparison to antidepressant medication. It is more difficult to prescribe, it is more expensive, quality in practice cannot be guaranteed, and it is not widely available outside major city centers. Automated Internet-delivered CBT (iCBT) is equally effective as face-to-face CBT[15], yet inherently more scalable and therefore offers a more efficient use of scarce public health resources. It is as easy to prescribe iCBT as it is to prescribe medication, and the fidelity of treatment is guaranteed across service providers.

One big advantage of iCBT compared to face-face CBT and antidepressant medication is that RCTs are relatively simple and quick to do. If patients can be screened over the Internet and do not have to be seen in person, then a large trial can be finished within six months of institutional review board approval. A rapid cycling research model would allow us to search for specific-treatments-for-specific-group pairings by running several trials at the same time. We would no longer have to wait 15 years for a null result.

In conclusion, Perlis emphasizes "the need for changes in how depression care is delivered, measured, and used to inform future practice". We agree: the uptake and quality of mental health care that is delivered in the community needs immediate improvement. While disorder-specific targeted treatments are yet to be developed, iCBT is an evidence-based treatment that can be used to reduce the burden associated with depression and anxiety disorders now.

Gavin Andrews, Megan J. Hobbs
Clinical Research Unit for Anxiety and Depression, School of Psychiatry, University of New South Wales, Sydney, Australia

Both authors contributed to the substantive content of this piece.

1. Kruijshaar ME, Barendregt J, Vos T et al. Eur J Epidemiol 2005;20:103-11.
2. Vos T, Flaxman AD, Naghavi M et al. Lancet 2013;380:2163-96.
3. Goldberg DP, Andrews G, Krueger RF et al. Psychol Med 2009;39:2043-59.
4. Perlis RH. World Psychiatry 2016;15:228-35.
5. Harris MG, Hobbs MJ, Burgess PM et al. Med J Australia 2015;202:185-9.
6. Hobbs MJ. In: Wright JD (ed). International encyclopedia of social and behavioral science, 2nd ed. Oxford: Elsevier, 2015:339-43.
7. Mahli G, Bassett D, Boyce P et al. Aust N Z J Psychiatry 2015;49:1087-206.
8. Andrews G, Dean K, Genderson M et al. Management of mental disorders, 5th ed. Sydney: Amazon.com, 2014.
9. Cuijpers P, Sijbrandij M, Koole S et al. Clin Psychol Rev 2014;34:130-40.
10. Grinker RR, Miller J, Sabshin M et al. The phenomena of depressions. New York: Harper & Row, 1961.
11. Andrews G, Neilson M, Hunt C et al. Br J Psychiatry 1990;157:13-8.
12. van Schaik DJF, Klijn AFJ, van Hout HPJ et al. Gen Hosp Psychiatry 2004;26:184-9.
13. Karyotaki E, Smit Y, Holdt Henningsen K et al. J Affect Dis 2016;194:144-52.
14. Montejo AL, Llorca G, Izquierdo JA et al. J Clin Psychiatry 2001;62(Suppl. 3):10-21.
15. Andrews G, Cuijpers P, Craske MG et al. PLoS One 2010;5:e13196.

DOI:10.1002/wps.20364

Toward precision medicine for depression: admitting ignorance and focusing on failures

I first recognized what R. Perlis[1] calls "artisanal practice" as a medical student in Liberia. Witch doctors – basically *the* local primary and psychiatric care providers at the time – regularly engaged in "throwing the bones". Chicken bones, often in a bag but sometimes in hand, were shaken and thrown on the ground, resulting in a unique pattern which served as the basis for recommendation(s) offered to each "patient". "Throwing the bones" was common and well accepted. So much so that it was rare to have dehydrated neonates arrive at the hospital without dung spread over their depressed fontanelles, courtesy of the "doctor". Personalized care – maybe; precision medicine – not so much.

But let me also defend some important aspects of personalized care. Organizing treatment steps based on reasonable Hippocratic principles and logic is appropriate[2]. Avoiding medications that lower seizure thresholds in patients at risk, especially when safer options exist, is smart. My guess is that maybe 7-15% of depressed patients should avoid certain medications for medical reasons. In sum, let's recognize that we too are throwing the bones when selecting among antidepressant medications for specific patients, while we avoid certain treatments in selected patients for safety reasons.

To advance the precision of treatment selection, I agree with Perlis that we must provide systematic patient education, become adroit in the implementation of measurement-based care (to enhance efficacy and tolerability, and to reduce variations that interfere with signal detection) and shed our information about how to choose among treatments. Let's leave the bones in the bag! We desperately need more knowledge about

how to accomplish our clinical tasks, including treatment selection. However, we seem to get entrenched in our beliefs and routines, and our own administrative, reimbursement and legal cultures. I'll bet that the Liberian "doctors" are still able change practices more easily than we can!

Perlis highlights the issue of slow adoption with his experience in pharmacogenetic testing research. Clinicians are moved almost entirely by what impacts their patients' outcomes, despite evidence of cost-effectiveness. There is still a paucity of research capable of changing the minds of clinicians and patients. Uptake and changes in practice would speed up if we had more research that focused on questions pivotal to clinician-patient decisions that result in clear evidence of benefit to substantial numbers of patients[3]. Issues in implementation would be clarified and uptake facilitated by addressing specific questions, such as: when in the course of treatment steps and with which medications is pharmacogenetic testing useful? Or, can we identify which patients have treatment-resistant depression at the outset?[4]

Let's assume that we have engineered consistent high-quality, measurement-based care, and have electronic health records and a cadre of educated and collaborative patients. Having somehow set this table to aggressively pursue precision medicine, the question becomes: do any of our prior successes in matching treatments and patients suggest a preferred path forward?

One major focus might be on identifying with a high degree of certainty which patients are very likely to not respond or succeed (i.e., to go after treatment failures). Depression is not unchecked cancer, with its generally predictable downhill and often terminal course. Success is an exception in cancer without treatment. Therefore, in cancer treatment research, a focus on success makes sense. Even after a successful cancer treatment begins to fail, we can learn from these failures. Depression, on the other hand, is a heterogeneous syndrome that has a highly variable course which is affected by changes in support, stresses, comorbidities and substances to name but a few. Adding to these challenges is the fact that only a small proportion of the "successes" will be specifically responding to the medication.

By focusing on depressed persons whose treatments have failed, we can learn which features of our patients or their treatments are contributing to the failures. An example of this in another area of study would be the pool of anemic patients who have been non-responsive to iron. This group would be enriched in patients with B12 deficiency. This B12 deficient subset might be easier to detect, especially in large patient samples and with the use of machine learning.

As a further illustration of the potential value of a focus on failures, consider how atypical depression grew out of a recognition that some depressed patients, often with atypical features, fared poorly with tricyclic antidepressants but succeeded with monoamine oxidase inhibitors[5].

Perlis' own work to define risk factors for treatment-resistant depression[4] also illustrates how a failure focus can be productive. His results indicated that there is a meaningful proportion of treatment-resistant patients (maybe 25%) who can be specifically identified by the measures used. Uher et al[6] also hit pay dirt with a failure focus, finding that anhedonic depressed patients do poorly with serotonin/noradrenaline reuptake blockers. Hedonically-impaired patients with treatment-resistant depression may have a dysfunctional mesolimbic dopamine system. Fawcett et al[7] recently found higher doses of adjunctive pramipexole to be associated with substantial and largely sustained benefits to treatment-resistant depression patients with severely impaired interest/activity.

Finally, to advance precision medicine, do we really need to wait to change psychiatric practices broadly? Culture changes are led by the few; almost never by the many. Multi-site registries that engage only those providers who are willing to make the changes above could generate large numbers of subjects for computations that involve large numbers of variables. I suspect that even randomization after the first step (though not essential) would be feasible in such registries and might well speed up discovery, given providers that possess the requisite curiosity and humility.

In conclusion, I largely concur with the challenges raised by R. Perlis in moving into the precision medicine space. These problems are all solvable as they are all man-made. Certainly better patient education, widespread use of measurement-based care and a willingness to throw away those bones are essential next steps for a coalition of the willing. A focus on failures may be a fertile field to till.

A. John Rush
Duke-National University of Singapore, Singapore

1. Perlis RH. World Psychiatry 2016;15:228-35.
2. Preskorn SH. Outpatient management of depression: a guide for the primary-care practitioner, 2nd ed. Caddo: Professional Communications, 1999.
3. Rush AJ. J Clin Psychiatry 2015;76:1366-72.
4. Perlis RH. Biol Psychiatry 2013;74:7-14.
5. Stern SL, Rush AJ, Mendels J. Am J Psychiatry 1980;137:545-52.
6. Uher R, Perlis RH, Henigsberg N et al. Psychol Med 2012;42:967-80.
7. Fawcett J, Rush AJ, Vukelich J et al. Am J Psychiatry 2016;173:107-11.

DOI:10.1002/wps.20365

Can we at least learn to fail faster?

For clinicians and (more important) patients, the current trial-and-error process of finding an effective depression treatment is frustrating and discouraging. Our ability to accurately match individual patients with specific medications is embarrassingly poor[1]. And, given the delayed symptomatic response to most depression treatments, the cycle time for each trial-and-error is as long as two months. It is therefore not surprising that many patients starting depres-

sion treatment become discouraged and never return.

As R. Perlis[2] clearly describes, more accurate prediction or personalized treatment selection is not yet in sight. It may not even be just over the horizon. Much of the research that claims to support personalization of treatment is really more relevant to general prediction of depression outcome or general prediction of treatment response than to selection of specific treatments for individuals[1]. I refer to this mis-application of evidence as "trying to answer a four-group question with a two-group research design".

Stated statistically, personalized or precision treatment selection depends on interaction effects rather than main effects. If we hope to detect interactions rather than just main effects, research to support precision medicine for depression will certainly require much larger samples than we are accustomed to. More important, selection of and testing for promising interactions or moderators will likely require a clearer understanding of treatment mechanisms and more precise measures of outcome.

While accurate prediction of treatment success may be off in the distance, we are probably closer to faster detection of depression treatment failure. And "failing faster" would be a significant improvement on the current state. Even though depression treatment guidelines often advise waiting six weeks or more to assess the effectiveness of antidepressant medication, evidence from placebo-controlled trials consistently demonstrates separation between active medication and placebo as early as seven days[3]. Even more promising, direct assessment of the neuropsychological "building blocks" of depression may allow even more rapid discrimination of treatment success or failure – identifying treatments unlikely to work earlier than traditional clinical measures.

For example, C. Harmer and colleagues at Oxford have shown that biased processing of emotional information (measured by a computerized task resembling a video game) can change within hours of a first dose of antidepressant medication[4]. We may soon welcome the day when we tell patients: "Download this app, take this pill tonight, and send me your test results in the morning. We can decide tomorrow if this medication is worth continuing". That scenario would be a dramatic improvement over our current advice to "take this medication for a month, and we can decide then if it was worth the wait".

The National Institute of Mental Health's Research Domain Criteria (RDoC) scheme[5] helps to reveal the connection between these two goals (precision prediction of treatment success and rapid detection of treatment failure). Under the RDoC scheme, we hope to resolve the heterogeneous category of depression into more crisply defined components or building blocks. Any individual case of depression would represent some admixture of more fundamental elements such as decreased sensitivity to reward, impaired executive function, and overvaluation of negative emotional stimuli.

Following this scheme, performance-based assessment of those RDoC components could facilitate advances in both directions: faster detection of treatment failure and more accurate prediction of treatment success. Stated statistically, discovery of mediators (processes that explain or account for the success of any specific treatment) will inform the discovery of moderators (pre-treatment characteristics identifying individuals for whom that treatment will be successful). Ultimately, this "experimental medicine" approach would also facilitate the development of more specific (and more effective) new treatments.

I expect that advances in precision medicine for depression will likely come sooner from neuropsychology than from genomics.

Gregory E. Simon
Group Health Research Institute, Seattle, WA, USA

1. Simon GE, Perlis RH. Am J Psychiatry 2010; 167:1445-55.
2. Perlis R. World Psychiatry 2016;15:228-35.
3. Taylor MJ, Freemantle N, Geddes JR et al. Arch Gen Psychiatry 2006;63:1217-23.
4. Harmer CJ, Goodwin GM, Cowen PJ. Br J Psychiatry 2009;195:102-8.
5. Insel T, Cuthbert B, Garvey M et al. Am J Psychiatry 2010;167:748-51.

DOI:10.1002/wps.20366

How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence

Pim Cuijpers[1,2], Ioana A. Cristea[3], Eirini Karyotaki[1,2], Mirjam Reijnders[1,2], Marcus J.H. Huibers[1,2,4]

[1]Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [2]EMGO Institute for Health and Care Research, Amsterdam, The Netherlands; [3]Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania; [4]Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

We report the current best estimate of the effects of cognitive behavior therapy (CBT) in the treatment of major depression (MDD), generalized anxiety disorder (GAD), panic disorder (PAD) and social anxiety disorder (SAD), taking into account publication bias, the quality of trials, and the influence of waiting list control groups on the outcomes. In our meta-analyses, we included randomized trials comparing CBT with a control condition (waiting list, care-as-usual or pill placebo) in the acute treatment of MDD, GAD, PAD or SAD, diagnosed on the basis of a structured interview. We found that the overall effects in the 144 included trials (184 comparisons) for all four disorders were large, ranging from g=0.75 for MDD to g=0.80 for GAD, g=0.81 for PAD, and g=0.88 for SAD. Publication bias mostly affected the outcomes of CBT in GAD (adjusted g=0.59) and MDD (adjusted g=0.65), but not those in PAD and SAD. Only 17.4% of the included trials were considered to be high-quality, and this mostly affected the outcomes for PAD (g=0.61) and SAD (g=0.76). More than 80% of trials in anxiety disorders used waiting list control groups, and the few studies using other control groups pointed at much smaller effect sizes for CBT. We conclude that CBT is probably effective in the treatment of MDD, GAD, PAD and SAD; that the effects are large when the control condition is waiting list, but small to moderate when it is care-as-usual or pill placebo; and that, because of the small number of high-quality trials, these effects are still uncertain and should be considered with caution.

Key words: Cognitive behavior therapy, major depression, generalized anxiety disorder, panic disorder, social anxiety disorder, meta-analysis, publication bias, quality of trials, waiting list control groups

(World Psychiatry 2016;15:245–258)

Every year almost 20% of the general population suffers from a common mental disorder, such as depression or an anxiety disorder[1]. These conditions not only result in personal suffering for patients and their families, but also in huge economic costs, in terms of both work productivity loss and health and social care expenditures[2-6].

Several evidence-based treatments are available for common mental disorders, including pharmacological and psychological interventions. Many patients receive pharmacological treatments, and these numbers are increasing in high-income countries[7]. Psychological treatments are equally effective in the treatment of depression[8] and anxiety disorders[9-11]. However, they are less available or accessible[12], especially in low- and middle-income countries. At the same time, about 75% of patients prefer psychotherapy over the use of medication[13].

The most extensively tested form of psychotherapy is cognitive behavior therapy (CBT). Dozens of trials and several meta-analyses have shown that CBT is effective in treating depression[8,14] and anxiety disorders[9-11]. However, in recent years, it has become clear that the effects of CBT and other psychotherapies have been considerably overestimated due to at least three reasons.

The first reason is publication bias[15,16]. This refers to the tendency of authors to submit, or journals to accept, manuscripts for publication based on the direction or strength of the study's findings[17]. There is considerable indirect evidence of publication bias in psychotherapy research, based on excess publication of small studies with large effect sizes[16]. Moreover, there is also direct evidence of publication bias: a recent study found that almost one quarter of trials of psychotherapy for adult depression funded by the US National Institutes of Health were not published[15]. After adding the effect sizes of these unpublished trials to those of the published ones, the mean effect size for psychotherapy dropped by more than 25%.

The second reason why the effects of psychotherapies have been overestimated is that the quality of many trials is suboptimal. In a meta-analysis of 115 trials of psychotherapy for depression, only 11 met all basic indicators of quality, and the effect sizes of these trials were considerably smaller than those of lower quality ones[18]. However, that meta-analysis only included trials up to 2008, and since then many new studies have been conducted. Because more recent trials are typically of a better quality than older ones, it is not known what the current best estimate of the effect size of CBT is after taking these newer studies into account.

A third reason why the effects of psychotherapy have been overestimated is that many trials have used waiting list control groups. Although all control conditions in psychotherapy trials have their own problems[19,20], the improvement found in patients on waiting lists has been found to be lower than that expected on the basis of spontaneous remission[19]. It has been suggested, therefore, that waiting list is in fact a "nocebo" (the opposite of a placebo; an inert treatment that appears to cause an adverse effect) and that trials using it considerably overestimate the effects of psychological treatments[21]. Other control conditions, such as care-as-usual and pill placebo, can allow a better estimate of the true effect size of CBT.

In the present paper, we report the most up-to-date and accurate estimate of the effects of CBT in the treatment of

major depression (MDD), generalized anxiety disorder (GAD), panic disorder (PAD) and social anxiety disorder (SAD), taking into account the three above-mentioned major problems of the existing psychotherapy research: publication bias, low quality of trials, and the nocebo effect of waiting list control groups.

METHODS

Identification and selection of studies

We searched four major bibliographic databases (PubMed, PsycINFO, Embase and the Cochrane database of randomized trials) by combining terms (both MeSH terms and text words) indicative of psychological treatment and either SAD (social phobia, social anxiety, public-speaking anxiety), GAD (worry, generalized anxiety), or PAD with or without agoraphobia (panic, panic disorder), with filters for randomized controlled trials. We also checked the references of earlier meta-analyses on psychological treatments for the included disorders. The deadline for the searches was August 14, 2015.

For the identification of trials of CBT for MDD, we used an existing database[22] updated to January 2016 by combining terms indicative of psychological treatment and depression (both MeSH terms and text words).

We included randomized trials in which CBT was directly compared with a control condition (waiting list, care-as-usual or pill placebo) in adults with MDD, GAD, PAD or SAD. Only trials in which recruited subjects met diagnostic criteria for the disorder according to a structured diagnostic interview – such as the Structured Clinical Interview for DSM (SCID), the Composite International Diagnostic Interview (CIDI) or the Mini International Neuropsychiatric Interview (MINI) – were included.

In addition to any therapy in which cognitive restructuring was one of the core components, we also included purely behavioral therapies, i.e., trials of behavioral activation for depression and exposure for anxiety disorders. We included therapies that used individual, group and guided self-help formats, but excluded self-guided therapies without any professional support, because their effects have been found to be considerably smaller than other formats[23]. Studies on therapies delivering only (applied) relaxation were excluded, as were studies on eye movement desensitization and reprocessing (EMDR), interpersonal or psychodynamic therapy, virtual reality therapy, transdiagnostic therapies, as well as studies in which CBT was combined with pill placebo.

In order to keep heterogeneity as low as possible, we included only studies using waiting list, care-as-usual or pill placebo control groups. Care-as-usual was defined broadly as anything patients would normally receive, as long as it was not a structured type of psychotherapy. Psychological placebo conditions were not included, because they have considerable effects on depression[24] and probably also on anxiety disorders[19]. Comorbid mental or somatic disorders were not used as an exclusion criterion. Studies on inpatients and on adolescents or children (below 18 years of age) were excluded, as were studies recruiting patients with other types of depressive disorders than MDD (dysthymia or minor depression). We also excluded maintenance studies, aimed at people who already had a partial or complete remission after an earlier treatment, and studies that did not report sufficient data to calculate standardized effect sizes. Studies in English, German and Dutch were considered for inclusion.

Quality assessment and data extraction

We assessed the quality of included studies using four criteria of the "risk of bias" assessment tool developed by the Cochrane Collaboration[25]. Although "risk of bias" and quality are not synonyms[25], the former can be seen as an indicator of the quality of studies. The four criteria were: adequate generation of allocation sequence; concealment of allocation to conditions; blinding of assessors; and dealing with incomplete outcome data (this was assessed as positive when intention-to-treat analyses were conducted, meaning that all randomized patients were included in the analyses). The assessment of the quality of included studies was conducted by two independent researchers, and disagreements were solved through discussion.

We also coded participant characteristics (disorder, recruitment method, target group); characteristics of the psychotherapies (treatment format, number of sessions); and general characteristics of the studies (country where the study was conducted, year of publication).

Meta-analyses

For each comparison between a psychotherapy and a control condition, the effect size indicating the difference between the two groups at post-test was calculated (Hedges' g). Effect sizes of 0.8 can be assumed to be large, while effect sizes of 0.5 are moderate, and effect sizes of 0.2 are small[26]. Effect sizes were determined by subtracting (at post-test) the average score of the psychotherapy group from the average score of the control group, and dividing the result by the pooled standard deviation. Because some studies had relatively small sample sizes, we corrected the effect size for small sample bias[27]. If means and standard deviations were not reported, we calculated the effect size using dichotomous outcomes, and if these were not available either, we used other statistics (such a t or p value).

In order to calculate effect sizes, we used all measures examining depressive symptoms, such as the Beck Depression Inventory (BDI[28] or BDI-II[29]) and the Hamilton Rating Scale for Depression[30], or anxiety symptoms, such as the Beck Anxiety Inventory[31], the Penn State Worry Questionnaire[32], the Fear Questionnaire[33], and the Liebowitz Social Anxiety Scale[34]. We did not use measures of mediators, dysfunctional thinking, quality of life or generic severity. To calculate pooled mean effect sizes, we used the Comprehensive Meta-Analysis (CMA)

	MDD	GAD	PAD	SAD	Total
Pubmed	4,562	757	947	457	6,723
Cochrane	5,072	1,831	1,597	1,083	9,583
PsycINFO	2,530	558	484	329	3,901
Embase	4,243	596	914	815	6,568
Total	16,407	3,742	3,942	2,684	26,775
⇓					
After removal of duplicates	13,384	2,267	2,310	1,619	19,580
⇓					
Full-text retrieved	1,885	196	546	330	2,957
⇓					
Excluded:					
– Secondary papers	390	36	63	53	542
– No diagnosis	138	44	147	58	387
– No control group	392	59	217	143	811
– No CBT	211	22	55	23	311
– Other reason	414	15	34	17	480
Total	1,831	172	516	294	2,813
⇓					
Included in meta-analysis	54	24	30	36	144

Figure 1 Flow chart of inclusion of trials. MDD – major depression, GAD – generalized anxiety disorder, PAD – panic disorder, SAD – social anxiety disorder, CBT – cognitive behavior therapy

software (version 3.3.070). Because we expected considerable heterogeneity among the studies, we employed a random effects pooling model in all analyses.

Numbers-needed-to-treat (NNT) were calculated using the formulae provided by Furukawa[35], in which the control group's event rate was set at a conservative 19% (based on the pooled response rate of 50% reduction of symptoms across trials of psychotherapy for depression)[36]. As a test of homogeneity of effect sizes, we calculated the I^2 statistic (a value of 0 indicates no observed heterogeneity, and larger values indicate increasing heterogeneity, with 25 as low, 50 as moderate, and 75 as high heterogeneity)[37]. We calculated 95% confidence intervals around I^2 using the non-central chi-squared-based approach within the Heterogi module for Stata[38,39].

We conducted subgroup analyses according to the mixed effects model, in which studies within subgroups are pooled with the random effects model, while tests for significant differences between subgroups are conducted with the fixed effects model. For continuous variables, we used meta-regression analyses to test whether there was a significant relationship between the continuous variable and the effect size, as indicated by a Z value and an associated p value. Multivariate meta-regression analyses, with the effect size as the dependent variable, were conducted using CMA.

We tested for publication bias by inspecting the funnel plot on primary outcome measures and by Duval and Tweedie's trim and fill procedure[40], which yields an estimate of the effect size after the publication bias has been taken into account. We also conducted Egger's test of the intercept to quantify the bias captured by the funnel plot and to test whether it was significant.

RESULTS

Selection and inclusion of trials

After examining a total of 26,775 abstracts (19,580 after removal of duplicates), we retrieved 2,957 full-text papers for further consideration. We excluded 2,813 of the retrieved papers. The PRISMA flow chart describing the inclusion process and the reasons for exclusion is presented in Figure 1. A total of 144 trials met inclusion criteria for this meta-analysis: 54 on MDD, 24 on GAD, 30 on PAD, and 36 on SAD.

Characteristics of included trials

The 144 trials included a total of 184 comparisons between CBT and a control condition (63 comparisons for MDD, 31 for GAD, 42 for PAD, and 48 for SAD). A total of 11,030 patients were enrolled (6,229 in the CBT groups, 2,469 in the waiting list control groups, 1,823 in the care-as-usual groups and 509 in the pill placebo groups). A total of 113 trials were aimed at adults in general and 31 at other more specific target groups. Eighty trials recruited patients (also) from the community, 51 recruited exclusively from clinical populations, and 13 used other recruitment methods. Sixty-seven trials were conducted in North America, 14 in the UK, 36 in other European countries, 15 in Australia, 4 in East Asia, and 8 in other geographic areas. Of all included trials, 44 (30.6%) were conducted in 2010 or later.

CBT was delivered in individual format in 87 comparisons, in group format in 53, in guided self-help format in 35, and in a mixed or another format in 9. The number of treatment sessions ranged from one to 25.

Quality assessment

Sixty trials reported an adequate sequence generation, while the other 84 did not. A total of 46 trials reported allocation to conditions by an independent (third) party. Seventy trials reported blinding of outcome assessors and 57 conducted intention-to-treat analyses. Only 25 trials (17.4%) met all four quality criteria, 62 met two or three criteria, and the remaining 57 met one or none of the criteria. Of the trials conducted in 2010 or later, 29.5% were rated as high-quality, compared to 12.0% of the older studies.

Effects of CBT on MDD

The pooled effect size of the 63 comparisons between CBT and control conditions in MDD[41-94] was g=0.75 (95% CI: 0.64-0.87), with high heterogeneity (I^2=71). This effect size corresponds to a NNT of 3.86. Studies using a waiting list control group had significantly (p=0.002) larger effect sizes (g=0.98; 95% CI: 0.80-1.17) than those using care-as-usual (g=0.60; 95% CI: 0.45-0.75) and pill placebo control groups (g=0.55; 95% CI: 0.28-0.81) (Table 1 and Figure 2).

Table 1 Effects of cognitive behavior therapy for major depression (MDD), generalized anxiety disorder (GAD), panic disorder (PAD) and social anxiety disorder (SAD) compared to control conditions

		N	g	95% CI	p	I^2	95% CI	p	NNT
MDD									
All control conditions	All studies	63	0.75	0.64-0.87	<0.001	71	62-77		3.86
	High-quality studies	11	0.73	0.46-1.00	<0.001	78	56-86		3.98
	Adjusted for publication bias	71	0.65	0.53-0.78		76	69-80		4.55
Type of control	Waiting list	28	0.98	0.80-1.17	<0.001	68	50-77	0.002	2.85
	Care-as-usual	30	0.60	0.45-0.75	<0.001	69	54-78		4.99
	Pill placebo	5	0.55	0.28-0.81	<0.001	45	0-78		5.51
High-quality studies	Waiting list	6	0.93	0.49-1.37	<0.001	82	56-90	0.06	3.02
	Care-as-usual	5	0.43	0.16-0.70	0.002	46	0-79		7.29
GAD									
All control conditions	All studies	31	0.80	0.67-0.93	<0.001	33	0-56		3.58
	High-quality studies	9	0.82	0.60-1.04	<0.001	46	0-73		3.49
	Adjusted for publication bias	42	0.59	0.44-0.75		62	44-72		5.08
Type of control	Waiting list	24	0.85	0.72-0.99	<0.001	13	0-47	<0.001	3.35
	Care-as-usual	4	0.45	0.26-0.64	<0.001	0	0-68		6.93
	Pill placebo	3	1.32	0.83-1.81	<0.001	0	0-73		2.08
High-quality studies	Waiting list	8	0.88	0.67-1.10	<0.001	33	0-69	0.05	3.22
	Care-as-usual	1	0.45	0.08-0.83	0.02	0			6.93
PAD									
All control conditions	All studies	42	0.81	0.59-1.04	<0.001	77	69-82		3.53
	High-quality studies	4	0.61	0.27-0.96	0.001	26	0-75		4.89
Type of control	Waiting list	33	0.96	0.70-1.23	<0.001	77	67-82	<0.001	2.92
	Care-as-usual	4	0.27	−0.12 to 0.65	0.17	31	0-77		12.25
	Pill placebo	5	0.28	0.03-0.54	0.03	8	0-67		11.77
High-quality studies	Waiting list	4	0.61	0.27-0.96	0.001	26	0-75		4.89
SAD									
All control conditions	All studies	48	0.88	0.74-1.03	<0.001	64	50-73		3.22
	High-quality studies	8	0.76	0.47-1.06	<0.001	71	25-84		3.80
Type of control	Waiting list	40	0.98	0.83-1.14	<0.001	64	47-73	<0.001	2.85
	Care-as-usual	3	0.44	0.12-0.77	0.01	23	0-79		7.11
	Pill placebo	5	0.47	0.24-0.70	<0.001	0	0-64		6.59
High-quality studies	Waiting list	5	1.00	0.61-1.40	<0.001	71	0-87	0.03	2.79
	Care as usual	2	0.30	−0.04 to 0.64	0.08	0			10.91
	Pill placebo	1	0.57	0.20-0.93	0.002	0			5.29

NNT – Number-needed-to-treat

Only 11 of the 63 studies were rated as being high-quality. The effect size in these studies was similar to that in the total pool (g=0.73; 95% CI: 0.46-1.00; I^2=78). No high-quality study used a pill placebo control group. The difference between waiting list and care-as-usual among the high-quality studies was not significant (p=0.06), but this may be related to the small number of those studies.

Egger's test indicated considerable asymmetry of the funnel plot (intercept: 1.54; 95% CI: 0.59-2.50; p=0.001). Duval and Tweedie's trim and fill procedure also indicated considerable publication bias (number of imputed studies: 8; adjusted effect size: g=0.65; 95% CI: 0.53-0.78; I^2=76). For high-quality studies, no indication for publication bias was found (but this may again be related to the small number of those studies).

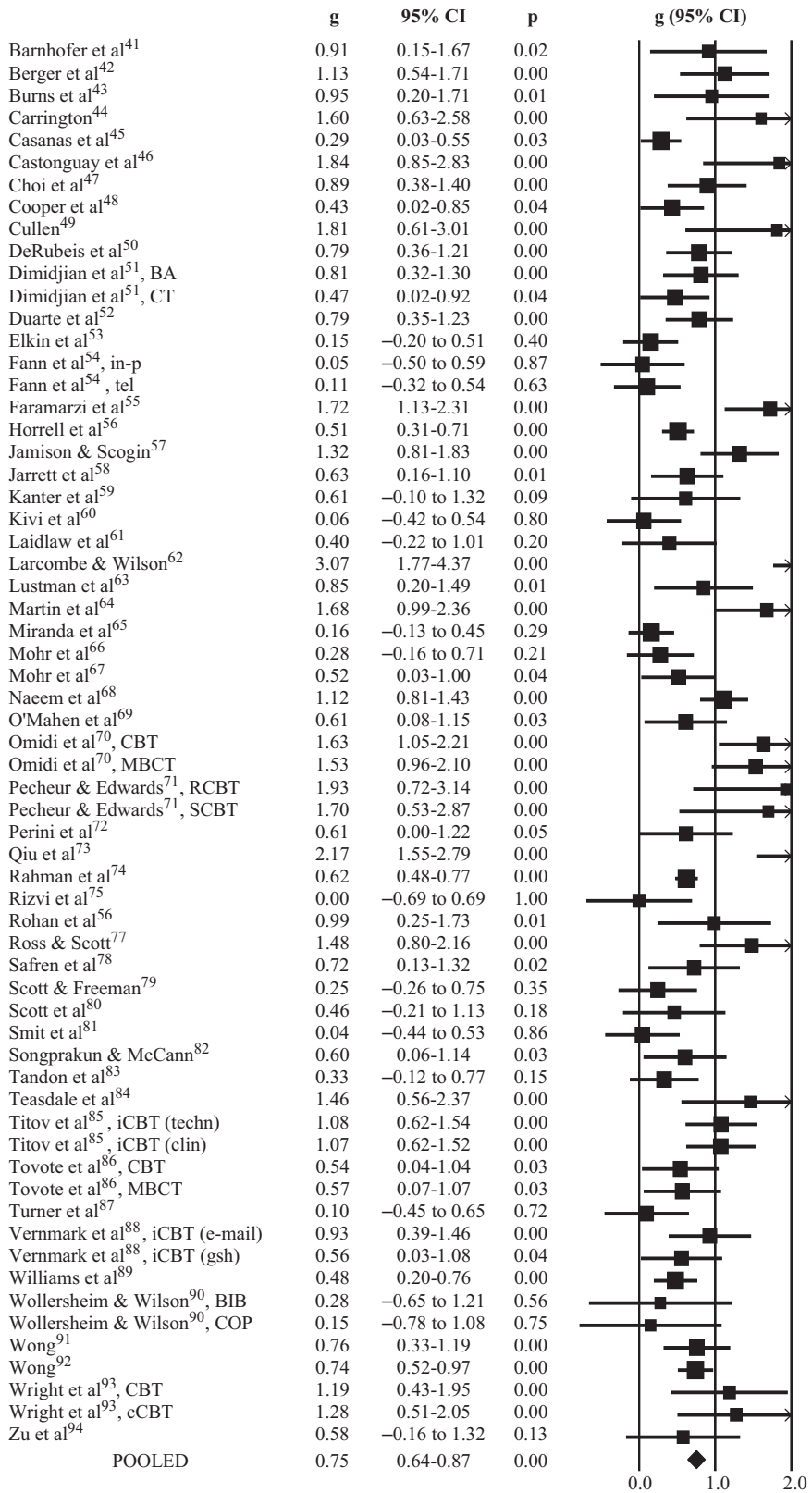

Figure 2 Effects of cognitive behavior therapy (CBT) for major depression compared to control conditions: forest plot.
BA – behavioral activation, CT – cognitive therapy, in-p – in person, tel – telephone, MBCT – mindfulness based CBT, RCBT – religious CBT, SCBT – secular CBT, iCBT – Internet-delivered CBT, techn – supported by a technician, clin – supported by a clinician, e-mail – supervised by e-mail, gsh – guided self-help format, BIB – bibliotherapy, COP – coping, cCBT – computerized CBT

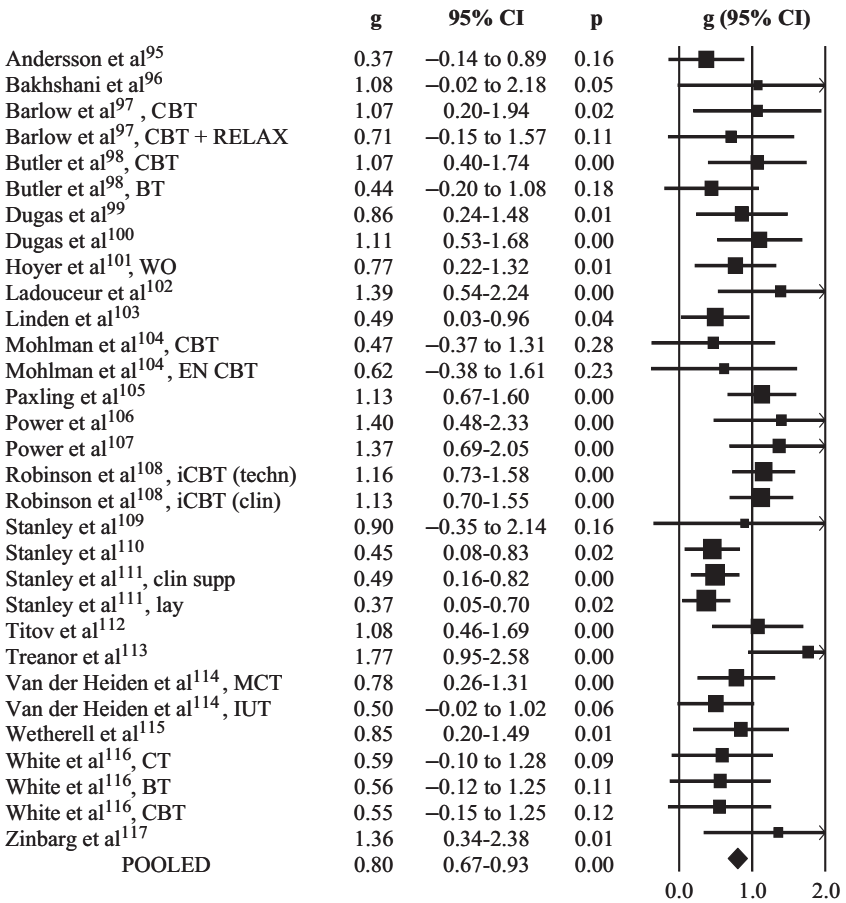

Figure 3 Effects of cognitive behavior therapy (CBT) for generalized anxiety disorder compared to control conditions: forest plot.
RELAX – relaxation, BT – behavior therapy, WO – worry exposure, EN CBT – enhanced CBT, iCBT – Internet-delivered CBT, techn – technician assistance, clin – clinician assistance, clin supp – supported by a clinician, lay – lay provider, MCT – metacognitive therapy, IUT – intolerance-of-uncertainty therapy, CT – cognitive therapy

Effects of CBT on GAD

The pooled effect size of the 31 comparisons between CBT and control conditions in GAD[95-117] was g=0.80 (95% CI: 0.67-0.93; NNT=3.58), with low to moderate heterogeneity (I^2=33) (Table 1 and Figure 3). The vast majority of studies (24 of 31) used a waiting list control group. Studies using a pill placebo control group (g=1.32) had a significantly (p<0.001) larger effect than those using a waiting list (g=0.85) or care-as-usual control group (g=0.45). The number of studies using pill placebo (N=3) and care-as-usual control groups (N=4) was very small, however (Table 1 and Figure 3).

Only 9 of the 31 studies were rated as high-quality, and 8 of these used a waiting list control group, so the effects of care-as-usual and pill placebo among high-quality studies could not be estimated.

Egger's test was significant (intercept: 1.60; 95% CI: 0.38-2.83; p=0.006). Duval and Tweedie's trim and fill procedure resulted in an adjusted effect size of g=0.59 (95% CI: 0.44-0.75; I^2=62; number of imputed studies: 11). For high-quality studies, no indication for publication bias was found (but this may again be related to the small number of those studies).

Effects of CBT on PAD

The 42 comparisons between CBT and control conditions in PAD[118-147] resulted in a pooled effect size of g=0.81 (95% CI: 0.59-1.04; I^2=77; NNT=3.53). In the vast majority of the comparisons (N=33), a waiting list control condition was used. The difference between studies using a waiting list (g=0.96) and either care-as-usual (g=0.27) or pill placebo (g=0.28) was significant (p<0.001). The four comparisons of CBT versus care-as-usual even indicated a non-significant effect size (g=0.27; 95% CI: −0.12 to 0.65; p=0.17) (Table 1 and Figure 4).

The four high-quality studies all used a waiting list control group and resulted in an effect size of g=0.61 (95% CI: 0.27-0.96).

Although Egger's test indicated significant asymmetry of the funnel plot (intercept: 3.62; 95% CI: 0.90-6.34; p=0.005), Duval and Tweedie's trim and fill procedure did not indicate any missing studies and therefore the adjusted and unadjusted effect sizes were the same. In the four high-quality studies, no indication for publication bias was found.

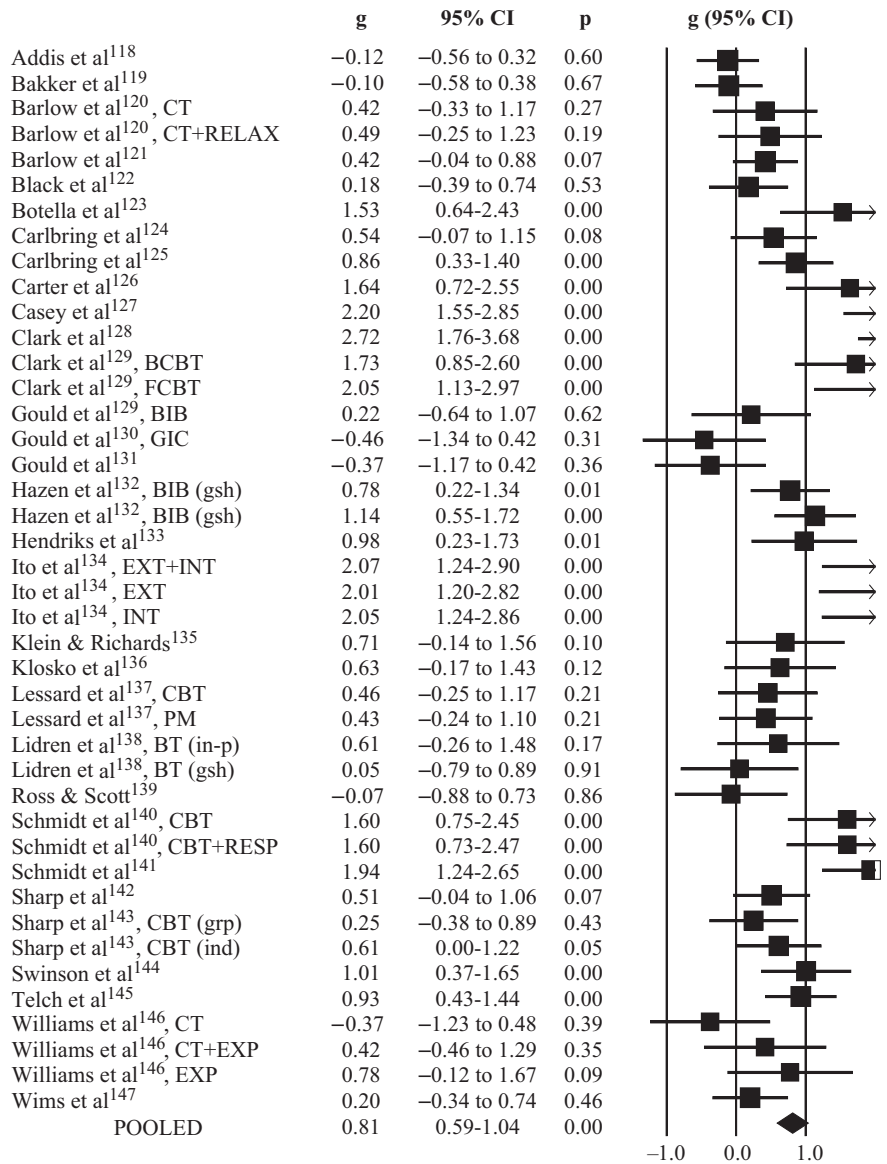

Figure 4 Effects of cognitive behavior therapy (CBT) for panic disorder compared to control conditions: forest plot.
CT – cognitive therapy, RELAX – relaxation, BCBT – brief CBT, FCBT – full CBT, BIB – bibliotherapy, GIC – guided imaginal coping, gsh – guided self help, grp – group format, EXT – external cues, INT – interoceptive, PM – panic management, in-p – in person, RESP – respiratory training, ind – individual format, EXP – exposure

Effects of CBT on SAD

The 48 comparisons between CBT and a control condition[148-183] resulted in a pooled effect size of g=0.88 (95% CI: 0.74-1.03; I^2=64; NNT=3.22). Again, the large majority of studies used a waiting list control group (N=40), with only three using care-as-usual and five pill placebo. The studies using a waiting list control group resulted in significantly (p<0.001) larger effect sizes (g=0.98) than those using a pill placebo (g=0.47) or care-as-usual control group (g=0.44) (Table 1 and Figure 5).

Only eight studies were rated as high-quality, and five of these used a waiting list control group. This implies that for SAD there are not enough high-quality studies to assess the effects of CBT compared to care-as-usual or pill placebo.

Egger's test pointed at significant asymmetry of the funnel plot (intercept: 2.46; 95% CI: 0.96-3.96; p=0.001), but Duval and Tweedie's trim and fill procedure did not indicate missing studies and the adjusted and unadjusted effect sizes were the same.

Multivariate meta-regression analyses

We conducted four separate analyses, for each disorder, with the effect size as the dependent variable and characteristics of

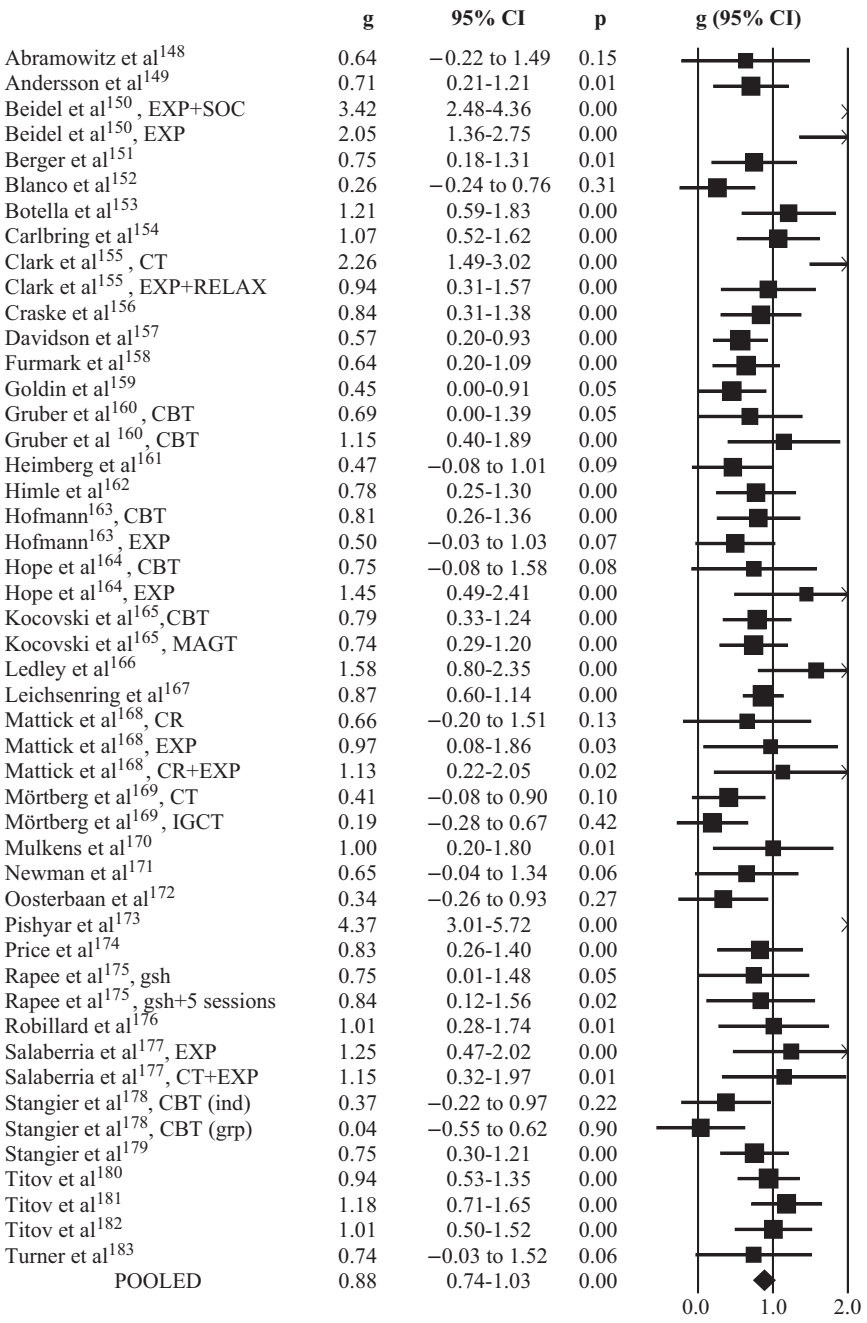

Figure 5 Effects of cognitive behavior therapy (CBT) for social anxiety disorder compared to control conditions: forest plot.
EXP – exposure, SOC – social skills, CT – cognitive therapy, RELAX – relaxation, cCBT – computerized CBT, MAGT – mindfulness acceptance group therapy, CR – cognitive restructuring, IGCT – intensive group CT, gsh – guided self help, ind – individual format, grp – group format

the participants (adults in general or more specific populations), the intervention (format and number of sessions) and the study in general (type of control group, quality and geographic area) as predictors. As shown in Table 2, very few predictors were significant in these analyses, possibly because of the relatively small number of studies per disorder and the relatively large number of predictors.

DISCUSSION

In this study, we aimed to establish the most up-to-date and accurate estimate of the effects of CBT in the treatment of MDD, GAD, PAD and SAD. We also aimed to examine whether the problems of publication bias, low quality of trials, and the use of waiting list control groups have an impact on the effect

Table 2 Standardized regression coefficients of characteristics of studies on cognitive behavior therapy for major depression (MDD), generalized anxiety disorder (GAD), panic disorder (PAD) and social anxiety disorder (SAD) compared to control conditions

		MDD			GAD			PAD			SAD		
		Coeff	**SE**	**p**	**Coeff**	**SE**	**p**	**Coeff**	**SE**	**p**	**Coeff**	**SE**	**p**
Quality of trial		−0.05	0.07	0.46	−0.01	0.07	0.94	−0.09	0.11	0.43	−0.01	0.09	0.92
Control condition	Waiting list	Ref.			Ref.			Ref.			Ref.		
	Care-as-usual	−0.43	0.15	**0.01**	−0.30	0.38	0.43	−0.61	0.41	0.69	−0.67	0.46	0.15
	Pill placebo	−0.44	0.30	0.15	0.60	0.40	0.15	−0.67	0.34	0.05	−0.53	0.29	0.08
Adults vs. specific target groups		0.01	0.17	0.95	−0.43	0.28	0.14	−0.07	0.38	0.85	0.67	0.75	0.38
Format	Individual	Ref.			Ref.			Ref.			Ref.		
	Group	−0.23	0.21	0.28	−0.17	0.23	0.47	0.28	0.31	0.37	−0.06	0.25	0.83
	Guided self-help	−0.32	0.23	0.16	0.06	0.28	0.84	−0.36	0.30	0.24	−0.06	0.36	0.86
	Mixed/other	−0.28	0.28	0.32	0.04	0.30	0.89	0.48	0.68	0.48	0.20	0.45	0.65
Number of sessions		−0.01	0.02	0.67	−0.01	0.02	0.60	0.06	0.04	0.13	0.04	0.03	0.19
Geographic area	North America	Ref.			Ref.			Ref.			Ref.		
	Europe	−0.02	0.19	0.92	−0.51	0.19	**0.01**	0.65	0.25	**0.01**	−0.13	0.24	0.59
	Australia	0.31	0.29	0.29	−0.19	0.30	0.52	0.37	0.54	0.49	0.40	0.31	0.20
	Other	0.47	0.22	**0.04**	−0.78	0.66	0.25	1.58	0.48	**0.003**			

Significant p values are highlighted in bold prints

sizes. We found that the overall effects for all four disorders were large, ranging from g=0.75 for MDD to g=0.80 for GAD, g=0.81 for PAD, and g=0.88 for SAD.

The first problem, publication bias, mostly affected the outcomes of CBT for GAD and MDD. For GAD, it was estimated that about one quarter of the studies were missing and, after adjusting for these missing studies, the effect size dropped from g=0.80 to g=0.59. For MDD, 14% of the studies were missing, and the pooled effect size dropped from g=0.75 to g=0.65. However, this was a relatively small drop compared to that reported in other studies on publication bias in psychotherapies for MDD[15,18,184]. This may be due to the fact that we used more stringent inclusion criteria for this meta-analysis (only patients meeting diagnostic criteria for MDD; only waiting list, treatment-as-usual or pill placebo control groups; only CBT). In PAD and SAD, we found few indications of publication bias.

The second problem we aimed to examine was the quality of trials. We found that the methodological quality in most studies was low or unknown. We evaluated the quality by the Cochrane "risk of bias" assessment tool, and found that across all disorders only 25 trials (17.4%) were rated as high-quality. The effect size was lower in high-quality studies for PAD (g=0.61 compared to g=0.81 in all studies) and SAD (g=0.76 compared to g=0.88 in all studies). We did not find strong indications that the quality of trials was associated with the effect size in MDD and GAD. Although we did not find a strong association between effect size and quality of trials for all disorders, the small number of high-quality studies still means that the overall effect sizes we found for all four disorders are uncertain.

The third problem we aimed to examine was the influence of waiting list control groups on the effects of CBT. We found that the vast majority of studies for the three anxiety disorders used a waiting list control group (77.4% of the comparisons for GAD, 78.6% for PAD, and 83.3% for SAD). In MDD, the number of studies using care-as-usual and pill placebo control conditions was larger, but still 44.4% (28 out of 63) of the included studies used a waiting list control group. This means that much of the evidence on the effects of CBT is based on the use of waiting list control groups. As indicated earlier, improvements found in patients on waiting lists are lower than can be expected on the basis of spontaneous remission[19,185]. Waiting list is probably a "nocebo"[21], considerably overestimating the effects of psychological treatments. This was confirmed in our meta-analysis, in which we found for each of the disorders that studies with a waiting list control group resulted in significantly higher effect sizes than those with a care-as-usual or pill placebo control group.

The few studies on anxiety disorders that used care-as-usual or pill placebo control groups indicated small to moderate

effect sizes. In the four studies comparing CBT for PAD with care-as-usual, the effect size was even non-significant (p=0.17). Furthermore, because of the small number of studies, and the even smaller number of high-quality studies, the effects of CBT in anxiety disorders are quite uncertain.

An exception to the small to moderate effects of CBT in anxiety disorders was the group of studies comparing CBT to pill placebo for GAD. These studies resulted in a very large effect size (g=1.32). However, because of the small number of trials and the low quality of all three of them, these results should be considered with caution.

One reason to conduct this meta-analysis was to examine whether the quality of trials has increased in recent years. Indeed, 29.5% of the studies conducted in 2010 or later were rated as high-quality, while that was true for only 12.0% of the older studies. Furthermore, 52.0% of all high-quality studies were conducted in 2010 or later. This is likely to have led to a more accurate estimate of effect sizes.

The present study has several strengths, including the broad scope of the meta-analyses, covering four common mental disorders, the rigorous selection and assessment of the trials, and their relatively large number.

One possible limitation is that we used strict inclusion criteria, only focusing on trials in which patients met diagnostic criteria for the disorder according to a structured interview and trials in which either a waiting list, care-as-usual or pill placebo control group was used. We did not include studies in which, for example, generic counselling was used as a control condition. This may contribute to explain the small number of trials comparing CBT with control conditions other than waiting lists, especially in anxiety disorders and among the sets of high-quality studies. Furthermore, care-as-usual control groups can vary considerably depending on the country and the treatment setting where the therapy is offered, and may therefore be too heterogeneous to allow a reliable assessment of the effects across studies. Finally, we only focused on short-term outcomes, because only few studies reported long-term outcomes and the follow-up periods differed considerably.

On the basis of our data, we conclude that CBT is probably effective in the treatment of MDD, GAD, PAD and SAD, and that the effects are large when compared to waiting list control groups, but small to moderate when compared to more conservative control groups, such as care-as-usual and pill placebo. Because of the small number of high-quality studies, these effects are still uncertain and should be considered with caution.

REFERENCES

1. Steel Z, Marnane C, Iranpour C et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. Int J Epidemiol 2014;43:476-93.
2. 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.
3. Gustavsson A, Svensson M, Jacobi F et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011;21:718-79.
4. Hu T-W. Perspectives: an international review of the national cost estimates of mental illness, 1990-2003. J Mental Health Policy Econ 2006;9:3-13.
5. Chisholm D, Sweeny K, Sheehan P et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. Lancet Psychiatry 2016;3:415-24.
6. Bloom DE, Cafiero E, Jané-Llopis E et al. The global economic burden of non-communicable diseases. Geneva: World Economic Forum, 2011.
7. Olfson M, Marcus SC. National patterns in antidepressant medication treatment. Arch Gen Psychiatry 2009;66:848-56.
8. Cuijpers P, Sijbrandij M, Koole SL et al. The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. World Psychiatry 2013;12:137-48.
9. Acarturk C, Cuijpers P, van Straten A et al. Psychological treatment of social anxiety disorder: a meta-analysis. Psychol Med 2009;39:241-54.
10. Cuijpers P, Sijbrandij M, Koole S et al. Psychological treatment of generalized anxiety disorder: a meta-analysis. Clin Psychol Rev 2014;34:130-40.
11. Sánchez-Meca J, Rosa-Alcázar AI, Marín-Martínez F et al. Psychological treatment of panic disorder with or without agoraphobia: a meta-analysis. Clin Psychol Rev 2010;30:37-50.
12. Olfson M, Marcus SC. National trends in outpatient psychotherapy. Am J Psychiatry 2010;167:1456-63.
13. McHugh RK, Whitton SW, Peckham AD et al. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. J Clin Psychiatry 2013;74:595-602.
14. Barth J, Munder T, Gerger H et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med 2013;10:e1001454.
15. Driessen E, Hollon SD, Bockting CLH et al. Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of US National Institutes of Health-funded trials. PLoS One 2015;10:e0137864.
16. Cuijpers P, Smit F, Bohlmeijer E et al. Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. Br J Psychiatry 2010;196:173-8.
17. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA 1990;263:1385-9.
18. Cuijpers P, van Straten A, Bohlmeijer E et al. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med 2010;40:211-23.
19. Mohr DC, Spring B, Freedland KE et al. The selection and design of control conditions for randomized controlled trials of psychological interventions. Psychother Psychosom 2009;78:275-84.
20. Cuijpers P, Cristea IA. How to prove that your therapy is effective, even when it is not: a guideline. Epidemiol Psychiatr Sci 2015;28:1-8.
21. Furukawa TA, Noma H, Caldwell DM et al. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta Psychiatr Scand 2014;130:181-92.
22. Cuijpers P, van Straten A, Warmerdam L et al. Psychological treatment of depression: a meta-analytic database of randomized studies. BMC Psychiatry 2008;8:36.
23. Cuijpers P, Donker T, Johansson R et al. Self-guided psychological treatment for depressive symptoms: a meta-analysis. PLoS One 2011;6: e21274.
24. Cuijpers P, Driessen E, Hollon SD et al. The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. Clin Psychol Rev 2012;32:280-91.
25. Higgins JPT, Altman DG, Gotzsche PC et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
26. Cohen J. Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale: Erlbaum, 1988.
27. Hedges LV, Olkin I. Statistical methods for meta-analysis. Orlando: Academic Press, 1985.
28. Beck AT, Ward CH, Mendelson M et al. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561-71.
29. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: manual. San Antonio: Psychological Corporation, 1996.
30. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56-62.
31. Beck AT, Epstein N, Brown G et al. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56:893-7.
32. Meyer TJ, Miller ML, Metzger RL et al. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990;28:487-95.

33. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. Behav Res Ther 1979;17:263-7.
34. Liebowitz MR. Social phobia. In: Klein DF (ed). Modern trends in pharmacopsychiatry. Berlin: Karger, 1987:141-73.
35. Furukawa TA. From effect size into number needed to treat. Lancet 1999; 353:1680.
36. Cuijpers P, Karyotaki E, Weitz E et al. The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. J Affect Disord 2014;159:118-26.
37. Higgins JPT, Thompson SG, Deeks JJ et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
38. Ioannidis JPA, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. BMJ 2007;335:914-6.
39. Orsini N, Bottai M, Higgins J et al. Heterogi: Stata module to quantify heterogeneity in a meta-analysis. Statistical Software Components S449201. Boston: Boston College Department of Economics, 2005.
40. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455-63.
41. Barnhofer T, Crane C, Hargus E et al. Mindfulness-based cognitive therapy as a treatment for chronic depression: a preliminary study. Behav Res Ther 2009;47:366-73.
42. Berger T, Hämmerli K, Gubser N et al. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011;40:251-66.
43. Burns A, O'Mahen H, Baxter H et al. A pilot randomised controlled trial of cognitive behavioural therapy for antenatal depression. BMC Psychiatry 2013;13:33.
44. Carrington CH. A comparison of cognitive and analytically oriented brief treatment approaches to depression in black women. College Park: University of Maryland, 1979.
45. Casanas R, Catalan R, del Val JL et al. Effectiveness of a psychoeducational group program for major depression in primary care: a randomized controlled trial. BMC Psychiatry 2012;12:230.
46. Castonguay LG, Schut AJ, Aikens DE et al. Integrative cognitive therapy for depression: a preliminary investigation. J Psychother Integr 2004;14:4-20.
47. Choi I, Zou J, Titov N et al. Culturally attuned Internet treatment for depression amongst Chinese Australians: a randomised controlled trial. J Affect Disord 2012;136:459-68.
48. Cooper PJ, Murray L, Wilson A et al. Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. Br J Psychiatry 2003;182:412-9.
49. Cullen JM. Testing the effectiveness of behavioral activation therapy in the treatment of acute unipolar depression. Dissertation, Western Michigan University, Kalamazoo, MI, 2002.
50. DeRubeis RJ, Hollon SD, Amsterdam JD et al. Cognitive therapy vs medications in the treatment of moderate to severe depression. Arch Gen Psychiatry 2005;62:409-16.
51. Dimidjian S, Hollon SD, Dobson KS et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. J Consult Clin Psychol 2006;74:658-70.
52. Duarte PS, Miyazaki MC, Blay SL et al. Cognitive-behavioral group therapy is an effective treatment for major depression in hemodialysis patients. Kidney Int 2009;76:414-21.
53. Elkin I, Shea MT, Watkins JT et al. National Institute of Mental Health treatment of depression collaborative research program: general effectiveness of treatments. Arch Gen Psychiatry 1989;46:971-82.
54. Fann JR, Bombardier CH, Vannoy S et al. Telephone and in-person cognitive behavioral therapy for major depression after traumatic brain injury: a randomized controlled trial. J Neurotrauma 2015;32:45-57.
55. Faramarzi M, Alipor A, Esmaelzadeh S et al. Treatment of depression and anxiety in infertile women: cognitive behavioral therapy versus fluoxetine. J Affect Disord 2008;108:159-64.
56. Horrell L, Goldsmith KA, Tylee AT et al. One-day cognitive-behavioural therapy self-confidence workshops for people with depression: randomised controlled trial. Br J Psychiatry 2014;204:222-33.
57. Jamison C, Scogin F. The outcome of cognitive bibliotherapy with depressed adults. J Consult Clin Psychol 1995;63:644-50.
58. Jarrett RB, Schaffer M, McIntire D et al. Treatment of atypical depression with cognitive therapy or phenelzine: a double-blind, placebo-controlled trial. Arch Gen Psychiatry 1999;56:431-7.
59. Kanter JW, Santiago-Rivera AL, Santos MM et al. A randomized hybrid efficacy and effectiveness trial of behavioral activation for Latinos with depression. Behav Ther 2015;46:177-92.
60. Kivi M, Eriksson MC, Hange D et al. Internet-based therapy for mild to moderate depression in Swedish primary care: short term results from the PRIM-NET randomized controlled trial. Cogn Behav Ther 2014;43: 289-98.
61. Laidlaw K, Davidson K, Toner H et al. A randomised controlled trial of cognitive behaviour therapy vs treatment as usual in the treatment of mild to moderate late life depression. Int J Geriatr Psychiatry 2008;23: 843-50.
62. Larcombe NA, Wilson PH. An evaluation of cognitive-behaviour therapy for depression in patients with multiple sclerosis. Br J Psychiatry 1984; 145:366-71.
63. Lustman PJ, Griffith LS, Freedland KE et al. Cognitive behavior therapy for depression in type 2 diabetes mellitus: a randomized, controlled trial. Ann Intern Med 1998;129:613-21.
64. Martin PR, Aiello R, Gilson K et al. Cognitive behavior therapy for comorbid migraine and/or tension-type headache and major depressive disorder: an exploratory randomized controlled trial. Behav Res Ther 2015;73:8-18.
65. Miranda J, Chung JY, Green BL et al. Treating depression in predominantly low-income young minority women: a randomized controlled trial. JAMA 2003;290:57-65.
66. Mohr DC, Boudewyn AC, Goodkin DE et al. Comparative outcomes for individual cognitive-behavior therapy, supportive-expressive group psychotherapy, and sertraline for the treatment of depression in multiple sclerosis. J Consult Clin Psychol 2001;69:942-9.
67. Mohr DC, Duffecy J, Ho J et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013;8: e70086.
68. Naeem F, Sarhandi I, Gul M et al. A multicentre randomised controlled trial of a carer supervised culturally adapted CBT (CaCBT) based self-help for depression in Pakistan. J Affect Disord 2014;156:224-7.
69. O'Mahen H, Himle JA, Fedock G et al. A pilot randomized controlled trial of cognitive behavioral therapy for perinatal depression adapted for women with low incomes. Depress Anxiety 2013;30:679-87.
70. Omidi A, Mohammadkhani P, Mohammadi A et al. Comparing mindfulness based cognitive therapy and traditional cognitive behavior therapy with treatments as usual on reduction of major depressive disorder symptoms. Iran Red Crescent Med J 2013;15:142-6.
71. Pecheur DR, Edwards KJ. A comparison of secular and religious versions of cognitive therapy with depressed Christian college students. J Psychol Theol 1984;12:45-54.
72. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. Aust N Z J Psychiatry 2009;43:571-8.
73. Qiu J, Chen W, Gao X et al. A randomized controlled trial of group cognitive behavioral therapy for Chinese breast cancer patients with major depression. J Psychosom Obstet Gynecol 2013;34:60-7.
74. Rahman A, Malik A, Sikander S et al. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. Lancet 2008;372:902-9.
75. Rizvi SJ, Zaretsky A, Schaffer A et al. Is immediate adjunctive CBT more beneficial than delayed CBT in treating depression? A pilot study. J Psychiatr Pract 2015;21:107-13.
76. Rohan KJ, Roecklein KA, Tierney Lindsey K et al. A randomized controlled trial of cognitive-behavioral therapy, light therapy, and their combination for seasonal affective disorder. J Consult Clin Psychol 2007; 75:489-500.
77. Ross M, Scott M. An evaluation of the effectiveness of individual and group cognitive therapy in the treatment of depressed patients in an inner city health centre. J R Coll Gen Pr 1985;35:239-42.
78. Safren SA, Gonzalez JS, Wexler DJ et al. A randomized controlled trial of cognitive behavioral therapy for adherence and depression (CBT-AD) in patients with uncontrolled type 2 diabetes. Diabetes Care 2014;37:625-33.
79. Scott AI, Freeman CP. Edinburgh primary care depression study: treatment outcome, patient satisfaction, and cost after 16 weeks. BMJ 1992; 304:883-7.
80. Scott C, Tacchi MJ, Jones R et al. Acute and one-year outcome of a randomised controlled trial of brief cognitive therapy for major depressive disorder in primary care. Br J Psychiatry 1997;171:131-4.

81. Smit A, Kluiter H, Conradi HJ et al. Short-term effects of enhanced treatment for depression in primary care: results from a randomized controlled trial. Psychol Med 2006;36:15-26.
82. Songprakun W, McCann TV. Effectiveness of a self-help manual on the promotion of resilience in individuals with depression in Thailand: a randomised controlled trial. BMC Psychiatry 2012;12:1.
83. Tandon SD, Leis JA, Mendelson T et al. Six-month outcomes from a randomized controlled trial to prevent perinatal depression in low-income home visiting clients. Matern Child Health J 2014;18:873-81.
84. Teasdale JD, Fennell MJ, Hibbert GA et al. Cognitive therapy for major depressive disorder in primary care. Br J Psychiatry 1984;144:400-6.
85. Titov N, Andrews G, Davies M et al. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5:e10939.
86. Tovote KA, Fleer J, Snippe E et al. Individual mindfulness-based cognitive therapy and cognitive behavior therapy for treating depressive symptoms in patients with diabetes: results of a randomized controlled trial. Diabetes Care 2014;37:2427-34.
87. Turner A, Hambridge J, Baker A et al. Randomised controlled trial of group cognitive behaviour therapy versus brief intervention for depression in cardiac patients. Aust N Z J Psychiatry 2013;47:235-43.
88. Vernmark K, Lenndin J, Bjärehed J et al. Internet administered guided self-help versus individualized e-mail therapy: a randomized trial of two versions of CBT for major depression. Behav Res Ther 2010;48:368-76.
89. Williams C, Wilson P, Morrison J et al. Guided self-help cognitive behavioural therapy for depression in primary care: a randomised controlled trial. PLoS One 2013;8:e52735.
90. Wollersheim JP, Wilson GL. Group treatment of unipolar depression: a comparison of coping, supportive, bibliotherapy, and delayed treatment groups. Prof Psychol Res Pract 1991;22:496-502.
91. Wong DFK. Cognitive behavioral treatment groups for people with chronic depression in Hong Kong: a randomized wait-list control design. Depress Anxiety 2008;25:142-8.
92. Wong DFK. Cognitive and health-related outcomes of group cognitive behavioural treatment for people with depressive symptoms in Hong Kong: randomized wait-list control study. Aust N Z J Psychiatry 2008;42:702-11.
93. Wright JH, Wright AS, Albano AM et al. Computer-assisted cognitive therapy for depression: maintaining efficacy while reducing therapist time. Am J Psychiatry 2005;162:1158-64.
94. Zu S, Xiang Y-T, Liu J et al. A comparison of cognitive-behavioral therapy, antidepressants, their combination and standard treatment for Chinese patients with moderate-severe major depressive disorders. J Affect Disord 2014;152:262-7.
95. Andersson G, Paxling B, Roch-Norlund P et al. Internet-based psychodynamic versus cognitive behavioral guided self-help for generalized anxiety disorder: a randomized controlled trial. Psychother Psychosom 2012; 81:344-55.
96. Bakhshani NM, Lashkaripour K, Sadjadi SA. Effectiveness of short term cognitive behavior therapy in patients with generalized anxiety disorder. J Med Sci 2007;7:1076-81.
97. Barlow DH, Rapee RM, Brown TA. Behavioral treatment of generalized anxiety disorder. Behav Ther 1992;23:551-70.
98. Butler G, Fennell M, Robson P et al. Comparison of behavior therapy and cognitive behavior therapy in the treatment of generalized anxiety disorder. J Consult Clin Psychol 1991;59:167-75.
99. Dugas MJ, Brillon P, Savard P et al. A randomized clinical trial of cognitive-behavioral therapy and applied relaxation for adults with generalized anxiety disorder. Behav Ther 2010;41:46-58.
100. Dugas MJ, Ladouceur R, Léger E et al. Group cognitive-behavioral therapy for generalized anxiety disorder: treatment outcome and long-term follow-up. J Consult Clin Psychol 2003;71:821-5.
101. Hoyer J, Beesdo K, Gloster AT et al. Worry exposure versus applied relaxation in the treatment of generalized anxiety disorder. Psychother Psychosom 2009;78:106-15.
102. Ladouceur R, Dugas MJ, Freeston MH et al. Efficacy of a cognitive-behavioral treatment for generalized anxiety disorder: evaluation in a controlled clinical trial. J Consult Clin Psychol 2000;68:957-64.
103. Linden M, Zubraegel D, Baer T et al. Efficacy of cognitive behaviour therapy in generalized anxiety disorders. Psychother Psychosom 2005; 74:36-42.
104. Mohlman J, Gorenstein EE, Kleber M et al. Standard and enhanced cognitive-behavior therapy for late-life generalized anxiety disorder: two pilot investigations. Am J Geriatr Psychiatry 2003;11:24-32.
105. Paxling B, Almlöv J, Dahlin M et al. Guided internet-delivered cognitive behavior therapy for generalized anxiety disorder: a randomized controlled trial. Cogn Behav Ther 2011;40:159-73.
106. Power KG, Jerrom DWA, Simpson RJ et al. A controlled comparison of cognitive-behaviour therapy, diazepam and placebo in the management of generalized anxiety. Behav Psychother 1989;17:1-14.
107. Power KG, Simpson RJ, Swanson V et al. A controlled comparison of cognitive-behaviour therapy, diazepam, and placebo, alone and in combination, for the treatment of generalised anxiety disorder. J Anxiety Disord 1990;4:267-92.
108. Robinson E, Titov N, Andrews G et al. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5:e10942.
109. Stanley MA, Hopko DR, Diefenbach GJ et al. Cognitive-behavior therapy for late-life generalized anxiety disorder in primary care: preliminary findings. Am J Geriatr Psychiatry 2003;11:92-6.
110. Stanley MA, Wilson NL, Novy DM et al. Cognitive behavior therapy for generalized anxiety disorder among older adults in primary care: a randomized clinical trial. JAMA 2009;301:1460-7.
111. Stanley MA, Wilson NL, Amspoker AB et al. Lay providers can deliver effective cognitive behavior therapy for older adults with generalized anxiety disorder: a randomized trial. Depress Anxiety 2014;31:391-401.
112. Titov N, Andrews G, Robinson E et al. Clinician-assisted Internet-based treatment is effective for generalized anxiety disorder: randomized controlled trial. Aust N Z J Psychiatry 2009;43:905-12.
113. Treanor M, Erisman SM, Salters-Pedneault K et al. Acceptance-based behavioral therapy for GAD: effects on outcomes from three theoretical models. Depress Anxiety 2011;28:127-36.
114. van der Heiden C, Muris P, van der Molen HT. Randomized controlled trial on the effectiveness of metacognitive therapy and intolerance-of-uncertainty therapy for generalized anxiety disorder. Behav Res Ther 2012;50:100-9.
115. Wetherell JL, Gatz M, Craske MG. Treatment of generalized anxiety disorder in older adults. J Consult Clin Psychol 2003;71:31-40.
116. White J, Keenan M, Brooks N. Stress control: a controlled comparative investigation of large group therapy for generalized anxiety disorder. Behav Psychother 1992;20:97-113.
117. Zinbarg RE, Lee JE, Yoon KL. Dyadic predictors of outcome in a cognitive-behavioral program for patients with generalized anxiety disorder in committed relationships: a 'spoonful of sugar' and a dose of nonhostile criticism may help. Behav Res Ther 2007;45:699-713.
118. Addis ME, Hatgis C, Krasnow AD et al. Effectiveness of cognitive-behavioral treatment for panic disorder versus treatment as usual in a managed care setting. J Consult Clin Psychol 2004;72:625-35.
119. Bakker A, van Dyck R, van Balkom AJ. Paroxetine, clomipramine, and cognitive therapy in the treatment of panic disorder. J Clin Psychiatry 1999;60:831-8.
120. Barlow DH, Craske MG, Cerny JA et al. Behavioral treatment of panic disorder. Behav Ther 1989;20:261-82.
121. Barlow DH, Gorman JM, Shear MK et al. Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: a randomized controlled trial. JAMA 2000;283:2529-36.
122. Black DW, Wesner R, Bowers W et al. A comparison of fluvoxamine, cognitive therapy, and placebo in the treatment of panic disorder. Arch Gen Psychiatry 1993;50:44-50.
123. Botella C, García-Palacios A, Villa H et al. Virtual reality exposure in the treatment of panic disorder and agoraphobia: a controlled study. Clin Psychol Psychother 2007;14:164-75.
124. Carlbring P, Westling BE, Ljungstrand P et al. Treatment of panic disorder via the Internet: a randomized trial of a self-help program. Behav Ther 2001;32:751-64.
125. Carlbring P, Bohman S, Brunt S et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. Am J Psychiatry 2006;163:2119-25.
126. Carter MM, Sbrocco T, Gore KL et al. Cognitive-behavioral group therapy versus a wait-list control in the treatment of African American women with panic disorder. Cogn Ther Res 2003;27:505-18.
127. Casey LM, Newcombe PA, Oei TP. Cognitive mediation of panic severity: the role of catastrophic misinterpretation of bodily sensations and panic self-efficacy. Cogn Ther Res 2005;29:187-200.
128. Clark DM, Salkovskis PM, Hackmann A et al. A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. Br J Psychiatry 1994;164:759-69.

129. Clark DM, Salkovskis PM, Hackmann A et al. Brief cognitive therapy for panic disorder: a randomized controlled trial. J Consult Clin Psychol 1999;67:583-9.
130. Gould RA, Clum GA, Shapiro D. The use of bibliotherapy in the treatment of panic: a preliminary investigation. Behav Ther 1993;24:241-52.
131. Gould RA, Clum GA. Self-help plus minimal therapist contact in the treatment of panic disorder: a replication and extension. Behav Ther 1995;26:533-46.
132. Hazen AL, Walker JR, Eldridge GD. Anxiety sensitivity and treatment outcome in panic disorder. Anxiety 1996;2:34-9.
133. Hendriks G-J, Keijsers GPJ, Kampman M et al. A randomized controlled study of paroxetine and cognitive-behavioural therapy for late-life panic disorder. Acta Psychiatr Scand 2010;122:11-9.
134. Ito LM, De Araujo LA, Tess VLC et al. Self-exposure therapy for panic disorder with agoraphobia: randomised controlled study of external v. interoceptive self-exposure. Br J Psychiatry 2001;178:331-6.
135. Klein B, Richards JC. A brief Internet-based treatment for panic disorder. Behav Cogn Psychother 2001;29:113-7.
136. Klosko JS, Barlow DH, Tassinari R et al. A comparison of alprazolam and behavior therapy in treatment of panic disorder. J Consult Clin Psychol 1990;58:77-84.
137. Lessard M-J, Marchand A, Pelland M-È et al. Comparing two brief psychological interventions to usual care in panic disorder patients presenting to the emergency department with chest pain. Behav Cogn Psychother 2012;40:129-47.
138. Lidren DM, Watkins PL, Gould RA et al. A comparison of bibliotherapy and group therapy in the treatment of panic disorder. J Consult Clin Psychol 1994;62:865-9.
139. Ross CJ, Davis TM, Macdonald GF. Cognitive-behavioral treatment combined with asthma education for adults with asthma and coexisting panic disorder. Clin Nurs Res 2005;14:131-57.
140. Schmidt NB, Trakowski JH, Staab JP. Extinction of panicogenic effects of a 35% CO_2 challenge in patients with panic disorder. J Abnorm Psychol 1997;106:630-8.
141. Schmidt NB, McCreary BT, Trakowski JJ et al. Effects of cognitive behavioral treatment on physical health status in patients with panic disorder. Behav Ther 2003;34:49-63.
142. Sharp DM, Power KG, Simpson RJ et al. Fluvoxamine, placebo, and cognitive behaviour therapy used alone and in combination in the treatment of panic disorder and agoraphobia. J Anxiety Disord 1996;10:219-42.
143. Sharp DM, Power KG, Swanson V. A comparison of the efficacy and acceptability of group versus individual cognitive behaviour therapy in the treatment of panic disorder and agoraphobia in primary care. Clin Psychol Psychother 2004;11:73-82.
144. Swinson RP, Fergus KD, Cox BJ et al. Efficacy of telephone-administered behavioral therapy for panic disorder with agoraphobia. Behav Res Ther 1995;33:465-9.
145. Telch MJ, Lucas JA, Schmidt NB et al. Group cognitive-behavioral treatment of panic disorder. Behav Res Ther 1993;31:279-87.
146. Williams SL, Falbo J. Cognitive and performance-based treatments for panic attacks in people with varying degrees of agoraphobic disability. Behav Res Ther 1996;34:253-64.
147. Wims E, Titov N, Andrews G et al. Clinician-assisted Internet-based treatment is effective for panic: a randomized controlled trial. Aust N Z J Psychiatry 2010;44:599-607.
148. Abramowitz JS, Moore EL, Braddock AE et al. Self-help cognitive-behavioral therapy with minimal therapist contact for social phobia: a controlled trial. J Behav Ther Exp Psychiatry 2009;40:98-105.
149. Andersson G, Carlbring P, Holmström A et al. Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: a randomized controlled trial. J Consult Clin Psychol 2006;74:677-86.
150. Beidel DC, Alfano CA, Kofler MJ et al. The impact of social skills training for social anxiety disorder: a randomized controlled trial. J Anxiety Disord 2014;28:908-18.
151. Berger T, Hohl E, Caspar F. Internet-based treatment for social phobia: a randomized controlled trial. J Clin Psychol 2009;65:1021-35.
152. Blanco C, Heimberg RG, Schneier FR et al. A placebo-controlled trial of phenelzine, cognitive behavioral group therapy, and their combination for social anxiety disorder. Arch Gen Psychiatry 2010;67:286-95.
153. Botella C, Gallego MJ, Garcia-Palacios A et al. An Internet-based self-help treatment for fear of public speaking: a controlled trial. Cyberpsychol Behav Soc Netw 2010;13:407-21.
154. Carlbring P, Gunnarsdóttir M, Hedensjö L et al. Treatment of social phobia: randomised trial of internet-delivered cognitive-behavioural therapy with telephone support. Br J Psychiatry 2007;190:123-8.
155. Clark DM, Ehlers A, Hackmann A et al. Cognitive therapy versus exposure and applied relaxation in social phobia: a randomized controlled trial. J Consult Clin Psychol 2006;74:568-78.
156. Craske MG, Niles AN, Burklund LJ et al. Randomized controlled trial of cognitive behavioral therapy and acceptance and commitment therapy for social phobia: outcomes and moderators. J Consult Clin Psychol 2014;82:1034-48.
157. Davidson JR, Foa EB, Huppert JD et al. Fluoxetine, comprehensive cognitive behavioral therapy, and placebo in generalized social phobia. Arch Gen Psychiatry 2004;61:1005-13.
158. Furmark T, Carlbring P, Hedman E et al. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. Br J Psychiatry 2009;195:440-7.
159. Goldin PR, Ziv M, Jazaieri H et al. Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. J Consult Clin Psychol 2012;80:1034-40.
160. Gruber K, Moran PJ, Roth WT et al. Computer-assisted cognitive behavioral group therapy for social phobia. Behav Ther 2001;32:155-65.
161. Heimberg RG, Liebowitz MR, Hope DA et al. Cognitive behavioral group therapy vs phenelzine therapy for social phobia: 12-week outcome. Arch Gen Psychiatry 1998;55:1133-41.
162. Himle JA, Bybee D, Steinberger E et al. Work-related CBT versus vocational services as usual for unemployed persons with social anxiety disorder: a randomized controlled pilot trial. Behav Res Ther 2014;63:169-76.
163. Hofmann SG. Cognitive mediation of treatment change in social phobia. J Consult Clin Psychol 2004;72:392-9.
164. Hope DA, Heimberg RG, Bruch MA. Dismantling cognitive-behavioral group therapy for social phobia. Behav Res Ther 1995;33:637-50.
165. Kocovski NL, Fleming JE, Hawley LL et al. Mindfulness and acceptance-based group therapy versus traditional cognitive behavioral group therapy for social anxiety disorder: a randomized controlled trial. Behav Res Ther 2013;51:889-98.
166. Ledley DR, Heimberg RG, Hope DA et al. Efficacy of a manualized and workbook-driven individual treatment for social anxiety disorder. Behav Ther 2009;40:414-24.
167. Leichsenring F, Salzer S, Beutel ME et al. Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: a multicenter randomized controlled trial. Am J Psychiatry 2013;170:1-9.
168. Mattick RP, Peters L, Clarke JC. Exposure and cognitive restructuring for social phobia: a controlled study. Behav Ther 1989;20:3-23.
169. Mörtberg E, Clark DM, Sundin Ö et al. Intensive group cognitive treatment and individual cognitive therapy vs. treatment as usual in social phobia: a randomized controlled trial. Acta Psychiatr Scand 2007;115:142-54.
170. Mulkens S, Bögels SM, de Jong PJ et al. Fear of blushing: effects of task concentration training versus exposure in vivo on fear and physiology. J Anxiety Disord 2001;15:413-32.
171. Newman MG, Hofmann SG, Trabert W et al. Does behavioral treatment of social phobia lead to cognitive changes? Behav Ther 1994;25:503-17.
172. Oosterbaan DB, van Balkom AJ, Spinhoven P et al. Cognitive therapy versus moclobemide in social phobia: a controlled study. Clin Psychol Psychother 2001;8:263-73.
173. Pishyar R, Harris LM, Menzies RG. Responsiveness of measures of attentional bias to clinical change in social phobia. Cogn Emot 2008;22:1209-27.
174. Price M, Anderson PL. The impact of cognitive behavioral therapy on post event processing among those with social anxiety disorder. Behav Res Ther 2011;49:132-7.
175. Rapee RM, Abbott MJ, Baillie AJ et al. Treatment of social phobia through pure self-help and therapist-augmented self-help. Br J Psychiatry 2007; 191:246-52.
176. Robillard G, Bouchard S, Dumoulin S et al. Using virtual humans to alleviate social anxiety: preliminary report from a comparative outcome study. Stud Health Technol Inf 2010;154:57-60.
177. Salaberria K, Echeburua E. Long-term outcome of cognitive therapy's contribution to self-exposure in vivo to the treatment of generalized social phobia. Behav Modif 1998;22:262-84.
178. Stangier U, Heidenreich T, Peitz M et al. Cognitive therapy for social phobia: individual versus group treatment. Behav Res Ther 2003;41:991-1007.

179. Stangier U, Schramm E, Heidenreich T et al. Cognitive therapy vs interpersonal psychotherapy in social anxiety disorder: a randomized controlled trial. Arch Gen Psychiatry 2011;68:692-700.
180. Titov N, Andrews G, Schwencke G et al. Shyness 1: distance treatment of social phobia over the Internet. Aust N Z J Psychiatry 2008;42:585-94.
181. Titov N, Andrews G, Schwencke G. Shyness 2: treating social phobia online: replication and extension. Aust N Z J Psychiatry 2008;42:595-605.
182. Titov N, Andrews G, Choi I et al. Shyness 3: randomized controlled trial of guided versus unguided Internet-based CBT for social phobia. Aust N Z J Psychiatry 2008;42:1030-40.
183. Turner SM, Beidel DC, Jacob RG. Social phobia: a comparison of behavior therapy and atenolol. J Consult Clin Psychol 1994;62:350-8.
184. Cuijpers P, Berking M, Andersson G et al. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Can J Psychiatry 2013;58:376-85.
185. Cuijpers P, Cristea IA. What if a placebo effect explained all the activity of depression treatments? World Psychiatry 2015;14:310-1

DOI:10.1002/wps.20346

Ultra high risk status and transition to psychosis in 22q11.2 deletion syndrome

Maude Schneider[1,2], Marco Armando[1,3], Maria Pontillo[3], Stefano Vicari[3], Martin Debbané[1,4,5], Frauke Schultze-Lutter[6], Stephan Eliez[1,7]

[1]Developmental Imaging and Psychopathology Lab, Geneva, Switzerland; [2]Center for Contextual Psychiatry, Department of Neuroscience, KU Leuven, Leuven, Belgium; [3]Child and Adolescence Neuropsychiatry Unit, Department of Neuroscience, Bambino Gesù Children Hospital, Rome, Italy; [4]Developmental Clinical Psychology Unit, Faculty of Psychology, University of Geneva, Geneva, Switzerland; [5]Research Department of Clinical, Educational and Health Psychology, University College London, London, UK; [6]University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland; [7]Department of Genetic Medicine and Development, School of Medicine, University of Geneva, Geneva, Switzerland

The 22q11.2 deletion syndrome (22q11DS) is characterized by high rates of psychotic symptoms and schizophrenia, making this condition a promising human model for studying risk factors for psychosis. We explored the predictive value of ultra high risk (UHR) criteria in a sample of patients with 22q11DS. We also examined the additional contribution of socio-demographic, clinical and cognitive variables to predict transition to psychosis within a mean interval of 32.5 ± 17.6 months after initial assessment. Eighty-nine participants with 22q11DS (age range: 8-30 years; mean 16.1 ± 4.7) were assessed using the Structured Interview for Psychosis-Risk Syndromes. Information on Axis I diagnoses, internalizing and externalizing symptoms, level of functioning and IQ was also collected. At baseline, 22 (24.7%) participants met UHR criteria. Compared to those without a UHR condition, they had a significantly lower functioning, more frequent anxiety disorders, and more severe psychopathology. Transition rate to psychosis was 27.3% in UHR and 4.5% in non-UHR participants. Cox regression analyses revealed that UHR status significantly predicted conversion to psychosis. Baseline level of functioning was the only other additional predictor. This is the first study investigating the predictive value of UHR criteria in 22q11DS. It indicates that the clinical path leading to psychosis is broadly comparable to that observed in other clinical high-risk samples. Nevertheless, the relatively high transition rate in non-UHR individuals suggests that other risk markers should be explored in this population. The role of low functioning as a predictor of transition to psychosis should also be investigated more in depth.

Key words: 22q11.2 deletion syndrome, schizophrenia, clinical high risk state, ultra high risk criteria, transition to psychosis, level of functioning

(World Psychiatry 2016;15:259–265)

In the past 20 years, there has been an increasing interest in people presenting with potentially prodromal symptoms of psychosis, i.e. with a clinical high risk state. Criteria have been developed to identify that high risk state: the ultra high risk (UHR) and the basic symptom criteria[1,2]. A recent meta-analysis reported an ability of UHR criteria to detect transition to psychosis within two years in 20% of individuals in clinical samples[3]. Yet, although persons with a clinical high risk state have a significantly increased risk of developing psychosis, many will not develop a psychotic disorder. The specificity of clinical high risk assessments hence remains relatively low[4].

Studies of genetic syndromes associated with increased risk of schizophrenia have become increasingly important. Among these genetic conditions, 22q11.2 deletion syndrome (22q11DS) is particularly valuable[5]. This syndrome is characterized in most cases by a microdeletion of 3 million base pairs on chromosome 22 band q11, and has an estimated prevalence of 1:2.000-4.000 live births[6]. From a clinical perspective, 22q11DS is associated with high rates of psychiatric disorders, especially schizophrenia[7]. While 23 to 45% of affected adolescents report transient psychotic experiences[8-11], up to 40% of affected adults are diagnosed with a psychotic disorder[7]. Moreover, 22q11DS was found in 0.3 to 2.0% of patients with schizophrenia[12-14], with rates of up to 5.7% in patients with childhood-onset schizophrenia[15]. Taken together, these findings indicate that 22q11DS is a highly relevant genetic risk factor for schizophrenia and the most promising human model for studying risk factors and states at risk for schizophrenia[5].

Several studies have investigated prodromal symptoms in patients with 22q11DS, reporting rates between 45 and 56% for UHR symptoms and between 10 and 21% for UHR criteria (including frequency and onset/worsening requirements)[8,10,11,16-18]. Armando et al[8] compared the symptom profile of UHR patients with (N=30) vs. without (N=81) 22q11DS and found no significant group difference in positive symptoms, while negative symptoms were more severe in patients with 22q11DS.

Yet, few studies have prospectively investigated risk factors for psychosis in the 22q11DS population. Gothelf et al[19] found that anxiety disorder and lower full-scale IQ at baseline, and a greater decline in verbal IQ were the best predictors of transition to psychosis. In line with these findings, Vorstman et al[20] reported that an early cognitive decline, particularly in verbal IQ, was a robust predictor of psychosis. Finally, a recent study highlighted the role of poor premorbid adjustment during childhood and adolescence in the development of UHR symptoms and full-blown psychosis[21]. No study, however, has longitudinally examined the predictive value of UHR criteria in people with 22q11DS.

We investigated prospectively, in a large group of patients with 22q11DS over an average period of 32 months, the value of UHR criteria as well as of other relevant variables as predictors of conversion to psychosis. We hypothesized that the predictive value of UHR criteria would be comparable to that found in other clinical samples, albeit expecting an overall higher transition rate, given the higher prevalence of psychotic

disorders in 22q11DS. Secondly, we expected that low baseline verbal IQ, the presence of an anxiety disorder, and low baseline level of functioning would increase the predictive accuracy in addition to the presence of an UHR condition.

METHODS

Participants

We included 89 participants (56 from Geneva and 33 from Rome) with a genetically confirmed 22q11DS diagnosis, aged between 8 and 30 years (mean 16.1 ± 4.7) at baseline. Having a psychotic disorder at baseline was an exclusion criterion. Children were assessed from the age of 8 onwards, as previous studies reported the presence of psychotic symptoms in young children with 22q11DS[9]. Participants were followed-up over a mean period of 32.5 ± 17.6 months (range: 12-85).

Participants from Geneva were recruited through advertisements in patient associations or word of mouth; those from Rome were referred from the Genetic Clinical Unit of the Bambino Gesù Hospital or recruited through advertisement in patient associations. Written informed consent from the participants and their parents was collected at both sites under protocols approved by local institutional ethical review boards.

Compared to the Geneva cohort, participants from Rome were younger (mean age 14.3 ± 5.1 vs. 17.1 ± 4.2 years, $t=2.89$, $p=0.005$) and had a higher full-scale IQ (84.5 ± 10.9 vs. 72.2 ± 10.0, $t=-5.44$, $p<0.001$), while the gender distribution at baseline was similar (females: 51.5% vs. 55.4%, $\chi^2=0.12$, $p=0.725$).

Assessments

All participants completed the Structured Interview for Psychosis-Risk Syndromes[22] to assess the severity of positive, negative, disorganization and general symptoms, as well as the presence of UHR symptoms (any P1-P5 ≥3) and UHR criteria (i.e., attenuated psychotic symptoms (APS), brief limited intermittent symptoms (BLIPS), or genetic risk and functional decline (GRFD) criteria). We also explored the rate of participants meeting criteria for perceptive (P4) and non-perceptive (P1, P2, P3 or P5) APS or BLIPS[23]. For the global assessment of functioning, the Childhood Global Assessment Scale (CGAS)[24] or the Global Assessment of Functioning (GAF) was used.

The presence of any Axis I DSM-IV psychiatric disorder was assessed using structured clinical interviews. In both cohorts, the Structured Clinical Interview for Axis I DSM-IV (SCID-I)[25] was administered to adult participants and their parents. In Geneva, parents of participants below 18 years completed the Diagnostic Interview for Children and Adolescents - Revised (DICA-IV)[26] and diagnoses were confirmed with the participants. The psychotic disorders supplement of the Schedule for Affective Disorders and Schizophrenia for School-Age Children - Present and Lifetime version (K-SADS-PL)[27] was also administered. In Rome, the K-SADS-PL, including the psychotic disorders supplement, was used for children and adolescents.

Intellectual functioning was assessed by trained psychologists at both time points using the Wechsler Intelligence Scale for Children - third edition (WISC-III)[28] or the Wechsler Adult Intelligence Scale - third edition (WAIS-III)[29]. Verbal IQ, performance IQ and full-scale IQ were used as indicators of intellectual functioning.

Parents of all participants completed the Child Behavior Checklist (CBCL)[30] or the Adult Behavior Checklist (ABCL)[31]. Internalizing, externalizing and total problems T-scores were used as global measures of the severity of psychopathology.

Statistical analyses

All statistical analyses were performed using SPSS version 21. Cross-sectional group comparisons between participants with and without UHR criteria at baseline were performed using independent t-tests, Mann-Whitney U, or χ^2 tests. Because of the variable interval between baseline and follow-up, we conducted a Cox regression analysis to determine whether the presence of any UHR state at baseline (UHR status) significantly predicted conversion to psychosis. Furthermore, we used Cox regression analyses to examine whether perceptive and non-perceptive APS/BLIPS were both predictors of conversion to psychosis. In case of two significant models, both predictors were entered in a stepwise Cox regression analysis.

We then examined the contribution of additional variables at baseline to improve the UHR-based prediction model. Potential predictors included: demographic characteristics (age, gender), clinical parameters (presence of any anxiety, affective or behavioral disorder, CGAS/GAF scores, CBCL/ABCL internalizing, externalizing and total problems T-scores, and severity of positive, negative, disorganization and general symptoms), and cognitive variables (verbal, performance and full-scale IQ).

A multiple step approach was adopted in order to derive a parsimonious model. First, each predictor was entered individually in a series of Cox regression analyses and selected for further analyses when the Wald statistic was significant at a liberal level ($p<0.05$). Next, each selected predictor was entered in a multiple Cox regression analysis, with UHR status always included as a predictor. Variables were further selected if the Wald statistic was significant ($p<0.05$) for both variables (UHR status and the additional predictor), indicating that the additional predictor contributed to the improvement of prediction without decreasing the predictive value of UHR status. If more than one predictor met the above-mentioned criteria, all predictors were analyzed together using a forward and backward Cox regression analysis to exclude effects of blocking. The maximum number of predictors entering the final model was limited to a 1:5 ratio of number of predictors to event. The proportional hazard assumption was tested at each step prior to each Cox regression following the procedure described by Kleinbaum and Klein[32].

Table 1 Descriptive characteristics of non-ultra high risk (UHR–) and ultra high risk (UHR+) participants at baseline

	UHR− (N=67)	UHR+ (N=22)	Statistics
Age (years), mean±SD (range)	15.9 ± 4.9 (8-30)	16.6 ± 4.0 (9-24)	t=–0.616, p=0.539
Gender, N females (%)	38 (56.7%)	10 (45.4%)	χ^2=0.845, p=0.358
Any antipsychotics, N (%)	0 (0%)	5 (22.7%)	**χ^2=16.134, p<0.001**
SIPS P1, median (range)	1.00 (0-3)	3.00 (0-4)	**U=1217.00, p<0.001**
SIPS P2, median (range)	1.00 (0-3)	2.50 (1-5)	**U=1174.00, p<0.001**
SIPS P3, median (range)	0.00 (0-2)	0.00 (0-3)	U=857.50, p=0.145
SIPS P4, median (range)	0.00 (0-4)	3.00 (0-6)	**U=1298.00, p<0.001**
SIPS P5, median (range)	0.00 (0-3)	2.00 (0-5)	**U=1027.00, p=0.002**
Any Axis I diagnosis, N (%)	41 (61.2%)	16 (72.7%)	χ^2=0.957, p=0.328
Any anxiety disorder, N (%)	20 (29.9%)	13 (59.1)	**χ^2=6.069, p=0.014**
Any mood disorder, N (%)	15 (22.4%)	5 (22.7%)	χ^2=0.001, p=0.974
Any behavioral disorder, N (%)	17 (25.4%)	6 (27.3%)	χ^2=0.031, p=0.860
Any substance use, N (%)	1 (1.5%)	0 (0%)	χ^2=0.332, p=0.564
CGAS/GAF score, mean±SD	63.0 ± 11.3	56.6 ± 8.8	**U=484.50, p=0.016**
CGAS/GAF score <70, N (%)	45 (67.2%)	19 (86.4%)	χ^2=3.022, p=0.082
Verbal IQ, mean ± SD	81.1 ± 13.2	77.3 ± 10.8	t=1.227, p=0.223
Performance IQ, mean ± SD	77.1 ± 11.9	73.9 ± 14.1	t=1.070, p=0.288
Full-scale IQ, mean ± SD	77.6 ± 11.9	74.1 ± 11.9	t=1.208, p=0.230
CBCL/ABCL internalizing T-score, mean ± SD	62.7 ± 11.0	70.3 ± 9.4	**U=1036.50, p=0.004**
CBCL/ABCL externalizing T-score, mean ± SD	53.8 ± 9.8	60.9 ± 10.8	**U=968.50, p=0.009**
CBCL/ABCL total problems T-score, mean ± SD	62.1 ± 10.7	69.5 ± 11.5	**U=998.00, p=0.013**

SIPS – Structured Interview for Psychosis-Risk Syndromes, CGAS – Childhood Global Assessment Scale, GAF – Global Assessment of Functioning, CBCL – Child Behavior Checklist, ABCL – Adult Behavior Checklist

Significant differences are highlighted in bold prints

RESULTS

Baseline characteristics

Twenty-two (24.7%) participants met UHR criteria at baseline (UHR+) and 67 (75.3%) did not (UHR−). Compared to UHR− participants, UHR+ were more frequently under antipsychotic medication and diagnosed with an anxiety disorder, had more severe positive symptoms (all P subscales except grandiosity) and internalizing and externalizing symptoms, and had a lower functioning at baseline, although not a higher rate of functional deficit (CGAS/GAF score <70) (Table 1). The most frequent UHR condition was APS (N=15; 68.2%), followed by GRFD (N=6; 27.3%) and BLIPS (N=2; 9.1%); only one participant (4.2%) met criteria for both APS and GRFD. Among the 17 participants with APS or BLIPS, four (23.5%) presented with perceptive, six (35.3%) with non-perceptive, and seven (41.2%) with both perceptive and non-perceptive APS/BLIPS.

In addition, 10 (11.2%) participants experienced UHR symptoms − six on item P4 (perceptual abnormalities/hallucinations), two on item P2 (suspiciousness/persecutory ideas), and two on several items − but failed to meet frequency (N=6) or both frequency and onset/worsening requirements (N=4). Altogether, the prevalence of UHR symptoms (regardless of frequency and onset/worsening requirements) was 36.0%.

Outcome

Altogether, nine (10.1%) participants had converted to psychosis at follow-up, four being minors (<18 years) and five adults at baseline (Table 2). The six UHR+ converters included three of those diagnosed with APS (out of 15, 20.0%), both of those diagnosed with BLIPS (100%), and one of those diagnosed with GRFD (out of six, 16.7%). Of the three false-negative cases (i.e., UHR− participants at baseline who converted to psychosis), one had reported UHR symptoms at baseline that did not meet frequency and worsening/onset criteria. Five participants (all UHR+) were receiving antipsychotic medication at baseline; three of them converted to psychosis at follow-up.

None of the ten participants who remitted from UHR status (i.e., UHR+ at baseline, but UHR− at follow-up) had received antipsychotic medication at either baseline or follow-up. Furthermore, four out of 89 (4.5%) participants had a new onset of UHR criteria at follow-up.

Table 2 Outcome at follow-up for non-ultra high risk (UHR–) and ultra high risk (UHR+) participants

UHR− participants (N=67)	
UHR− at follow-up, N (%)	60 (89.6%)
UHR+ at follow-up, N (%)	4 (6.0%)
Psychotic disorder at follow-up, N (%)	3 (4.5%)
UHR+ participants (N=22)	
UHR− at follow-up, N (%)	10 (45.5%)
UHR+ at follow-up, N (%)	6 (27.3%)
Psychotic disorder at follow-up, N (%)	6 (27.3%)

Predictors of conversion to psychosis

UHR status at baseline was a significant predictor of transition to psychosis (β=1.823, SE=0.733, Wald (df=1) = 6.181, p=0.013; Exp(β) = 6.188, 95% CI: 1.471-26.033). Furthermore, presence of both perceptive APS/BLIPS (β=1.644, SE=0.737, Wald (df=1) = 4.975, p=0.026; Exp(β) = 5.178, 95% CI: 1.221-21.961) and non-perceptive APS/BLIPS (β=3.397, SE=0.876, Wald (df=1) = 15.021, p<0.001; Exp(β) = 29.868, 95% CI: 5.360-166.432) significantly predicted transition to psychosis. When both variables were entered in a stepwise Cox regression analysis, only the presence of non-perceptive APS/BLIPS remained in the final model.

With regard to additional predictors, CGAS/GAF at baseline remained the only significant predictor after the two selection steps. The final model (Table 3), including UHR status and CGAS/GAF as predictors, was highly significant (−2LL=48.768, χ^2(df=2) = 15.329, p<0.001). Cumulative hazard rates of the model were 0.015 at two years, 0.024 at three years and 0.113 at four years.

DISCUSSION

UHR symptoms and criteria

Altogether, 32 (36.0%) participants reported at least one UHR symptom (APS or BLIPS) regardless of the frequency and onset/worsening requirements. Twenty-two of them (24.7%) fully met UHR criteria (i.e., including frequency and onset/worsening requirements). Both rates are broadly consistent with previous studies in 22q11DS, reporting rates between 45 and 56% for UHR symptoms and between 10 and 21% for UHR criteria[10,11,16-18,33]. Thus, our findings confirm that patients with 22q11DS are at increased risk of experiencing attenuated symptoms of psychosis, regardless of transition to psychosis[23,34]. Indeed, recent estimates from the general population were between 7.3 and 9.9% for lifetime UHR symptoms and between 0.4 and 1.3% for current UHR criteria.

In the present sample, APS was the most prevalent UHR condition (68.2%), followed by GRFD (27.3%) and BLIPS (9.1%). While the preponderance of APS and the low frequency of BLIPS is consistent with findings from other clinical high risk populations[3,35-37], GRFD was more frequent than in most clinical UHR samples[3].

In line with earlier findings[9], patients meeting UHR criteria exhibited lower level of functioning and higher levels of internalizing and externalizing symptoms, and were more frequently diagnosed with an anxiety disorder. These findings highlight that the presence of a UHR status in 22q11DS, similarly to other clinical populations[3], is in itself a condition that aggravates the clinical picture and, consequently, requires clinical attention, irrespective of any potential future transition to psychosis.

Although a European guidance on early intervention in clinical high risk states does not recommend the use of antipsychotics as first line treatment in patients with UHR[38], nearly a quarter of UHR+ but none of UHR− participants were receiving antipsychotic medication at baseline. This practice might have been linked to the treating clinicians' awareness of UHR symptoms and increased risk of psychosis in this population, and could be interpreted as a psychosis-prevention approach. However, antipsychotics might have been also prescribed for other behavioral problems, such as severe anxiety or externalizing symptoms, in this more symptomatic group.

Outcome

We observed a transition rate to psychosis of 27.3% among UHR+ participants, which is comparable to previous reports in other clinical samples[3,39]. Furthermore, with only 4.5% of UHR− participants developing psychosis, we found that the UHR status significantly predicted transition to psychosis in this specific subgroup of patients. However, in light of the increased risk of psychosis recognized in this population, it might be surprising that the transition rates were "only" in line with those of other clinical samples and not considerably higher[40].

Table 3 Final Cox regression model

Predictors	β	SE	Wald (df)	p	Exp(β)	95% CI
Any UHR criteria at baseline	1.544	0.748	4.266 (1)	0.039	4.685	1.082-20.286
Baseline CGAS/GAF	−0.086	0.030	8.209 (1)	0.004	0.903	0.865-0.973

UHR – ultra high risk, CGAS – Childhood Global Assessment Scale, GAF – Global Assessment of Functioning

Several reasons may explain this finding. First, the mean age of UHR+ participants in our sample was 16 years, with 23.6% being 12 years or younger and 53.9% being 15 years or younger. In children and younger adolescents, the clinical significance as well as the psychosis-predictive value of UHR criteria, especially of APS, was reported to be significantly lower than in individuals aged 16 or older[3-5,41]. Thus, the rather high proportion of participants below 16 years might have lowered the overall transition rate considerably. Second, the remission rate from UHR status in non-converters (62.5%) was in the upper range of what is typically described[42], suggesting that UHR symptom fluctuation is very common in this population[19]. Several characteristics of 22q11DS, such as intellectual disability and heightened anxiety levels, and their impact on adaptive functioning and everyday living skills, might indicate that stress sensitivity significantly influences variability in symptom severity[43,44]. Importantly, the interplay of these factors in daily life will have to be examined in future studies as they carry potentially crucial clinical implications.

Although the number of participants meeting criteria for each specific UHR condition (APS, BLIPS and GRFD) remains limited, we observed that a higher percentage of participants with BLIPS (100%) converted to psychosis, followed by APS (20.0%), and GRFD (16.7%). This pattern falls in line with results from a recent meta-analysis reporting the highest transition risk for BLIPS, and the lowest for GRFD[3]. Yet, while our numbers were consistent with the pooled transition rate for APS reported in that meta-analysis (17.4% at two and 29.1% at three years), the reported pooled transition rates for BLIPS (46.6% at two and 51.8% at three years) and GRFD (1.9% at two and 1.4% at three years) were lower in that meta-analysis than those found in the present sample[3].

We found that non-perceptive APS/BLIPS were a stronger predictor of transition to psychosis than perceptive APS/BLIPS. This finding is in line with previous reports of a low clinical significance of perceptive APS in children and adolescents from the general population[23,45,46]. This highlights the need for further studies examining UHR criteria and symptoms in relation to age in patients with 22q11DS.

The rate of false negative cases (i.e., UHR− participants who converted to psychosis) was higher (4.5%) than that reported in recent meta-analyses (0.9-1.6%) on patients seeking help at specialized early psychosis detection services[3,4]. This corroborates the fact that 22q11DS is a psychosis-risk condition in itself and linked to a higher baseline probability to develop psychosis. It also highlights the need of investigating other potential risk markers (e.g., clinical or cognitive) in order to improve the detection of patients who will convert to psychosis.

Additional predictors of transition to psychosis

In addition to UHR criteria, only lower baseline level of functioning, but not the presence of a functional deficit (CGAS/GAF score <70), significantly increased the predictive value of the model in our sample. This finding is in line with several studies that also identified lower functioning scores as a relevant predictor of the onset of psychosis in UHR[41,47-49] as well as 22q11DS samples[21,50]. Yet, contrary to other findings in 22q11DS[19,20], baseline verbal IQ and presence of an anxiety disorder did not increase the predictive value of the model. However, it should be noted that these previous studies never included UHR status as a baseline predictor. Hence, it is possible that including UHR status as a predictor reduced the variance explained by these other factors. Another potential explanation of this finding is that anxiety and verbal IQ decline precede or co-occur with the onset of UHR symptoms but do not predict transition to psychosis. Future studies examining the temporal dynamics of these different risk factors would help testing these hypotheses.

Strengths and limitations

This is the first study examining the predictive value of UHR criteria in patients with 22q11DS, which constitutes an important first step towards prevention of psychosis in this population. However, the relatively small sample size prevented a more detailed analysis of clinical outcomes or of interactions between variables. This limitation should be considered in light of the low prevalence of the syndrome, and is generally found in all longitudinal studies on 22q11DS[19].

A second limitation is the variable interval between the two assessments, which has been taken into account by the use of Cox regression analyses. This also deals with the fact that the true survival time is unknown in such studies (i.e., some participants are still likely to develop psychosis after the second assessment)[32].

A third limitation is the variance in age of the participants. This might have influenced some of the results, although it has been reported that the mean age of onset of psychosis in 22q11DS is lowered[7,19]. Furthermore, neither our assessment nor our sample size allowed a detailed analysis of treatment effects. However, this neglect of treatment effects beyond the prescription of antipsychotics is rather the rule than the exception in naturalistic UHR follow-up studies.

CONCLUSIONS

Our findings suggest that the psychopathological path leading to transition to psychosis in 22q11DS is broadly comparable to that observed in other clinical high risk samples, and confirm that 22q11DS can serve as a good human model for studying risk factors for psychosis.

The relatively high percentage of false negatives (i.e., UHR– participants who converted to psychosis) highlights that our efforts should now focus on investigating other possible, more subtle, risk markers – such as cognitive deficits and basic symptoms – to increase the sensitivity of our predictive model[2]. The role of low functioning as a predictor of transition to

psychosis should also be investigated more in depth by distinguishing different areas of functioning.

ACKNOWLEDGEMENTS

This work was supported by research grants from the Swiss National Science Foundation (grant nos. 324730_121996 and 324730_144260) and The National Centre of Competence in Research "Synapsy - The Synaptic Bases of Mental Diseases". The Brain and Behavior Research Foundation (formerly NARSAD) (grant no. 21278), the Swiss National Science Foundation (grant no. 162006) and the National Institute of Mental Health International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome (grant no. 5U01MH101722-02) also supported the present work. The authors would like to thank all the families who contributed to the study as well as the family associations (Aidel 22, Génération 22, Connect 22, Relais 22, and Creaf 22) for their ongoing support. Special thanks go to S. Menghetti from the Developmental Imaging and Psychopathology Lab for coordinating the project in Geneva and C. Di Gilio for leading the Genetic Unit in Rome. M. Schneider and M. Armando contributed equally to this work; F. Schultze-Lutter and S. Eliez share last authorship.

REFERENCES

1. 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.
2. Schultze-Lutter F, Debbané M, Theodoridou A et al. Revisiting the basic symptom concept: toward translating risk symptoms for psychosis into neurobiological targets. Front Psychiatry 2016;7:9.
3. Schultze-Lutter F, Michel C, Schmidt SJ et al. EPA guidance on the early detection of clinical high risk states of psychoses. Eur Psychiatry 2015;30: 405-16.
4. Fusar-Poli P, Cappucciati M, Rutigliano G et al. At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction. World Psychiatry 2015;14:322-32.
5. Insel TR. Rethinking schizophrenia. Nature 2010;468:187-93.
6. Bassett AS, McDonald-McGinn DM, Devriendt K et al. Practical guidelines for managing patients with 22q11.2 deletion syndrome. J Pediatr 2011;159:332-9.
7. Schneider M, Debbané M, Bassett AS et al. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. Am J Psychiatry 2014;171:627-39.
8. Armando M, Girardi P, Vicari S et al. Adolescents at ultra-high risk for psychosis with and without 22q11 deletion syndrome: a comparison of prodromal psychotic symptoms and general functioning. Schizophr Res 2012; 139:151-6.
9. Debbané M, Glaser B, David MK et al. Psychotic symptoms in children and adolescents with 22q11.2 deletion syndrome: neuropsychological and behavioral implications. Schizophr Res 2006;84:187-93.
10. Schneider M, Van der Linden M, Glaser B et al. Preliminary structure and predictive value of attenuated negative symptoms in 22q11.2 deletion syndrome. Psychiatry Res 2012;196:277-84.
11. Stoddard J, Niendam T, Hendren R et al. Attenuated positive symptoms of psychosis in adolescents with chromosome 22q11.2 deletion syndrome. Schizophr Res 2010;118:118-21.
12. Arinami T. Analyses of the associations between the genes of 22q11 deletion syndrome and schizophrenia. J Hum Genet 2006;51:1037-45.
13. Shprintzen RJ, Karayiorgou M, Morris MA et al. Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. Proc Natl Acad Sci USA 1995;92:7612-6.
14. Stefansson H, Rujescu D, Cichon S et al. Large recurrent microdeletions associated with schizophrenia. Nature 2008;455:232-6.
15. Sporn A, Addington A, Reiss AL et al. 22q11 deletion syndrome in childhood onset schizophrenia: an update. Mol Psychiatry 2004;9:225-6.
16. Rockers K, Ousley O, Sutton T et al. Performance on the Modified Card Sorting Test and its relation to psychopathology in adolescents and young adults with 22q11.2 deletion syndrome. J Intellect Disabil Res 2009;53:665-76.
17. Shapiro DI, Cubells JF, Ousley OY et al. Prodromal symptoms in adolescents with 22q11.2 deletion syndrome and schizotypal personality disorder. Schizophr Res 2011;129:20-8.
18. Tang SX, Yi JJ, Calkins ME et al. Psychiatric disorders in 22q11.2 deletion syndrome are prevalent but undertreated. Psychol Med 2013;44:1-11.
19. Gothelf D, Schneider M, Green T et al. Risk factors and the evolution of psychosis in 22q11.2 deletion syndrome: a longitudinal 2-site study. J Am Acad Child Adolesc Psychiatry 2013;52:1192-203.
20. Vorstman JAS, Breetvelt EJ, Duijff SN et al. Cognitive decline preceding the onset of psychosis in patients with 22q11.2 deletion syndrome. JAMA Psychiatry 2015;72:377-85.
21. Radoeva PD, Fremont W, Antshel KM et al. Longitudinal study of premorbid adjustment in 22q11.2 deletion (velocardiofacial) syndrome and association with psychosis. Dev Psychopathol (in press).
22. McGlashan T, Walsh BC, Woods SW. The psychosis-risk syndrome: handbook for diagnosis and follow-up. New York: Oxford University Press, 2010.
23. Schimmelmann BG, Michel C, Martz-Irngartinger A et al. Age matters in the prevalence and clinical significance of ultra-high-risk for psychosis symptoms and criteria in the general population: findings from the BEAR and BEARS-kid studies. World Psychiatry 2015;14:189-97.
24. Shaffer D, Gould MS, Brasic J et al. A children's global assessment scale (CGAS). Arch Gen Psychiatry 1983;40:1228-31.
25. Spitzer RL, Williams JB, Gibbon M et al. The Structured Clinical Interview for DSM-III-R (SCID). I: history, rationale, and description. Arch Gen Psychiatry 1992;49:624-9.
26. Reich W. Diagnostic interview for children and adolescents (DICA). J Am Acad Child Adolesc Psychiatry 2000;39:59-66.
27. 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.
28. Wechsler D. The Wechsler Intelligence Scale for Children – third edition: administration and scoring manual. San Antonio: Psychological Corporation, 1991.
29. Wechsler D. Wechsler Adult Intelligence Scale – III: administration and scoring manual. San Antonio: Psychological Corporation, 1997.
30. Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms and profiles. Burlington: University of Vermont, Research Center for Children, Youth, and Families, 2001.
31. Achenbach TM, Rescorla LA. Manual for the ASEBA adult forms and profiles. Burlington: University of Vermont, Research Center for Children, Youth and Families, 2003.
32. Kleinbaum DG, Klein M. Survival analysis: a self-learning text. New York: Springer, 2005.
33. Schneider M, Schaer M, Mutlu AK et al. Clinical and cognitive risk factors for psychotic symptoms in 22q11.2 deletion syndrome: a transversal and longitudinal approach. Eur Child Adolesc Psychiatry 2013;23:425-36.
34. Schultze-Lutter F, Michel C, Ruhrmann S et al. Prevalence and clinical significance of DSM-5-attenuated psychosis syndrome in adolescents and young adults in the general population: the Bern Epidemiological At-Risk (BEAR) study. Schizophr Bull 2014;40:1499-508.
35. Hartmann JA, Yuen HP, McGorry PD et al. Declining transition rates to psychotic disorder in "ultra-high risk" clients: investigation of a dilution effect. Schizophr Res 2016;170:130-6.
36. 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.
37. Nelson B, Yuen K, Yung AR. Ultra high risk (UHR) for psychosis criteria: are there different levels of risk for transition to psychosis? Schizophr Res 2011;125:62-8.
38. 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.
39. 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.
40. Fusar-Poli P, Schultze-Lutter F. Predicting the onset of psychosis in patients at clinical high risk: practical guide to probabilistic prognostic reasoning. Evid Based Ment Health 2016;19:10-5.
41. Cornblatt BA, Carrión RE, Auther A et al. Psychosis prevention: a modified clinical high risk perspective from the Recognition and Prevention (RAP) program. Am J Psychiatry 2015;172:986-94.
42. Simon AE, Velthorst E, Nieman DH et al. Ultra high-risk state for psychosis and non-transition: a systematic review. Schizophr Res 2011;132:8-17.
43. Angkustsiri K, Leckliter I, Tartaglia N et al. An examination of the relationship of anxiety and intelligence to adaptive functioning in children with chromosome 22q11.2 deletion syndrome. J Dev Behav Pediatr 2012;33:713-20.

44. Beaton EA, Simon TJ. How might stress contribute to increased risk for schizophrenia in children with chromosome 22q11.2 deletion syndrome? J Neurodev Disord 2011;3:68-75.
45. Bartels-Velthuis AA, Jenner JA, van de Willige G et al. Prevalence and correlates of auditory vocal hallucinations in middle childhood. Br J Psychiatry 2010;196:41-6.
46. Bartels-Velthuis AA, van de Willige G, Jenner JA et al. Course of auditory vocal hallucinations in childhood: 5-year follow-up study. Br J Psychiatry 2011;199:296-302.
47. Cannon TD, Cadenhead K, Cornblatt B et al. Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry 2008;65:28-37.
48. Seidman LJ, Giuliano AJ, Meyer EC et al. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. Arch Gen Psychiatry 2010; 67:578-88.
49. Thompson A, Nelson B, Yung A. Predictive validity of clinical variables in the "at risk" for psychosis population: international comparison with results from the North American Prodrome Longitudinal Study. Schizophr Res 2011;126:51-7.
50. Yuen T, Chow EWC, Silversides CK et al. Premorbid adjustment and schizophrenia in individuals with 22q11.2 deletion syndrome. Schizophr Res 2013;151:221-5.

DOI:10.1002/wps.20347

"Prolonged grief disorder" and "persistent complex bereavement disorder", but not "complicated grief", are one and the same diagnostic entity: an analysis of data from the Yale Bereavement Study

Paul K. Maciejewski[1-3], Andreas Maercker[4], Paul A. Boelen[5,6], Holly G. Prigerson[1,3]

[1]Center for Research on End-of-Life Care, Weill Cornell Medicine, New York, NY, USA; [2]Department of Radiology, Weill Cornell Medicine, New York, NY, USA; [3]Department of Medicine, Weill Cornell Medicine, New York, NY, USA; [4]Department of Psychology, Division of Psychopathology, University of Zurich, Switzerland; [5]Department of Clinical Psychology, Utrecht University, Utrecht, The Netherlands; [6]Arq Psychotrauma Expert Group, Diemen, The Netherlands

There exists a general consensus that prolonged grief disorder (PGD), or some variant of PGD, represents a distinct mental disorder worthy of diagnosis and treatment. Nevertheless, confusion remains over whether different names and proposed symptom criteria for this disorder identify the same or different diagnostic entities. This study aimed to determine whether PGD, complicated grief (CG), and persistent complex bereavement disorder (PCBD) as described by the DSM-5 are substantively or merely semantically different diagnostic entities. Data were derived from the Yale Bereavement Study, a longitudinal community-based study of bereaved individuals funded by the US National Institute of Mental Health, designed explicitly to evaluate diagnostic criteria for disordered grief. The results suggested that the difference between PGD and PCBD is only semantic. The level of agreement between the original PGD test, a new version of the PGD test proposed for ICD-11 and the PCBD test was high (pairwise kappa coefficients = 0.80-0.84). Their estimates of rate of disorder in this community sample were similarly low (~10%). Their levels of diagnostic specificity were comparably high (95.0-98.3%). Their predictive validity was comparable. In contrast, the test for CG had only moderate agreement with those for PGD and PCBD; its estimate of rate of disorder was three-fold higher (~30%); its diagnostic specificity was poorer, and it had no predictive validity. We conclude that PGD, PCBD and proposed ICD-11, but not CG, symptom-diagnostic tests identify a single diagnostic entity. Ultimately, brief symptom-diagnostic tests, such as the one proposed here for ICD-11, may have the greatest clinical utility.

Key words: Prolonged grief disorder, complicated grief, persistent complex bereavement disorder, DSM-5, ICD-11, diagnostic specificity, predictive validity

(World Psychiatry 2016;15:266–275)

Over the past two decades, there has been increasing awareness and conclusive research demonstrating that prolonged grief disorder (PGD)[1] – intense, prolonged symptoms of grief, coupled with some form of functional impairment beyond 6 months post-loss – constitutes a distinct mental disorder. Numerous studies have demonstrated that symptoms of grief are distinct from symptoms of depression and anxiety[2-7]; that PGD is distinct from other mental disorders, including major depressive disorder, generalized anxiety disorder and post-traumatic stress disorder[1,8,9]; and that PGD, independent of other mental disorders, is associated with significant suffering and enduring functional impairments[1,3,9-11].

In light of extensive, convincing evidence in support of PGD as a new diagnostic category, the World Health Organization (WHO) has moved to introduce PGD, and the American Psychiatric Association has moved to introduce what appears to be a version of PGD, into their respective diagnostic classification systems (ICD-11 and DSM-5). However, despite these advances, and perhaps due to lack of unanimity in terminology and conceptualization of the disorder, there continues to be confusion about PGD and its relationships to normative grief and to other mental disorders.

In recent years, competing characterizations and symptom-diagnostic tests have been proposed for what would appear to be a single disorder of grief. The primary opposition has been between the notion of "prolonged grief disorder", introduced by Prigerson et al[1] and proposed for adoption in a shortened version by the ICD-11[12,13], and the notion of "complicated grief" (CG), which has historical roots in the concept of depression as a bereavement-related "complication"[14] and has been reproposed by Shear et al[15]. Presented with these two main alternatives, the DSM-5[16] introduced yet a third diagnostic concept, i.e. "persistent complex bereavement disorder" (PCBD), that appears a compromise between "prolonged" and "complicated" grief. It is unclear whether DSM-5's PCBD is essentially PGD, CG or another diagnostic entity altogether.

Semantic differences between PGD, CG and PCBD hinge on the response to the central question: "is all grief normal?".

For proponents of PGD, the answer to this question is: "no, not all grief is normal; in particular, prolonged, unresolved, intense grief is not normal". From the PGD perspective, grief symptoms in themselves are neither atypical nor pathological. PGD is characterized by normal symptoms of grief that remain too intense for too long. That is, all symptoms of grief are normal, but some combination of their severity and their duration is not. For PGD, the pathology is in the time course of the symptoms, not in the symptoms *per se*.

For proponents of CG, the answer to the question is: "yes, all grief is normal; but, there are complications (mental disorders) in bereavement aside from grief that merit clinical attention".

From this perspective, neither symptoms nor processes of grief are ever intrinsically pathological[17]. For CG, the pathology is attributed to factors other than grief, e.g. bereavement-related depression or trauma, that interfere with otherwise normal grief processes.

The DSM-5 designation "persistent complex bereavement disorder" omits the term "grief" altogether, which avoids pathologizing any form of grief and thereby leaves safe the assumption that all grief is normal. The assertion that the course of grief, in itself, can be pathological in some instances, i.e., that some grieving processes are inherently abnormal, separates PGD from both CG and PCBD.

Despite semantic differences, the proposed symptom-diagnostic tests for PGD, CG and PCBD may point to a single underlying diagnostic entity. The items included in these tests are almost entirely derived from a common set of instruments, i.e., the Inventory of Complicated Grief (ICG)[18] and its revisions. There is considerable evidence that items in these instruments represent a unidimensional underlying construct[1,18,19]. Nevertheless, the items in each diagnostic formulation constitute unique criteria sets. There may be substantive differences between symptom-diagnostic tests for PGD, CG and PCBD that pose the risk of diluting the assessment of what is, at its core, a pure grief construct.

To date, only the original symptom-diagnostic test for PGD has been validated empirically. In a US National Institute of Mental Health (NIMH)-funded study designed explicitly to evaluate diagnostic criteria for disordered grief, with data collected in a community-based sample (which is essential for distinguishing between normal and pathological grief reactions), Prigerson et al[1] established the construct validity, diagnostic sensitivity and specificity, and predictive validity of diagnostic criteria for PGD.

In contrast, Shear et al introduced the proposed test for CG in a review article[15], without an empirical evaluation or validation. The proposed test for CG, which includes multiple items not included in the ICG, was informed by a post-hoc analysis of ICG data[20] collected in highly comorbid, treatment-seeking, patient samples, ill-suited for drawing distinctions between normal and pathological grief, recruited for studies that were not designed for the purpose of evaluating diagnostic criteria for CG.

The symptom-diagnostic test for PCBD is proposed in an appendix to DSM-5[16]. The proposed ICD-11 characterization of PGD presents its core diagnostic features[13], but the symptoms included in this narrative proposal have yet to be reduced explicitly to a symptom-diagnostic test, i.e., there is no specification of how many of these symptoms need to be present to satisfy the symptom criterion.

In the present investigation, we aimed to compare proposed symptom-diagnostic tests for PGD (both the original version[1] and a new one consistent with the core diagnostic features of PGD as proposed for ICD-11[13]), for CG[15] and for PCBD[16]. We restricted our focus to an examination of tests for meeting the symptom criterion for grief disorder, as opposed to the time from loss and impairment criteria, because of the central role that the symptom criterion plays in the conceptualization, definition and recognition of the disorder.

Given legitimate concerns about pathologizing normal grief reactions, we prioritized diagnostic specificity above diagnostic sensitivity, favoring tests that minimize "false positives" (i.e., normal grief reactions diagnosed as mental illness) and thereby reduce the likelihood of over-diagnosis and over-treatment. Furthermore, since short tests and simple algorithms are preferred in clinical practice[21,22] and lead to higher reliabilities in routine care[23], we considered the brevity and simplicity of each symptom-diagnostic test for grief disorder to be indicative of its potential ease of use and clinical utility.

METHODS

Study sample

Data were obtained from the Yale Bereavement Study (YBS), a NIMH-funded investigation designed to evaluate consensus criteria[24] for disordered grief. The YBS was a longitudinal, interview-based study of community-dwelling bereaved individuals. It was approved by the institutional review boards of all participating sites. Written informed consent was obtained from all study participants. Interviews were conducted by master's degree-level interviewers trained by YBS investigators. Interviewers were required to demonstrate nearly perfect agreement (kappa $\geq$0.90) with the YBS investigators for diagnoses of psychiatric disorders and PGD in five pilot interviews before being permitted to interview for the study. The YBS study is described in greater detail elsewhere[1].

YBS participants (N=317) completed an initial baseline interview at an average of 6.3 $\pm$ 7.0 months post-loss; first follow-up interviews (N=296, 93.4% of participants) at an average of 10.9 $\pm$ 6.1 months post-loss; and second follow-up interviews (N=263, 83.0% of participants) at an average of 19.7 $\pm$ 5.8 months post-loss. For analysis, data were restructured into more uniform time periods (0-6 months, 6-12 months, and 12-24 months post-loss).

The average age of participants was 61.8 $\pm$ 18.7 years. The majority of participants were female (73.7%), white (95.3%), educated beyond high school (60.4%), and spouses of the deceased (83.9%).

The present study sample (N=268; 84.5% of YBS participants) included participants interviewed at least once within 6-12 months post-loss and who provided sufficient information to evaluate PGD, CG, PCBD and proposed ICD-11 tests for grief or bereavement disorder.

Grief symptoms (items)

Grief and bereavement-related symptoms (items) were assessed with the rater-version of the Inventory of Complicated

Grief - Revised (ICG-R)[25], a structured interview designed to assess a wide variety of potential grief and bereavement-related symptoms, using five-point scales to represent increasing levels of symptom severity.

The ICG-R is a modification of the ICG[18] that includes all the symptoms proposed by the consensus panel[24] and additional symptoms enabling the testing of alternative diagnostic algorithms[26].

The ICG-R and the original ICG have proven to be reliable and valid[18,25]. Based on prior work[24,25], a symptom was considered present if rated "4" or "5", and absent if rated "1", "2" or "3" on its five-point scale.

Symptom-diagnostic tests

The focus of the present investigation is restricted to symptom-diagnostic tests for grief disorder (and not other tests or criteria for disorder, e.g. timing or impairment criteria).

Each of the tests under examination has two components, one including items that capture the essence of the syndrome (hereafter, referred to as "category A" items) and another including items that collectively capture the severity of the syndrome (hereafter, referred to as "category B" items).

Each of the tests described below was assessed at 6-12 months post-loss.

Prolonged grief disorder (PGD) test

The PGD symptom-diagnostic test examined here is identical to the one introduced by Prigerson et al[1]. It includes eleven items represented directly in the ICG-R. A positive test indicates endorsement of at least one of two category A items and at least five of nine category B items.

Complicated grief (CG) test

Formally, the proposed CG symptom-diagnostic test[15] consists of twelve (four category A and eight category B) items. However, several of these items contain multiple elements and therefore could be met in multiple ways. For example, the item "experiencing intense emotional or physiological reactivity to memories of the person who died or to reminders of the loss" could be met four ways, yet it is presented as a single item.

Nine of the twelve CG test items can be, and were, represented directly by one or more ICG-R items. Two CG test items, i.e. "troubling rumination" and "emotional or physiological reactivity", can be, and were, approximated by ICG-R items. The CG test "troubling rumination" item (i.e., "frequent troubling rumination about circumstances or consequences of the death, such as concerns about how or why the person died or about not being able to manage without their loved one, thoughts of having let the deceased person down, and others") was approximated by the ICG-R "preoccupation" item (i.e., "do you ever have trouble doing the things you normally do because you are thinking about [the person who died] so much?"). The CG test "emotional or physiological reactivity" item (i.e., "experiencing intense emotional or physiological reactivity to memories of the person who died or to reminders of the loss") was approximated by the ICG-R "memories upset you" item (i.e., "do memories of [the person who died] ever upset you?"). One CG test item contained an element of survivor guilt, which can be, and was, represented directly by the ICG-R "survivor guilt" item, and an element of suicidal ideation, which was represented by a positive screen for suicidal ideation using the Yale Evaluation of Suicidality[27].

Because we decided to use the ICG-R "preoccupation" item to represent the CG test "troubling rumination" item, and to avoid a double counting of this symptom, we chose to count this item only once as "troubling rumination" and not also doubly as "preoccupation". Whether this item was counted as "preoccupation" (in category A) or "troubling rumination" (in category B) had no impact on results of the CG test in the present sample. Therefore, in the present investigation, a positive CG test indicates endorsement of at least one of three category A items (i.e., excluding the fourth, operationally redundant, "preoccupation" item) and at least two of eight category B items.

DSM-5 persistent complex bereavement disorder (PCBD) test

The proposed PCBD symptom-diagnostic test[16] consists of sixteen (four category A and twelve category B) items.

Thirteen of the sixteen PCBD test items can be, and were, represented directly by one or more ICG-R items. Two PCBD test items can be, and were, approximated by ICG-R items. The PCBD test "difficulty in positive reminiscing about the deceased" item was approximated by the ICG-R "do memories of [the person who died] ever upset you?" item. The PCBD test "maladaptive appraisals about oneself in relation to the deceased or the death (e.g., self-blame)" item was approximated by the ICG-R "do you feel at all guilty for surviving, or that it is unfair that you should live when [the person who died] died?" item. One PCBD test item reflects suicidal ideation and was represented by a positive screen for suicidal ideation using the Yale Evaluation of Suicidality.

In the present study, and consistent with the DSM-5 proposal[16], a positive PCBD test indicates endorsement of at least one of four category A items and at least six of twelve category B items.

ICD-11 prolonged grief disorder (ICD-11) test

An "ICD-11 version" of the PGD symptom-diagnostic test was constructed based on a narrative proposal for the diagnostic assessment of PGD for ICD-11[13]. This narrative proposal includes seven (two category A and five category B) items that are represented directly in the ICG-R and that have been found to be informative and unbiased in the empirical evaluation of items presented in Prigerson et al[1].

The proposal did not include specification of a symptom threshold, i.e. a minimum number of items (symptoms) required to satisfy the symptom criterion. Therefore, we conducted a receiver operating characteristic (ROC) analysis[28] to determine an optimum symptom threshold.

Based on the results of this analysis, in the present study, a positive "ICD-11" test indicates endorsement of at least one of two category A items and at least three of five category B items. Presenting with at least three of five category B items was associated with a sensitivity of 83.3% and a specificity of 96.2%. Presenting with at least two of five items yielded lower specificity (sensitivity = 100%, specificity = 87.0%), while presenting with at least four of five items yielded much lower sensitivity (sensitivity = 60.0%, specificity = 99.6%).

Criterion standard to evaluate diagnostic properties of tests

The criterion standard used to establish absence or presence of grief disorder in the present sample is the one developed, employed and described in detail in Prigerson et al[1].

Construction of this criterion standard combined elements of clinical judgment, reflected in raters' diagnoses of disordered grief, with sophisticated measurement techniques. Employing methods from item response theory[29], scores from a two-parameter logistic (2-PL) item response model (IRM) for grief intensity – based on twelve informative unbiased ICG-R items (symptoms) – were used to order individuals based on the severity of their grief symptoms. An optimum minimum symptom severity threshold "cutoff" score, representing a metric boundary between cases and non-cases of disordered grief, was then determined by varying this "cutoff" score to find a point of maximum agreement between rater diagnoses of disordered grief and cases identified by means of grief intensity scores.

Outcomes employed to evaluate predictive validity of positive tests

Potential adverse outcomes following from disordered grief, i.e. subsequent other mental disorders, suicidal ideation, functional impairment, and low quality of life, were each assessed at 12-24 months post-loss.

Mental disorders were assessed using the Structured Clinical Interview for DSM-IV (SCID) Non-Patient Version[30]. They included generalized anxiety disorder, post-traumatic stress disorder and major depressive disorder. Research has supported the reliability and validity of SCID diagnoses[31].

Positive responses to one or more of the four Yale Evaluation of Suicidality screening questions were categorized as having suicidal ideation.

The Established Populations for Epidemiological Studies of the Elderly[32] measured performance of activities of daily living[33] and physical functioning[34]. Individuals with at least "some difficulty" with at least one of the fourteen tasks (e.g., bathing) were considered functionally impaired in order to make the measure sensitive to impairment in a highly functioning sample.

Scores less than 5 (below the lowest quartile) on the Medical Outcomes Short-Form[35] indicated inferior quality of life.

Statistical analysis

Pairwise agreement between tests was assessed and evaluated using kappa statistics[36,37]. The diagnostic sensitivity and specificity of each test was evaluated in relation to the criterion standard. The predictive validity of each symptom-diagnostic test (evaluated between 6 to 12 months post-loss) was examined using logistic regression models for the examined outcomes (evaluated between 12 to 24 months post-loss) within strata defined by the absence/presence of other mental disorders at the time of the test. Suicidal ideation was not considered to be a potential outcome for either the CG or PCBD tests, because each of these tests included suicidal ideation as an item.

RESULTS

Table 1 presents the items employed in each test. Of the combined total of twenty items, the PGD test uses eleven, the CG test eighteen, the PCBD test fifteen, and the ICD-11 test seven.

CG employs two items previously reported to be biased[1]: loneliness (reported to be biased with respect to gender, relationship to diseased, and time from loss) and inability to care (reported to be biased with respect to relationship to diseased). It also uses three items (envy, upsetting memories, and drawn to places) previously reported to be uninformative[1], and one item (suicidal ideation) that might be characterized as a correlate or consequence of prolonged, intense grief rather than a symptom of grief.

PCBD employs one reportedly biased item (loneliness)[1], one reportedly uninformative item (upsetting memories)[1], and one item (suicidal ideation) better characterized as a correlate or consequence of prolonged, intense grief[1].

Neither PGD nor ICD-11 employs reportedly biased or uninformative items, and neither employs correlates or consequences of prolonged, intense grief as items.

Figure 1 displays the point prevalence rate of disorder at 6-12 months post-loss for each test. The prevalence rates for PGD, CG, PCBD and ICD-11 were, respectively, 11.9% (95% CI: 8.1%-15.8%), 30.2% (95% CI: 24.7%-35.7%), 14.2% (95% CI: 10.0%-18.4%), and 12.7% (95% CI: 8.7%-16.7%). There were no statistically significant pairwise differences in prevalence rates between PGD, PCBD and ICD-11 diagnoses (all pairwise p values >0.05), while the prevalence rate for CG diagnosis was significantly higher than those for PGD, PCBD and ICD-11 diagnoses (all pairwise p values <0.001).

Table 1 Items included in symptom-diagnostic tests

	PGD		CG		PCBD		ICD-11	
Symptom (Item)	**Category**	**Item**	**Category**	**Item**	**Category**	**Item**	**Category**	**Item**
Yearning	A	1	A	1	A	1, 2	A	1
Preoccupation	A	2	B	1	A	3, 4	A	2
Part of yourself died	B	1			B	11	B	2
Disbelief; Trouble accepting death	B	2	B	2	B	1	B	1
Avoidance of reminders	B	3	B	8	B	6		
Hard to trust others	B	4	B	5	B	8		
Anger; Bitterness	B	5	B	4	B	4	B	3
Difficulty moving on	B	6			B	12	B	5
Numbness	B	7	B	3	B	2		
Life empty, meaningless, unfulfilling	B	8	A	2	B	10		
Stunned	B	9	B	3	B	2		
Loneliness			A	2	B	9		
Survivor guilt			A	3	B	5	B	4
Suicidal ideation			A	3	B	7		
Inability to care			B	5				
Envious of others without loss			B	5				
Symptoms of deceased			B	6				
Hear or see deceased			B	6				
Memories upset you			B	7	B	3		
Drawn to places			B	8				

PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

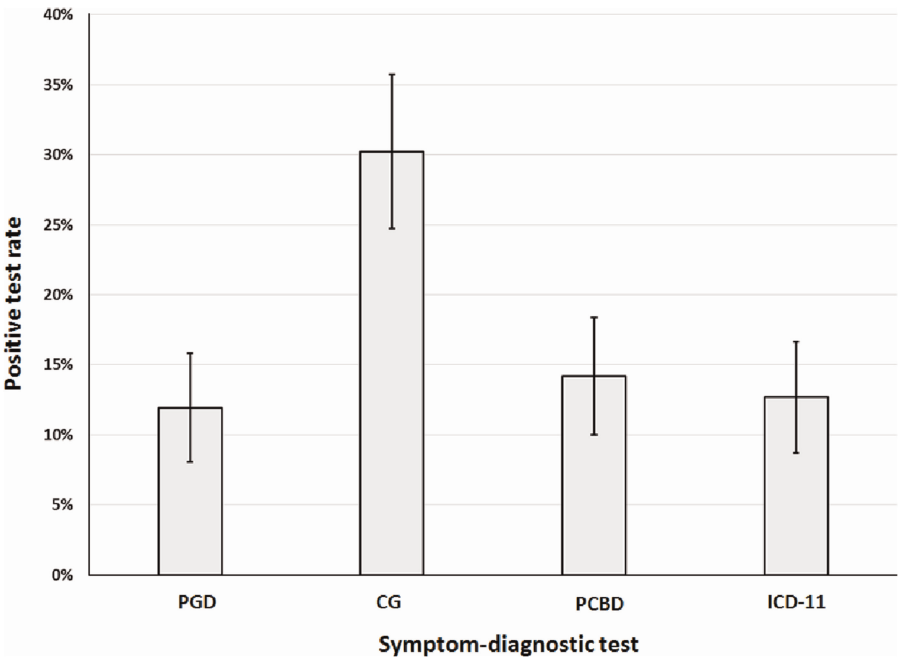

Figure 1 Positive symptom-diagnostic test rates (N=268). PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

Table 2 Pairwise agreement (kappa) between symptom-diagnostic tests (N=268)

Test	PGD	CG	PCBD	ICD-11
PGD	**1.00**			
CG	0.48	**1.00**		
PCBD	**0.80**	0.55	**1.00**	
ICD-11	**0.83**	0.50	**0.84**	**1.00**

PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

Kappa values indicating almost perfect agreement are highlighted in bold prints

Table 2 presents pairwise agreement between the four tests. The PGD, PCBD and ICD-11 tests were in almost perfect agreement with each other (with pairwise kappa ranging from 0.80 to 0.84). The CG test was in moderate agreement with each of the other tests (with pairwise kappa ranging from 0.48 to 0.55).

Table 3 displays properties of each test, and in particular each test's diagnostic specificity, in relation to the criterion standard. The PGD, PCBD and ICD-11 tests had high and comparable diagnostic specificity, with values of 98.3%, 95.0%, and 96.2%, respectively. The CG test had 78.6% diagnostic specificity. The positive predictive value of the CG test was 37.0%, considerably lower than those for the PGD (87.5%), PCBD (68.4%), and ICD-11 (73.5%) tests. Figure 2 highlights the tradeoff between diagnostic sensitivity and diagnostic specificity for each of the four tests.

Tables 4 and 5 present an examination of the predictive validity of each of the four tests in terms of four subsequent (12-24 months post-loss) adverse outcomes, i.e., other mental disorders (major depressive disorder, post-traumatic stress disorder or generalized anxiety disorder), suicidal ideation, functional impairment, and low quality of life, stratified by absence/presence of concurrent (6-12 month post-loss) mental disorders.

Among individuals without other mental disorders at 6-12 months post-loss (Table 4), positive PGD tests were significantly associated with other mental disorders (RR=4.40, p=0.048), suicidal ideation (RR=3.06, p=0.017), functional impairment (RR=2.08, p<0.001), and low quality of life (RR=3.40, p<0.001) at 12-24 months post-loss. Positive PCBD tests were associated with low quality of life (RR=2.68, p=0.006) at 12-24 months post-loss; and positive ICD-11 tests were associated with suicidal ideation (RR=5.04, p<0.001), functional impairment (RR=2.07, p<0.001), and low quality of life (RR=3.23, p<0.001) at 12-24 months post-loss.

Among individuals with other mental disorders at 6-12 months post-loss (Table 5), positive PGD and ICD-11 tests were each significantly related to other mental disorders (PGD: RR=4.00, p=0.0.039; ICD-11: RR=4.64, p=0.022) at 12-24 months post-loss.

Positive CG tests were not significantly associated with other mental disorders, functional impairment and low quality of life at 12-24 months post-loss, either in the absence (Table 4) or in the presence (Table 5) of concurrent (6-12 months post-loss) mental disorders.

DISCUSSION

This study aimed to determine whether the differences between PGD, CG and PCBD are substantive or merely semantic. Our results indicate that there is no substantive difference between PGD and PCBD. The high level of agreement between the PGD, PCBD and proposed ICD-11 tests; their similarly low estimates of rate of disorder (~10%) in this community population; their comparably high levels of diagnostic specificity, and their comparable predictive validity, all suggest that PGD and PCBD identify the same diagnostic entity. Therefore, the difference between PGD and PCBD is mainly semantic. In contrast, the CG test had only moderate agreement with the PGD, PCBD and proposed ICD-11 tests, a three-fold higher estimate of rate of disorder (~30%) in this community sample, much poorer diagnostic specificity, and no predictive validity. Therefore, the difference between PGD and PCBD on the one hand, and CG on the other, is substantive.

Given that PGD and PCBD tests identify the same diagnostic entity, the main difference between PGD (proposed for adoption in ICD-11) and PCBD (introduced in DSM-5) is in the meaning of terms used to describe this same entity. The primary opposition is between use of the term "grief" and use

Table 3 Diagnostic sensitivity and specificity of the tests in relation to the criterion standard (N=268)

Test	True positive	False positive	True negative	False negative	Positive predictive value	Negative predictive value	Sensitivity	Specificity
PGD	28	4	234	2	87.5%	99.2%	93.3%	98.3%
CG	30	51	187	0	37.0%	100.0%	100.0%	78.6%
PCBD	26	12	226	4	68.4%	98.3%	86.7%	95.0%
ICD-11	25	9	229	5	73.5%	97.9%	83.3%	96.2%

PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

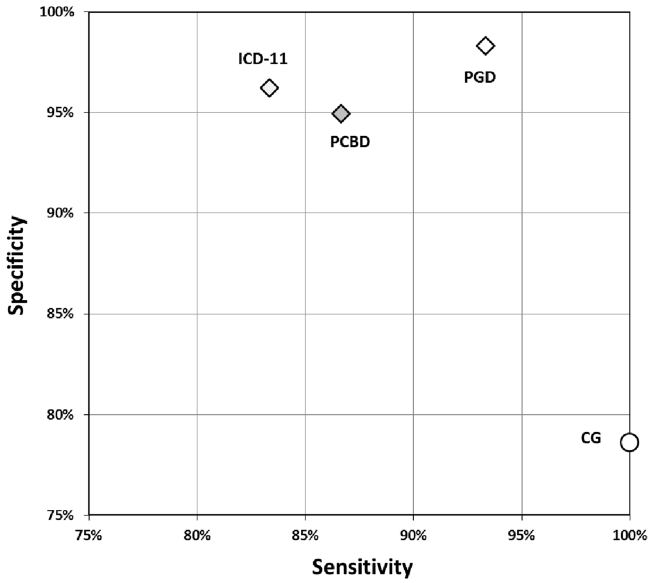

Figure 2 Symptom-diagnostic test specificity versus sensitivity (N=268). PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

of the term "bereavement" in the name of the disorder. Grief is deep mental anguish, a process of the psyche. Bereavement is an event, the loss of a valued loved-one due to death. Grief is a mental entity; bereavement is not. At face value, there is no mental entity identified in the name PCBD. How can the name of a mental disorder not identify a mental entity?

The use of the term "complex" in the name PCBD is also somewhat obfuscating. The PGD construct is fairly simple to understand: individuals who are "stuck" in a state of intense grief for a long time have PGD. If the underlying disorder is not difficult to understand, then what is "complex" about PCBD? The name PCBD has no clear meaning and should be abandoned by the DSM in favor of PGD. Even if the DSM retains this name, researchers, clinicians and the general public should understand that there is no substantive difference between what the DSM calls PCBD and what the ICD calls PGD.

Disagreement between the CG test on the one hand, and the PGD and PCBD tests on the other, combined with the CG test's limited specificity (78.6%), poor positive predictive value (only 37.0%), and lack of predictive validity, suggest that the CG test is not a valid indicator of a grief-specific disorder. Indeed, in the current study sample, a majority of individuals with positive CG tests had negative PGD (original version), PCBD, and PGD (version proposed for ICD-11) tests. For this reason, treatment studies based on samples defined in terms of the CG may be of questionable value for a grief-specific disorder devoid of the CG "contaminants".

The fact that one test includes a different set of items than another test does not necessarily imply that the two tests are grounded in different constructs or identify different disorders. Tests for PGD (both the original version and the one proposed for ICD-11) and PCBD are different but essentially equivalent measures of a single, underlying attribute, i.e. intense grief, and should be viewed as such. The notion that symptoms of grief are normal but that a combination of their high intensity and long duration is abnormal reconciles the belief that all grief symptoms are normal, but not all grieving processes are normal. This view, rooted in the uni-dimensionality of the underlying grief construct, is in opposition to the notion that some symptoms are normal and others are atypical and abnormal, i.e., that pathology is expressed in the form of atypical symptoms. Current and future alternative symptom-diagnostic tests should be evaluated in terms of specificity, accuracy, parsimony, and perhaps in reference to external validity; not in terms of whether or not individual items on the test define the pathology.

Inclusion of biased items and external correlates of PGD (e.g., suicidal thoughts) in a criteria set for grief disorder is questionable on psychometric and conceptual grounds. The tests for CG and PCBD contain items that were previously

Table 4 Predictive validity of symptom-diagnostic tests in the absence of other mental disorders (N=213)

Test (6-12 months post-loss)	Outcome (12-24 months post-loss)							
	Other mental disorders		Suicidal ideation		Functional impairment		Low quality of life	
	RR	p	RR	p	RR	p	RR	p
PGD	**4.40**	**0.048**	**3.06**	**0.017**	**2.08**	**0.001**	**3.40**	**0.001**
CG	2.90	0.101	–	–	0.98	0.926	1.08	0.834
PCBD	3.52	0.097	–	–	1.61	0.058	**2.68**	**0.006**
ICD-11	3.52	0.097	**5.04**	**0.001**	**2.07**	**0.001**	**3.23**	**0.001**

PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

Other mental disorders considered were major depressive disorder, post-traumatic stress disorder and generalized anxiety disorder

Suicide ideation is not considered as a potential outcome for CG and PCBD, because they include suicidal ideation as an item

Statistically significant values are highlighted in bold prints

Table 5 Predictive validity of symptom-diagnostic tests in the presence of other mental disorders (N=27)

Test (6-12 months post-loss)	Outcome (12-24 months post-loss)							
	Other mental disorders		Suicidal ideation		Functional impairment		Low quality of life	
	RR	p	RR	p	RR	p	RR	p
PGD	**4.00**	**0.039**	2.00	0.121	0.80	0.480	1.03	0.930
CG	3.14	0.221	–	–	0.86	0.655	0.86	0.655
PCBD	3.44	0.065	–	–	0.69	0.228	0.88	0.697
ICD-11	**4.64**	**0.022**	1.67	0.203	0.93	0.816	1.19	0.586

PGD – prolonged grief disorder test (original version), CG – complicated grief test, PCBD – persistent complex bereavement disorder test, ICD-11 – prolonged grief disorder test (ICD-11 proposed version)

Other mental disorders considered were major depressive disorder, post-traumatic stress disorder and generalized anxiety disorder

Suicide ideation is not considered as a potential outcome for CG and PCBD, because they include suicidal ideation as an item

Statistically significant values are highlighted in bold prints

identified to be biased[1]. In particular, the loneliness item included in both of these tests has been reported to be biased not only with respect to the bereaved individual's gender and relationship to the deceased, but also with respect to time from loss. Although inclusion of one or even a few biased items in a multi-item test does not necessarily mean that the test as a whole is biased, inclusion of biased items opens the possibility that some groups of individuals may be misdiagnosed by the test due to misinterpretation of the severity of their symptoms. For example, for bereaved spouses, loneliness is a moderate symptom, whereas for bereaved non-spouses, loneliness is a significantly more severe symptom of grief. Inclusion of the loneliness item in a diagnostic test for a disorder of grief makes it more likely that a bereaved spouse would be mistakenly diagnosed with that disorder due to a misinterpretation of the severity of his/her loneliness symptom. The tests for CG and PCBD also include an external correlate or consequence of PGD, i.e. suicidal thoughts, as an item. Suicidality may be related to grief disorder, but to include it as a symptom that represents grief is to misunderstand what grief is, and to confound the essence of the syndrome with its consequences.

In order to include an "ICD-11 version" of a symptom-diagnostic test for PGD in the present analysis, we needed to specify a symptom threshold. The current narrative proposal[13] for an ICD-11 version of PGD does not make this specification. In an effort to develop diagnostic guidelines that accommodate flexible exercise of clinical judgment, the WHO discourages methods of diagnostic assessment that employ arbitrary thresholds and "pick lists" of items, but supports the use of symptom thresholds that have been established empirically[38]. Based on results of the ROC analysis in this study, presenting with at least three of the proposed five accessory symptoms represents an optimum balance of diagnostic sensitivity and specificity in relation to our criterion standard. For this reason we recommend that future ICD-11 research diagnostic criteria include this "at least three of five" accessory symptom rule for diagnosing cases of PGD. The brief, five-item "ICD-11 version" of the PGD test also has the advantage that short tests have over longer ones for ease of use and clinical utility[21-23].

The present study evaluates the performance of symptom-diagnostic tests for grief disorder applied within a period of 6 to 12 months post-loss. This is consistent with empirical evidence that presence of enduring, intense grief beyond 6 months post-loss is predictive of subsequent mental disorders, suicidal ideation, functional impairment, and worse quality of life[1], and with proposed diagnostic criteria for PGD[1,13] and CG[15]. However, it is inconsistent with the DSM-5 specification that PCBD ought not to be diagnosed within 12 months post-loss. In our view, this DSM-5 "time from loss" criterion is not only arbitrary but also contrary to published empirical research findings. In the present study, the PCBD test applied within 6 to 12 months post-loss had near perfect agreement with PGD tests, had high specificity and sensitivity with respect to our criterion standard, and was predictive of subsequent (i.e., 12 to 24 month) worse quality of life. Based on these findings, the PCBD symptom-diagnostic test applied within 6 to 12 months post-loss is an empirically valid test for disorder notwithstanding the DSM-5's arbitrary "at least 12 months' time from loss" criterion for PCBD.

The present investigation has a few limitations that warrant some consideration. One limitation is that some ICG-R items employed in the present analysis may not have mapped exactly onto some items in the proposed CG and PCBD tests. More formal instruments to assess CG and PCBD have been introduced only recently[39,40]. These have yet to be established and validated in general community settings. The fact that our proxy PCBD symptom-diagnostic test had high diagnostic specificity and sensitivity, as well as some predictive validity, suggests that some imprecision in our representation of some PCBD items did not undermine the validity of the overall PCBD test appreciably. Given the properties of the PGD and PCBD

tests, future refinements in conceptualization and wording of items might be expected to make marginal improvements in what are already highly reliable and valid tests.

Another limitation of the present study is that the YBS sample represents mainly elderly, white widows living in a relatively small region of the US, whose spouses died primarily from natural causes. Future studies ought to examine whether and the extent to which properties of PGD and PGBD tests and items differ with respect to the bereaved individual's age, gender, race, ethnicity, relationship to the deceased, and geographic or cultural setting, as well as with respect to circumstances of the lost loved-one's death.

The present study has a number of strengths. Most importantly, the YBS was designed explicitly to evaluate diagnostic criteria for disordered grief. YBS instrumentation included an extensive battery of grief items sufficient to compare the four symptom-diagnostic tests included in the present analysis. YBS data were collected in a community sample, allowing us to evaluate methods of diagnostic assessment that are intended to discriminate between normal and disordered grief. Finally, the YBS's longitudinal design allowed us to examine the predictive validity of positive symptom-diagnostic tests for disordered grief.

In conclusion, the PGD, PCBD and proposed ICD-11 PGD symptom-diagnostic tests identify a single, common diagnostic entity. Therefore, the main differences between PGD and PCBD are semantic, not substantive. The test for CG is incongruous with those for PGD and PCBD, has a poorer diagnostic specificity and no predictive validity. Clinical and scientific communities ought to recognize that PGD and PCDB are substantively the same disorder, and ought to work toward a common understanding of that disorder and adopt useful ways to recognize it clinically. The term "prolonged grief disorder" captures the essence of the disorder, facilitates understanding it, and thereby supports clinical judgment in its diagnostic assessment.

ACKNOWLEDGEMENTS

P.K. Maciejewski and H.G. Prigerson were supported by research grants MD007652 from the National Institute of Minority Health and Health Disparities, MH56529 from the National Institute of Mental Health, and CA106370 and CA197730 from the National Cancer Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. A. Maercker is chair of the ICD-11 Working Group on Disorders Specifically Associated with Stress that developed the PGD proposal for the WHO. Unless specifically stated, the views expressed in this paper are those of this author and do not represent the official policies or positions of the WHO.

REFERENCES

1. Prigerson HG, Horowitz MJ, Jacobs SC et al. Prolonged grief disorder: psychometric validation of criteria proposed for DSM-V and ICD-11. PLoS Med 2009;6:e1000121.
2. Prigerson HG, Frank E, Kasl SV et al. Complicated grief and bereavement-related depression as distinct disorders: preliminary empirical validation in elderly bereaved spouses. Am J Psychiatry 1995;152:22-30.
3. Prigerson HG, Bierhals AJ, Kasl SV et al. Complicated grief as a disorder distinct from bereavement-related depression and anxiety: a replication study. Am J Psychiatry 1996;153:1484-6.
4. Boelen PA, van den Bout J, de Keijser J. Traumatic Grief as a disorder distinct from bereavement-related depression and anxiety: a replication study with bereaved mental health care patients. Am J Psychiatry 2003;160:1339-41.
5. Boelen PA, van den Bout J. Complicated grief, depression, and anxiety as distinct postloss syndromes: a confirmatory factor analysis study. Am J Psychiatry 2005;162:2175-7.
6. Golden AM, Dalgleish T. Is prolonged grief distinct from bereavement-related posttraumatic stress? Psychiatry Res 2010;178:336-41.
7. Spuij M, Reitz E, Prinzie P et al. Distinctiveness of symptoms of prolonged grief, depression, and post-traumatic stress in bereaved children and adolescents. Eur Child Adolesc Psychiatry 2012;21:673-9.
8. Horowitz MJ, Siegel B, Holen A et al. Diagnostic criteria for complicated grief disorder. Am J Psychiatry 1997;154:904-10.
9. Silverman GK, Jacobs SC, Kasl SV et al. Quality of life impairments associated with diagnostic criteria for traumatic grief. Psychol Med 2000;30:857-62.
10. Prigerson HG, Bierhals AJ, Kasl SV et al. Traumatic grief as a risk factor for mental and physical morbidity. Am J Psychiatry 1997;154:616-23.
11. Boelen PA, Prigerson HG. The influence of symptoms of prolonged grief disorder, depression, and anxiety on quality of life among bereaved adults. Eur Arch Psychiatry Clin Neurosci 2007;257:444-52.
12. Maercker A, Brewin CR, Bryant RA et al. Proposals for mental disorders specifically associated with stress in the International Classification of Diseases-11. Lancet 2013;381:1683-5.
13. 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.
14. Zisook S, Shuchter SRX. Uncomplicated bereavement. J Clin Psychiatry 1993;54:365-72.
15. Shear MK, Simon N, Wall M et al. Complicated grief and related bereavement issues for DSM-5. Depress Anxiety 2011;28:103-17.
16. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington: American Psychiatric Publishing, 2015.
17. Zisook S, Pies R, Corruble E. When is grief a disease? Lancet 2012;379:1590.
18. Prigerson HG, Maciejewski PK, Reynolds CF et al. Inventory of Complicated Grief: a scale to measure maladaptive symptoms of loss. Psychiatry Res 1995;59:65-79.
19. Boelen PA, van den Bout J, de Keijser J et al. Reliability and validity of the Dutch version of the Inventory of Traumatic Grief. Death Stud 2003;27:227-47.
20. Simon NM, Wall MM, Keshaviah A et al. Informing the symptom profile of complicated grief. Depress Anxiety 2011;28:118-26.
21. 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 Psychol 2013;48:177-93.
22. 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.
23. First MB, Pincus HA, Levine JB et al. Clinical utility as a criterion for revising psychiatric diagnoses. Am J Psychiatry 2004;161:946-54.
24. Prigerson HG, Shear MK, Jacobs SC et al. Consensus criteria for traumatic grief. A preliminary empirical test. Br J Psychiatry 1999;174:67-73.
25. Prigerson HG, Jacobs S. Traumatic grief as a distinct disorder: a rationale, consensus criteria, and a preliminary empirical test. In: Stroebe MS, Hansson RO, Stroebe W et al (eds). Handbook of bereavement research: consequences, coping, and care. Washington: American Psychological Association, 2001:613-45.
26. Horowitz MJ, Siegel B, Holen A et al. Diagnostic criteria for complicated grief disorder. Am J Psychiatry 1997;154:904-10.
27. Latham AE, Prigerson HG. Suicidality and bereavement: complicated grief as psychiatric disorder presenting greatest risk for suicidality. Suicide Life Threat Behav 2004;34:350-62.
28. Metz CE. Basic principles of ROC analysis. Semin Nucl Med 1978;8:283-98.
29. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of items response theory. Newbury Park: Sage, 1991.
30. First MB, Spitzer RL, Gibbon M et al Structured Clinical Interview for DSM-IV Axis I Disorders, Non-Patient Version (SCID-I/NP). New York: Biometrics Research Department, New York State Psychiatric Institute, 1995.
31. Williams JBW, Gibbon M, First MB et al. The Structured Clinical Interview for DSM-III-R (SCID): II. Multisite test-retest reliability. Arch Gen Psychiatry 1992;49:630-6.
32. Cornoni-Huntley J, Ostfeld AM, Taylor JO et al. Established populations for epidemiologic studies of the elderly: study design and methodology. Aging Clin Exp Res 1993;5:27-37.
33. Katz S, Downs TD, Cash HR et al. Progress in the development of an index of ADL. Gerontologist 1970;10:20-30.

34. Nagi SZ. An epidemiology of disability among adults in the United States. Milbank Mem Fund Q 1976;54:439-67.
35. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. Med Care 1992;30: 473-82.
36. Cohen J. A coefficient of agreement for nominal scales. Ed Psychol Meas 1960;20:37-46.
37. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
38. 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.
39. Bui E, Mauro C, Robinaugh DJ et al. The Structured Clinical Interview for Complicated Grief: reliability, validity, and exploratory factor analysis. Depress Anxiety 2015;32:485-92.
40. Lee SA. The Persistent Complex Bereavement Inventory: a measure based on the DSM-5. Death Stud 2015;39:399-410.

DOI:10.1002/wps.20348

Community mental health care worldwide: current status and further developments

Graham Thornicroft, Tanya Deb, Claire Henderson

Centre for Global Mental Health, King's College London, Institute of Psychiatry, Psychology & Neuroscience, London SE5 8AF, UK

This paper aims to give an overview of the key issues facing those who are in a position to influence the planning and provision of mental health systems, and who need to address questions of which staff, services and sectors to invest in, and for which patients. The paper considers in turn: a) definitions of community mental health care; b) a conceptual framework to use when evaluating the need for hospital and community mental health care; c) the potential for wider platforms, outside the health service, for mental health improvement, including schools and the workplace; d) data on how far community mental health services have been developed across different regions of the world; e) the need to develop in more detail models of community mental health services for low- and middle-income countries which are directly based upon evidence for those countries; f) how to incorporate mental health practice within integrated models to identify and treat people with comorbid long-term conditions; g) possible adverse effects of deinstitutionalization. We then present a series of ten recommendations for the future strengthening of health systems to support and treat people with mental illness.

Key words: Community mental health care, mental health services, low- and middle-income countries, evidence-based interventions, schools, workplace, chronic care model, deinstitutionalization

(World Psychiatry 2016;15:276–286)

Most people in the world who have mental illnesses receive no treatment[1,2]. This "treatment gap" is increasingly appreciated worldwide[3-6]. The World Health Organization (WHO) published in 2010 the first edition of its Mental Health Gap Action Programme (mhGAP) Implementation Guide[7-9], which contains case finding and treatment guidelines for nine categories of mental and neurological disorders that have a major global public health impact.

This evidence-based approach is now being put into practice in over 90 countries worldwide. But what pattern of services and what systems of care best support the provision of the quality and quantity of treatment and care required for people with mental illnesses in the different scenarios (not only high- vs. low- and middle-income countries, but also high- vs. low-resource areas within countries)? That question is addressed in this paper, which focuses on the current status and new developments of community mental health care worldwide.

DEFINING COMMUNITY MENTAL HEALTH CARE

Our definition of community mental health care highlights several fundamental issues.

First, community mental health care encompasses: a) a population approach, b) viewing patients in a socio-economic context, c) individual as well as population-based prevention, d) a systemic view of service provision, e) open access to services, f) team-based services, g) a long-term, longitudinal, life-course perspective, and h) cost-effectiveness in population terms[10]. It also includes a commitment to social justice by addressing the needs of traditionally underserved populations, such as ethnic minorities, homeless persons, children and adolescents, and immigrants, and to provision of services where those in need are located and in a fashion that is acceptable as well as accessible[11].

Second, community mental health care focuses not only upon people's deficits and disabilities (an illness perspective), but also upon their strengths, capacities and aspirations (a recovery perspective). Services and supports thus aim to enhance a person's ability to develop a positive identity, to frame the illness experience, to self-manage the illness, and to pursue personally valued social roles[12].

Third, community mental health care includes the community in a broadly defined sense. As a corollary of the second point, it emphasizes not just the reduction or management of environmental adversity, but also the strengths of the families, social networks, communities and organizations that surround people who experience mental illnesses[13].

Fourth, community mental health care melds evidence-based medicine and practical ethics. A scientific approach to services prioritizes using the best available data on the effectiveness of interventions. At the same time, people who experience mental illnesses have the right to understand their illnesses (to the extent that professionals understand them), to consider the available options for interventions and whatever information is available on their effectiveness and side effects, and to have their preferences included in a process of shared decision making[14,15].

Thus, we define community mental health care as comprising the principles and practices needed to promote mental health for a local population by: a) addressing population needs in ways that are accessible and acceptable; b) building on the goals and strengths of people who experience mental illnesses; c) promoting a wide network of supports, services and resources of adequate capacity; and d) emphasizing services that are both evidence-based and recovery-oriented[16].

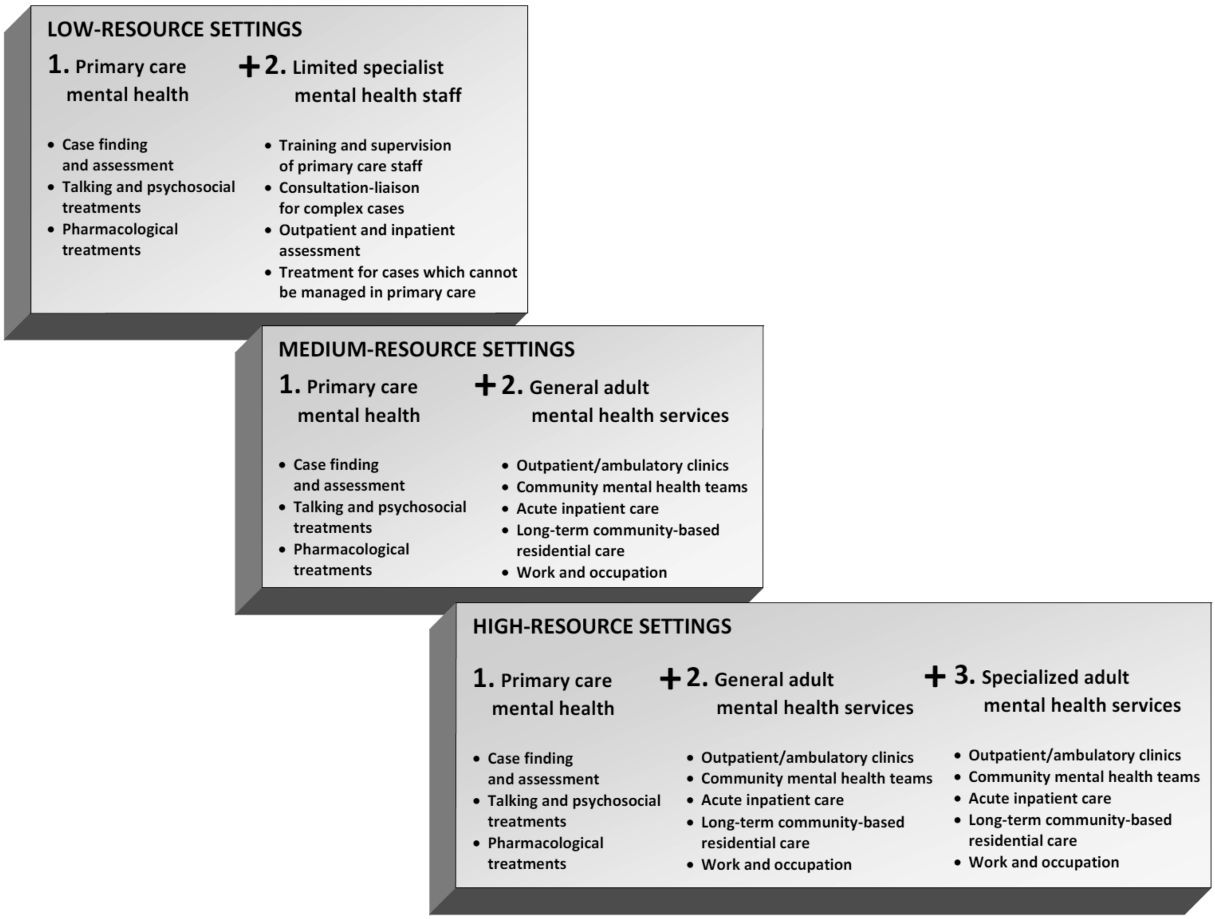

Figure 1 Balanced care model in relation to high-, middle- and low-income countries

A CONCEPTUAL FRAMEWORK FOR COMMUNITY MENTAL HEALTH CARE

The available evidence[17-20] suggests that a model of care including solely hospital based provision (usually inpatient and outpatient care) will be insufficient to provide access for people facing barriers to care, and to offer continuity of follow-up for those with longer-term disability. At the same time, there is not strong evidence that community-based services alone can offer the brief spells of intense treatment sometimes needed during mental health crises. The balanced care model has been formulated as a conceptual framework for providing both hospital and community based services[18].

Yet, it is clear that high-income countries have about 200 times more financial resources for their mental health services than do low-income countries[21]. Many low-income countries in sub-Saharan Africa, for example, have only about one psychiatrist for every million people (Chad, Eritrea and Liberia each have only one psychiatrist in the entire country), compared with 137 per million in the US[22]. So, a single global model of care simply cannot apply. The balanced care model, therefore, applies somewhat differently to countries which are classified by the World Bank Group[23] as high-, middle- or low-income countries (see Figure 1) and, if utilized, needs to be carefully considered for minor or major adaptation in any particular site or country.

The balanced care model suggests that, in low-income countries or sites, most of the available mental health provision should be invested in staff for primary health care and community settings[24]. The roles of these staff include case finding and assessment, brief talking and psychosocial treatments, and pharmacological treatments[25,26]. The very limited numbers of specialist mental health care staff (usually in the capital city and sometimes also in regional centers) are only able to provide training and supervision of primary care staff, consultation-liaison for complex cases, and outpatient and inpatient assessment and treatment for cases which cannot be managed in primary care[27,28].

In middle-income settings, the balanced care model indicates including as investment priorities, in addition to a continuing emphasis upon primary care, five key elements of general adult mental health services: a) outpatient/ambulatory clinics[29]; b) community mental health teams[30-33]; c) acute inpatient care, even though there continues to be relatively

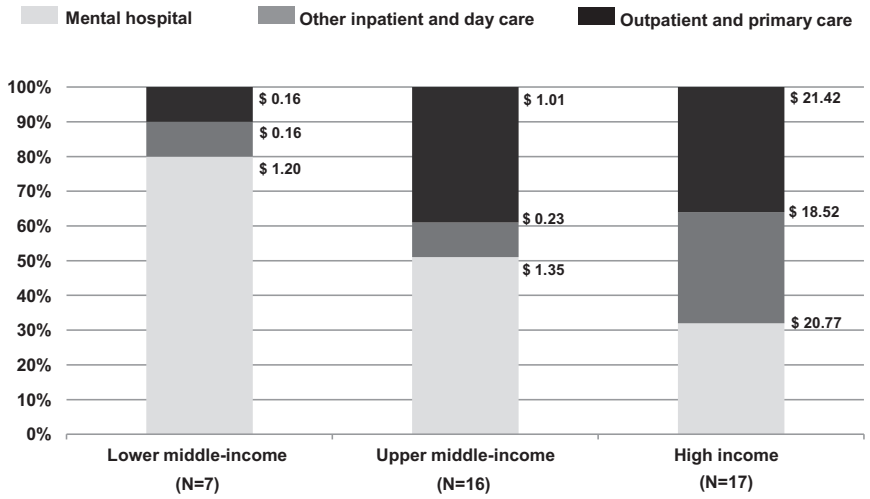

Figure 2 Global distribution of mental health expenditure per capita, by health setting (data from the WHO Atlas[21])

weak evidence about several aspects of inpatient care or highly supported alternative settings[34-38]; d) long-term community-based residential care, with an appropriate range of support[16]; and e) options for work and occupation[39].

In high-income settings, in addition to primary care services and to the provision of general adult mental health services, the balanced care model implies that a series of specialized services should be provided, as resources allow (see Figure 1). These services will need to be provided in the same five categories as set out for middle-income countries.

COMMUNITY SERVICES PROVIDED ACROSS WIDER PLATFORMS OF CARE

Some interventions intended to improve mental health may be better provided from outside health services. The recent edition of the Disease Control Priorities Manual (DSP-3) sets out the arguments for this case[40]. The bases for these wider types of intervention are sometimes called "platforms", and two are particularly relevant here: population-level and community-level platforms. Examples of the former include legislation, regulations, and public information campaigns, and examples of the latter include schools, workplaces, and neighborhoods/community groups[41,42]. A recent review[42], based upon the best available evidence from low- and middle-income settings, has shown which such interventions are most cost-effective.

At the population level, interventions which are evidence based include: laws and regulations to reduce demand for alcohol use (enforcement of blood alcohol limits for drivers, alcohol taxation, advertising bans, minimum drinking age[43,44]); laws and regulations to restrict access to means of self-harm/suicide[45]; child protection laws[46]; laws promoting conditional cash transfers in order to alleviate poverty[47]; and mass public awareness campaigns[48-51].

At the community level, interventions of known effectiveness include: integrating mental health promotion strategies (e.g., stress reduction and awareness of alcohol and drug misuse) into occupational health and safety policies[52]; universal and targeted socio-emotional learning school programs for vulnerable children[50,53]; mental health awareness school programs[54,55]; methods for the identification and case detection of children with mental disorders in schools[56]; early child enrichment/preschool educational programs[57]; parenting programs for children aged 2-14 years[58]; gender equity and/or economic empowerment programs for vulnerable groups[59]; and training of gatekeepers (including community health workers, police and social workers) in identification of young people with mental disorders, including self-harm[60].

THE EXTENT OF COMMUNITY MENTAL HEALTH SERVICE DEVELOPMENT

There is a vast variability worldwide in the development of community mental health services[61]. The most comprehensive global source of information in this respect is the WHO World Mental Health Atlas[21], which summarizes the key characteristics of national mental health systems across the world, and is periodically updated. The most recent edition (2014) includes data from 171 of the 194 member states of the United Nations.

Figure 2 shows the proportional expenditure for mental hospital, other inpatient and day care, and outpatient and primary care services, across lower middle-income, upper middle-income and high-income countries. This clearly illustrates the very large differences in absolute spending, and also the differing relative expenditure across the three service categories, reinforcing the point that relatively little of the small mental health budgets in low- and middle-income countries is spent outside inpatient care[21].

Several important trends emerge from the WHO Atlas. Compared with the results from the 2011 survey, globally, there was a slight decrease (5%) in the number of mental hospitals, and a larger reduction in the number of mental hospital beds, which fell by nearly 30%, with a more substantial decrease (45%) in the Region of the Americas. At the same time, there was an increase of over 20% in the rate of admissions to mental hospitals, indicating an increasing bed turnover rate and decreasing average length of stay[21].

At the global level, the number of beds available in psychiatric wards in general hospitals increased by 60% between 2011 and 2014. In the Western Pacific Region, in particular, psychiatric beds in general hospitals increased more than 8-fold since 2011.

The WHO Atlas does not contain data allowing conclusions on whether reducing number of beds in psychiatric hospitals is associated with greater expenditure on community services.

DEVELOPING COMMUNITY MENTAL HEALTH MODELS IN LOW- AND MIDDLE-INCOME COUNTRIES

The work of the WPA Task Force on the Steps, Obstacles and Mistakes to Avoid in the Implementation of Community Mental Health Care reveals more detailed patterns in the development of community mental health services in recent years[11,62]. This work combined a review of the relevant literature with detailed consultation processes in many regions of the world to identify challenges and solutions in implementing community based models of mental health care. A series of regional papers describe the findings in detail[63-68]. Table 1 summarizes the main challenges which were identified and gives examples of approaches through which progress has sometimes been made.

The continuing lack of trained mental health practitioners is a substantial issue that affects most countries of the world[21]. In response to this, alternative approaches have been implemented which allocate duties previously reserved for psychiatrists or psychiatric nurses to non-specialized staff. This redistribution of clinical tasks is usually referred to as task shifting or task sharing[69], and has been applied to a range of health conditions, including HIV/AIDS[70], epilepsy[71], surgery[72], hypertension and diabetes[73].

There is now emerging evidence that this approach can be a cost-effective method to provide treatment and care for people with depression[74-76], psychosis[77,78], and perinatal psychiatric disorders[79]. One part of this new approach is to provide training using clear and relevant guidance that staff can apply directly in the clinical situation, such as the WHO mhGAP Intervention Guide[7]. But training alone is insufficient, and it is increasingly clear that ongoing supervision is likely to be necessary to support staff to begin to apply the guidelines, and to gain and maintain clinical competence[80]. The costs of such supervision, therefore, need to be included in the core resources necessary to make community care sustainable[78].

The new cadre of staff includes front-line health care workers, such as community health workers, and posts between the traditional roles of nurse and doctor, such as the clinical officer or medical officer[81-84]. Such staff are often recruited from the local area, and will have rich understanding of the sociocultural context[82-84].

This reconceptualization of the role of the psychiatrist requires first of all a new training curriculum, one that emphasizes the public health need for psychiatrists to work both directly in secondary and tertiary services, and to act as *multipliers* by potentiating the capacity of primary care staff to detect and treat people with mental illness[4,85,86]. It has been suggested[87] that in high-income countries this capacity (in particular in the treatment of people with major depressive disorder) may well be enhanced by changes in the organization and function of health care teams, such as those already being used to improve outcomes in other chronic diseases. Responsibility for active follow-up should be given to a case manager (for example, a practice nurse); adherence to treatment and patient outcomes should be regularly monitored; treatment plans should be frequently adjusted when patients do not improve; and the case manager and primary care physician should have the possibility to consult and refer to a psychiatrist when necessary.

Flexible and accessible working relationships between the primary care doctor, the case manager and a mental health specialist are considered essential to allow most patients with mental disorders to access more effective treatment in primary care, as well as the minority needing ongoing specialist care to be identified and referred. The adaptation of the ideas behind this model to low- and middle-income countries is still to be investigated.

INTEGRATING CARE FOR PEOPLE WITH COMORBID LONG-TERM CONDITIONS

It is becoming increasingly recognized that chronic physical and mental conditions are often comorbid. For example, among patients with diabetes, hypertension, tuberculosis and HIV/AIDS, the rates of anxiety and depression are at least double those of the general population[88]. The common co-occurrence of these diseases in one person can interfere with the treatment regimen for a particular condition; for example, adherence to treatment for tuberculosis or antiretroviral therapy for HIV/AIDS is significantly undermined by the presence of untreated depression among these patients[89,90].

At the same time, in many low- and middle-income countries, primary care staff are trained to identify and treat physical but not mental conditions. The growing evidence of how commonly such comorbidities occur, and the inadequate health care system response to them, clearly indicates the need for structural change in how care is provided.

Within the context of increasingly strong calls to address the social determinants of health[91] and to move towards universal health coverage, few countries will be able to respond

Table 1 Obstacles, challenges, lessons learned and solutions in implementing community-oriented mental health care

Obstacles and challenges		Examples of lessons learned and solutions
Society	Disregard for, or violation of, human rights of people with mental illness	• Oversight by: civil society and service user groups, government inspectorates, international non-governmental organizations (NGOs), professional associations
	Stigma and discrimination, reflected in negative attitudes of health staff	• Encourage consumer and family and carer involvement in policy making, medical training, service provision (e.g., board member, consumer provider), service evaluation (consumer satisfaction survey)
	Need to address different models of abnormal behavior	• Traditional and faith-based paradigms need to be amalgamated, blended, or aligned as much as possible with medical paradigms
Government	Low priority given by government to mental health	• Government task force on mental illness • Establish cross-party political support for the national policy and implementation • Effective advocacy on mental health gap, global burden of disease, impact of mental health conditions, cost-effectiveness of interventions
	Absence or inappropriate mental health policy	• Advocate for and formulate policy based upon widespread consultation with the full range of stakeholder groups
	Old or inappropriate mental health legislation	• Create powerful lobby and rationale for mental health law
	Inadequate financial resources in relation to population level needs	• Recruit key political and governance champions to advocate for adequate funding of initiatives
	Lack of alignment between payment methods, services and outcomes	• Provide small financial incentives for valued outcomes • Create categories of reimbursement consistent with system strategy
	Need to address infrastructure	• Government to plan and finance efficient use of buildings, essential supplies and electronic information systems
	Need to address structure of community-oriented service system	• Design the mental health system from local primary care to regional care to central specialty care and fill in gaps with new resources as funding grows
	Inadequate human resources for delivery of mental health care	• Task sharing to non-traditional staff cadres such as community health workers and health extension workers
	Brain drain and failure to retain staff	• United Nations agencies/international NGOs to optimize sustainability of their projects
	Non-sustainable, parallel programs by international NGOs	• Close relations with ministries and other stakeholders and international NGOs • Mental health plan in place so NGOs can help achieve these goals sustainably
Organization of health system	Need to design, monitor, and adjust organization of mental health system	• Set implementation plan with clear coordination between services • Prioritization of target groups, especially people with severe mental illness
	Lack of a feasible mental health program or non-implementation of mental health program	• Make program highly practical by identifying resources available, tasks to be completed, allocation of responsibilities, timescales, reporting and accountability arrangements, progress monitoring/evaluation systems
	Need to specify developmental phases	• Planners and professional leaders to design five- and ten-year plans
	Poor utilization of existing mental health facilities	• Improve awareness of benefits of facilities and services • Inbuilt monitoring quality of care, especially process and outcome phases
	Need to include non-medical services	• Include families, faith-based social services, NGOs, housing services, vocational services, peer-support services, and self-help services. All stakeholders involved in designing system
	Lack of multi-sectoral collaboration, e.g. including traditional healers, housing, criminal justice, or education sectors	• Development of clear policy/implementation plan by all stakeholders • Collaborate with other local service to identify and help people with mental illness • Familiarization sessions between practitioners in the Western and local traditions
	Poor availability of psychotropic medication	• Drug revolving funds, public-private partnerships

Table 1 Obstacles, challenges, lessons learned and solutions in implementing community-oriented mental health care *(continued)*

Obstacles and challenges		Examples of lessons learned and solutions
Professionals and practitioners	Need for leadership	• Psychiatrists and other professionals need to be involved as experts in planning, education, research, and overcoming inertia and resistance in the current environment
	Difficulty sustaining in-service training/adequate supervision	• Training of the trainers by staff from other regions or countries • Shifting of some psychiatric functions to trained and available practitioners
	High staff turnover and burnout, or low staff morale	• Emphasize career-long continuing training programs • Training of supervisors
	Poor quality of care/concern about staff skills	• Ongoing training and supervision • Encourage and reward quality by awards and similar processes
	Professional resistance, e.g. to community-oriented care and service user involvement	• Government and professional societies promote the importance of community-oriented care and service user involvement • Develop training in recovery-oriented psychosocial rehabilitation as part of training of new psychiatrists, including at medical schools in low- and middle-income countries
	Dearth of relevant research to inform cost-effective services and lack of data on mental health service evaluation	• More funding on research, for both qualitative and quantitative evidence of successfully implemented examples of community-oriented care
	Failure to address disparities (e.g., by ethnic, economic groups)	• All key stakeholders involved; advocacy for under-represented groups to develop policies and implementation plans
Users, families, and other advocates	Need for advocacy	• Users and other advocates may be involved in all aspects of social change, planning, lobbying the government, monitoring the development and functioning of the service system, and improving the service system
	Need for self-help and peer support services	• Users to lead these movements
	Need for shared decision making	• Users and other advocates must demand at all levels that the system shift to value the goals of users and families and that shared decision making become the norm

effectively to the future health and economic burden that mental disorders and other chronic diseases will pose simply by pursuing "business as usual" approaches. Rather, health systems need new approaches that are capable of mounting an effective, integrated and efficient response to the prevention and management of mental disorders and other chronic conditions.

In order to progressively reform or transform health systems so that they are better equipped to deal with the kinds of health problems that increasingly dominate the demands put upon them, an integrated model of chronic disease prevention and management is called for. Such an approach has already been articulated in the form of the "chronic care model", which was initially developed by US health service researchers and practitioners[92,93], and subsequently adapted to the international level by the WHO in its Innovative Care for Chronic Conditions Framework (ICCCF)[94].

This framework sets out critical principles and requirements for coordinated care, i.e., that it should be community-based, person-centered, and system-oriented. It has been shown to be effective in improving patient outcomes and patient satisfaction across a range of chronic conditions in high-income settings[95,96]. Yet, few examples to date have shown its successful implementation in low- and middle-income countries. We do have ongoing and completed examples of certain elements in India, Ethiopia and especially South Africa, where we can find perhaps the most ambitious effort to date to reform or "re-engineer" the entire health system towards chronic care[97-99].

The chronic care model codifies a number of systemic changes associated with quality improvements in chronic illness care, including: support of service users to manage themselves ("self-management support"); support of clinical decision making through guidelines; clear delineation of clinical roles and responsibilities; improved clinical information systems and service coordination; and collaboration with community groups[93]. The successful outcomes achieved by this model with hypertension and diabetes have led mental health service researchers and practitioners to apply it to mental disorders such as depression and anxiety, and evidence is growing of the effectiveness of the ICCCF approach[88,95,96,100-106].

One advantage of such an integrated care approach, to be empirically tested in future, is that it may be more effective in providing physical health care to people with severe mental illness, and so diminish the high levels of premature mortality in the latter group, which may lead to 20-30 years less life expectancy[107-110].

POSSIBLE ADVERSE EFFECTS OF DEINSTITUTIONALIZATION

Deinstitutionalization has taken place for over half a century in many high-income countries worldwide[111]. Although supported by both the WHO[112] and the WPA[11], this process has been subjected to a number of criticisms. Commentators have claimed a series of adverse effects, in particular high numbers of mentally ill people who are in prison, are homeless or are neglected. There has even been a recent call to "bring back the asylum"[113]. This contention has been advanced particularly where there have been concerns that reduced bed numbers, for example from hospital "downsizing" or closures, have not been accompanied by commensurate increases in the numbers of appropriately supported residential places in the community[114,115].

These objections to community care have been examined in a recent study which reviewed the consequences of reducing the number of beds for long-term psychiatric patients[116]. The authors of this review focused upon cohort studies of people with severe mental disorders who were discharged from psychiatric hospitals following an admission of one year or longer, and in whom data were analyzed at the individual level. They concluded that, contrary to the results of ecological studies, instances of homelessness, incarceration or suicide among those discharged were rare.

Indeed, where bed reduction is done responsibly, it has been shown that the overall costs of community-based care are similar to those of hospital-based services for long-term patients, while the quality of life and satisfaction among individuals receiving residential care in the community are higher compared to those in hospital[117-119]. On the other hand, where hospital closures are intended to be primarily cost-cutting exercises, without proper replacement by services in the community, then it is clear that the quantity and quality of care will suffer and may well lead to adverse outcomes for the people concerned, including the risk for "transinstitutionalization"[120,121].

IMPLICATIONS OF THE EVIDENCE BASE FOR DEVELOPING COMMUNITY MENTAL HEALTH CARE

The foregoing discussion raises profound questions about why treatment and care for people with established mental illnesses, as well as evidence-based methods to prevent mental illness, have remained a low investment priority for governments in most countries worldwide, indeed a level of disregard that has been described as structural or systemic discrimination[122,123]. What has been learned since the mid-20th century, when deinstitutionalization first gained momentum in some high-income countries? We frame this closing section of our paper in terms of a series of recommendations, based upon the lessons learned.

We consider the greatest challenge in mental health care to be the degree of disregard shown to the fact that the large majority of people with mental illness worldwide receive no treatment[124]. To scale up services to the quantum required necessarily means providing most services not in specialist care settings, but in primary, community health care services, and in population-level and community-level platforms as discussed above.

Proposal 1. Central and regional governments should measure the treated percentage of people with mental illness (coverage) and set specific targets to increase coverage over set time periods.

It is unacceptable that governments continue to allow people with all types of mental illness to die about 10 years before others in their communities[125], and people with severe mental illness to die 15-30 years earlier, in countries at all resource levels[107,109,110,126,127]. Taking this issue seriously means reducing cardiovascular and pulmonary as well as suicide risk factors, again tasks that are more feasible in primary and community care settings.

Proposal 2. Health care services need to recognize the far lower life expectancy among people with mental disorders, and develop and evaluate new methods to reduce this health disparity.

It is clear that stigma and discrimination act as a pervasive influence that affects all levels of planning and implementation of treatments and services related to mental health. Yet, there is now an evidence base that contact-based interventions are effective to reduce stigma[48,128-130]. The implication for community mental health is the need for population-level and community-level platforms to use contact-based interventions to reduce stigma and discrimination.

Proposal 3. Evidence-based interventions need to be provided in the long term at the population and community levels to reduce stigma and discrimination experienced by people with mental illness.

Part of the explanation for the mental health gap is that the services provided are often seen by people with mental illness and their carers as being inaccessible or unacceptable. Indeed, scaling up mental health care means paying attention both to the quantity and to the quality of care available[131,132]. While the question of institutionalization has usually been described within hospital settings, human rights issues also need to be quality assured within community mental health services[133].

Proposal 4. Mental health staff should provide care that service users (and their family members) find accessible and acceptable.

The available evidence shows that a reasonable portfolio of mental health services, for example for a district or for a

Table 2 References to mental health in Sustainable Development Goals (SDGs)[144]

Mental health is included in the Principles of the SDGs (formally called the "Declaration")

- To **promote physical and mental health** and well-being, and to extend life expectancy for all, we must achieve universal health coverage and access to quality health care (Paragraph 7)
- We are committed to the prevention and treatment of non-communicable diseases, **including behavioral, developmental and neurological disorders**, which constitute **a major challenge for sustainable development (**Paragraph 26)

Mental health is included within Goal 3 in three targets

- By 2030, reduce by one third premature mortality from non-communicable diseases through prevention and treatment and **promote mental health** and well-being (Target 3.4)
- Strengthen the **prevention and treatment of substance abuse**, including narcotic drug abuse and harmful use of alcohol (Target 3.5)
- **Achieve universal health covera**ge, including financial risk protection, **access to quality essential health-care services** and access to safe, effective, quality and affordable essential medicines and vaccines for all (Target 3.8)

region, will need to include provision of both (limited) inpatient care, and a range of outpatient and community services, according to the resources available[17].

Proposal 5. Mental health care should consist of a careful balance of hospital and community care, with most care provided at or near to people's homes.

Value for money in providing treatments to people with mental illness means both investing in evidence-based care, and disinvesting in harmful, ineffective or less-effective interventions. At present, in countries of all resource level, understanding of how to implement good practice is not well developed[134-136].

Proposal 6. Mental health planners, both in times of economic growth and recession, should invest in treatments known to be effective, and disinvest from treatments known to be ineffective or even harmful.

There is a particular need to pay attention to how far people with mental illness control their own treatment and care plans, as in most countries worldwide forms of involuntary or compulsory treatment are commonly practiced. The United Nations Convention on the Rights of Persons with Disability sets out a framework which can be used to improve the respect of human rights of people with mental illness (referred to in this context as person with psychosocial disabilities)[137]. Within both hospital- and community-based services, an important issue is how far patients/consumers actively participate in treatment through joint decision making processes.

Proposal 7. Health care staff and service users should develop and evaluate methods to improve shared decision making.

In several countries, high- and low-income ones, a wide range of health care practitioners from non-Western traditions provide health care related interventions[138], yet there are a number of challenges at present to an integrated approach, namely: a) the pathways to such practitioners for people with mental illnesses have not been documented in a systematic way; b) the methods of assessment and case formulation are rarely described, nor how far Western and non-Western staff cross-refer patients; c) the numbers of people receiving such care (and so their contribution to overall treatment coverage) is unknown; d) the nature of the interventions delivered is sometimes not described; e) the outcomes of care may not have been examined by scientific methods; and so f) the cost-effectiveness of such treatments is frequently unknown.

Indeed, official statements of mental health policy, for example in national mental health plans, rarely even acknowledge the existence of the non-state funded health care providers and sectors. In our view, therefore, a great deal now needs to be done to clarify these issues and to find effective methods to bring non-Western health care staff into a wider and integrated mental health care system[139-141]. More and more detailed work is needed to identify the relative strengths of these various approaches, and how Western and non-Western tradition practitioners can form providers' networks, including cross-referral patterns, for the benefit of patients.

Proposal 8. Health care practitioners (of Western and non-Western traditions) should take practical steps to see each other as partners in an integrated system that increases the total amount of mental health care available, while ensuring that only effective and acceptable treatments are provided.

Many reports from service users and service user advocacy groups highlight that therapeutic pessimism from health care staff, whether hospital or community based, can itself be a factor promoting worse clinical outcome[142]. The social movement related to recovery has identified this feature of mental health staff, in particular, as hindering clinical progress[143].

Proposal 9. Mental health services should develop dedicated programs for recovery: this implies that staff understand an individual's personal recovery goals and fully support their achievement.

Mental health has recently been given a greater relative importance by the United Nations, as it has been clearly referred to within the Sustainable Development Goals (SDGs), and their related targets and indicators[144-148] (see Table 2). In the period until 2030, the development of global mental health will be advanced by embedding mental health initiatives, as far as possible, into wider SDG-related investments, so as to improve mental health both directly and indirectly.

Proposal 10. Developments to improve mental health will be enhanced by: a) increasing mental health care delivery; b) strengthening health systems (particularly providing integrated care for people with long-term conditions); c) investing in platforms to deliver population-level and community-level interventions; and d) embedding evidence-based measures into global SDG-related activities that will promote mental health and prevent mental illness.

ACKNOWLEDGEMENTS

G. Thornicroft is supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South London at King's College London Foundation Trust. He acknowledges financial support from the UK Department of Health via the NIHR Biomedical Research Centre and Dementia Unit, awarded to South London and Maudsley National Health System (NHS) Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust. He is also supported by the European Union Seventh Framework Programme (FP7/2007-2013) Emerald project. The views expressed here are those of the authors and not necessarily those of the NHS, the NIHR or the UK Department of Health.

REFERENCES

1. Wang PS, Aguilar-Gaxiola S, Alonso J et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO World Mental Health Surveys. Lancet 2007;370:841-50.
2. Thornicroft G. Most people with mental illness are not treated. Lancet 2007;370:807-8.
3. Demyttenaere K, Bruffaerts R, Posada-Villa J et al. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. JAMA 2004;291:2581-90.
4. Saxena S, Thornicroft G, Knapp M et al. Resources for mental health: scarcity, inequity, and inefficiency. Lancet 2007;370:878-89.
5. Collins PY, Patel V, Joestl SS et al. Grand challenges in global mental health. Nature 2011;475:27-30.
6. Collins PY, Saxena S. Action on mental health needs global cooperation. Nature 2016;532:25-7.
7. World Health Organization. mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings. Geneva: World Health Organization, 2010.
8. Dua T, Barbui C, Clark N et al. Evidence-based guidelines for mental, neurological, and substance use disorders in low- and middle-income countries: summary of WHO recommendations. PLoS Med 2011;8: e1001122.
9. Barbui C, Dua T, Van Ommeren M et al. Challenges in developing evidence-based recommendations using the GRADE approach: the case of mental, neurological, and substance use disorders. PLoS Med 2010;7: e1000322.
10. Thornicroft G, Tansella M. Better mental health care. Cambridge: Cambridge University Press, 2009.
11. Thornicroft G, Alem A, Antunes Dos Santos R et al. WPA guidance on steps, obstacles and mistakes to avoid in the implementation of community mental health care. World Psychiatry 2010;9:67-77.
12. Slade M. Personal recovery and mental illness. A guide for mental health professionals. Cambridge: Cambridge University Press, 2009.
13. Warner R. Recovery from schizophrenia: psychiatry and political economy. Hove: Brunner-Routledge, 2004.
14. Drake RE, Bajraktari I, Tansella M. Technology and behavioural health: an implementation challenge. Epidemiol Psychiatr Sci 2014;23:313-5.
15. Slade M, Amering M, Farkas M et al. Uses and abuses of recovery: implementing recovery-oriented practices in mental health systems. World Psychiatry 2014;13:12-20.
16. Thornicroft G, Szmukler G, Mueser K et al. Oxford textbook of community mental health. Oxford: Oxford University Press, 2011.
17. Thornicroft G, Tansella M. Components of a modern mental health service: a pragmatic balance of community and hospital care – Overview of systematic evidence. Br J Psychiatry 2004;185:283-90.
18. Thornicroft G, Tansella M. The balanced care model for global mental health. Psychol Med 2013;43:849-63.
19. Thornicroft G, Tansella M. The balanced care model: the case for both hospital- and community-based mental healthcare. Br J Psychiatry 2013; 202:246-8.
20. Thornicroft G, Tansella M, Law A. Steps, challenges and lessons in developing community mental health care. World Psychiatry 2008;7:87-92.
21. World Health Organization. World mental health atlas 2014. Geneva: World Health Organization, 2014.
22. Miller G. Mental health in developing countries. The unseen: mental illness's global toll. Science 2006;311:458-61.
23. World Bank. World Bank list of economies. Washington: World Bank, 2014.
24. Desjarlais R, Eisenberg L, Good B et al. World mental health. Problems and priorities in low income countries. Oxford: Oxford University Press, 1995.
25. Eaton J. Ensuring access to psychotropic medication in sub-Saharan Africa. Afr J Psychiatry 2008;11:179-81.
26. Beaglehole R, Bonita R. Global public health: a scorecard. Lancet 2008; 372:1988-96.
27. Saxena S, Maulik P. Mental health services in low and middle income countries: an overview. Curr Opin Psychiatry 2003;16:437-42.
28. Alem A. Community-based vs. hospital-based mental health care: the case of Africa. World Psychiatry 2002;1:99-100.
29. Becker T, Koesters M. Psychiatric outpatient clinics. In: Thornicroft G, Szmukler GI, Mueser KT et al (eds). Oxford textbook of community mental health. Oxford: Oxford University Press, 2011:179-91.
30. Rodriguez JJ. Mental health care systems in Latin America and the Caribbean. Int Rev Psychiatry 2010;4:317-24.
31. Thornicroft G, Becker T, Holloway F et al. Community mental health teams: evidence or belief? Br J Psychiatry 1999;175:508-13.
32. Tyrer S, Coid J, Simmonds S et al. Community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality. Cochrane Database Syst Rev 2000;2:CD000270.
33. Dieterich M, Irving CB, Park B et al. Intensive case management for severe mental illness. Cochrane Database Syst Rev 2010;10:CD007906.
34. Holloway F, Sederer L. Inpatient treatment. In: Thornicroft G, Szmukler GI, Mueser KT et al (eds). Oxford textbook of community mental health. Oxford: Oxford University Press, 2011:223-31.
35. Johnstone P, Zolese G. Systematic review of the effectiveness of planned short hospital stays for mental health care. BMJ 1999;318:1387-90.
36. Lasalvia A, Tansella M. Acute in-patient care in modern, community-based mental health services. Where and how? Epidemiol Psichiatr Soc 2010;19:275-81.
37. Lelliott P, Bleksley S. Improving the quality of acute inpatient care. Epidemiol Psichiatr Soc 2010;19:287-90.
38. Lloyd-Evans B, Slade M, Jagielska D et al. Residential alternatives to acute psychiatric hospital admission: systematic review. Br J Psychiatry 2009; 195:109-17.
39. Crowther RE, Marshall M, Bond GR et al. Helping people with severe mental illness to obtain work: systematic review. BMJ 2001;322:204-8.
40. Patel V, Chisholm D, Parikh R et al. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. Lancet 2016;387:1672-85.

41. Wahlbeck K. Public mental health: the time is ripe for translation of evidence into practice. World Psychiatry 2015;14:36-42.
42. Petersen I, Evans-Lacko S, Semrau M et al. Promotion, prevention and protection: interventions at the population- and community-levels for mental, neurological and substance use disorders in low- and middle-income countries. Int J Ment Health Syst 2016;10:30.
43. Rehm J, Chisholm D, Room R et al. Alcohol. In: Jamison D, Breman J, Measham A et al (eds). Disease control priorities in developing countries, 2nd ed. New York: Oxford University Press, 2006:887-906.
44. Rehm J, Mathers C, Popova S et al. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 2009;373:2223-33.
45. World Health Organization. Preventing suicide: a global imperative. Geneva: World Health Organization, 2014.
46. Petersen I, Lund C, Stein DJ. Optimizing mental health services in low-income and middle-income countries. Curr Opin Psychiatry 2011;24:318-23.
47. Lund C, De Silva M, Plagerson S et al. Poverty and mental disorders: breaking the cycle in low-income and middle-income countries. Lancet 2011;378:1502-14.
48. Thornicroft G, Mehta N, Clement S et al. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. Lancet 2016;387:1123-32.
49. Clement S, Lassman F, Barley E et al. Mass media interventions for reducing mental health-related stigma. Cochrane Database Syst Rev 2013;7: CD009453.
50. Knapp M, McDaid D, Parsonage M. Mental health promotion and prevention: the economic case. London: Personal Social Services Research Unit, London School of Economics and Political Science, 2011.
51. Evans-Lacko S, Henderson C, Thornicroft G et al. Economic evaluation of the anti-stigma social marketing campaign in England 2009-2011. Br J Psychiatry 2013;202(Suppl. 55):s95-101.
52. Probst TM, Gold D, Caborn J. A preliminary evaluation of SOLVE: addressing psychosocial problems at work. J Occup Health Psychol 2008;13:32-42.
53. Durlak JA, Weissberg RP, Dymnicki AB et al. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. Child Dev 2011;82:405-32.
54. Rahman A, Mubbashar MH, Gater R et al. Randomised trial of impact of school mental-health programme in rural Rawalpindi, Pakistan. Lancet 1998;352:1022-5.
55. Kelly CM, Mithen JM, Fischer JA et al. Youth mental health first aid: a description of the program and an initial evaluation. Int J Ment Health Syst 2011;5:4.
56. Barry MM, Clarke AM, Jenkins R et al. A systematic review of the effectiveness of mental health promotion interventions for young people in low and middle income countries. BMC Public Health 2013;13:835.
57. Kagitcibasi C, Sunar D, Bekman S et al. Continuing effects of early enrichment in adult life: the Turkish Early Enrichment Project 22 years later. J Appl Dev Psychol 2009;30:764-79.
58. Rahman A, Iqbal Z, Roberts C et al. Cluster randomized trial of a parent-based intervention to support early development of children in a low-income country. Child Care Health Dev 2009;35:56-62.
59. Ssewamala FM, Han CK, Neilands TB. Asset ownership and health and mental health functioning among AIDS-orphaned adolescents: findings from a randomized clinical trial in rural Uganda. Soc Sci Med 2009;69: 191-8.
60. Chibanda D, Mesu P, Kajawu L et al. Problem-solving therapy for depression and common mental disorders in Zimbabwe: piloting a task-shifting primary mental health care intervention in a population with a high prevalence of people living with HIV. BMC Public Health 2011;11:828.
61. Thornicroft G, Semrau M, Alem A et al. Global mental health: putting community care into practice. London: Wiley-Blackwell, 2011.
62. Thornicroft G, Alem A, Drake RE et al. Community mental health: putting policy into practice globally. London: Wiley-Blackwell, 2011.
63. Hanlon C, Wondimagegn D, Alem A. Lessons learned in developing community mental health care in Africa. World Psychiatry 2010;9:185-9.
64. Semrau M, Barley EA, Law A et al. Lessons learned in developing community mental health care in Europe. World Psychiatry 2011;10:217-25.
65. Drake RE, Latimer E. Lessons learned in developing community mental health care in North America. World Psychiatry 2012;11:47-51.
66. McGeorge P. Lessons learned in developing community mental health care in Australasia and the South Pacific. World Psychiatry 2012;11:129-32.
67. Thara R, Padvamati R. Community mental health care in South Asia. World Psychiatry 2013;12:176-7.
68. Ito H, Setoya Y, Suzuki Y. Lessons learned in developing community mental health care in East and Southeast Asia. World Psychiatry 2012;11:186-90.
69. McPake B, Mensah K. Task shifting in health care in resource-poor countries. Lancet 2008;372:870-1.
70. Shumbusho F, van GJ, Lowrance D et al. Task shifting for scale-up of HIV care: evaluation of nurse-centered antiretroviral treatment at rural health centers in Rwanda. PLoS Med 2009;6:e1000163.
71. Kengne AP, Fezeu L, Awah PK et al. Task shifting in the management of epilepsy in resource-poor settings. Epilepsia 2010;51:931-2.
72. Chu K, Rosseel P, Gielis P et al. Surgical task shifting in Sub-Saharan Africa. PLoS Med 2009;6:e1000078.
73. Labhardt ND, Balo JR, Ndam M et al. Task shifting to non-physician clinicians for integrated management of hypertension and diabetes in rural Cameroon: a programme assessment at two years. BMC Health Serv Res 2010;10:339.
74. Patel V, Weiss HA, Chowdhary N et al. Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): a cluster randomised controlled trial. Lancet 2010;376:2086-95.
75. Patel VH, Kirkwood BR, Pednekar S et al. Improving the outcomes of primary care attenders with common mental disorders in developing countries: a cluster randomized controlled trial of a collaborative stepped care intervention in Goa, India. Trials 2008;9:4.
76. Chatterjee S, Chowdhary N, Pednekar S et al. Integrating evidence-based treatments for common mental disorders in routine primary care: feasibility and acceptability of the MANAS intervention in Goa, India. World Psychiatry 2008;7:39-46.
77. Chatterjee S, Leese M, Koschorke M et al. Collaborative community based care for people and their families living with schizophrenia in India: protocol for a randomised controlled trial. Trials 2011;12:12.
78. Chatterjee S, Naik S, John S et al. Effectiveness of a community-based intervention for people with schizophrenia and their caregivers in India (COPSI): a randomised controlled trial. Lancet 2014;383:1385-94.
79. Rahman A, Fisher J, Bower P et al. Interventions for common perinatal mental disorders in women in low- and middle-income countries: a systematic review and meta-analysis. Bull World Health Organ 2013;91:593-601.
80. Kohrt BA, Jordans MJ, Rai S et al. Therapist competence in global mental health: development of the ENhancing Assessment of Common Therapeutic factors (ENACT) rating scale. Behav Res Ther 2015;69:11-21.
81. Sodhi S, Banda H, Kathyola D et al. Supporting middle-cadre health care workers in Malawi: lessons learned during implementation of the PALM PLUS package. BMC Health Serv Res 2014;14(Suppl. 1):S8.
82. Kakuma R, Minas H, van Ginneken N et al. Human resources for mental health care: current situation and strategies for action. Lancet 2011;378: 1654-63.
83. Ssebunnya J, Kigozi F, Kizza D et al. Integration of mental health into primary health care in a rural district in Uganda. Afr J Psychiatry 2010;13: 128-31.
84. Balaji M, Chatterjee S, Koschorke M et al. The development of a lay health worker delivered collaborative community based intervention for people with schizophrenia in India. BMC Health Serv Res 2012;12:42.
85. Kigozi F, Ssebunnya J. The multiplier role of psychiatrists in low income settings. Epidemiol Psychiatr Sci 2014;23:123-7.
86. Prince M, Patel V, Saxena S et al. No health without mental health. Lancet 2007;370:859-77.
87. Von Korff M, Goldberg D. Improving outcomes in depression. The whole process of care needs to be enhanced. BMJ 2001;323:948-9.
88. Oni T, McGrath N, BeLue R et al. Chronic diseases and multi-morbidity – a conceptual modification to the WHO ICCC model for countries in health transition. BMC Public Health 2014;14:575.
89. Joska JA, Obayemi A Jr, Cararra H et al. Severe mental illness and retention in anti-retroviral care: a retrospective study. AIDS Behav 2014;18: 1492-500.
90. Gaynes BN, Pence BW, Atashili J et al. Changes in HIV outcomes following depression care in a resource-limited setting: results from a pilot study in Bamenda, Cameroon. PLoS One 2015;10:e0140001.
91. World Health Organization. Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. Geneva: World Health Organization, 2008.

92. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74:511-44.
93. Wagner EH, Austin BT, Davis C et al. Improving chronic illness care: translating evidence into action. Health Aff 2001;20:64-78.
94. World Health Organization. Innovative Care for Chronic Conditions: building blocks for action. Geneva: World Health Organization, 2002.
95. Tsai AC, Morton SC, Mangione CM et al. A meta-analysis of interventions to improve care for chronic illnesses. Am J Manag Care 2005;11:478-88.
96. Coleman K, Austin BT, Brach C et al. Evidence on the Chronic Care Model in the new millennium. Health Aff 2009;28:75-85.
97. Semrau M, Evans-Lacko S, Alem A et al. Strengthening mental health systems in low- and middle-income countries: the Emerald programme. BMC Med 2015;13:79.
98. Lund C, Tomlinson M, De Silva M et al. PRIME: a programme to reduce the treatment gap for mental disorders in five low- and middle-income countries. PLoS Med 2012;9:e1001359.
99. Von Korff M, Gruman J, Schaefer J et al. Collaborative management of chronic illness. Ann Intern Med 1997;127:1097-102.
100. World Health Organization and Calouste Gulbenkian Foundation. Integrating the response to mental disorders and other chronic diseases in health care systems. Geneva: World Health Organization, 2014.
101. Hipgrave DB, Alderman KB, Anderson I et al. Health sector priority setting at meso-level in lower and middle income countries: lessons learned, available options and suggested steps. Soc Sci Med 2014;102:190-200.
102. Ku GM, Kegels G. Adapting chronic care models for diabetes care delivery in low-and-middle-income countries: a review. World J Diabetes 2015;6:566-75.
103. Maher D, Harries AD, Zachariah R et al. A global framework for action to improve the primary care response to chronic non-communicable diseases: a solution to a neglected problem. BMC Public Health 2009;9:355.
104. Mahomed O, Asmail S, Freeman M. An integrated chronic disease management model: a diagonal approach to health system strengthening in South Africa. J Health Care Poor Underserved 2014;25:1723-9.
105. Mahomed OH, Asmall S. Development and implementation of an integrated chronic disease model in South Africa: lessons in the management of change through improving the quality of clinical practice. Int J Integr Care 2015;15:e038.
106. Nuno R, Coleman K, Bengoa R et al. Integrated care for chronic conditions: the contribution of the ICCC Framework. Health Policy 2012;105:55-64.
107. Wahlbeck K, Westman J, Nordentoft M et al. Outcomes of Nordic mental health systems: life expectancy of patients with mental disorders. Br J Psychiatry 2011;199:453-8.
108. Thornicroft G. Physical health disparities and mental illness: the scandal of premature mortality. Br J Psychiatry 2011;199:441-2.
109. Lawrence D, Hancock KJ, Kisely S. The gap in life expectancy from preventable physical illness in psychiatric patients in Western Australia: retrospective analysis of population based registers. BMJ 2013;346:f2539.
110. 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 (in press).
111. Thornicroft G, Bebbington P. Deinstitutionalisation – from hospital closure to service development. Br J Psychiatry 1989;155:739-53.
112. World Health Organization. World health report 2001. Mental health: new understanding, new hope. Geneva: World Health Organization, 2001.
113. Sisti DA, Segal AG, Emanuel EJ. Improving long-term psychiatric care: bring back the asylum. JAMA 2015;313:243-4.
114. Lamb RH, Bachrach LL. Some perspectives on deinstitutionalization. Psychiatr Serv 2001;52:1039-45.
115. Leff J, Thornicroft G, Coxhead N et al. The TAPS Project. 22: A five-year follow-up of long-stay psychiatric patients discharged to the community. Br J Psychiatry 1994;165(Suppl. 25):13-7.
116. Winkler P, Barrett B, McCrone P et al. Deinstitutionalised patients, homelessness and imprisonment: systematic review. Br J Psychiatry 2016;208:421-8.
117. Salisbury TT, Thornicroft G. Deinstitutionalisation does not increase imprisonment or homelessness. Br J Psychiatry 2016;208:412-3.
118. Knapp M, Beecham J, Anderson J et al. The TAPS project. 3: Predicting the community costs of closing psychiatric hospitals. Br J Psychiatry 1990;157:661-70.
119. Taylor TL, Killaspy H, Wright C et al. A systematic review of the international published literature relating to quality of institutional care for people with longer term mental health problems. BMC Psychiatry 2009;9:55.
120. Prins SJ. Does transinstitutionalization explain the overrepresentation of people with serious mental illnesses in the criminal justice system? Community Ment Health J 2011;47:716-22.
121. Fakhoury W, Priebe S. Deinstitutionalization and reinstitutionalization: major changes in the provision of mental healthcare. Psychiatry 2007;6: 313-6.
122. Corrigan PW, Markowitz FE, Watson AC. Structural levels of mental illness stigma and discrimination. Schizophr Bull 2004;30:481-91.
123. Thornicroft G. Shunned: discrimination against people with mental illness. Oxford: Oxford University Press, 2006.
124. Kohn R, Saxena S, Levav I et al. Treatment gap in mental health care. Bull World Health Organ 2004;82:858-66.
125. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. World Psychiatry 2014;13:153-60.
126. Thornicroft G. Premature death among people with mental illness. BMJ 2013;346:f2969.
127. Fekadu A, Medhin G, Kebede D et al. Excess mortality in severe mental disorders: a 10-year population-based cohort study in rural Ethiopia. Br J Psychiatry 2015;206:289-96.
128. Mehta N, Clement S, Marcus E et al. Systematic review of evidence for effective interventions to reduce mental health related stigma and discrimination: medium and long-term effectiveness and interventions in low- and middle-income countries. Br J Psychiatry 2015;207:377-84.
129. Corrigan PW, Morris SB, Michaels PJ et al. Challenging the public stigma of mental illness: a meta-analysis of outcome studies. Psychiatr Serv 2012;63:963-73.
130. Griffiths KM, Carron-Arthur B, Parsons A et al. Effectiveness of programs for reducing the stigma associated with mental disorders. A meta-analysis of randomized controlled trials. World Psychiatry 2014;13:161-75.
131. Eaton J, McCay L, Semrau M et al. Scale up of services for mental health in low-income and middle-income countries. Lancet 2011;378:1592-603.
132. Gaebel W, Grossimlinghaus I, Heun R et al. European Psychiatric Association (EPA) guidance on quality assurance in mental healthcare. Eur Psychiatry 2015;30:360-87.
133. Drew N, Funk M, Tang S et al. Human rights violations of people with mental and psychosocial disabilities: an unresolved global crisis. Lancet 2011;378:1664-75.
134. Bauer MS, Williford WO, Dawson EE et al. Principles of effectiveness trials and their implementation in VA Cooperative Study #430: 'Reducing the efficacy-effectiveness gap in bipolar disorder'. J Affect Disord 2001;67:61-78.
135. Thornicroft G, Lempp H, Tansella M. The place of implementation science in the translational medicine continuum. Psychol Med 2011;41:2015-21.
136. Thornicroft G. Evidence-based mental health care and implementation science in low- and middle-income countries. Epidemiol Psychiatr Sci 2012;21:241-4.
137. United Nations. Convention on the Rights of Persons with Disabilities. New York: United Nations, 2006.
138. Campbell-Hall V, Petersen I, Bhana A et al. Collaboration between traditional practitioners and primary health care staff in South Africa: developing a workable partnership for community mental health services. Transcult Psychiatry 2010;47:610-28.
139. Abbo C, Ekblad S, Waako P et al. The prevalence and severity of mental illnesses handled by traditional healers in two districts in Uganda. Afr Health Sci 2009;9(Suppl. 1):S16-22.
140. Ngoma MC, Prince M, Mann A. Common mental disorders among those attending primary health clinics and traditional healers in urban Tanzania. Br J Psychiatry 2003;183:349-55.
141. Shankar BR, Saravanan B, Jacob KS. Explanatory models of common mental disorders among traditional healers and their patients in rural south India. Int J Soc Psychiatry 2006;52:221-33.
142. Starcevic V. Overcoming therapeutic pessimism in hypochondriasis. Am J Psychother 2002;56:167-77.
143. Slade M. 100 ways to support recovery. London: Rethink, 2009.
144. United Nations. The 2030 Agenda for Global Action and the Sustainable Development Goals. New York: United Nations, 2015.
145. Thornicroft G, Patel V. Including mental health among the new sustainable development goals. BMJ 2014;349:g5189.
146. Gureje O, Thornicroft G. Health equity and mental health in post-2015 sustainable development goals. Lancet Psychiatry 2015;2:12-4.
147. Votruba N, Eaton J, Prince M et al. The importance of global mental health for the Sustainable Development Goals. J Ment Health 2014;23:283-6.
148. Minas H, Tsutsumi A, Izutsu T et al. Comprehensive SDG goal and targets for non-communicable diseases and mental health. Int J Ment Health Syst 2015;9:12.

DOI:10.1002/wps.20349

The four basic components of psychoanalytic technique and derived psychoanalytic psychotherapies

Four aspects jointly determine the very essence of psychoanalytic technique: interpretation, transference analysis, technical neutrality, and countertransference analysis.

Interpretation is the verbal communication by the analyst of the hypothesis of an unconscious conflict that seems to have dominantly emerged now in the patient's communication in the therapeutic encounter. In general, interpretation of a defense or a defensive relationship initiates the interpretative process, followed by the interpretation of the context, or the impulsive relationship against which the defense was erected, and the analysis of the motivation for this defensive process.

Interpretative interventions may be classified into: a) clarification, by which the analyst attempts to clarify what is consciously going on in the patient's mind; b) confrontation, that is, tactful bringing into awareness nonverbal aspects of the patient's behavior; and c) interpretation proper, the analyst's proposed hypothesis of the unconscious meaning that relates all these aspects of the patient's communication to each other.

This condensing hypothesis is interpretation "in the here and now", to be followed or completed with interpretation "in the there and then", that is, the genetic aspects of interpretation that refer to the patient's past, and link the unconscious aspects of the present with the unconscious aspects of the past.

Transference may be defined as the unconscious repetition in the here and now of pathogenic conflicts from the past, and the analysis of transference is the main source of specific change brought about by psychoanalytic treatment.

The classical concept of transference analysis has been expanded significantly by the concept of the analysis of the "total transference" proposed by the Kleinian approach[1]. This involves a systematic analysis of the transference implications of the patient's total verbal and nonverbal manifestations in the hours as well as the patient's direct and implicit communicative efforts to influence the analyst in a certain direction, and the consistent exploration of the transference implications of material from the patient's external life that, at any point, he/she brings into the session.

The inclusion of a systematic consideration of the patient's total functioning at the point of the activation of a predominant transference points to an important implicit consequence of transference interpretation, i.e., the analysis of character. Defensive characterological patterns tend to become dominant transference resistances and lend themselves to systematic analysis leading to characterological modification. This is a significant effect of psychoanalytic treatment, surprisingly underemphasized in the literature.

Technical neutrality tends to be misinterpreted as a recommendation for an analyst's distant, uninvolved attitude, "a mirror to the patient's presentations". In essence, it simply refers to the analyst's not taking sides in the patient's activated internal conflicts, remaining equidistant, as A. Freud[2] put it, from the patient's id, ego, and super ego, and from his/her external reality. Technical neutrality, in addition, implies the analyst's not attempting to influence the patient with his/her own value systems. S. Freud's early metaphor of the analyst as a "mirror" clearly was questioned by himself, and he protested against a view of analytic objectivity as "disgruntled indifference"[3].

Technical neutrality also implies the concept of "abstinence", in the sense that the analytic relationship should not be utilized for the gratification of libidinal or aggressive impulses of the patient or the analyst. In contrast, technical neutrality does not imply the concept of "anonymity", a questionable development in psychoanalytic thinking in the 1950s, importantly related, in my view, to authoritarian pressures within psychoanalytic education, and the related institutionally fostered idealization of the training analyst, who should not show any usual personal human characteristic to the patient. This idealization of the analyst has been sharply criticized in recent years, particularly by the relational school.

Technical neutrality implies a natural and sincere approach to the patient within general socially appropriate behavior, as part of which the analyst avoids all references or focus upon his/her own life interests or problems. The analyst cannot avoid that personal features emerge in the treatment situation, and do become the source of transference reactions. The patient's realistic reaction to realistic aspects of the analyst's behavior should not be considered a transference reaction: not everything is transference! Maintaining the definition of transference as an inappropriate reaction to the reality presented by the analyst, that reflects the activation of the patient's unconscious conflicts, should differentiate transference from other patient's realistic reactions to natural, as well as idiosyncratic, aspects of the treatment situation.

Countertransference is the analyst's total, moment-to-moment emotional reaction to the patient and to the particular material that the patient presents. The contemporary view of countertransference is that of a complex formation co-determined by the analyst's reaction to the patient's transference, to the reality of the patient's life, to the reality of the analyst's life, and to specific transference dispositions activated in the analyst as a reaction to the patient and his/her material.

Under ordinary circumstances, countertransference mostly is determined by the vicissitudes of the transference, and as such, the analyst's emotional reactions may fluctuate significantly within each session. In contrast to acute fluctuations of the countertransference, chronic distortions of the analyst's internal attitude toward the patient usually indicate significant difficulties in the analyst's understanding of the transference. They often point to a stalemate in the analytic situation that the analyst may need to resolve outside the actual times of analytic sessions

with the patient, through self-exploration or consultation. Serious characterological difficulties of the analyst may contribute to such chronic countertransference distortions, but most frequently they relate to more limited difficulties in his/her understanding and interpretations and are related to particular developments in the transference[4].

Full internal tolerance of countertransference reactions, including regressive fantasies about specific relations with the patient, may be followed by the analyst's internal exploration of the meanings of his/her reaction in terms of the present transference situation, and thus prepare the road for transference analysis.

This is an overall outline of the basic aspects that, I suggest, essentially define psychoanalytic technique, and that may be applied to the analysis of various developments in the analytic situation, such as the analysis of dreams, character, acting out, and repetition compulsion, all of which, in the end, will culminate in transference analysis.

Otto F. Kernberg
Personality Disorders Institute, New York Presbyterian Hospital, Westchester Division, and Weill Medical College of Cornell University, New York, NY, USA

1. Joseph B. Int J Psychoanal 1985;66:447-54.
2. Freud A. The ego and the mechanisms of defense. New York: International Universities Press, 1936.
3. Freud S. Letter to Oskar Pfister of 10/22/1927. In: Meng H, Freud EL (eds). Psychoanalysis and faith: the letters of Sigmund Freud and Oskar Pfister. New York: Basic Books, 1963.
4. Kernberg OF. In: vanLuyn B, Akhtar S, Livesley J (eds). Severe personality disorders: major issues in everyday practice. Cambridge: Cambridge University Press, 2007:42-58.

DOI:10.1002/wps.20368

Functional remediation: the pathway from remission to recovery in bipolar disorder

Bipolar disorder is not just a mood disorder. Patients nowadays do not just want *to feel well*, they want *to do well* because they want *to be well*. This is equivalent to say that the critical endpoint is not anymore mere *improvement*, nor even *remission*, but *recovery*. The current therapeutic armamentarium, consisting of traditional drugs such as lithium, plus anticonvulsants, antipsychotics and, in some cases, antidepressants, has made remission an achievable goal for many patients with bipolar disorder. Illness-focused psychological interventions, such as psychoeducation, have helped many to stay well for longer periods of time, and in some cases, indefinitely. But many patients with bipolar disorder stay there, more or less *feeling well*, but not *doing well* at all. Many take their medicines, after having learnt that stopping them leads to relapse and misery and, in addition, more medication, but are unable to get their jobs back or to finish their studies. Many live on the ashes of what used to be their social life before everything was gone with the fire of the illness.

For a long time, the assumption was that recovery was difficult due to social factors, stigma and discrimination. And those are indeed powerful reasons for many to feel socially disabled. But we also learnt that the illness itself carries an increased vulnerability to stress and cognitive difficulties, which were historically neglected, and that those problems persist over time beyond clinical remission.

Functional remediation is an intervention that aims to fill the gap between remission and recovery. Obviously inspired by traditional neurocognitive remediation techniques, such as those that have worked well in brain damage and other neuropsychiatric conditions, its major feature is that it focuses on functioning rather than cognition[1].

The intervention has, therefore, a neurocognitive and psychosocial background including modeling techniques, role playing, self-instructions, verbal instructions, and positive reinforcement, together with metacognition, with objective functioning as the main target. It includes education on cognitive deficits and their impact on daily life, and provides strategies to manage deficiencies across several cognitive domains, such as attention, memory and executive functions. The family and caregivers can also be involved in the process to facilitate the practice of these strategies at home and for reinforcement[2].

Functional remediation is not a mere sensible proposal. It is manualized and evidence-based. The first randomized, controlled trial to test it has been published[3] and is now being replicated. The primary outcome was the improvement in global, clinician-rated measure of psychosocial functioning. A total of 268 outpatients were enrolled across 10 academic sites in Spain. After 21 weekly group sessions, functional remediation improved aspects related to work functioning and interpersonal abilities, increasing personal autonomy and reducing financial dependence.

The intervention works for patients with bipolar I and bipolar II disorder as well, and the positive effects last at least 6 months beyond the final session of the program[4]. In its current format, it is intended for late-stage bipolar disorder, but with some modifications it could be tailored to enhance cognitive reserve[5] and prevent further progression of cognitive and functional impairment in patients at early stages. Hence, there is great potential in designing an intervention combining psychoeducation and functional remediation with focus on early stages and prevention of further morbidity and mortality.

As Insel[6] has questioned, is it realistic to expect conditions as complex as psychotic, mood or anxiety disorders to respond to a singular intervention? Bipolar disorder, perhaps the most polymorphic and complex of all psychiatric conditions, clearly

needs a multidisciplinary and integrative approach, combining the best of drug therapy, biophysical techniques, and psychosocial interventions.

A common criticism that is made to sophisticated and lengthy psychotherapies is that they are difficult to implement in a community-care based system and may not be cost-effective. There have been several attempts to reduce the length and intensity of evidence-based psychoeducational packages, but most of those are unpublished because they failed. There is often a "wishful thinking" background in those aiming at designing an intervention that is effective and brief. It would be like learning a second language or to play a musical instrument with only a few sessions.

Cost-effectiveness is an issue but, if one counts indirect costs, it is likely that any intervention that works is actually cost-effective, especially when occupational outcome is concerned. There is some hypocrisy and discrimination in restricting access to sophisticated psychotherapies when access to complex and very expensive medical procedures, such as transplantation, is granted for most patients in the Western world. The paradox is that you can have a liver transplantation if you are 69 years old and abstinent for 3 months, but you cannot have access to the (psycho)therapy that will keep you abstinent for the rest of your life. Once again, there is no health without mental health.

Functional remediation is not just a fashionable therapy for bipolar disorder. Across the 21 sessions, the patients are walked through plenty of practical challenges and exercises that help them in improving their interpersonal, social and occupational skills. A major strength of this approach is that it fills the gap between neurocognitive processes and social skills, bringing in neuroscience to the traditional scope of social therapies. Hence, changes in the ability to deactivate the default mode network under neurocognitive challenges are expected in bipolar patients who have received this sort of therapy, and studies are ongoing to confirm that.

It is happening in many fields within psychiatry that traditional outcomes, such as psychotic, depressive, manic or anxious symptoms, are being replaced or perhaps upgraded with other targets that are more closely correlated with functioning[7]. Neurocognitive symptoms are the best example. Conditions such as major depression, which were never the focus of neuropsychological assessment except to exclude patients at risk of dementia, are now being studied using not only mood, but also processing speed, executive function and memory as primary outcomes[8]. Neuroimaging and neuropsychological assessments, among other biomarkers, will be increasingly incorporated into clinical trials. Clinical staging will become part of routine assessment[9]. The growing interest in distal outcomes such as functioning as opposed to quality of life or symptoms will run in parallel with molecular and translational psychopathology[10] and the explosion of personalized medicine as applied to mental health.

Functional remediation is a novel psychosocial intervention that has been found to improve the outcome of patients with bipolar disorder. In contrast with patient and family psychoeducation, cognitive-behavioral therapy, and interpersonal social rhythm therapy, the focus of this intervention is not improvement of mood or relapse prevention, but psychosocial adjustment. It proved to be effective in reducing global disability and enhancing interpersonal and occupational functioning. Albeit considered a therapy for late-stage, functionally impaired bipolar patients, there is huge interest in tailoring it for the prevention of cognitive and psychosocial impairment in recently diagnosed patients, following the principle that prevention is better than cure. The final aim is to allow people with bipolar disorder not only *to feel well*, but *to do well* and *to be well*. Getting closer.

Eduard Vieta, Carla Torrent
Bipolar Unit, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Catalonia, Spain

1. Martínez-Arán A, Torrent C, Solé B et al. Clin Pract Epidemiol Ment Health 2011;7:112-6.
2. Vieta E, Torrent C, Martínez-Arán A. Functional remediation for bipolar disorder. Cambridge: Cambridge University Press, 2014.
3. Torrent C, Bonnin Cdel M, Martínez-Arán A et al. Am J Psychiatry 2013; 170:852-9.
4. Bonnin CM, Torrent C, Arango A et al. Br J Psychiatry 2016;208:87-93.
5. Forcada I, Mur M, Mora E et al. Eur Neuropsychopharmacol 2015;25:214-22.
6. Insel TR. World Psychiatry 2015;14:151-3.
7. Martinez-Aran A, Vieta E. Eur Neuropsychopharmacol 2015;25:151-7.
8. Solé B, Jiménez E, Martinez-Aran A et al. Eur Neuropsychopharmacol 2015;25:231-47.
9. McGorry P, Keshavan M, Goldstone S et al. World Psychiatry 2014;13:211-23.
10. Vieta E. Acta Psychiatr Scand 2014;129:323-7.

DOI:10.1002/wps.20351

Mindfulness-based cognitive therapy for relapse prophylaxis in mood disorders

Relapse and recurrence are common and debilitating aspects of major depressive disorder. Furthermore, the risk of developing a chronic course of illness increases with each successive episode and, even among patients who achieve clinical remission, residual depressive symptoms are commonly reported. Maintenance antidepressant monotherapy is effective as long as it is continued, yet in practice side effect burden, tachyphylaxis, safety concerns and premature discontinuation can combine to push non-compliance rates as high as 40%[1]. Alternatives to long-term antidepressant monotherapy, especially those addressing mood outcomes in a broader context of well-being, may appeal to patients wary of continued intervention.

Studies have shown that, for a number of recovered depressed patients, mild dysphoria activates patterns of ruminative self-focus that can maintain and intensify the dysphoric state[2]. The task of relapse prevention, therefore, can be to pre-empt the establishment of these dysfunctional patterns. Mindfulness-based cognitive therapy (MBCT) was designed to achieve this aim by teaching formerly depressed patients how to be more aware of negative thoughts and feelings at times of potential relapse/recurrence, and to respond to those thoughts and feelings in ways that allow them to disengage from treating them as facts or identifying them with one's sense of self-worth. In order to increase its potential cost-efficiency, this strategy was designed as a group skills training approach rather than as an individual psychological therapy.

The MBCT program[3] integrates the practice of mindfulness meditation with the tools of cognitive therapy (CT). A significant component consists of formal meditation exercises such as the body scan, sitting and walking meditations, as well as mindful movement in the form of gentle yoga and stretching. The generalization of mindfulness skills to aspects of everyday life is supported through informal practices such as mindful eating; noticing body sensations, affect and thoughts during pleasant and unpleasant experiences; as well as taking a mindful approach to aspects of one's daily routine which are typically completed on "automatic pilot". A novel aspect is the addition of the "three-minute breathing space", a brief centering meditation exercise designed for use during times of emotional challenges or stress. The CT components include psychoeducation about depressive symptoms and discussion of the cognitive model, including automatic thoughts, and how thoughts are impacted by situations and moods. Participants are also encouraged to identify activities that generate a sense of pleasure or mastery, to be implemented during times of low mood.

The first four sessions of the 8-week program provide a framework for patients to learn to approach present moment experiences in a non-judgmental way. This message is conveyed tacitly through the formal meditation practices, which promote learning to focus (and re-focus as needed) attentional resources to anchors such as the breath and bodily sensations. This process facilitates the ability to observe the structure of one's internal experience as it arises in a given moment, with the intention not to judge the content, knowing that the "judgment" or "reaction" component of one's experience can be more detrimental than the raw experience itself. The skill to deconstruct experience in this way is then applied to depression, using exercises from CT that underscore how reactions to given situations can be colored by thought and interpretation. Thus, the understanding is cultivated that thoughts are not facts, and that thoughts, feelings and body sensations are often transient and dynamic aspects of experience.

In the fourth session, psychoeducation specific to depressive illness is formally introduced. In addition to information surrounding the nature of commonly discussed depressive symptoms (neurovegetative and mood), the types of negative thinking that are associated with depression are explored. Thus, individuals are encouraged to build upon their ability to detect the early warning signs of relapse, and to identify their unique "relapse signatures".

The latter four sessions of the program emphasize the development of a thoughtful and flexible response style for dealing with the signs and symptoms of relapse. The theme "thoughts are not facts" is the focus in the sixth session, which employs a CT exercise to illustrate how readily mood can impact thoughts. In the seventh session, relapse prevention strategies drawn from CT are discussed. The groundwork is laid for an individualized relapse prevention plan for each participant that includes the involvement of family members in an early warning system, keeping a list of highly effective pleasure and mastery activities, as well as noting familiar automatic thoughts and cognitive themes that have preceded relapse in the past.

Randomized controlled trials evaluating MBCT efficacy have found it to be superior to treatment as usual[4] and to perform as well as continuation antidepressant pharmacotherapy[5] in preventing depression relapse/recurrence. These outcomes are supported by a meta-analysis[6] reporting a relative risk reduction of 34% for those receiving MBCT. Of particular interest is that patients with recurrent depression (three or more past episodes) are more likely to benefit from treatment than those who have experienced only one or two episodes of illness.

In a recent study of 424 patients who were on a therapeutic dose of maintenance antidepressant pharmacotherapy, one half continued on this therapeutic regimen, while the other half was randomized to MBCT and discontinued their medication[7]. There were no differences in relapse/recurrence rates between the two groups (47% antidepressant vs. 44% MBCT) over a two year follow-up.

These findings, among others, have supported the adoption of MBCT within a broader matrix of mental health treatments for mood disorders. For example, the UK National Institute of Health and Care Excellence (NICE) Guidelines for preventing depressive recurrence include a recommendation to provide MBCT for patients who have experienced more than two prior depressive episodes.

It is surprising that relatively little is known about how MBCT prevention effects occur. According to one recent review, the most reliable pattern of change predicting outcome in MBCT is bivariate in nature: increases in mindfulness and metacognitive awareness of emotions are matched by decreases in rumination and worry[8]. These findings are consistent with qualitative interviews of patients, who describe developing a different type of relationship to sad moods, rather than their elimination altogether.

Expanding MBCT's public health impact will require addressing two outstanding issues. First, MBCT faces challenges to dissemination that are common to all psychotherapeutic treatments, including service costs, waiting lists, travel time and a shortage of trained therapists. Web-based psychological interventions offer one solution to many of these barriers. Mindful mood balance (MMB) is an online treatment which

provides high fidelity and widespread access to the core benefits of the in-person MBCT program[9]. Second, a clear understanding of the type and amount of practice required to achieve positive clinical outcomes still eludes the field. Perhaps the most reliable finding is that program benefits have been associated with formal (30-40 min) compared to informal (3-5 min) mindfulness practice[10]. As the evidence base evolves, it can be expected that the establishment of competency standards for clinicians working within the MBCT model will yield more targeted recommendations regarding optimum levels of practice density.

Zindel V. Segal, Le-Anh Dinh-Williams
Program in Psychological Clinical Science, University of Toronto Scarborough, Toronto, Ontario, Canada

1. Simon GE, Von Korff M, Rutter CM et al. Arch Gen Psychiatry 2001;58:395-401.
2. Segal ZV, Kennedy S, Gemar M et al. Arch Gen Psychiatry 2006;63:749-55.
3. Segal ZV, Williams JM, Teasdale JD. Mindfulness-based cognitive therapy for depression. New York: Guilford, 2013.
4. Teasdale JD, Segal ZV, Williams JM et al. J Consult Clin Psychol 2000;68: 615-23.
5. Segal ZV, Bieling P, Young T et al. Arch Gen Psychiatry 2010;67:1256-64.
6. Piet J, Hougaard E. Clin Psychol Rev 2011;31:1032-40.
7. Kuyken W, Hayes R, Barrett B et al. Lancet 2015;386:63-73.
8. van der Velden AM, Kuyken W, Wattar U et al. Clin Psychol Rev 2015;37:26-39.
9. Boggs JM, Beck A, Felder JN et al. J Med Internet Res 2014;16:e87.
10. Crane C, Crane RS, Eames C et al. Behav Res Ther 2014;63:17-24.

DOI:10.1002/wps.20352

Bodily distress disorder in ICD-11: problems and prospects

Classifying the disorders associated with burdensome somatic concerns has been a challenging exercise in psychiatric nosology[1]. The classifications of these conditions in ICD-10 and DSM-IV have not fared much better than earlier attempts[2]. Even though not exactly identical, these classifications were broadly similar and criticisms of either system are therefore generally applicable to both. Among the most salient criticisms are those relating to their utility in routine clinical practice. These include the rarity of the major categories of the group, both in the community and in general clinical practice, as well as the evidence suggesting poor diagnostic reliability[3].

A central feature of the definition of these disorders, that the symptoms are not due to physical or medical causes, has been criticized for being unreliable and for posing a fundamental nosological problem: defining a disorder on the basis of the absence of a feature rather than the presence of a problem[4]. Labels assigned to burdensome somatic preoccupations that have come to be seen as pejorative create another problem for clinical utility. Some patients object to the term "somatoform", which they think may imply that their symptoms are of doubtful clinical importance and are "in their heads" or not real. Furthermore, the notion that the symptoms are medically unexplained is often rejected by patients as essentially an issue of detection.

As part of the activities designed to lead to the approval of ICD-11 by the World Health Assembly in 2018, the World Health Organization, through its International Advisory Group[5], constituted the Somatic Distress and Dissociative Disorders Working Group, which, among other tasks, was asked to propose changes to the section on somatoform disorders in ICD-10. The Working Group has proposed a new and much simplified category of bodily distress disorder, which replaces all of ICD-10 categories within the group of somatoform disorders (F45.0) and, to a large extent, neurasthenia (F48.0), bringing these together under a single category. The only ICD-10 somatoform condition excluded from BDD is hypochondriasis (F45.2).

In the proposed new classification, bodily distress disorder is defined as "characterized by the presence of bodily symptoms that are distressing to the individual and excessive attention directed toward the symptoms, which may be manifest by repeated contact with health care providers. If a medical condition is causing or contributing to the symptoms, the degree of attention is clearly excessive in relation to its nature and progression. Excessive attention is not alleviated by appropriate clinical examination and investigations and appropriate reassurance. Bodily symptoms and associated distress are persistent, being present on most days for at least several months, and are associated with significant impairment in personal, family, social, educational, occupational or other important areas of functioning. Typically, the disorder involves multiple bodily symptoms that may vary over time. Occasionally there is a single symptom – usually pain or fatigue – that is associated with the other features of the disorder" (this is the proposed brief definition for bodily distress disorder; for the format of ICD-11 diagnostic guidelines, see First et al[6]).

Responding to the same set of criticisms, the DSM-5 created a new grouping called Somatic Symptom and Related Disorders, in which the prototypic condition is somatic symptom disorder. Even though this diagnosis can be given to a condition with "one or more somatic symptoms", it nevertheless requires that "excessive thoughts, feelings, or behaviors are related to the somatic symptoms or associated health concerns". Specifically, for a diagnosis of somatic symptom disorder, at least one of three psychological criteria should be present: health anxiety, disproportionate and persistent concerns about the medical seriousness of the symptoms, and excessive time and energy devoted to the symptoms or health concerns.

In both the proposed bodily distress disorder and somatic symptom disorder, the most fundamental revision has been the abolition of the distinction between medically explained and medically unexplained somatic complaints. On the other

hand, there are now specific psychological criteria that need to be fulfilled before the diagnosis can be given. The revised classifications thus address the problem of defining somatoform disorders on the basis of the absence of a feature (a physical or medical cause) by specifying the features that must be present, such as distress and excessive thoughts and behaviors[7].

Dropping the criterion of "medically unexplained" is not without its consequences and has been criticized in somatic symptom disorder. It has been argued that patients with medical conditions and with a justifiable reason for somatic complaints may receive an inappropriate psychiatric diagnosis, with the possibility of associated stigma[8]. The specification in bodily distress disorder that "if a medical condition is causing or contributing to the symptoms, the degree of attention is clearly excessive in relation to its nature and progression" is meant to address this concern.

A single somatic symptom may lead to a diagnosis of bodily distress disorder or somatic symptom disorder. A good justification for this revision is that a single symptom, for example pain, may sometimes be as bothersome as multiple somatic symptoms. However, the point has been made that this lowering of the threshold for the diagnosis may lead to an inappropriate labeling of apparently healthy persons as having a psychological disorder[8]. This concern is addressed in bodily distress disorder by the requirement that other features, in particular associated psychological features, as well as significant functional impairment, be present before the diagnosis is given. Also, further information is provided in the proposed diagnostic guidelines that seeks to delineate mild bodily distress disorder from normal somatic concerns which may exist in the community and do not require clinical attention.

One of the important differences between the proposed ICD-11 and the DSM-5 approaches is the name of the prototype disorder. While the DSM-5 has retained the word "somatic", the proposed ICD category has avoided this term altogether. While no label can prevent completely the risk of negative connotations and misinterpretations, a more descriptive label that avoids the term "somatic" might prove more acceptable to both patients and primary care clinicians.

While the DSM-5 has retained hypochondriasis (or health anxiety) within the cluster of Somatic Symptom and Related Disorders, the current proposal for ICD-11 has placed hypochondriasis within the grouping of Obsessive-Compulsive and Related Disorders. The position of DSM-5 is supported by evidence suggesting a high co-occurrence of hypochondriasis with somatization disorder as well as shared cognitive perceptual styles between the two conditions. On the other hand, the position of the ICD-11 Working Group is supported by findings associating repetitive cognition and behaviors as well as task-related neural activation patterns on brain imaging with hypochondriasis[1]. Also, there is evidence that, unlike somatization disorders, hypochondriasis responds to some treatments used for obsessive-compulsive and related disorders[9].

The new proposals for bodily distress disorder are being systematically tested in the field studies conducted as part of the ICD revision process. These studies include Internet-based approaches, in which a large number of clinicians participate through the Global Clinical Practice Network (http://gcp.network), as well field studies conducted in clinical settings. It is hoped that the findings from the field studies will provide opportunities for a further strengthening of the utility and validity of the classification of burdensome somatic concerns in ICD-11 prior to its approval by the World Health Assembly.

Oye Gureje[1], Geoffrey M. Reed[2]
[1]WHO Collaborating Centre for Research and Training in Mental Health, Neuroscience and Substance Abuse, Department of Psychiatry, University of Ibadan, Ibadan, Nigeria; [2]World Health Organization, Geneva, Switzerland and School of Psychology, Universidad Nacional Autónoma de México, Mexico City, Mexico

O. Gureje is Chair of the WHO Working Group on Somatic Distress and Dissociative Disorders. G.M. Reed is a member of the WHO Secretariat, Department of Mental Health and Substance Abuse. The views expressed herein are those of the authors and, except as specifically indicated, do not represent the official positions of the WHO Working Group, the International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders or the WHO.

1. Gureje O. Curr Opin Psychiatry 2015;5:345-9.
2. Mayou R, Kirmayer LJ, Simon G et al. Am J Psychiatry 2005;162:847-55.
3. Creed F, Gureje O. Int Rev Psychiatry 2012;24:556-67.
4. Voigt K, Nagel A, Meyer B et al. J Psychosom Res 2010;68:403-14.
5. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. World Psychiatry 2011;10:86-92.
6. First MB, Reed GM, Saxena S et al. World Psychiatry 2015;14:82-90.
7. Rief W, Martin A. Annu Rev Clin Psychol 2014;10:339-67.
8. Frances A. J Nerv Ment Dis 2013;201:530-1.
9. Greeven A, Van Balkom AJ, Visser S et al. Am J Psychiatry 2007;164:91-9.

DOI:10.1002/wps.20353

Psychodynamic therapy of obsessive-compulsive disorder: principles of a manual-guided approach

Obsessive-compulsive disorder (OCD) is a chronic disabling disorder characterized by recurrent obsessions and uncontrolled compulsions. Recent research suggests that OCD is more common than assumed before[1]. Cognitive-behavioral therapy and selective serotonin reuptake inhibitors have been shown to be equally efficacious in OCD[2], but with rates between 50% and 60% for response and 25% or below for remission[3]. Thus, further development of efficacious treatments is required.

Despite the long clinical tradition of describing and treating OCD from a psychodynamic perspective, no evidence-based psychodynamic treatment exists. Recent research on anxiety disorders, however, suggests that manual-guided short-term psychodynamic therapy (STPP) may be a promising approach[4]. Building on STPP for anxiety disorders, a model of STPP for OCD was developed which is based on Luborsky's supportive-expressive therapy[5]. The treatment consists of twelve modules which include both the characteristic elements of supportive-expressive therapy (i.e., focus on the core conflictual relationship theme, CCRT, and on the helping alliance) and additional disorder-specific treatment elements. In the following the treatment is briefly described.

At the beginning of treatment, the CCRT associated with the symptoms of OCD is assessed. A CCRT encompasses three components: a wish (W, e.g. aggressive or sexual impulses), a response from others (RO, e.g. to be condemned), and a response of the self (RS, e.g. obsessions and/or compulsions)[5]. Focusing on the CCRT, the therapist relates the patient's OCD symptoms (RS) to his or her wishes (or impulses and affects, W) and to the (expected) responses by others (RO). The CCRT is presented to the patient as his or her "OCD formula". This formula allows patients to understand their pattern of anxiety and OCD reactions. It translates the patient's symptoms into (internal and external) interpersonal relationships.

Enhancing the patient's cognitive and emotional understanding of his or her symptoms and of the underlying CCRT represents the expressive (interpretive) element of SE therapy[5]. An expressive intervention addressing the CCRT for Shakespeare's Lady Macbeth's compulsive washing may be[6]: "As we have seen your compulsive washing (RS) is related to your aggression, the murder of Duncan (W), and to your feelings of guilt (internalized RO). By your compulsive washing rituals, you are trying to make your deed undone and to get relief from your guilt feelings... By washing your hands again and again, you are replacing moral purity by physical cleanness".

During treatment, the CCRT and its components are worked through in present and past relationships, including the "here and now" relationship with the therapist. Consistent with available evidence[7], working through the CCRT can be expected to improve the patients' understanding of their conflicts, to reduce their OCD symptoms and to help them in developing more adaptive behaviors (RS). Both within and between sessions, patients are asked to work on their OCD formula, that is to monitor their emotions including their bodily components and to identify the components of the CCRT that lead to anxiety and OCD. Doing so, patients may achieve a better understanding and awareness of their OCD symptoms and a sense of control (i.e., not being helpless towards OCD), the latter being of particular importance for OCD patients.

Establishing a secure therapeutic alliance is regarded as the central ingredient of the supportive element of the intervention. Luborsky[5] has formulated several principles for establishing a secure alliance, e.g. conveying a sense of understanding and acceptance or recognizing the patient's growing ability to work on his or her problems in the same way the therapist does.

In order to tailor the treatment specifically to OCD, we integrated disorder-specific treatment elements that proved to be clinically helpful in OCD into the manual-guided model of STPP[8]. They encompass, for example:

- Differentiating between thinking and acting (e.g., "If you have sexual wishes towards these young women, this does not imply that you have actually committed adultery").
- Mitigating the rigid and hyper-strict super-ego (conscience) typically characteristic of OCD patients[8] (e.g., by not condemning the patient for his or her sexual or aggressive impulses; by encouraging the patient to resist against the super-ego's strict demands[7]). The super-ego can be regarded a part of the RO component of the CCRT.
- Freud's original recommendation to induce OCD patients to face the feared situation and to use the aroused experiences to work on the underlying conflict[9], in other words on the CCRT. The therapist may do so by saying, for example: "When you have these sexual (aggressive, etc.) thoughts towards young women, you get afraid that something terrible will happen to your wife. By carrying out your rituals you are trying to prevent this. We need to work on your expectation which entails not performing your rituals and tolerating the fear – and ultimately see what happens".

Further modules include: a) informing the patient about the disorder and the treatment, b) addressing ambivalence and setting treatment goals, c) establishing an encouraging inner dialogue, d) addressing (potential) non-response and resistance, and e) focusing on termination and relapse prevention.

We are planning to test the presented approach in a randomized controlled trial.

Falk Leichsenring, Christiane Steinert
Department of Psychosomatics and Psychotherapy, University of Giessen, Giessen, Germany

1. Jacobi F, Hofler M, Siegert J et al. Int J Methods Psychiatr Res 2014;23: 304-19.
2. Romanelli RJ, Wu FM, Gamba R et al. Depress Anxiety 2014;31:641-52.
3. Foa EB, Liebowitz MR, Kozak MJ et al. Am J Psychiatry 2005;162:151-61.
4. Keefe JR, McCarthy KS, Dinger U et al. Clin Psychol Rev 2014;34:309-23.
5. Luborsky L. Principles of psychoanalytic psychotherapy. Manual for supportive-expressive treatment. New York: Basic Books, 1984.
6. Freud S. Obsessions and phobias. London: Hogarth Press, 1962/1895.
7. Crits-Christoph P, Luborsky L. In: Luborsky L, Crits-Christoph P (eds). Understanding transference: the CCRT method. New York: Basic Books, 1990:133-46.
8. Lang H. The inhibited rebel. Structure, psychodynamics and therapy of subjects with obsessive-compulsive disorders. Stuttgart: Klett-Cotta, 2015.
9. Freud S. Lines of advance in psycho-analytic therapy. London: Hogarth Press, 1955/1919.

DOI:10.1002/wps.20339

The social defeat hypothesis of schizophrenia: issues of measurement and reverse causality

Eleven years ago, two of us[1] published the social defeat hypothesis of schizophrenia, in an attempt to find a common denominator for several schizophrenia risk factors. The hypothesis posits that the long-term experience of being excluded from the majority group leads to an increased baseline activity and/or sensitization of the mesolimbic dopamine system, putting the individual at increased risk for the disorder[1,2].

The hypothesis may explain to a certain degree why a history of migration, membership of a disadvantaged ethnic minority group (e.g., African-American ethnicity), urban upbringing, low IQ, childhood trauma, drug abuse, hearing impairment, homosexuality[3,4], and perhaps also autism are schizophrenia risk factors.

We noted that the experience of defeat is neither a specific nor a sufficient or necessary risk factor for schizophrenia, and that other factors, including genetic vulnerability, co-participate in determining the nature of the outcome. Interestingly, neuroreceptor imaging studies reported evidence of dopamine sensitization in non-psychotic subjects with hearing impairment or with a history of childhood trauma, thus supporting the hypothesis[5,6].

However, there are at least two good reasons to criticize the hypothesis. First, it is difficult to measure social defeat in humans, because assessments based on interviews or questionnaires are biased by a tendency to give socially desirable replies. Second, one could argue that many children who go on to develop schizophrenia exhibit motor, cognitive and social impairments and that social defeat, therefore, is not a causal factor, but a consequence of a disorder in neurodevelopment, already present before the onset of psychosis and mainly driven by genetic factors.

As for the first issue, we recognize that the social defeat hypothesis is based on an interpretation of group comparisons (e.g., migrants versus natives, deaf subjects versus normal hearing individuals) and that we do not know with certainty whether individuals who develop schizophrenia are more "defeated" than others. This situation entails the risk of an ecological fallacy, which would be the case if, for example, successful migrants were found to be at equal risk of schizophrenia as non-successful migrants. However, we contend that the social defeat hypothesis is the most viable interpretation of the available data. The pattern of findings for ethnic minorities in Europe, for example, shows the highest risks for the least successful and most discriminated groups: African-Caribbeans and Black Africans in the UK, Inuit in Denmark and Moroccan-Dutch in the Netherlands.

As to the second point of criticism, we agree that schizophrenia likely "begins" long before the onset of psychosis. Studies of the Philadelphia Neurodevelopmental Cohort, for example, have shown that individuals aged 11 to 21 years who endorse psychotic symptoms (but do not meet the criteria for schizophrenia) are cognitively delayed, have a diminished whole brain grey matter volume, and grey matter volume deficits in frontal, temporal and parietal cortex[7]. It is true that these individuals are more likely to develop schizophrenia than others. However, given the fact that about 16% of all cohort members endorse psychotic symptoms, it is also evident that the majority will *not* develop the disorder and that motor, cognitive, social or anatomic impairments are merely risk factors or risk indicators of disorder, not hallmarks.

We propose that the epidemiology of schizophrenia supports a role for social exclusion, because it is unlikely that the genes that contribute to a defective neurodevelopment also code for migration, disadvantaged ethnic minority status, urban upbringing, low IQ, childhood trauma, drug abuse, homosexuality, hearing loss and autism. The social defeat hypothesis offers a more parsimonious explanation for this pattern of findings and deserves further development and testing.

First, since only two studies examined the risk of schizophrenia among individuals with a non-heterosexual orientation, further investigations of this topic are required. The hypothesis can also be tested in various other discriminated groups, such as those who are physically less attractive, who harbor a congenital or acquired handicap, a gender identity disorder, etc..

Second, it is important to examine whether "defeated" individuals who develop schizophrenia differ from other defeated subjects in the way they cope with defeat. Are they more likely to deny the very occurrence of defeat or do they attribute their problems to external causes? If they deny any problem, can implicit association tests reveal that they are implicitly aware of an inferior position?

Third, it is possible to conduct experiments in the laboratory. One can expose individuals to a negative evaluation or rejection and examine which subjects react by developing an increase in subclinical psychotic symptoms.

Fourth, any of these approaches can be examined in models of gene x social defeat interaction which, if apparent, would add to the validity of the notion of "defeat" underlying environmental effects.

Finally, using neuroreceptor imaging, dopamine function can be compared between non-psychotic members of excluded and non-excluded groups. For example, a prospective study of dopamine function among migrants, shortly after arrival and after an interval of several years, and a comparison of the results to those obtained from a native control group, would be most informative.

In sum, the hypothesis seems to provide many promising avenues for investigating epidemiological patterns that are still lacking a satisfactory explanation.

Jean-Paul Selten[1,2], Jim van Os[1,3], Elizabeth Cantor-Graae[4]
[1]Department of Psychiatry and Psychology, University of Maastricht, Maastricht, The Netherlands; [2]Rivierduinen, Leiden, The Netherlands; [3]King's College London, King's Health Partners, Department of Psychosis Studies, Institute of Psychiatry, London, UK; [4]Division of Social Medicine and Global Health, Department of Clinical Sciences in Malmø, Lund University, Malmø, Sweden

1. Selten JP, Cantor-Graae E. Br J Psychiatry 2005;187:101-2.
2. Selten JP, van der Ven E, Rutten B et al. Schizophr Bull 2013;39:1180-6.
3. Chakraborty A, McManus S, Brugha TS et al. Br J Psychiatry 2011;198:143-8.
4. Gevonden MJ, Selten JP, Myin-Germeys I et al. Psychol Med 2014;44:421-33.
5. Gevonden MJ, Booij J, van den Brink W et al. JAMA Psychiatry 2014;71: 1364-72.
6. Oswald LM, Wand GS, Kuwabara H et al. Psychopharmacology 2014;231: 2417-33.
7. Satterthwaite TD, Wolf DH, Calkins ME et al. JAMA Psychiatry 2016;73: 515-24.

DOI:10.1002/wps.20369

Self-Help Plus (SH+): a new WHO stress management package

Consistent with its recommendations for stress management interventions[1], the World Health Organization (WHO) has developed a new psychological intervention for managing stress and coping with adversity. This new intervention is intended to be relevant for coping with any type of adversity, including chronic poverty, endemic community and gender-based violence, long-term armed conflict, and displacement. It is especially targeted towards places with enormous needs but limited humanitarian access, such as Syria and South Sudan.

Following exposure to adversity, rates of diverse mental health problems and non-pathological distress increase. At the same time, most people affected by adversity do not have access to effective mental health and psychosocial support[2]. Without mental health specialists on the ground, either for direct service delivery or for training and supervising non-specialists[3,4], new approaches need to be established that can be delivered without an extensive workforce for mental health.

SH+ was developed to address these needs. It does not require much time from experts for implementation: instead, it uses a guided self-help format and is delivered through a pre-recorded audio course, complemented with bibliotherapy. The potential of using a course to access hard-to-reach populations has been demonstrated previously[5]. Evidence for bibliotherapy is also promising[6]. Furthermore, research has found that guided self-help programs produce better results than "pure" (unguided) self-help, and the effects produced by guided self-help are surprisingly similar to face-to-face psychological treatment[7]. SH+ was designed to be relevant for large segments of adversity-affected populations: it is intended to be transdiagnostic, easily adaptable to different cultures and languages, and both meaningful and safe for people with and without mental disorders. The program was developed with experts in psychological care and global mental health, and colleagues in the humanitarian field. It underwent extensive peer-review, with 43 external experts reviewing the intervention.

The SH+ package has two components: a pre-recorded course and a self-help book. Pre-recorded audio material (locally adapted) is delivered across five 2-hour sessions and in groups of 20 to 30 people. The audio material imparts key information about stress management and guides participants through individual exercises and small group discussions. A written facilitator guide helps briefly trained non-specialist facilitators to conduct the course using these audio materials. To augment the course materials, an illustrated self-help book reviews all essential content and concepts. The book – inspired by an existing illustrated self-help guide[8] – contains more than 400 illustrations and conveys key points with minimal text. It was written to be useful both as a standalone product and as a key resource for those participating in the course.

The format of SH+ is innovative in that it seeks to ensure that key intervention components are delivered as intended through the use of pre-recorded audio, without the burden of extensive training and supervision. This mode of delivery holds promise for helping hard-to-reach populations: the package may be introduced in areas where a conventionally delivered mental health intervention would not be feasible (e.g., remote areas, or areas where humanitarian access is limited).

SH+ is based on acceptance and commitment therapy (ACT), a form of cognitive-behavioral therapy, with distinct features[9]. ACT is based on the concept that ongoing attempts to suppress unwanted thoughts and feelings can paradoxically make these problems worse. Instead, it emphasizes learning new ways to accommodate difficult thoughts and feelings – primarily through mindfulness approaches – without letting them dominate, while guiding people to take proactive steps towards living in a way that is consistent with their values. ACT has been shown to be useful for a range of mental health issues[10] and has been used successfully in a guided self-help format[11].

Components of the SH+ package are currently being piloted in Syria, with Syrian refugees in Turkey, and with South

Sudanese refugees in northern Uganda. Initial feedback has been positive. Funding has been secured for a full-scale randomized controlled trial to evaluate the SH+ course in Uganda later this year.

Following evaluation and any necessary revisions, the SH+ package may become part of WHO's growing collection of low-intensity psychological interventions. Thinking Healthy (for perinatal depression)[12] and Problem Management Plus (PM+; delivered in face-to-face sessions)[4] are the first two of this collection. Over the next five years, the WHO will design and rigorously test additional psychological interventions for different age groups and using varying delivery models. Mental health specialists will always be essential for supervision and for management of those for whom these interventions are insufficient. Yet these potentially scalable intervention programs may reduce reliance on scarce specialists, thereby hopefully making mental health care more widely available to those in need.

JoAnne E. Epping-Jordan[1], Russ Harris[2], Felicity L. Brown[3], Kenneth Carswell[4], Claudette Foley[2], Claudia García-Moreno[4], Cary Kogan[5], Mark van Ommeren[4]

[1]WHO Consultant, Seattle, WA, USA; [2]WHO Consultant, Melbourne, Australia; [3]Harvard T.H. Chan School of Public Health, Boston, MA, USA; [4]World Health Organization, Geneva, Switzerland; [5]University of Ottawa, Ottawa, Canada

The authors alone are responsible for the views expressed in this letter, that do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated. SH+ was developed through funds from United Nations (UN) Multi-Partner Trust Fund for UN Action Against Sexual Violence in Conflict, and the WHO Country Office in Syria. Testing with South Sudanese refugees in Uganda is supported by Enhanced Learning and Research for Humanitarian Assistance (ELHRA)'s Research for Health in Humanitarian Crises (R2HC) initiative funded by the UK Department for International Development and the Wellcome Trust.

1. Tol WA, Barbui C, van Ommeren M. JAMA 2013;310:477-8.
2. Tol WA, Barbui C, Galappatti A et al. Lancet 2011;378:1581-91.
3. World Health Organization and United Nations High Commissioner for Refugees. mhGAP Humanitarian Intervention Guide (mhGAP-HIG). Geneva: World Health Organization, 2015.
4. Dawson KS, Bryant RA, Harper M et al. World Psychiatry 2015;14:354-7.
5. Cuijpers P, Muñoz RF, Clarke GN et al. Clin Psychol Rev 2009;29:449-58.
6. den Boer PC, Wiersma D, Van den Bosch RJ. Psychol Med 2004;34:959-71.
7. World Health Organization. WHO mhGAP guideline update. Geneva: World Health Organization, 2015.
8. Harris R, Aisbett B. The illustrated happiness trap. Boston: Shambhala, 2013.
9. Hayes SC, Pistorello J, Levin M. Couns Psychol 2012;40:976-1002.
10. A-Tjak JGL, Davis ML, Morina N et al. Psychother Psychosom 2015;84:30-6.
11. Fledderus M, Bohlmeijer ET, Pieterse ME et al. Psychol Med 2012;42:485-95.
12. World Health Organization. Thinking Healthy: a manual for psychosocial management of perinatal depression (WHO generic field-trial version 1.0). Geneva: World Health Organization, 2015.

DOI:10.1002/wps.20355

Alarming increase of suicide in a remote Indigenous Australian population: an audit of data from 2005 to 2014

Early in 2016, a 10-year old Indigenous girl committed suicide in a remote desert community of Western Australia[1]. This tragedy triggered national and international media attention, followed by demands for increased resources to provide effective prevention of suicide in Indigenous Australians. Due to the absence of any reliable reference information regarding suicide trends, we conducted a de-identified retrospective audit of suicide deaths in this region between 2005 and 2014.

Contemporaneous publications have described a worldwide increase in youth suicide, now among the top five causes of mortality in the 15 to 19 year age group. Research in other developed countries, including Canada and the US, has identified an Indigenous amplification of this phenomenon that gained international attention[2]. In each of these countries, Indigenous populations have their own unique history of social and cultural turmoil and ongoing post-colonial discrimination and adjustment difficulties[3].

Indigenous suicide appears to have been virtually unknown in Australia for the first 100 years following European colonization of the country[4]. Our audit revealed that the Kimberley region now has one of the highest rates of suicide in the world (age-adjusted rate of 174 per 100,000 in 2014). This region forms the northern-most part of the State of Western Australia, comprising 500,000 km^2 of coastline to inland desert. Approximately one third of the 35,000 population are Indigenous Australians, scattered in 200 communities of varying size, many accessible only by air or dirt track, with poor living conditions and significantly low levels of education. Many have experienced complex trauma, including displacement of family, and have relatives who were forcibly taken into state or foster care as a child – the "stolen generations"[5,6]. As the community mental health care provider covering the region, the Kimberley Mental Health and Drug Service recorded details of suicides in a de-identified register for the years 2005-2014. These reports, which were made internally by staff and externally by other health services and the police, have now been audited.

Reported suicides in previous decades amongst Indigenous Australians in the region were as follows: 1 in the 1960s, 3 in the 1970s, 21 in the 1980s and 46 in the 1990s[7]. Our register recorded a total of 125 deaths during the period 2005-2014, which is likely to be an underestimate. Of these, 102 (81%) were identified as Indigenous Australians, and 91 (73%) were male. This provided an age adjusted suicide rate of 74 per 100,000 in the region, in contrast to 10.6 per 100,000 for the general Australian population in 2012, and 11.4 per 100,000 globally in the same year[2].

Of the 102 Indigenous suicides, 69 (67%) were less than 30 years old, and 28 (27%) were less than 20 years old. Indigenous child suicide increased dramatically during this period, with

suicides in those aged 14 or under increasing from 1 in the first five years (2005-2009) to 5 in the second five years (2010-2014).

When Indigenous suicides were stratified by month, there was a seasonal variation, with increased suicides during the "wet" tropical season. Only 30% of those who suicided had previously engaged with, or been referred to, the Kimberley Mental Health and Drug Service, suggesting that ICD-10 and DSM-5 diagnoses of mental disorders may not be a good predictor of Indigenous suicide. Instead, impulsivity (possibly due to alcohol and cannabis toxicity complicated by complex trauma) has been identified and correlated to increased rates of Indigenous suicide[7]. Hanging was the method of suicide in 88% of Indigenous cases.

Current responses to this problem, though well intentioned, are fragmented and funded by various government programs. A culturally informed, long-term, collaborative approach focusing on resilience in young people may hold the key to effective suicide prevention in the Kimberley region[7,8]. Cultural continuity factors identified in First Nations people in Canada have been associated with suicide prevention[8]. We recommend that further funding be focused on research and development of effective Indigenous youth resilience programs that bolster cultural identity.

Anita Campbell[1], Sivasankaran Balaratnasingam[2,3], Catherine McHugh[4], Aleksandar Janca[3], Murray Chapman[2,3]
[1]Kimberley Aboriginal Medical Services Council, Broome, WA, Australia; [2]Kimberley Mental Health and Drug Service, Broome, WA, Australia; [3]School of Psychiatry and Clinical Neurosciences, University of Western Australia, Perth, WA, Australia; [4]Kimberley Population Health Unit, Broome, WA, Australia

1. News.com.au. Calls for nation action after girl, 10, commits suicide. March 9, 2016.
2. World Health Organization. Preventing suicide: a global imperative. Geneva: World Health Organization, 2014.
3. Elliott-Farrelly T. Adv Ment Health 2004;3:138-45.
4. Australian Bureau of Statistics. 2011 Census quickstats. Canberra: Australian Bureau of Statistics, 2013.
5. Jones IH, Horne DJ. Soc Sci Med 1973;7:219-28.
6. Hunter E, Milroy H. Arch Suicide Res 2006;10:141-57.
7. Hunter E. Soc Sci Med 1991;33:661-71.
8. Chandler MJ, Lalonde CE, Sokol BW et al. Monogr Soc Res Child Dev 2003; 68:1-130.

DOI:10.1002/wps.20356

Pathological gambling: a behavioral addiction

Pathological gambling, also referred to as gambling disorder, has become the first recognized non-substance behavioral addiction in the DSM-5. In this classification, several disorders in the heterogeneous DSM-IV category of Impulse Control Disorders Not Elsewhere Classified were reclassified based on data gathered during the time of DSM-IV. However, the DSM-5 classification has generated controversy, with some academic opinion being in favor of leaving pathological gambling in the chapter of impulse control disorders (see, for example, Grant et al[1] in this journal).

Here we provide a summary of the arguments that support the classification of pathological gambling as an addictive disorder (the "pro" arguments) and address those arguments raised by colleagues who favor a different nosology (the "con" arguments). On the "pro" side, several commonalities between pathological gambling and substance use disorders can be highlighted. Among these commonalities are their similar neurobiological underpinnings of brain function and cognitive features[2]. They include similarities in aspects of reward processing between pathological gambling and substance use disorders which are distinct from impulse control disorders. While these latter disorders have rewarding aspects for the individual[1], this reward is based on negative reinforcement: people have a feeling of relief after the act. In sharp contrast, substance-induced addictions and gambling offer positive reinforcement, at least in the early stages of the disease process[2], when people report a "kick" or a state of "flow". Only at later stages, compulsive features and negative reinforcement predominate. Furthermore, an increased salience of stimuli linked to problematic behavior is a central feature shared by pathological gambling and substance use disorders. In both conditions, reward anticipation is dysfunctional irrespective of the type of reward. Evidence suggests that individuals with gambling or substance use disorders exhibit a hypo-responsive reward circuitry. These results support the view that dopaminergic dysfunction constitutes a common feature of both substance-related and behavioral addictions, although further research is warranted[2].

Moreover, pathological gambling and substance use disorders have similar diagnostic characteristics, and comorbidity rates are high[2]. There is overlap in pharmacological and behavioral treatments. Shared genetic vulnerabilities between pathological gambling and substance use disorders exist[3], and a co-aggregation of pathological gambling and substance use disorder in first-degree relatives of individuals with pathological gambling as compared to controls' relatives has been observed[4].

Arguments against a classification of pathological gambling as an addictive disorder, as for example outlined by Grant et al[1], can be refuted without the need of classifying pathological gambling as an impulse control disorder. One of the arguments put forward was that it is premature to consider pathological gambling as an addiction given the finding of shared genetic vulnerability factors between pathological gambling and major depression. We think that the existence of these shared factors can be explained otherwise, given that mood disorders are the second most common co-occurring disorders in pathological gambling, after substance use disorders. In addition, shared genetic liability also exists between substance dependence (e.g., nicotine[5], cocaine[6]) and depression.

Another argument put forward[1] is that no obvious clinical utility exists for categorizing pathological gambling as an addiction because treatment approaches other than those used in the

treatment of substance use disorders may be useful for that condition. Examples outlined are lithium and exposure therapies. However, lithium has the potential to reduce excessive gambling in all likelihood because of its effectiveness in treating comorbid bipolar symptoms rather than pathological gambling *per se*[7]. We agree that exposure therapies can help to reduce gambling urges in pathological gambling. However, this treatment approach has been also successfully used in substance use disorders and is effective in reducing drug- or drug cue-related urges[8].

Finally, when considering prevention, classification of pathological gambling can have a significant impact. While the onset and course of addictions can profoundly be influenced by preventive measures[9], this has not been shown for impulse control disorders.

In summary, the arguments put forward by Grant et al[1] are not sufficient to counter the classification of pathological gambling as an addictive disorder in DSM-5 and to justify a different classification in the upcoming ICD-11. Rather, the opposite holds true. Pathological gambling can best be understood as a "behavioral" addiction, in which the individual is not addicted to a rewarding chemical substance but to a behavior that is rewarding to him/her.

Karl Mann[1], Mira Fauth-Bühler[1], Susumu Higuchi[2], Marc N. Potenza[3], John B. Saunders[4]

[1]Department of Addictive Behavior and Addiction Medicine, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany; [2]National Hospital Organization Kurihama Medical and Addiction Center, Yokosuka, Kanagawa, Japan; [3]Departments of Psychiatry, Neurobiology and Child Study Center and CASAColumbia, Yale University School of Medicine, New Haven, CT, USA; [4]Centre for Youth Substance Abuse Research, University of Queensland, Brisbane, Australia; Disciplines of Psychiatry and Addiction Medicine, Faculty of Medicine, University of Sydney, Sidney, Australia

K. Mann and J.B. Saunders are members of the ICD-11 Working Group on Substance-Related and Addictive Disorders. The views expressed in this letter are not representative of the views of that Working Group. K. Mann and M. Fauth-Bühler contributed equally to this piece.

1. Grant JE, Atmaca M, Fineberg NA et al. World Psychiatry 2014;13:125-7.
2. Fauth-Bühler M, Mann K, Potenza MN. Addict Biol (in press).
3. Lang M, Leménager T, Streit F et al. Eur Psychiatry 2016;36:38-46.
4. Mann K, Leménager T, Zois E et al. Submitted for publication.
5. Edwards AC, Kendler KS. J Affect Disord 2012;15;142:90-7.
6. Arango-Lievano M, Kaplitt MG. Med Sci 2015;31:546-50.
7. Hollander E, Pallanti S, Allen A et al. Am J Psychiatry 2005;162:137-45.
8. Vollstädt-Klein S, Loeber S, Kirsch M et al. Biol Psychiatry 2011;69:1060-6.
9. Holder HD. Am J Addict 2001;10:1-15.

DOI:10.1002/wps.20373

WPA Position Statement on Gender Identity and Same-Sex Orientation, Attraction and Behaviours

Recent controversies in many countries suggest a need for clarity on same-sex orientation, attraction and behaviour (formerly referred to as homosexuality). Along with other international organizations, the WPA considers sexual orientation to be innate and determined by biological, psychological, developmental and social factors.

Over 50 years ago, Kinsey et al[1] documented a diversity of sexual behaviours among people. Surprisingly for the time, he described that for over 10% of individuals this included same-sex sexual behaviours. Subsequent population research has demonstrated that approximately 4% of people identify with a same-sex sexual orientation (e.g., gay, lesbian and bisexual orientations). Another 0.5% identify with a gender identity other than the gender assigned at birth (e.g., transgender)[2]. Globally, this equates to over 250 million individuals. There is a recognized need for moving towards a non-binary gender identity.

Psychiatrists have a social responsibility to advocate for a reduction in social inequalities for all individuals, including inequalities related to gender identity and sexual orientation.

Despite an unfortunate history of perpetuating stigma and discrimination, it has been decades since modern medicine abandoned pathologizing same-sex orientation and behaviour[3]. The World Health Organization (WHO) accepts same-sex orientation as a normal variant of human sexuality[4]. The United Nations Human Rights Council[5] values lesbian, gay, bisexual and transgender (LGBT) rights. In two major diagnostic and classification systems (ICD-10 and DSM-5), same-sex sexual orientation, attraction and behaviour are not seen as pathologies.

There is considerable research evidence to suggest that sexual behaviours and sexual fluidity depend upon a number of factors[6]. Furthermore, it has been shown conclusively that LGBT individuals have higher than expected rates of psychiatric disorders[7,8], and once their rights and equality are recognized these rates start to drop[9-12].

People with diverse sexual orientations and gender identities may have grounds for exploring therapeutic options to help them live more comfortably, reduce distress, cope with structural discrimination, and develop a greater degree of acceptance of their sexual orientation or gender identity. Such principles apply to any individual who experiences distress relating to an aspect of their identity, including heterosexual individuals.

The WPA believes strongly in evidence-based treatment. There is no sound scientific evidence that innate sexual orientation can be changed. Furthermore, so-called treatments of homosexuality can create a setting in which prejudice and discrimination flourish, and they can be potentially harmful[13]. The provision of any intervention purporting to "treat" something that is not a disorder is wholly unethical.

1. The WPA holds the view that lesbian, gay, bisexual and transgender individuals are and should be regarded as valued members of society, who have exactly the same rights and responsibilities as all other citizens. This includes equal access to health care and the rights and responsibilities that go along with living in a civilized society.
2. The WPA recognizes the universality of same-sex expression, across cultures. It holds the position that a same-sex sexual orientation *per se* does not imply objective psychological dysfunction or impairment in judgement, stability or vocational capabilities.
3. The WPA considers same-sex attraction, orientation and behaviour as normal variants of human sexuality. It recognizes the multi-factorial causation of human sexuality, orientation, behaviour and lifestyle. It acknowledges the lack of scientific efficacy of treatments that attempt to change sexual orientation and highlights the harm and adverse effects of such "therapies".
4. The WPA acknowledges the social stigma and consequent discrimination of people with same-sex sexual orientation and transgender gender identity. It recognizes that the difficulties they face are a significant cause of their distress and calls for the provision of adequate mental health support.
5. The WPA supports the need to decriminalize same-sex sexual orientation and behaviour and transgender gender identity, and to recognize LGBT rights to include human, civil and political rights. It also supports anti-bullying legislation; anti-discrimination student, employment and housing laws; immigration equality; equal age of consent laws; and hate crime laws providing enhanced criminal penalties for prejudice-motivated violence against LGBT people.
6. The WPA emphasizes the need for research on and the development of evidence-based medical and social interventions that support the mental health of lesbian, gay, bisexual and transgender individuals.

Dinesh Bhugra[1], Kristen Eckstrand[2], Petros Levounis[3], Anindya Kar[4], Kenneth R. Javate[5]
[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [2]Vanderbilt University School of Medicine, Nashville, TN, USA; [3]Addiction Institute of New York, New York, NY, USA; [4]Calcutta National Medical College & Hospital, Calcutta, India; [5]The Medical City, Manila, Philippines

1. Kinsey AC, Pomeroy CB, Martin CE. Sexual behavior in the male. Bloomington: Indiana University Press, 1948.
2. Gates GJ. How many people are lesbian, gay, bisexual and transgender? http://williamsinstitute.law.ucla.edu.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed. Washington: American Psychiatric Association, 1980.
4. World Health Organization. The ICD-10 classification of mental and behavioural disorders. Geneva: World Health Organization, 1992.
5. Office of the United Nations High Commissioner for Human Rights. Born free and equal. Sexual orientation and gender identity in international human rights law. New York and Geneva:

Office of the United Nations High Commissioner for Human Rights, 2012.
6. Ventriglio A, Kalra G, Bhugra D. Sexual minorities and sexual fluidity. Unpublished manuscript, 2016.
7. Levounis P, Drescher J, Barber ME. The LGBT casebook. Washington: American Psychiatric Publishing, 2012.
8. Kalra G, Ventriglio A, Bhugra D. Int Rev Psychiatry 2015;27:463-9.
9. Gonzales G. N Engl J Med 2014;370:1373-6.
10. Hatzenbuehler ML, Keyes KM, Hasin D. Am J Publ Health 2009;99:2275-81.
11. Hatzenbuehler ML, O'Cleingh C, Grasso C et al. Am J Publ Health 2012;102:285-91.
12. Padula WV, Heru S, Campbell JD. J Gen Intern Med 2016;31:394-401.
13. Rao TSS, Jacob KS. Ind J Psychiatry 2012;54:1-3.

DOI:10.1002/wps.20340

Improving education, policy and research in mental health worldwide: the role of the WPA Collaborating Centres

The WPA, within its 2014-2017 Action Plan[1], established a network of Collaborating Centres to develop innovative initiatives on education, policy and research in mental health. The aim of this network is to create repositories of information as well as offer practical advice and guidance on teaching, policy and research.

The WPA Collaborating Centres have been appointed by the WPA President and Executive Committee for a period of three years in the first instance, according to the following criteria: a) high scientific reputation at national and international levels; b) pre-eminent status in the country's health, research or academic structures; c) high quality of academic and research leadership; d) stability in terms of achievements, staff and resources; e) willingness to deliver the WPA Action Plan; f) clear and appropriate technical expertise.

The functions of the WPA Collaborating Centres are to: a) collect and disseminate information on mental health; b) provide training and links to clinical and research centers; c) support capacity building at country or regional level; d) conduct and coordinate educational and research activities with the support of the WPA[2].

The network includes now seven sites: the National Institute of Mental Health and Neurosciences (NIMHANS), Bangalore, India; the Department of Psychiatry, Chinese University of Hong Kong, Hong Kong; the Department of Psychiatry, University of Nairobi, Kenya; the Department of Psychiatry and Mental Health, University of Cape Town, South Africa; the Institute of Psychiatry, Faculty of Medicine, Ain Shams University, Cairo, Egypt; the Department of Psychiatry, Barts and London School of Medicine and Dentistry, Queen Mary University, London, UK; and the Department of Psychiatry, University of Naples SUN, Naples, Italy.

The network started its activities in 2016, adopting the principles of co-creation and the democratization of knowledge. In fact, mutual learning and exchanges are extremely important for developing new solutions that are sustainable and evidence based, and for providing better care for patients in times of economic constraints, shortage of skilled mental health professionals, and legal and policy obstacles to mental health care in all countries[3-5].

The Centers will provide opportunities for scholarship across high-, middle- and low-income countries, and will disseminate curricula, best clinical practice guidelines, shared policies and high impact research to improve patient care and public mental health. Another priority is to develop shared teaching and learning projects for medical students and psychiatric trainees[6,7]. In the future, the network will expand its aspirations by promoting social inclusion, protection of human rights in care environments, and adoption of effective complex biopsychosocial interventions in clinical practice[8,9].

Updates on the activities promoted by the WPA Collaborating Centres will be shared and disseminated through policy papers, educational activities and training programs.

Kamaldeep S. Bhui[1], Andrea Fiorillo[2], Dan Stein[3], Tarek Okasha[4], David Ndetei[5], Linda Lam[6], Santosh Chaturvedi[7], Mario Maj[2]

[1]Department of Psychiatry, Barts and London School of Medicine and Dentistry, Queen Mary University of London, London, UK; [2]Department of Psychiatry, University of Naples SUN, Naples, Italy; [3]University of Cape Town, Cape Town, South Africa; [4]Institute of Psychiatry, Faculty of Medicine, Ain Shams University, Cairo, Egypt; [5]Department of Psychiatry, University of Nairobi, Nairobi, Kenya; [6]Department of Psychiatry, Chinese University of Hong Kong, Hong Kong; [7]National Institute of Mental Health and Neurosciences (NIMHANS), Bangalore, India

1. Bhugra D. World Psychiatry 2014;13:328.
2. Kallivayalil RA. World Psychiatry 2015;14:374-5.
3. Shidhaye R, Lund C, Chisholm D. Int J Ment Health Syst 2015;30:40.
4. Patel V, Chisholm D, Parikh R et al. Lancet 2016;387:1672-85.
5. Patel V, Saxena S. N Engl J Med 2014;370:498-501.
6. Stanghellini G, Fiorillo A. World Psychiatry 2015; 14:107-8.
7. Baessler F, Riese F, Pinto da Costa M et al. World Psychiatry 2015;14:372-3.
8. Bhugra D. World Psychiatry 2015;14:254.
9. Patel V. World Psychiatry 2015;14:43-4.

DOI:10.1002/wps.20360

Acknowledgement
This publication has been partially supported by an unrestricted educational grant from Otsuka Pharmaceutical Italy S.r.l., which is hereby gratefully acknowledged.