World Psychiatry

OFFICIAL JOURNAL OF THE WORLD PSYCHIATRIC ASSOCIATION (WPA)

Volume 17, Number 1 *February 2018*

WILEY

IMPACT FACTOR: 26.561

ISSN 1723-8617

The World Psychiatric Association (WPA)

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

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

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

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

World Psychiatry

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

Research Reports containing unpublished data are welcome for submission to the journal.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

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

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

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

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

"If you can't measure it, you can't manage it" – essential truth, or costly myth?

In this issue of the journal, Kilbourne et al[1] make a reasoned and cogent argument for a more structured approach to the delivery of mental health care, with the aim of driving up the quality of care and its outcomes. They see mental health services as innovators in models of care delivery (i.e., community-based, multidisciplinary, and person-centred) but as laggards in learning from and adopting advances in monitoring and improving the quality of continuing care derived from other chronic disease disciplines. The paper advocates enhancements oriented to the structure proposed in Donabedian's framework, namely the organization of care, the clinical care processes, and the health care outcomes achieved.

The authors are right to highlight the importance of measurement, on the principle that "if you can't measure it, you can't manage it". Health management information systems are a core "building block" for well-functioning health systems. The purpose of those information systems is to routinely generate quality health information, and they are used directly for management decisions to improve health care delivery – supporting resource management, service monitoring, supervision and quality improvement. On the other hand, W. Edwards Deming[2] warned that the above principle could be "a costly myth". He could be right, and for several reasons.

First, intuitively obvious qualitative enhancements can be made without data either to diagnose the problem or confirm the benefits. Data can provide a mechanism to support incremental improvement, but the fundamental transformation towards a "learning health system" is cultural. Well-functioning health management information systems are just one essential component of an interactive suite of health system strengthening measures that may prove necessary developing: a peer-driven quality improvement culture; non-technical workforce skills (leadership, teamwork and communication); and person-centred care practices guided by values and preferences, with agreed care plans, and patients empowered for self-management.

Beyond evidence-based guidelines, consideration may also be given to the introduction of care pathways[3]. Every patient goes through a care process, and this varies among patients with particular conditions. Care pathways are about planning and managing those processes, in advance, for defined groups of individuals. Critically, this establishes explicit standards for care processes and outcomes, against which performance can then be judged. Not all health care activities lend themselves to this approach, since not all care is provided for a "well-defined patient group" and a "well-defined period of time". For continuing care of mental health conditions, pathways may need to be drawn up and delivered flexibly, contingent upon differing needs, clinical trajectories and treatment responses. A "stepped care" approach is often used, whereby a patient first receives the most effective, least invasive, least expensive and shortest form of assessment or intervention, escalated to the next level where necessary.

Second, problems can arise with the bureaucratization of the data process. All too often health management information systems are over-burdensome. Too many indicators are collected focusing on morbidity, basic service activity and routine risk assessment, with no obvious application to improvement in the quality of care or its outcomes. The problem is compounded when data are merely collected and then reported to higher levels of the health system for aggregation, analysis and centralized decision making, with no information used for improving performance and service delivery at the periphery. Such data systems do not fulfil the basic requirements of a health *management* information system. With the data collectors disengaged from the process, data quality is poor. These problems may be addressed through simplification and democratization.

Keeping things simple, there is much in common between care for hypertension, diabetes and chronic obstructive pulmonary disease on the one hand, and psychosis, epilepsy and depression on the other. Reduced to basics, people with these conditions need to be identified (detection and diagnosis), engaged on an agreed management plan (linkage to care), encouraged and supported to participate actively in the care process (adherence), and be retained in care (retention/drop out), having their treatment reviewed and revised to ensure optimal outcomes (treatment to target). At minimum, therefore, just five items of data need to be collected, although adherence and outcome monitoring need to be continuous across the episode of care.

Democratization of the data process involves two key elements: public and patient involvement in the design and governance of the system ("nothing about us without us"), and the ability to aggregate, analyze and use the data at every level of the health system, including facilities, teams and individual health professionals.

Smartphones or tablets, linked by mobile data to cloud servers, can promote the collection, aggregation, timely analysis and use of health management information systems data. After detection of a condition requiring continuing care, the app would generate a bespoke care pathway with follow-up appointments, and prompted actions and assessments (attendance, adherence, and outcome monitoring) to be carried out on each occasion. These basic health management information systems generate an electronic medical record for any health care professional providing care (promoting information and provider continuity), and a patient registry to track patients' progress.

Providers can target care toward patients with the greatest need (not adherent, not attending, not improving or meeting the

clinical targets defined by the program). Treatment adjustment may involve addressing barriers for patients with poor adherence, or modifying treatment, or considering referral for patients who do not improve despite adhering to care plans, until improvement occurs. Rapid feedback of aggregated data can be used, *inter alia*, to compare care quality and outcomes across health professionals, facilities and districts; to target supervision and support; to identify best performing professionals and facilities to mentor others; and to inform quality improvement initiatives with real time data to track effectiveness.

A "global perspective" is a bold undertaking. Kilbourne et al's paper cites examples from the US health care system, which is complex and particular in its financing models, and its highly fragmented nature. Fragmentation imposes challenges for quality improvement at the national level, where the reach of the state may be limited. At the same time, the authors correctly point out the barriers to implementing reforms in more unitary national or regional health services, such as those in the UK or Canada. Independent private or not-for-profit providers can be fleeter of foot.

From a global mental health perspective, the focus on equity is welcome, but data stratification should extend beyond quality and outcomes of care to include treatment coverage. The "treatment gap" for mental health care services is an affront to the fundamental right to health worldwide, particularly in low- and middle-income countries. It, too, needs to be measured to be reduced.

The important role of primary care also deserves more attention. In high-income countries, task-shifting (to lower levels of the health care system, supported to provide care through task-sharing with specialist services) can reduce costs through greater allocative efficiency, and may provide more holistic, integrated and person-centred care, particularly in the context of physical comorbidity. In resource poor and lower-income settings, task-shifting is understood to be an essential strategy for closing the treatment gap. In either type of setting, a more structured approach to care, supported by data used for quality improvement, can be an essential development.

Martin Prince
King's Global Health Institute, King's College London, London, UK

M. Prince receives salary support for his role as Director of the National Institute of Health Research (NIHR) Global Health Research Unit on Health System Strengthening in Sub-Saharan Africa, King's College London (GHR Unit: 16/136/54). This research was commissioned by the NIHR using official development assistance (ODA) funding. The views expressed in this paper are those of the author and not necessarily those of the UK National Health Service, the NIHR or the UK Department of Health.

1. Kilbourne AM, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
2. Edwards Deming W. The new economics for industry, government, education, 2nd ed. Cambridge: MIT Press, 1994.
3. Schrijvers G, van Hoorn A, Huiskes N. Int J Integr Care 2012;12:e192.
4. Bauer AM, Thielke SM, Katon W et al. Prev Med 2014;66:167-72.

DOI:10.1002/wps.20477

The impact of severe mental disorders and psychotropic medications on sexual health and its implications for clinical management

Angel L. Montejo[1], Laura Montejo[2], David S. Baldwin[3]

[1]Department of Nursing and Institute of Biomedicine of Salamanca, Neurosciences Area, University Hospital of Salamanca, Salamanca, Spain; [2]Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain; [3]Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

Sexual dysfunction often accompanies severe psychiatric illness and can be due to both the mental disorder itself and the use of psychotropic treatments. Many sexual symptoms resolve as the mental state improves, but treatment-related sexual adverse events tend to persist over time, and are unfortunately under-recognized by clinicians and scarcely investigated in clinical trials. Treatment-emergent sexual dysfunction adversely affects quality of life and may contribute to reduce treatment adherence. There are important differences between the various compounds in the incidence of adverse sexual effects, associated with differences in mechanisms of action. Antidepressants with a predominantly serotonergic activity, antipsychotics likely to induce hyperprolactinaemia, and mood stabilizers with hormonal effects are often linked to moderate or severe sexual dysfunction, including decreased libido, delayed orgasm, anorgasmia, and sexual arousal difficulties. Severe mental disorders can interfere with sexual function and satisfaction, while patients wish to preserve a previously satisfactory sexual activity. In many patients, a lack of intimate relationships and chronic deterioration in mental and physical health can be accompanied by either a poor sexual life or a more frequent risky sexual behaviour than in the general population. Here we describe the influence of psychosis and antipsychotic medications, of depression and antidepressant drugs, and of bipolar disorder and mood stabilizers on sexual health, and the optimal management of patients with severe psychiatric illness and sexual dysfunction.

Key words: Sexual health, sexual dysfunction, severe mental illness, psychosis, depression, bipolar disorder, antipsychotics, antidepressants, mood stabilizers, quality of life

(World Psychiatry 2018;17:3–11)

Psychosexual medicine and psychiatry are overlapping disciplines, and there is much interest among psychiatrists in improving their theoretical knowledge and clinical skills in addressing sexual dysfunction.

Adverse sexual effects are frequent with commonly prescribed psychotropic drugs, such as selective serotonin reuptake inhibitors (SSRIs) and prolactin-raising antipsychotics. Deterioration of libido, and arousal and orgasmic dysfunction are frequent disturbances, adversely affecting quality of life. Sexual dysfunction tends to be under-reported and under-recognized and systematic enquiries are needed to assess the incidence, severity and impairment associated with untoward sexual effects of psychotropic drugs.

Recent developments in the field include recognition of the beneficial effects of a healthy sexual life in patients with severe mental disorders; the need to incorporate this aspect in assessment and management within routine clinical practice[1]; a more in-depth understanding of the adverse effects of psychotropic drugs on sexual life; and more detailed guidelines about how to manage sexual dysfunction in these already deeply disadvantaged people.

PSYCHOSIS AND SEXUAL DYSFUNCTION

Influence of psychosis on sexuality

Disturbances in sexual functioning in patients with schizophrenia and related disorders may arise from multiple factors, including negative symptoms (apathy, avolition), depressive symptoms, and adverse effects of some antipsychotics[2]. People diagnosed with psychotic disorders often have unmet needs relating to sexuality and intimacy, which impact negatively on recovery and the ability to lead a fulfilling life. Psychosis tends to be a barrier to the expression of sexuality and intimacy[3].

It can be difficult to study sexuality in some cultures. However, a questionnaire study found a high frequency (70%) of sexual dysfunction in female patients with schizophrenia in India[4]. An investigation of sexual dysfunction in Chinese patients with schizophrenia found a similar frequency[5]. A Korean study found that sexual satisfaction was negatively correlated with length of illness in schizophrenic patients receiving risperidone[6].

Despite what many clinicians believe, adequate sexual expression can improve overall well-being, restore confidence and dignity, and allow patients with psychosis to overcome problems such as social disengagement and stigma. A study comparing sexual life in patients with psychosis and healthy controls found that sexual activity improved self-esteem, feelings of acceptance and additionally sleep, anxiety and mood in patients in a similar way as in controls[7]. Sexual relationships were considered highly relevant by the vast majority of patients, who were more concerned about affection and companionship than physical pleasure. Only 13% were able to maintain a steady partner and only 20% had coital activity, but more than half believed that sexual life was still important to them.

Some psychotic patients put their health at risk through sexually transmitted diseases, including HIV, by not using condoms[8]. This emphasizes the need to systematically evaluate po-

tentially risky behaviours in these patients, and provide education designed to promote safer sexual practices.

The presence of psychotic symptoms should not be incompatible with healthy sexual relationships. Whilst not all patients attach the same importance to sexual life, many young patients who previously had satisfactory sexual relationships are not prepared to lose this aspect of interpersonal functioning after diagnosis and start of pharmacological treatment. Many young male patients who drop out from antipsychotic medication report the onset of sexual dysfunction – especially erectile and orgasm problems in the short term and loss of desire over the longer term – as reasons for stopping treatment.

Influence of treatment of psychosis on sexuality

Sexual dysfunction is common during short- and long-term treatment with antipsychotics, and is associated with a considerable impact on quality of life in adult and adolescent patients[9]. Depending on the measurement method, it affects between 38 and 86% of patients[10-13], including remitted ones and those experiencing a first episode of schizophrenia[14,15].

Symptoms include decreased desire, difficulties in sexual arousal; problems with penile erection, vaginal lubrication and orgasm; and reduced sexual satisfaction. The most frequent complaints in clinical practice include orgasmic and erectile difficulties in the short term and decreased desire in the longer term. The most frequent pattern in male patients is the combination of lowered libido with erectile dysfunction, which is usually unacceptable[16,17].

Several factors are involved, including blockade of dopaminergic activity, hyperprolactinemia, and alpha-1 receptor blockade[18]. Hyperprolactinemia and related hypogonadism seems to be strongly implicated in sexual dysfunction, being sometimes accompanied by infertility, amenorrhea, gynecomastia and galactorrhoea[19,20]. Higher plasma prolactin levels are associated with higher rates of erectile and ejaculatory dysfunction in patients with a first episode of schizophrenia[16].

Dopamine-blocking and hyperprolactinaemia-inducing antipsychotics such as haloperidol, risperidone, paliperidone and amisulpride are more likely to be associated with decreased libido and/or arousal difficulties. By contrast, aripiprazole, quetiapine, olanzapine and ziprasidone have been linked to low rates of sexual dysfunction (16-27%) in open studies[21,22] and in meta-analyses[23]. A lower risk for prolactin elevation and sexual dysfunction was found with aripiprazole once-monthly when compared to long-acting paliperidone, this difference being associated with a greater improvement in quality of life[24].

Erectile problems with antipsychotic drugs may be specifically related to endothelial dysfunction linked to decreased nitric oxide production due to inhibition of endothelial nitric oxide synthetase[25] and vasoconstriction from beta 2-adrenergic effects[26].

Sexual dysfunction tends to be under-estimated in psychotic patients, for several reasons including lack of confidence in health providers, shame, cultural difficulties and lack of interest by psychiatrists. The extent of sub-optimal communication about sexuality in patients with a psychotic disorder assessed within routine clinical practice is considerable, affecting 50-73% of those with sexual dysfunction[13]. Lack of adequate discussion is more common in female patients, of whom 80% reported not to have discussed sexual function with their mental health care providers[27]. Cross-cultural factors are important, as a recent survey conducted in India found that the majority (73.2%) of professionals did not enquire about sexual problems in routine clinical settings, many admitting that they lacked expertise[28]. Furthermore, many patients with severe mental illness have received little sexual education, and have insufficient time allowed for the discussion of emotional relationships in general.

Reliable comparisons between antipsychotics are difficult, due to the wide variety of assessment techniques[29]. Only six questionnaires have been validated to assess sexual dysfunction in psychotic patients. Following a systematic review of psychometric and other properties, only the Antipsychotics and Sexual Functioning Questionnaire (ASFQ)[30], the Changes in Sexual Functioning Questionnaire (CSFQ)[31], and the Psychotropic-Related Sexual Dysfunction Questionnaire (PRSexDQ-SALSEX)[32] were found to address all aspects of sexual functioning, making them preferable for clinical practice and research[33].

Young men with psychosis consider impairment of sexual function as the most important adverse effect of antipsychotic medication affecting treatment adherence[34,35]. In a US-based nationwide survey of patients with schizophrenia, side effects relating to prolactin and other endocrine disturbance were significantly related to lower levels of treatment adherence[36]. Again, cross-cultural factors are probably important, as an investigation in India, using the PRSexDQ-SALSEX questionnaire, found that most patients (91.7%) reported good to fair tolerance of any sexual side effects[28].

Management of treatment-induced sexual dysfunction in psychotic patients

Decreasing the dosage, switching the antipsychotic, add-on strategies with a dopamine agonist, addition of aripiprazole, or use of a phosphodiesterase-5 (PDE-5) inhibitor have all shown some beneficial effects.

However, reducing antipsychotic dosage may sometimes engender relapse, so switching to another antipsychotic medication may be preferable in managing many patients with treatment-emergent sexual dysfunction. Switching to aripiprazole was found successful in several studies, improving delayed ejaculation/orgasm in some naturalistic settings[37], normalizing prolactin levels[38], and maintaining the clinical efficacy of previous treatment[39]. A careful switching protocol is needed to avoid the reappearance of troublesome psychotic symptoms[40]. Adjunctive aripiprazole reduces antipsychotic-induced hyper-

prolactinaemia[41] and sexual dysfunction[42]. When differing strategies were compared, switching to aripiprazole monotherapy was found superior to the addition of aripiprazole in patients with schizophrenia. Positive results have also been reported after switching to quetiapine or ziprasidone in 3 to 6-month prospective studies[43,44].

A Cochrane review of randomized controlled trials involving patients with schizophrenia and sexual dysfunction found that sildenafil can improve erectile function and sexual satisfaction when compared with placebo, and that switching to olanzapine and quetiapine may have a positive impact on sexual functioning in male and female patients[45].

A recent multidisciplinary consensus process concluded that switching an antipsychotic to a non-hyperprolactinaemic one is probably the best way to ameliorate antipsychotic-related sexual dysfunction, with aripiprazole being the first-line option[46]. Systematic screening for sexual dysfunction is strongly recommended[47]. Psychosocial interventions – i.e., psychoeducation, supportive psychotherapy and psychiatric rehabilitation – also play a crucial role, with the restoration of sexual function as an achievable recovery target[3]. Compounds with a lower frequency of sexual dysfunction should be considered as potential first-line options in psychotic patients with an active and satisfactory sexual life.

DEPRESSION AND SEXUAL DYSFUNCTION

Influence of depression on sexuality

Depressive symptoms are strongly associated with sexual difficulties and dissatisfaction, and screening for depression has been recommended in patients with sexual dysfunction and chronic illness[48]. Conversely, depressed patients should be screened for sexual dysfunction[49]. A longitudinal study found the prevalence of sexual problems in depressed individuals to be approximately twice the prevalence in controls (50% vs. 24%)[50].

Recurrent depressive disorder seems especially associated with sexual problems. For example, the US Study of Women's Health Across the Nation found that women with recurrent depressive episodes (but not those who experienced only a single episode) were more likely to report problems in sexual arousal, physical pleasure and emotional satisfaction, when compared to controls[51]. The Netherlands Mental Health Survey and Incidence Survey-2 found that the presence of 12-month mood disorders was associated with a significantly lower likelihood of reported sexual satisfaction[52].

Depression affects mood, energy, interest, capacity for pleasure, self-confidence and self-esteem, so it should be expected that depression lowers sexual interest and satisfaction; this effect seems more marked in younger patients[53]. Depressive symptoms commonly coexist with anxiety symptoms, which are also associated with reported sexual difficulties and dissatisfaction[54,55], and with obsessive-compulsive symptoms, themselves associated with loss of sexual pleasure and sexual dissatisfaction[56,57]. But depression can exert adverse effects on all aspects of the sexual response, including the ability to achieve and maintain penile erection, to attain adequate vaginal lubrication, and to achieve ejaculation or orgasm[58]. Most antidepressants can exert unwanted effects on sexual function and satisfaction, but the adverse effects of depression itself (and of comorbid mental or physical disorders and concomitant medication) are often overlooked when considering the management of patients with sexual dysfunction associated with antidepressant treatment.

Patients and health professionals can feel embarrassed to mention and discuss sexual symptoms, and consultation and recognition rates in primary medical care are low[51,59,60]. Unfortunately, reliance on spontaneous reports of sexual adverse events leads to a substantial under-estimate of sexual problems in depressed patients[61,62]. Screening and severity questionnaires can facilitate recognition and assessment, but cannot fully substitute for a comprehensive but sensitive assessment. The Arizona Sexual Experiences Scale (ASEX)[63], the CSFQ[31], the PRSexDQ-SALSEX[32] and the Sex Effects Scale (SexFX)[64] all have adequate key psychometric properties (validity, reliability and sensitivity to change) and have been recommended for assessing sexual function and satisfaction in depressed patients before and during antidepressant treatment[62].

Influence of treatment of depression on sexuality

It has proved difficult to accurately identify the incidence of treatment-emergent sexual dysfunction (encompassing both the worsening of pre-existing problems and the development of new sexual difficulties in previously untroubled patients) during antidepressant treatment. Two international studies of the prevalence of sexual dysfunction in depressed patients undergoing treatment with either an SSRI or serotonin-noradrenaline reuptake inhibitor (SNRI), which both accounted for self-reported sexual problems before starting treatment and the potential adverse effects of concomitant medication, found that 27-65% of female and 26-57% of male patients experienced either a worsening of pre-existing difficulties or the emergence of new sexual difficulties in the early weeks of treatment[65,66].

An early meta-analysis which included studies with differing designs (incorporating open-label, double-blind, cross-sectional and retrospective investigations) found that "treatment-emergent sexual dysfunction" was no more common with the antidepressants agomelatine, amineptine, bupropion, moclobemide, mirtazapine or nefazodone than with placebo. All other antidepressants were significantly more likely than placebo to be associated with "sexual dysfunction" (as a unitary category), and nearly all were significantly more likely than placebo to be associated with dysfunction in each phase of the sexual response[67]. Bupropion appears associated with a significantly lower rate of treatment-emergent sexual dysfunction than the SSRIs escitalopram, fluoxetine, paroxetine or sertra-

line[68], which may reflect the predominantly noradrenergic-dopaminergic mechanism of action of that drug[69].

A second meta-analysis, of 58 randomized controlled trials and five observational studies, found only minor differences between most antidepressants, although there were relative disadvantages for paroxetine and venlafaxine, and relative advantages for bupropion[70]. A systematic review of the relative efficacy and tolerability of mirtazapine and comparator antidepressants found the former to be less likely than other antidepressants to cause adverse sexual effects[71], possibly reflecting its antagonist effects at alpha-2 adrenergic and 5-HT2C receptors[72].

Some novel antidepressants may have a relatively low propensity for adverse effects on sexual function[73]. Randomized controlled trials with agomelatine suggest it has fewer adverse effects on sexual functioning than some other antidepressants, which is probably due to its antagonist effects at the 5-HT2C receptor, rather than the agonist effects at melatonin receptors[74-77], although the absence of effects on nitrergic relaxation of corpus cavernosum smooth muscle may also be relevant[78]. Vilazodone appears to have a low incidence of spontaneously reported adverse effects on sexual function, which may be related to partial agonist effects at the 5-HT1A receptor: it does not differ from placebo in improvement of sexual function during acute treatment of major depressive episodes, and the "number needed to harm" for sexual adverse effects has been estimated as 7 in men and 23 in women[79-81]. Treatment with the novel "multimodal" antidepressant vortioxetine is associated with a low incidence of reported adverse effects on sexual function in men (3-5%) and women (1-2%), which may relate to its antagonist effects at the 5-HT3 receptor, and to indirect effects in increasing the availability of dopamine and noradrenaline[82].

Risk factors for developing sexual dysfunction during antidepressant treatment include male gender, older age, lower academic achievement, absence of full-time employment, physical ill-health, multiple drug treatment, and troubled interpersonal relationships. Inter-individual variation in pharmacokinetic parameters may be important, as "poor metabolizer" status for cytochrome P450 2D6 contributes to sexual dysfunction with paroxetine[83,84], as does a genetic variation in P-glycoprotein which affects transfer of paroxetine across the blood-brain barrier[85].

Not all sexual effects of antidepressants are unwanted in all patients. Although behavioural approaches to premature ejaculation are effective in most patients[86], many men (including those without depression) troubled by persistent problems can benefit from treatment with either the tricyclic antidepressant clomipramine or SSRIs[87]. The short-acting SSRI dapoxetine is efficacious in treating premature ejaculation, with either daily dosing or "on demand" dosage[88]. It has similar efficacy to paroxetine, though it may be less well tolerated[89]. A systematic review of randomized placebo-controlled trials with trazodone (which has partial agonist effects at 5-HT1A receptors and antagonist effects at 5-HT2A and alpha-1 adrenergic receptors) indicates that it can be efficacious in reducing "psychogenic" erectile dysfunction, when prescribed at higher daily dosage (150-200 mg)[90].

Many patients experience treatment-emergent sexual dysfunction whilst taking an antidepressant[68], but in others the reduction of depressive symptoms through successful treatment can be accompanied by reported improvements in sexual desire and satisfaction[91,92]. Improvement in sexual function appears more common among patients who respond to antidepressant treatment[93].

The proportion of patients who stop treatment because of sexual problems is not established[94,95], nor is the time course of sexual dysfunction in patients who continue with antidepressant treatment[96].

Management of treatment-induced sexual dysfunction in depressed patients

Many interventions have been proposed for managing patients who report sexual dysfunction associated with antidepressants, but there are limited randomized controlled data evaluating the effectiveness and acceptability of psychological and pharmacological interventions[97], and no approach can be considered "ideal"[98,99].

When patients are concerned to preserve usual sexual functioning, choosing an antidepressant thought to have fewer sexual adverse effects is reasonable, when other considerations allow. However, some of these antidepressants have other side effects, limited availability, or questionable efficacy. Sexual side effects of some antidepressants may be dose-related, so reduction in daily dosage is commonly adopted as a first-line approach to management[100]. However, dosage reduction may contribute to depressive symptom relapse, and should only be considered when patients have achieved full remission, and after satisfactory completion of continuation treatment. Regular brief interruptions of treatment (so-called "drug holidays") have been proposed[101], but sexual function will improve in only a proportion of patients and with only some antidepressants: depressive symptoms may worsen, and troublesome discontinuation symptoms can emerge, making this approach potentially hazardous[101].

Many adjuvant interventions have been proposed for relieving sexual dysfunction associated with antidepressants, but few have been subjected to rigorous evaluation. Randomized placebo-controlled trials provide evidence of possible efficacy for bupropion and olanzapine[102], testosterone gel[103], and the PDE-5 inhibitors sildenafil (both in male and female patients[104,105]) and tadalafil[106]. Comparative studies are rare, but a placebo-controlled study found no evidence of efficacy for augmentation with mirtazapine or yohimbine in female patients[107]. Augmentation of antidepressants with aripiprazole can improve sexual interest and satisfaction in depressed women, independent of an improvement in depressive symptoms[108].

Switching from one antidepressant drug to another seems reasonable and is commonly adopted[103], but placebo-controlled

evidence of efficacy rests on a single study of switching from sertraline to (now withdrawn) nefazodone[97]. Switching from one drug to another may lead to discontinuation symptoms, and the replacement drug may prove less effective in controlling depressive symptoms. A single study found that regular exercise prior to sexual activity improved sexual desire and global sexual functioning in depressed women taking antidepressants[109].

Nitric oxide is involved in the physiology of the male and female sexual response. In men, nitric oxide in the corpus cavernosum of the penis binds to guanylate cyclase receptors, which results in increased levels of cyclic guanosine monophosphate (cGMP), leading to smooth muscle relaxation (vasodilation) in the intimal cushions of the helicine arteries, which in turn leads to vasodilation, increased blood flow into the spongy tissue of the penis, and subsequent erection. Sildenafil, tadalafil and vardenafil are potent and selective inhibitors of cGMP-specific PDE-5, which is responsible for degradation of cGMP in the corpus cavernosum, resulting in more cGMP and facilitation of erection[110]. In women, the role of nitric oxide and its interplay with estrogen is less well understood, but the PDE-5 inhibitor enhancement of nitric oxide-cGMP in non-adrenergic-non-cholinergic signalling for women seems similar to the effect in men, and nitric oxide release results in vasodilatation in clitoral and vaginal tissues[111].

A series of randomized placebo-controlled trials demonstrate that PDE-5 inhibitors are efficacious in resolving sexual dysfunction associated with antidepressants[104-106]. Studies of men with erectile dysfunction and depressive symptoms (but not undergoing antidepressant treatment) also show that prescription of PDE-5 inhibitors is often accompanied by a reduction in depressive symptoms, enhanced quality of life, and improved interpersonal relationships[112-114]. Furthermore, preclinical studies suggest that nitric oxide activity is an important vulnerability factor in the Flinders rat depressive phenotype[115], that passage of PDE-5 inhibitors across the blood-brain barrier can occur[116], and that sildenafil has antidepressant-like effects after central muscarinic receptor blockade[117]. PDE-5 inhibitors are often helpful when managing patients with sexual dysfunction associated with antidepressants, but side effects such as headache, dyspepsia and visual disturbances, and the need for cautious use in patients with cardiovascular disease, are all potential limitations.

BIPOLAR DISORDER AND SEXUAL DYSFUNCTION

Influence of bipolar disorder on sexuality

Bipolar disorder can involve sexual disturbances directly related to the illness phase. Male and female patients in manic or hypomanic episodes often experience hypersexuality, with an increased incidence of risky sexual behaviours[118]. By contrast, in depressive episodes, reduction of sexual desire is common. Overall, sexual dissatisfaction is often associated with bipolar disorder[52].

Patients with bipolar disorder tend to have more stable sexual partners and a more intense sexual activity than those with schizophrenia[119,120]. When compared to females, males with bipolar disorder tend to have more sexual partners and are more likely to have sexual intercourse with strangers[121]. Sexual dysfunction is a common residual symptom in euthymic patients with bipolar disorder, and has a significant negative impact on quality of life, similar to that of residual depressive symptoms and occupational stigma[122]. Moreover, impairment in desire, excitement and ability to achieve orgasm is significantly associated with suicide plans or a feeling that life is not worth living[123]. In addition, sexual dysfunction has been identified as a predictor of poor medication adherence[124].

A meta-analysis indicated a statistically significant association between a history of sexual abuse and a lifetime diagnosis of anxiety disorder, depression, eating disorders, sleep disorders and suicide attempts[125]. Unfortunately, no longitudinal studies assessing patients with bipolar disorder are available in this respect. Sexual aggression is common in youth with bipolar disorder, particularly in those with a lifetime history of comorbid post-traumatic stress disorder[126]. Prompt identification and treatment of these youth is highly needed.

Routine enquiries about sexual life, including questions about sexual drive during manic episodes, accompanied by simple psychoeducation, is highly recommended in bipolar patients to mitigate the physical, psychic and family consequences of promiscuous and risky sexual behaviour.

Influence of treatment of bipolar disorder on sexuality

Pharmacological management in bipolar disorder involves the use of lithium, anticonvulsants, antipsychotics, antidepressants and benzodiazepines, either in monotherapy or in combination. Sexual dysfunction is one of the most common side effects of these medications, has a high impact on quality of life, and is rated by patients as one of the most disabling problems.

Lithium is regarded as the first-line treatment in bipolar disorder, but several studies suggest some negative impact of this drug on sexual function, as it may reduce sexual desire, worsen erectile function and decrease sexual satisfaction[127,128]. Approximately one-third of patients receiving lithium experience sexual dysfunction, which usually involves more than a single domain, in both male and female patients[119]. Patients are significantly less likely to experience sexual intercourse, sexual fantasies, sexual desire, pleasure and satisfaction, and 30% of them attribute these problems to lithium treatment[129]. Despite this, it seems that lithium has a less pronounced adverse impact on sexual function compared to other treatments in bipolar disorder[130], especially antipsychotics[131]. The combination of benzodiazepines with lithium seems to be associated with an increased risk of sexual dysfunction, while this dysfunction does not appear to be related to serum lithium levels[132].

Anticonvulsants are often associated with sexual dysfunction in people with epilepsy (35-55% of patients)[133], but there is limited evidence of these adverse effects in patients with bipolar disorder[134-136].

Valproate may induce an increase of serum testosterone, androstenedione and dehydroepiandrosterone sulfate (DHEAS) concentrations, while prolactin levels typically remain within normal limits[137]. The increase in androgen levels is associated with a higher incidence of menstrual disorders and polycystic ovarian syndrome in women treated with this drug[138,139]. Decreased sexual desire and anorgasmia have also been described in bipolar women receiving valproate[140]. In men, valproate treatment may cause erectile dysfunction[141].

Carbamazepine is often associated with reduced levels of estradiol, progesterone and testosterone, and may cause hypogonadism, amenorrhea and decreased sexual function and sexual desire[129,142]. It may also increase sexual hormone-binding globulin (SHBG) concentration, leading to diminished bioactivity of testosterone and estradiol, and consequently reduced libido and erectile dysfunction[143].

Oxcarbazepine is not usually associated with changes in hormonal levels and sexual dysfunction[127], but there are occasional reports of anorgasmia and retrograde ejaculation[144,145]. Lamotrigine is not associated with sexual adverse effects in patients with bipolar disorder[146,147].

Management of treatment-induced sexual dysfunction in bipolar patients

There is little evidence about management of sexual dysfunction associated with mood stabilizers. Using the lowest effective dose of the drug, switching to alternatives, or some add-on strategies may be useful[148].

A small randomized placebo-controlled trial suggests that adjunctive aspirin (240 mg/day) may improve erectile dysfunction in patients undergoing lithium treatment[149]. There is currently no information on the potential utility of PDE-5 inhibitors such as sildenafil, but it seems reasonable to consider them based on clinical experience in other patients. There is some evidence that switching from enzyme-inducing (valproate, carbamazepine) to non-enzyme-inducing (oxcarbamazepine, lamotrigine) anticonvulsants can be beneficial[138].

In epileptic patients, switching to lamotrigine can be associated with an improvement in desire, pleasure, excitement and orgasm in women, but only in the pleasure dimension in men[150]. Addition of lamotrigine to carbamazepine or valproate can ameliorate sexual dysfunction in male patients[151].

CONCLUSIONS

Severe mental illness and many psychotropic drugs impair sexual function and reduce sexual satisfaction. Systematic enquiries in all patients about previous and current sexual life are needed to assess potential sexual dysfunction, and to manage it with the aims of preserving quality of life, maintaining emotional experiences and continuing partner relationships.

Treatments with fewer adverse sexual effects should be considered as potential first-line options in patients with severe mental illness interested in maintaining a sexual life. Managing treatment-emergent side effects adequately is crucial to facilitate compliance and achieve the best possible outcomes.

REFERENCES

1. Quinn C, Happell B. Talking about sexuality with consumers of mental health services. Perspect Psychiatr Care 2013;49:13-20.
2. Adam RL, Sidi H, Midin M et al. The role of atypical antipsychotics in sexuality: road to recovery in schizophrenia. Curr Drug Targets (in press).
3. de Jager J, McCann E. Psychosis as a barrier to the expression of sexuality and intimacy: an environmental risk? Schizophr Bull (in press).
4. Simiyon M, Chandra PS, Desai G. Sexual dysfunction among women with schizophrenia. A cross sectional study from India. Asian J Psychiatry 2016;24:93-8.
5. Hou CL, Zang Y, Rosen RC et al. Sexual dysfunction and its impact on quality of life in Chinese patients with schizophrenia treated in primary care. Compr Psychiatry 2016;65:116-21.
6. Lee JY, Kim SW, Lee YH et al. Factors associated with self-rated sexual function in Korean patients with schizophrenia receiving risperidone monotherapy. Hum Psychopharmacol 2015;30:416-24.
7. Montejo AL, Majadas S, Montejo L. Sexual and relational dysfunctions in people with schizophrenia. Eur Psychiatry 2014;29(Suppl. 1):1.
8. Gonzalez-Torres MA, Salazar MA, Inchausti L et al. Lifetime sexual behavior of psychiatric inpatients. J Sex Med 2010;7:3045-56.
9. Druyts E, Eapen S, Wu P et al. The risk of elevated prolactin levels in pediatric patients exposed to antipsychotics for the treatment of schizophrenia and schizophrenia spectrum disorders: protocol for a systematic review and meta-analysis. Syst Rev 2014;3:116.
10. Bobes J, Garcia-Portilla MP, Rejas J et al. Frequency of sexual dysfunction and other reproductive side-effects in patients with schizophrenia treated with risperidone, olanzapine, quetiapine, or haloperidol: the results of the EIRE study. J Sex Marital Ther 2003;29:125-47.
11. Chiesa A, Leucci V, Serretti A. Antipsychotics and sexual dysfunction: epidemiology, mechanisms and management. Clin Neuropsychiatry 2013; 10:31-6.
12. Uçok A, Incesu C, Aker T et al. Sexual dysfunction in patients with schizophrenia on antipsychotic medication. Eur Psychiatry 2007;22:328-33.
13. Montejo AL, Majadas S, Rico-Villademoros F et al. Frequency of sexual dysfunction in patients with a psychotic disorder receiving antipsychotics. J Sex Med 2010;7:3404-13.
14. Sathish Kumar SV, Sinha VK. Comparative study of sexual dysfunction and serum prolactin level associated with olanzapine, risperidone, and clozapine in patients with remitted schizophrenia. Indian J Psychiatry 2015;57:386-91.
15. Malik P, Kemmler G, Hummer M et al. Sexual dysfunction in first-episode schizophrenia patients: results from European First Episode Schizophrenia Trial. J Clin Psychopharmacol 2011;31:274-80.
16. Weiden PJ, Miller AL. Which side effects really matter? Screening for common and distressing side effects of antipsychotic medications. J Psychiatr Pract 2001;7:41-7.
17. Olfson M, Uttaro T, Carson WH et al. Male sexual dysfunction and quality of life in schizophrenia. J Clin Psychiatry 2005;66:331-8.
18. de Boer MK, Castelein S, Wiersma D et al. The facts about sexual (dys)function in schizophrenia: an overview of clinically relevant findings. Schizophr Bull 2015;41:674-86.
19. Haddad PM, Wieck A. Antipsychotic-induced hyperprolactinaemia: mechanisms, clinical features and management. Drugs 2004;64:2291-314.
20. Besnard I, Auclair V, Callery G et al. Antipsychotic-drug-induced hyperprolactinemia: physiopathology, clinical features and guidance. Encephale 2014;40:86-94.
21. Knegtering R, Castelein S, Bous H et al. A randomized open-label study of the impact of quetiapine versus risperidone on sexual functioning. J Clin Psychopharmacol 2004;24:56-61.

22. Knegtering H, Boks M, Blijd C et al. A randomized open-label comparison of the impact of olanzapine versus risperidone on sexual functioning. J Sex Marital Ther 2006;32:315-26.
23. Serretti A, Chiesa A. A meta-analysis of sexual dysfunction in psychiatric patients taking antipsychotics. Int Clin Psychopharmacol 2011;26:130-40.
24. Potkin SG, Loze JY, Forray C et al. Reduced sexual dysfunction with aripiprazole once-monthly versus paliperidone palmitate: results from QUALIFY. Int Clin Psychopharmacol 2017;32:147-54.
25. Montes de Oca P, Macotela Y, Nava G et al. Prolactin stimulates integrin-mediated adhesion of circulating mononuclear cells to endothelial cells. Lab Invest 2005;85:633-42.
26. Molinari C, Grossini E, Mary DA et al. Prolactin induces regional vasoconstriction through the beta2-adrenergic and nitric oxide mechanisms. Endocrinology 2007;148:4080-90.
27. Rosenberg KP, Bleiberg KL, Koscis J et al. A survey of sexual side effects among severely mentally ill patients taking psychotropic medications: impact on compliance. J Sex Marital Ther 2003;29:289-96.
28. Tharoor H, Kaliappan A, Gopal S. Sexual dysfunctions in schizophrenia: professionals and patients perspectives. Indian J Psychiatry 2015;57:85-7.
29. De Hert M, Detraux J, Peuskens J. Second-generation and newly approved antipsychotics, serum prolactin levels and sexual dysfunctions: a critical literature review. Expert Opin Drug Saf 2014;13:605-24.
30. de Boer MK, Castelein S, Bous J, et al. The Antipsychotics and Sexual Functioning Questionnaire (ASFQ): preliminary evidence for reliability and validity. Schizophr Res 2013;150:410-5.
31. Clayton AH, McGarvey EL, Clavet GJ. The Changes in Sexual Functioning Questionnaire (CSFQ): development, reliability, and validity. Psychopharmacol Bull 1997;33:731-45.
32. Montejo AL, Rico-Villademoros F. Psychometric properties of the Psychotropic-Related Sexual Dysfunction Questionnaire (PRSexDQ-SALSEX) in patients with schizophrenia and other psychotic disorders. J Sex Marital Ther 2008;34:227-39.
33. de Boer MK, Castelein S, Wiersma D et al. A systematic review of instruments to measure sexual functioning in patients using antipsychotics. J Sex Res 2014;51:383-9.
34. Kelly DL, Conley RR. Sexuality and schizophrenia: a review. Schizophr Bull 2004;30:767-79.
35. Bebbington PE, Angermeyer M, Azorin JM et al. Side-effects of antipsychotic medication and health-related quality of life in schizophrenia. Acta Psychiatr Scand 2009;119(Suppl. 438):22-8.
36. Dibonaventura M, Gabriel S, Dupclay L et al. A patient perspective of the impact of medication side effects on adherence: results of a cross-sectional nationwide survey of patients with schizophrenia. BMC Psychiatry 2012;12:20.
37. Montejo AL, Riesgo Y, Luque J et al. Observational, open-label, prospective multicenter study of sexual function in patients starting treatment with aripiprazole. Actas Esp Psiquiatr 2010;38:13-21.
38. Jeong HG, Lee MS, Lee HY et al. Changes in sexual function and gonadal axis hormones after switching to aripiprazole in male schizophrenia patients: a prospective pilot study. Int Clin Psychopharmacol 2012;27:177-83.
39. Lu ML, Shen WW, Chen CH. Time course of the changes in antipsychotic-induced hyperprolactinemia following the switch to aripiprazole. Prog Neuropsychopharmacol Biol Psychiatry 2008;32:1978-81.
40. Lee BH, Kim YK, Park SH. Using aripiprazole to resolve antipsychotic-induced symptomatic hyperprolactinemia: a pilot study. Prog Neuropsychopharmacol Biol Psychiatry 2006;30:714-7.
41. Meng M, Li W, Zhang S et al. Using aripiprazole to reduce antipsychotic-induced hyperprolactinemia: meta-analysis of currently available randomized controlled trials. Shanghai Arch Psychiatry 2015;27:4-17.
42. Fujioi J, Iwamoto K, Banno M et al. Effect of adjunctive aripiprazole on sexual dysfunction in schizophrenia: a preliminary open-label study. Pharmacopsychiatry 2017;50:74-8.
43. Montejo González AL, Rico-Villademoros F, Tafalla M et al. A 6-month prospective observational study on the effects of quetiapine on sexual functioning. J Clin Psychopharmacol 2005;25:533-8.
44. Montejo AL, Rico-Villademoros F, Daniel E et al. Changes in sexual function for outpatients with schizophrenia or other psychotic disorders treated with ziprasidone in clinical practice settings: a 3-month, prospective, observational study. J Clin Psychopharmacol 2008;28:568-70.
45. Schmidt HM, Hagen M, Kriston L et al. Management of sexual dysfunction due to antipsychotic drug therapy. Cochrane Database Syst Rev 2012;11:CD003546.
46. Montejo ÁL, Arango C, Bernardo M et al. Multidisciplinary consensus on the therapeutic recommendations for iatrogenic hyperprolactinemia secondary to antipsychotics. Front Neuroendocrinol 2017;45:25-34.
47. Basson R, Rees P, Wang R et al. Sexual function in chronic illness. J Sex Med 2010;7:374-88.
48. Nicolosi A, Laumann EO, Glasser DB et al. Sexual behavior and sexual dysfunctions after age 40: the global study of sexual attitudes and behaviors. Urology 2004;64:991-7.
49. Atlantis E, Sullivan T. Bidirectional association between depression and sexual dysfunction: a systematic review and meta-analysis. J Sex Med 2012;9:1497-507.
50. Angst J. Sexual problems in healthy and depressed persons. Int Clin Psychopharmacol 1998;13:S1-4.
51. Cyranowski JM, Bromberger J, Youk A et al. Lifetime depression history and sexual function in women at midlife. Arch Sex Behav 2004;33:539-48.
52. Vanwesenbeeck I, Have MT, de Graaf R. Associations between common mental disorders and sexual dissatisfaction in the general population. Br J Psychiatry 2014;205:151-7.
53. Hegeman JM, Kok RM, van der Mast RC et al. Phenomenology of depression in older compared with younger adults: meta-analysis. Br J Psychiatry 2012;200:275-81.
54. Lin C-F, Juang Y-Y, Wen J-K et al. Correlations between sexual dysfunction, depression, anxiety, and somatic symptoms among patients with major depressive disorder. Chang Gung Med J 2012;35:323-31.
55. Laurent SM, Simons AD. Sexual dysfunction in depression and anxiety: conceptualizing sexual dysfunction as part of an internalizing dimension. Clin Psychol Rev 2009;29:573-85.
56. Vulink NCC, Denys D, Bus L et al. Sexual pleasure in women with obsessive-compulsive disorder? J Affect Disord 2006;91:19-25.
57. Aksaray G, Yelken B, Kaptanoglu C et al. Sexuality in women with obsessive compulsive disorder. J Sex Marital Ther 2001;27:273-7.
58. Williams K, Reynolds MF. Sexual dysfunction in major depression. CNS Spectr 2006;11:19-23.
59. Nazareth I, Boynton P, King M. Problems with sexual function in people attending London general practitioners: cross sectional study. BMJ 2003;327:423-6.
60. Read S, King M, Watson J. Sexual dysfunction in primary medical care: prevalence, characteristics and detection by the general practitioner. J Publ Health Med 1997;19:387-91.
61. Rief W, Nestoriuc Y, von Lilienfeld-Toal A et al. Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: a systematic review and meta-analysis. Drug Saf 2009;32:1041-56.
62. Haberfellner EM. A review of the assessment of antidepressant-induced sexual dysfunction used in randomized, controlled clinical trials. Pharmacopsychiatry 2007;40:173-82.
63. McGahuey CA, Gelenberg AJ, Laukes CA et al. The Arizona Sexual Experience Scale (ASEX): reliability and validity. J Sex Marital Ther 2000;26:25-40.
64. Kennedy SH, Rizvi SJ, Fulton K et al. The Sex Effects Scale: pilot validation in a healthy population. Psychopharmacol Bull 2010;43:15-25.
65. Williams VSL, Baldwin DS, Hogue SL et al. Estimating the prevalence and impact of antidepressant-induced sexual dysfunction in 2 European countries: a cross-sectional patient survey. J Clin Psychiatry 2006;67:204-10.
66. Williams VSL, Edin HM, Hogue SL et al. Prevalence and impact of antidepressant-associated sexual dysfunction in three European countries: replication in a cross-sectional patient survey. J Psychopharmacol 2010;24:489-96.
67. Serretti A, Chiesa A. Treatment-emergent sexual dysfunction related to antidepressants: a meta-analysis. J Clin Psychopharmacol 2009;29:259-66.
68. Gartlehner G, Hansen RA, Morgan LC et al. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. Ann Intern Med 2011;155:772-85.
69. Demyttenaere K, Jaspers L. Bupropion and SSRI-induced side effects. J Psychopharmacol 2008;22:792-804.
70. Reichenpfader U, Gartlehner G, Morgan LC et al. Sexual dysfunction associated with second-generation antidepressants in patients with major depressive disorder: results from a systematic review with network meta-analysis. Drug Saf 2014;37:19-31.
71. Watanabe N, Omori IM, Nakagawa A et al. Mirtazapine versus other antidepressive agents for depression. Cochrane Database Syst Rev 2011;12:CD006528.

72. Benelli A, Frigeri C, Bertolini A et al. Influence of mirtazapine on the sexual behavior of male rats. Psychopharmacology 2004;171:250-8.
73. Montejo A, Majadas S, Rizvi SJ et al. The effects of agomelatine on sexual function in depressed patients and healthy volunteers. Hum Psychopharmacol 2011;26:537-42.
74. Montejo AL, Prieto N, Terleira A et al. Better sexual acceptability of agomelatine (25 and 50 mg) compared with paroxetine (20 mg) in healthy male volunteers. An 8-week, placebo-controlled study using the PRSEXDQ-SALSEX scale. J Psychopharmacol 2010;24:111-20.
75. Kennedy SH, Rizvi S, Fulton K et al. A double-blind comparison of sexual functioning, antidepressant efficacy, and tolerability between agomelatine and venlafaxine XR. J Clin Psychopharmacol 2008;28:329-33.
76. Hickie IB, Rogers NL. Novel melatonin-based therapies: potential advances in the treatment of major depression. Lancet 2011;378:621-31.
77. Montejo AL, Deakin JF, Gaillard R et al. Better sexual acceptability of agomelatine (25 and 50 mg) compared to escitalopram (20 mg) in healthy volunteers. A 9-week, placebo-controlled study using the PRSexDQ scale. J Psychopharmacol 2015;29:1119-28.
78. Gocmez S, Utkan T, Gacar N. Chronic administration of imipramine but not agomelatine and moclobemide affects the nitrergic relaxation of rabbit corpus cavernosum smooth muscle. Eur J Pharmacol 2013;714:442-7.
79. Schwartz TL, Siddiqui UA, Stahl SM. Vilazodone: a brief pharmacological and clinical review of the novel serotonin partial agonist and reuptake inhibitor. Ther Adv Psychopharmacol 2011;1:81-7.
80. Clayton AH, Kennedy SH, Edwards JB et al. The effect of vilazodone on sexual function during the treatment of major depressive disorder. J Sex Med 2013;10:2465-76.
81. Citrome L. Vilazodone for major depressive disorder: a systematic review of the efficacy and safety profile for this newly approved antidepressant - what is the number needed to treat, number needed to harm and likelihood to be helped or harmed? Int J Clin Pract 2012;66:356-68.
82. Baldwin DS, Hanumanthaiah VB. Vortioxetine in the treatment of major depressive disorder. Future Neurol (in press).
83. Zourkova AEH. Paroxetine-induced conversion of cytochrome P450 2D6 phenotype and occurence of adverse effects. Gen Physiol Biophys 2003; 22:103-13.
84. Zourkova A, Ceskova E, Hadasova E et al. Links among paroxetine-induced sexual dysfunctions, gender, and CYP2D6 activity. J Sex Marital Ther 2007;33:343-55.
85. Zourkova A, Slanar O, Jarkovsky J et al. MDR1 in paroxetine-induced sexual dysfunction. J Sex Marital Ther 2013;39:71-8.
86. Waldinger MD. Premature ejaculation: state of the art. Urol Clin North Am 2007;34:591-9.
87. Giuliano F, Hellstrom WJG. The pharmacological treatment of premature ejaculation. BJU Int 2008;102:668-75.
88. Hutchinson K, Cruickshank K, Wylie K. A benefit-risk assessment of dapoxetine in the treatment of premature ejaculation. Drug Saf 2012;35:359-72.
89. Jern P, Johansson A, Piha J et al. Antidepressant treatment of premature ejaculation: discontinuation rates and prevalence of side effects for dapoxetine and paroxetine in a naturalistic setting. Int J Impot Res 2015;27:75-80.
90. Fink HA, Macdonald R, Rutks IR et al. Trazodone for erectile dysfunction: a systematic review and meta-analysis. BJU Int 2003;92:441-6.
91. Baldwin D, Bridgman K, Buis C. Resolution of sexual dysfunction during double-blind treatment of major depression with reboxetine or paroxetine. J Psychopharmacol 2006;20:91-6.
92. Baldwin D, More RA, Brlley M. Resolution of sexual dysfunction during acute treatment of major depression with milnacipran. Hum Psychopharmacol 2008;23:527-32.
93. Clayton AH, Reddy S, Focht K et al. An evaluation of sexual functioning in employed outpatients with major depressive disorder treated with desvenlafaxine 50mg or placebo. J Sex Med 2013;10:768-76.
94. Werneke U, Northey S, Bhugra D. Antidepressants and sexual dysfunction. Acta Psychiatr Scand 2006;114:384-97.
95. Crawford AA, Lewis S, Nutt D et al. Adverse effects from antidepressant treatment: randomised controlled trial of 601 depressed individuals. Psychopharmacology 2014;231:2921-31.
96. Hu XH, Bull SA, Hunkeler EM et al. Incidence and duration of side effects and those rated as bothersome with selective serotonin reuptake inhibitor treatment for depression: patient report versus physician estimate. J Clin Psychiatry 2004;65:959-65.
97. Taylor MJ, Rudkin L, Hawton K. Strategies for managing antidepressant-induced sexual dysfunction: systematic review of randomised controlled trials. J Affect Disord 2005;88:241-54.
98. Clayton AH, Montejo AL. Major depressive disorder, antidepressants, and sexual dysfunction. J Clin Psychiatry 2006;67(Suppl. 6):33-7.
99. Baldwin DS, Palazzo MC, Masdrakis VG. Reduced treatment-emergent sexual dysfunction as a potential target in the development of new antidepressants. Depress Res Treat 2013;2013:256841.
100. Balon R, Segraves RT. Survey of treatment practices for sexual dysfunction(s) associated with anti-depressants. J Sex Marital Ther 2008;34:353-65.
101. Rothschild AJ. Selective serotonin reuptake inhibitor-induced sexual dysfunction: efficacy of a drug holiday. Am J Psychiatry 1995;152:1514-6.
102. Baldwin DS. Sexual dysfunction associated with antidepressant drugs. Expert Opin Drug Saf 2004;3:457-70.
103. Amiaz R, Pope HG Jr, Mahne T et al. Testosterone gel replacement improves sexual function in depressed men taking serotonergic antidepressants: a randomized, placebo-controlled clinical trial. J Sex Marital Ther 2011;37:243-54.
104. Nurnberg HG, Hensley PL, Gelenberg AJ et al. Treatment of antidepressant-associated sexual dysfunction with sildenafil - A randomized controlled trial. JAMA 2003;289:56-64.
105. Nurnberg HG, Hensley PL, Heiman JR et al. Sildenafil treatment of women with antidepressant-associated sexual dysfunction: a randomized controlled trial. JAMA 2008;300:395-404.
106. Segraves RT, Lee J, Stevenson R et al. Tadalafil for treatment of erectile dysfunction in men on antidepressants. J Clin Psychopharmacol 2007;27: 62-6.
107. Michelson D, Kociban K, Tamura R et al. Mirtazapine, yohimbine or olanzapine augmentation therapy for serotonin reuptake-associated female sexual dysfunction: a randomized, placebo controlled trial. J Psychiatr Res 2002;36:147-52.
108. Fava M, Dording CM, Baker RA et al. Effects of adjunctive aripiprazole on sexual functioning in patients with major depressive disorder and an inadequate response to standard antidepressant monotherapy: a post hoc analysis of 3 randomized, double-blind, placebo-controlled studies. Prim Care Companion CNS Disord 2011;13:PCC.10m00994.
109. Lorenz TA, Meston CM. Exercise improves sexual function in women taking antidepressants: results from a randomised crossover trial. Depress Anxiety 2014;31:188-95.
110. Andersson KE. Mechanisms of penile erection and basis for pharmacological treatment of erectile dysfunction. Pharmacol Rev 2011;63:811-59.
111. Musicki B, Liu T, Lagoda GA et al. Endothelial nitric oxide synthase regulation in female genital tract structures. J Sex Med 2009;6:247-53.
112. Nurnberg HG, Seidman SN, Gelenberg AJ et al. Depression, antidepressant therapies, and erectile dysfunction: Clinical trials of sildenafil citrate (Viagra) in treated and untreated patients with depression. Urology 2002; 60:58-66.
113. Rosen RC, Seidman SN, Menza MA et al. Quality of life, mood, and sexual function: a path analytic model of treatment effects in men with erectile dysfunction and depressive symptoms. Int J Impot Res 2004;16:334-40.
114. Kennedy SH, Dugre H, Defoy I. A multicenter, double-blind, placebo-controlled study of sildenafil citrate in Canadian men with erectile dysfunction and untreated symptoms of depression, in the absence of major depressive disorder. Int Clin Psychopharmacol 2011;26:151-8.
115. Wegener G, Harvey BH, Bonefeld B et al. Increased stress-evoked nitric oxide signalling in the Flinders sensitive line (FSL) rat: a genetic animal model of depression. Int J Neuropsychopharmacol 2010;13:461-73.
116. Liebenberg N, Harvey BH, Brand L et al. Chronic treatment with the phosphodiesterase type 5 inhibitors sildenafil and tadalafil display anxiolytic effects in Flinders Sensitive Line rats. Metab Brain Dis 2012;27:337-40.
117. Brink CB, Clapton JD, Eagar BE et al. Appearance of antidepressant-like effect by sildenafil in rats after central muscarinic receptor blockade: evidence from behavioural and neuro-receptor studies. J Neural Transm 2008;115:117-25.
118. Kopeykina I, Kim H-J, Khatun T et al. Hypersexuality and couple relationships in bipolar disorder: a review. J Affect Disord 2016;195:1-14.
119. Mazza M, Harnic D, Catalano V et al. Sexual behavior in women with bipolar disorder. J Affect Disord 2011;131:364-7.
120. Raja M, Azzoni A. Sexual behavior and sexual problems among patients with severe chronic psychoses. Eur Psychiatry 2003;18:70-6.
121. Downey J, Friedman RC, Haase E et al. Comparison of sexual experience and behavior between bipolar outpatients and outpatients without mood disorders. Psychiatry J 2016;2016:5839181.
122. Samalin L, de Chazeron I, Vieta E et al. Residual symptoms and specific functional impairments in euthymic patients with bipolar disorder. Bipolar Disord 2016;18:164-73.

123. Dell'Osso L, Carmassi C, Carlini M et al. Sexual dysfunctions and suicidality in patients with bipolar disorder and unipolar depression. J Sex Med 2009;6:3063-70.
124. Grover S, Ghosh A, Sarkar S et al. Sexual dysfunction in clinically stable patients with bipolar disorder receiving lithium. J Clin Psychopharmacol 2014;34:475-82.
125. Chen LP, Murad MH, Paras ML et al. Sexual abuse and lifetime diagnosis of psychiatric disorders: systematic review and meta-analysis. Mayo Clin Proc 2010;85:618-29.
126. Romero S, Birmaher B, Axelson D et al. Prevalence and correlates of physical and sexual abuse in children and adolescents with bipolar disorder. J Affect Disord 2009;112:144-50.
127. Fountoulakis KN, Kasper S, Andreassen O et al. Efficacy of pharmacotherapy in bipolar disorder: a report by the WPA section on pharmacopsychiatry. Eur Arch Psychiatry Clin Neurosci 2012;262:1-48.
128. Elnazer HY, Sampson A, Baldwin D. Lithium and sexual dysfunction: an under-researched area. Hum Psychopharmacol Clin Exp 2015;30:66-9.
129. Zuncheddu C, Carpiniello B. Sexual dysfunctions and bipolar disorder: a study of patients submitted to a long-term lithium treatment. Clin Ter 2006;157:419-24.
130. Dols A, Sienaert P, van Gerven H et al. The prevalence and management of side effects of lithium and anticonvulsants as mood stabilizers in bipolar disorder from a clinical perspective. Int Clin Psychopharmacol 2013; 28:287-96.
131. Kesebir S, Toprak B, Baykaran B et al. The level of awareness on sexually transmitted diseases of patients with bipolar mood disorder and patients with heroine dependence. Nöro Psikiyatr Arşivi 2014;51:242-7.
132. Ghadirian AM, Ananble L, Belanger MC. Lithium, benzodiazepines, and sexual function in bipolar patients. Am J Psychiatry 1992;149:801-5.
133. Kuba R, Pohanka M, Zákopcan J et al. Sexual dysfunctions and blood hormonal profile in men with focal epilepsy. Epilepsia 2006;47:2135-40.
134. Herzog AG, Drislane FW, Schomer D et al. Differential effects of antiepileptic drugs on sexual function and hormones in men with epilepsy. Neurology 2005;65:1016-20.
135. Szupera Z. The role of the antiepileptic drugs at the development of the sexual dysfunctions in male epileptic patients. Ideggyogy Sz 2007;60:4-13.
136. Yang Y, Wang X. Sexual dysfunction related to antiepileptic drugs in patients with epilepsy. Expert Opin Drug Saf 2016;15:31-42.
137. Pacchiarotti I, Murru A, Kotzalidis GD et al. Hyperprolactinemia and medications for bipolar disorder: systematic review of a neglected issue in clinical practice. Eur Neuropsychopharmacol 2015;25:1045-59.
138. La Torre A, Giupponi G, Duffy D et al. Sexual dysfunction related to psychotropic drugs: a critical review. Part III: mood stabilizers and anxiolytic drugs. Pharmacopsychiatry 2014;47:1-6.
139. Verrotti A, D'Egidio C, Mohn A et al. Antiepileptic drugs, sex hormones, and PCOS. Epilepsia 2011;52:199-211.
140. Schneck CD, Thomas MR, Gundersen D. Sexual side effects associated with valproate. J Clin Psychopharmacol 2002;22:532-4.
141. Verrotti A, Mencaroni E, Cofini M et al. Valproic acid metabolism and its consequences on sexual functions. Curr Drug Metab 2016;17:573-81.
142. Murru A, Popovic D, Pacchiarotti I et al. Management of adverse effects of mood stabilizers. Curr Psychiatry Rep 2015;17:66.
143. Pavone C, Giacalone N, Vella M et al. Relation between sexual dysfunctions and epilepsy, type of epilepsy, type of antiepileptic drugs: a prospective study. Urologia 2017;84:88-92.
144. Boora K, Chiappone K, Dubovsky SL. Oxcarbazepine-induced reversible anorgasmia and ejaculatory failure: a case report. Prim Care Companion J Clin Psychiatry 2009;11:173-4.
145. Calabrò RS, Italiano D, Pollicino P et al. Oxcarbazepine-related retrograde ejaculation. Epilepsy Behav 2012;25:174-5.
146. Harden CL. Sexual dysfunction in women with epilepsy. Seizure 2008;17: 131-5.
147. Bowden CL, Asnis GM, Ginsberg LD et al. Safety and tolerability of lamotrigine for bipolar disorder. Drug Saf 2004;27:173-84.
148. Yogarajah M, Mula M. Sexual dysfunction in epilepsy and the role of anti-epileptic drugs. Curr Pharm Des (in press).
149. Saroukhani S, Emami-Parsa M, Modabbernia A et al. Aspirin for treatment of lithium-associated sexual dysfunction in men: randomized double-blind placebo-controlled study. Bipolar Disord 2013;15:650-6.
150. Gil-Nagel A, López-Muñoz F, Serratosa JM et al. Effect of lamotrigine on sexual function in patients with epilepsy. Seizure 2006;15:142-9.
151. Husain AM, Carwille ST, Miller PP et al. Improved sexual function in three men taking lamotrigine for epilepsy. South Med J 2000;93:335-6.

DOI:10.1002/wps.20509

Insight in schizophrenia spectrum disorders: relationship with behavior, mood and perceived quality of life, underlying causes and emerging treatments

Paul H. Lysaker[1], Michelle L. Pattison[2], Bethany L. Leonhardt[3], Scott Phelps[4], Jenifer L. Vohs[5]

[1]Roudebush VA Medical Center, Indiana University School of Medicine, Indianapolis, IN, USA; [2]College of Applied Behavioral Sciences, University of Indianapolis, Indianapolis, IN, USA; [3]Indiana University School of Medicine, Eskenazi Health-Midtown Community Mental Health, Indianapolis, IN, USA; [4]Wesleyan University, Middletown, CT, USA; [5]Indiana University School of Medicine, Indianapolis, IN, USA

Poor insight in schizophrenia is prevalent across cultures and phases of illness. In this review, we examine the recent research on the relationship of insight with behavior, mood and perceived quality of life, on its complex roots, and on the effects of existing and emerging treatments. This research indicates that poor insight predicts poorer treatment adherence and therapeutic alliance, higher symptom severity and more impaired community function, while good insight predicts a higher frequency of depression and demoralization, especially when coupled with stigma and social disadvantage. This research also suggests that poor insight may arise in response to biological, experiential, neuropsychological, social-cognitive, metacognitive and socio-political factors. Studies of the effects of existing and developing treatments indicate that they may influence insight. In the context of earlier research and historical models, these findings support an integrative model of poor insight. This model suggests that insight requires the integration of information about changes in internal states, external circumstances, others' perspectives and life trajectory as well as the multifaceted consequences and causes of each of those changes. One implication is that treatments should, beyond providing education, seek to assist persons with schizophrenia to integrate the broad range of complex and potentially deeply painful experiences which are associated with mental illness into their own personally meaningful, coherent and adaptive picture.

Key words: Insight, schizophrenia, treatment adherence, therapeutic alliance, antipsychotic medication, depression, quality of life, neurocognition, social cognition, metacognition, stigma, psychotherapy, recovery

(World Psychiatry 2018;17:12–23)

The concept of insight into psychiatric disorders has long referred to the awareness of illness. In 1882, Pick[1] defined insight as a patient's recognition of "the pathological aspect of his mental processes, or some part of them, more or less clearly". For Pick, insight always involved a varying "degree of lucidity", with the weakest form of insight referred to as "illness-feeling", and the strongest full-fledged form of insight referred to as "illness-insight", denoting a cognitive process of conscious reflection and reason.

In 1934, Lewis[2] defined insight into mental illness as "a correct attitude to a morbid change in oneself" and "a matter of judgment", which could be achieved by inference, or by "what might be called secondary evidence of change in oneself – a lessened capacity". Other authors[3] further clarified that the judgments about having a mental illness were more than matters of perception and could be influenced by a patient's particular "culture and personality".

During the Second World War, there was a shift from understanding insight as a matter of varying degrees of denial to a failure to correctly perceive specific psychopathological elements[4]. Poor insight, from this time to the present, began to be equated with anosognosia, or the failure in neurological conditions to apprehend problems obvious to others, such as gross losses in mobility, hearing or speech. This view, which is referenced in the DSM-5, positions poor insight primarily as a barrier to treatment adherence. Indeed, it continues to be commonly proposed that recovery from mental illness occurs when persons achieve insight and accept they are ill and successfully engage in pharmacological treatment, which leads to symptom remission and the achievement of psychosocial milestones[5].

Turning to contemporary research, poor insight in schizophrenia has come to encompass more than general unawareness, including unawareness of symptoms, treatment need, the consequences of illness and alterations in cognitive processes. It has been found across cultures[6,7], and in early as well as later[8-11], and in acute as well as non-acute phases of illness[12].

Studies of insight, however, have increasingly suggested that this construct has a complex rather than linear relationship with health. Furthermore, it is associated with phenomena ranging from discrete cortical activity to broader self-understanding to larger social structures[13], all bearing on how we should conceptualize and treat poor insight within the field of mental health.

We summarize here the recent literature in this area, focusing on three issues: a) the effects of insight on behavior, mood and perceived quality of life; b) the complex roots of insight; and c) the effects of existing and emerging treatments on insight. We further delineate an emerging integrative model of insight, which not only helps clarify its paradoxical effects, but also has important implications for the development of more effective interventions.

The literature reviewed was found through a search of PubMed, Scopus, PsycINFO and EBSCO, including the search terms "insight", "awareness of illness", "self-reflection", "treatment adherence", "psychosis", and "schizophrenia". Articles published in English with publication dates of 2014 and onward were included in the review.

We considered two forms of insight: clinical insight, which involves awareness of symptoms, need for treatment, and psychosocial consequences of the disorder[14], and cognitive insight, which involves the capacity for self-reflectiveness and resistance to excessive certainty[15,16]. If not otherwise specified, the population studied included adults with a schizophrenia spectrum disorder.

EFFECTS OF INSIGHT ON BEHAVIOR, MOOD AND PERCEIVED QUALITY OF LIFE

Insight and treatment engagement

Clinical insight is associated with antipsychotic medication acceptance and attitudes

As noted above and elsewhere[10,13,17], poor clinical insight in schizophrenia has long been associated with negative attitudes towards taking antipsychotic medication and the decision to decline pharmacological treatments. Consistent with this, Misdrahi et al[18] and Lincoln et al[19] found that persons with poor insight were more likely to refuse to take antipsychotic medication, while others have found that groups with lower levels of adherence across cultures have consistently tended to have poorer clinical insight[20-23].

The relationship between insight and non-adherence may also occur early in the illness, as suggested by studies linking poor insight to a longer duration of untreated psychosis[24,25].

Concerning attitudes towards antipsychotic medication, Hui et al[24] found that poor insight predicted more negative appraisals of medication but not necessarily less adherence among psychosis patients. Others have similarly found that positive views of antipsychotic medication were related to better insight in general[26-29], regardless of gender[30].

Examining the relationship of insight and adherence over time, Zhou et al[31] found that poor insight was predictive of medication discontinuation at one-year follow-up. Czobor et al[32] also reported that poor insight predicted less adherence at 6 and 12 months follow-up in drug trials conducted in first-episode psychosis, while Abdel et al[33] found that insight predicted levels of post-discharge adherence.

Sui et al[34] examined the link of insight with non-adherence in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study, a nationwide public health-focused clinical trial of antipsychotic medication, and found that patients with poor insight who were most likely to discontinue treatment also had relatively more severe levels of psychopathology.

Of note, paradoxical findings have also been reported. Noordraven et al[35] found that better compliance was linked with the convergence of poor insight and high motivation among patients taking depot medication. A summary of the literature published between 2005 and 2015 found that poor insight was a predictor of poor adherence in 20 of 26 studies reviewed[36].

Clinical insight is associated with therapeutic alliance

Studying interpersonal elements of treatment, Lincoln et al[37] reported that lack of insight predicted drop out from cognitive behavior therapy in patients with psychosis. Poor insight was linked to poorer therapeutic alliance with a prescribing psychiatrist in schizophrenia and bipolar disorder.

Lack of insight has been reported to predict lower patient assessment of therapeutic alliance in cognitive therapy[38] and poorer clinician assessments of therapeutic alliance in general[39,40]. Better clinical insight has also been found to be associated with greater satisfaction with inpatient services[41].

Insight and outcome

Poorer clinical insight is associated with heightened symptoms

Poor insight has also long been linked to poorer clinical outcomes[10,13,17,42], with the joint possibilities that symptoms cause poor insight, and that poor insight causes less treatment adherence which may result in more symptoms.

Literature supporting this includes studies correlating poor clinical insight with overall symptom severity[43-45], positive symptoms[46,47], negative symptoms[44,48,49], disorganization symptoms [46,47,50-52], and excitement symptoms[46,47,53]. Of these studies, one was specific to first-episode psychosis[49], while others examined patients with comorbid trauma[49,52].

Evidence that insight influences symptoms includes findings that better insight predicts reduced risk of relapse[54] and improvement in positive and negative symptoms[55]. A better insight was also related to improvement in positive, negative and excitement symptoms in two first-episode psychosis samples[55,56].

Higher levels of symptoms at baseline also predicted lesser improvements in clinical insight over time in individuals with psychosis[57] and schizophrenia[58]. Of note, Kumar et al[45] found that greater levels of psychopathology did not predict poorer insight at later assessment time points. In addition, a study by Cobo et al[59] highlighted the importance of individual differences by demonstrating that the relationship of insight and symptoms is also influenced by gender, with insight having stronger links with disorganization and positive symptoms in men than women.

Poorer cognitive insight is associated with heightened symptoms

Kimhy et al[60] found that poor cognitive insight was linked with delusions in schizophrenia patients, but not in patients identified as at risk for developing schizophrenia. Others have found that poorer cognitive insight was linked with delusions in a high-risk group[61].

In a four-year longitudinal study, better cognitive insight in the domain of self-reflection was found to predict less severe

levels of overall psychopathology in first-episode psychosis[62]. Better cognitive insight at baseline was reported to predict greater improvements in cognition following cognitive remediation[63].

Other studies, however, have not found an association among cognitive insight and symptoms[64].

Poorer clinical insight is associated with poorer community functioning

Insight has also been proposed to have a bidirectional relationship with community function. Consistent with this, better clinical insight was correlated with better personal and social skills[65] and prosocial behavior[66]. Montemagni et al[67] found that clinical insight predicted the capacity for social connectedness and may mediate the relationship of negative symptoms to social function. Tastet et al[68] reported that better insight predicted a range of indices of higher levels of community function, including frequency of social contact and perceived social support.

Poor insight has also been found to predict a greater likelihood of involuntary legal commitment to treatment[69] and hospital readmission in first-episode psychosis[70].

Johnson et al[43] found that function and insight were only related to one another five years after a baseline assessment, suggesting that poor insight might have been a means of coping with, rather than a cause of poorer function. Chong et al[71] reported that baseline insight predicted function at six but not 12-month follow up in first-episode psychosis.

Cannavo et al[72] documented that improvement in psychosocial function was associated with significant improvements in clinical insight, while Klaas et al[73] found that good clinical insight had a positive effect upon social function, but that changes in social function did not influence insight in first-episode psychosis.

Of note, poor insight has been related to poorer premorbid function[74], and thus links with function may reflect pre-existing deficits.

Good clinical insight is associated with depression and poorer self-reported quality of life

In what may be referred to as the "insight paradox", earlier research suggested that good clinical insight may produce distress and compromise well-being[75]. Consistent with this, studies of diverse samples continue to find good insight to be correlated with depression and related constructs such as hopelessness[29,52,74,76-79].

Examining this relationship over time, research has also suggested that insight may cause depression. Evidence of this includes findings that good insight prospectively predicted higher levels of demoralization and that increased depressive symptoms were linked to improvements in insight in both first-episode psychosis and acute schizophrenia[39,56].

Good clinical insight has also been found to predict lower self-reported quality of life[80], with that relationship mediated by depression, suggesting that good insight may trigger depression, which may then degrade quality of life[80]. Schrank et al[81] found mutual and causal relationships in path analyses, with clinical insight promoting depression and self-stigma, and negative but not positive symptoms affecting that relationship.

The effects of insight on depression also appear to be moderated by a range of different factors. Replicating earlier research[75], several groups have found that the link between insight and depression was mediated by self-stigma[82-84]. Others have suggested that, beyond stigma, a generally negative appraisal of one's future[85] and tendency to ruminate[79] influence the effects of insight on mood.

Of note, some studies examining the relationship of clinical insight and depression in schizophrenia have failed to find significant relationships between these variables[86]. A meta-analysis of 50 cross-sectional studies found a weak relationship between clinical insight and depression, but reported that longitudinal studies suggested a stronger, more complex relationship which plays out differently depending upon the circumstances of the individual person's life[87]. This is consistent with clinical observations that, with insight, persons may discover negative things about their lives which are surprising and affect each individual differently[88].

Good cognitive insight is associated with depression and poorer self-reported quality of life

Greater levels of cognitive insight have also been linked to heightened levels of depression[77,86], as well as lower levels of self-reported subjective recovery[89] and subjective quality of life[90]. Phelan et al[91] suggested a complex relationship in which higher self-certainty was linked to better interviewer-rated quality of life only when patients had greater levels of psychopathology. By contrast, greater self-reflectivity in this study was related to better quality of life.

A meta-analysis of 17 studies suggested that the self-reflection dimension of cognitive insight, but not the self-certainty dimension, was responsible for the associations with depression and emotional distress[92]. These findings indicate that how one thinks about oneself may be most proximally linked to the path from insight to depression. Of note, in a non-clinical sample of 420 undergraduates, greater levels of cognitive insight were similarly linked to depression, suggesting that this phenomenon may be in effect beyond the boundaries of psychosis[93].

Good clinical and cognitive insight are associated with heightened suicidality

Consistent with the "insight paradox", clinical and cognitive insight have also been identified as potential risk factors for suicide.

Higher levels of clinical insight have been reported in a sample who had attempted versus one who had never attempted suicide[94]. Lopez-Moriñigo et al[95] reported that both

cognitive and clinical insight were directly linked to previous suicidality in inpatients with first-episode psychosis. Barrett et al[96] suggested that the relationship of insight and suicidality in first-episode psychosis patients may vary over time, with insight at baseline increasing and insight at one-year follow-up decreasing the risk of suicidality.

Other studies have found that the relationship of insight and suicidality may be fully mediated by depression[97].

THE COMPLEX ROOTS OF POOR INSIGHT IN SCHIZOPHRENIA

Multiple distinct phenomena which contribute to poor insight

Anomalous experiences as a root of poor insight

One challenge to the emergence of insight is the incomprehensibility and potentially traumatic nature of the unusual experiences which characterize schizophrenia[10]. From a medically oriented model, it has been proposed, for instance, that poor insight could reflect difficulties in making sense of the experience of positive and negative symptoms[10]. Evidence supporting this view includes the above-mentioned studies finding links between symptom severity and poor insight, especially prospectively[57].

From a phenomenological tradition, it has also been proposed that self-disturbance is not reducible to symptoms and leads persons to fail to accept a consensually valid reality and consequently fail to acknowledge mental illness[98]. Thus, poor clinical and cognitive insight could emerge in part when self-experience is simply impossible to understand.

Abnormalities in brain function as a root of poor insight

Focusing on brain function rather than experience, it has also been proposed that, in line with the anosognosia model, dysfunction in basic cortical activities may directly preclude the achievement of insight. Consistent with this, studies employing both structural[99] and functional[100-105] neuroimaging have implicated a number of areas and networks in the development and maintenance of insight.

Structural neuroimaging studies have linked poor insight to reduced volume in several areas, including frontal[104,106,107], temporal[107,108], parietal[104,107] and occipital[107] regions, the thalamus[101], basal ganglia[104] and cerebellum[107], in both first-episode psychosis[109] and chronic schizophrenia[107].

Reduced cortical thickness has been observed in those with poorer insight in the superior temporal gyrus, parahippocampus and insula[107] as well as in the ventral lateral prefrontal cortex in schizophrenia, and in other frontal, parietal and temporal areas in first-episode psychosis[110]. Structural changes in white matter integrity, suggesting impaired connectivity, have also been found in chronic[111] and early phase psychosis[100,112].

Compared to patients in a later phase of illness, individuals with first-episode psychosis may undergo more dramatic and accelerated gray matter loss in prefrontal, medial temporal and orbitofrontal regions[113], areas believed to be linked to poor insight. For instance, gray matter reduction in the cerebellum, prefrontal and temporal regions was reported in first-episode psychosis and associated with impaired insight[114].

In terms of functional imaging studies, converging evidence has now linked poor insight to activity in central midline structures, including basal ganglia[104], prefrontal cortex[103,104,115-117], cingulate cortex[103], insula[117], inferior parietal lobule[103] and the precuneus[116]. Functional resting state studies also suggest that insight may be related to the default mode network, particularly in the left hemisphere[102].

Despite these positive findings, some studies have failed to identify potential neural substrates associated with insight[118-120]. A number of methodological factors could contribute to these mixed findings, including inconsistent definitions of insight, varying instruments used to measure insight, and heterogeneous sample characteristics. The areas and networks linked to poor insight have also been implicated in various cognitive functions, including attention, executive functioning and memory. Further clarification of these links is needed to elucidate the possible neural substrates of insight.

Neurocognition as a root of poor insight

A third challenge to insight, related to abnormalities in brain function, are deficits in multiple neurocognitive domains. For instance, it is possible that deficits in attention, memory and executive functioning compromise individuals' abilities to access relevant memories about the experience of mental illness, to distinguish more from less salient aspects of those memories, to place those memories in sequence and to understand causal links[10,13].

Poor insight has been correlated with difficulties in set shifting in acute and non-acute phases of schizophrenia[121,122], the ability to change behavior following feedback on an executive function task[123], complex motor sequencing[124], and the ability to recall autobiographical details about negative events[125].

Bhagyavathi et al[126] reported a path analysis in which neurocognition affected insight, which then affected psychosocial function in schizophrenia. Vohs et al[11], by contrast, found clinical insight to be related to premorbid intelligence but not executive function in first-episode psychosis. Cernis et al[127] identified a subgroup with a premorbid IQ estimated as 120 or greater and found that they had better clinical insight than peers with typical levels of intellectual function. To explain these inconsistencies, it has been proposed that the relationship between insight and neurocognition may depend upon the phase of life[128,129] and that good neurocognitive function may be a necessary but not a sufficient condition for clinical insight[130].

Examining the relationship between neurocognition and clinical insight over time, Quee et al[57] reported that better baseline neurocognitive function predicted greater improvement in insight, while Chan et al[58] found that higher levels of perseveration at baseline predicted less improvement in insight in later assessments. Of note, in both of these studies, the proportion of the variance in insight accounted for by neurocognition was quite modest.

Consistent with an earlier meta-analysis[131] and systematic review[13], a more recent meta-analysis has found a very weak link between neurocognition and insight[132]. Other published studies continue to fail to find a link between insight and multiple facets of neurocognition in patients in varying phases of illness[47,51,133,134].

Cognitive insight also appears to be at best weakly related to neurocognition[132]. Higher levels of self-certainty have been correlated with poorer premorbid IQ, while lower levels of self-reflectivity were linked with poorer executive function in first-episode psychosis[135]. De Vos et al[123] reported that cognitive insight was related to performance on implicit process in a working priming 2-back task and the ability to change behavior following feedback on an executive function task. Ohmuro et al[136] found that higher self-certainty was related to poorer executive function in a group at risk for psychosis. Vohs et al[11] found no relationship between executive function and either dimension of cognitive insight in first-episode psychosis.

Social cognition as a root of poor insight

A fourth potential barrier to insight are deficits in social cognition, or the processes which allow people to grasp the meanings of social interactions and mental experiences of others[137], including theory of mind, affect recognition and attributional style[138]. Deficits in social cognition have been proposed to contribute to poor clinical insight when they block the opportunity to use the perspectives of others to understand past and present evidence of mental illness[139]. Consistent with this, Sanchez-Torres et al[51] reported that poor lifetime insight in psychosis was related to poorer social cognition.

Clinical insight has been linked to theory of mind in two studies, with that relationship persisting in the first[140], but not the second study[141], after controlling for symptom severity. Chan et al[48] found that global insight was linked to the ability to detect social *faux pas*. Zhang et al[64] found that people with schizophrenia classified as having high insight had better social cognitive abilities than those classified as having low insight, and that their social cognitive abilities were roughly equivalent to those of healthy controls.

Vohs et al[11] reported that clinical insight was linked with theory of mind abilities but not emotion recognition in first-episode psychosis. One study reported that cognitive insight was related to theory of mind tasks[141], although such a finding has not been replicated[11,64,140]. A meta-analysis has recently found a clinically significant but modest link between theory of mind and insight[142].

Metacognition as a root of poor insight

A fifth potential barrier to insight is metacognition, or the processes which allow persons to be aware of and form integrated and complex ideas about the self and others. While social cognition focuses on the correct detection of a discrete thought or feeling of another person, metacognition focuses on the integration of those details into a coherent whole, which varies more in terms of complexity rather than accuracy[143,144]. Metacognitive deficits thus could hypothetically limit a person's abilities to recognize changes in their mental states and see disruption across the larger course of their lives[139].

Consistent with this, a recent review[145] suggested that greater metacognitive capacity is related to higher levels of both clinical and cognitive insight, and that its contribution to both forms of insight is independent of symptom severity and neurocognitive function and may also protect against the emergence of depression in the wake of the development of insight.

Using other paradigms, poorer organization skills when reflecting upon oneself and less complex personal narratives have also been linked with poorer insight in schizophrenia[146,147]. Vohs et al[11] reported robust correlations between metacognitive capacity and multiple domains of clinical insight, as well as the self-certainty domain of cognitive insight, in first-episode psychosis. These findings are consistent with imaging studies which suggest that insight is related to cortical regions and circuits that may support processes necessary for metacognition, including self-consciousness and self-referential processing[102].

Social and political factors as a root of poor insight

Social issues which are beyond individuals have also been proposed to affect the development of insight. As reviewed above, stigma may result in insight leading to despair, and individuals may consider a meaningful life no longer a viable possibility[39,81]. Other studies have also found a direct link between insight and stigma[148], though others have failed to replicate that[149].

Beyond stigma, many people diagnosed with psychosis report that the diagnosis itself, especially when provided by institutions which appear to embrace stigma, is experienced as invalidating their potential identity and as an effort to control them in a paternalistic manner[150-152]. Indeed, independent studies have reported that clinical insight and stigma interact in a manner that results in the experience of reduced meaning[153] and self-clarity in life[154].

The ideas that persons form about what is and is not an illness are also affected by the larger social group, including family and culture. Clinical insight has been found to vary by cultural background[155,156]. Zisman-Ilani et al[157] similarly reported that insight among parents of adults with serious mental illness varied according to cultural backgrounds. Macgregor et al[158] found that the insight into the need for treatments by

parents of adult people with schizophrenia was affected by how frequently the parents and patients interacted and by the neurocognitive function of the parent. Raffard et al[159] found that better cognitive insight in the parents of adult patients was related to better clinical insight in the patients themselves.

An integrative model of insight

On the basis of the above review, an integrated model can be proposed in which insight in serious mental illness requires first the integration of multiple streams of information, including awareness of changes in internal states, external circumstances, the views of others and the larger trajectory of life. It then also requires the integration of information about the multifaceted consequences of each of those changes and their potential causes[139].

This model suggests that insight involves far more than a single or uniform level of unawareness and allows for multiple interacting causes. As summarized above, research continues to support this model, suggesting that poor insight could occur when persons cannot construct an account of their psychiatric challenges as a consequence of alterations in basic biological processes or structures, odd experiences, a loss of contact with the perspectives of others, the collapse of an integrated sense of self, and/or socio-political factors.

EFFECTS OF EXISTING AND EMERGING TREATMENTS ON INSIGHT

Applications of existing treatments to address insight

Psychological interventions which can affect insight

Earlier reviews established that poor insight was not a matter to be addressed through simple education[13]. Since then, some existing psychological interventions have shown promise in improving insight in psychosis.

Cognitive behavior therapy for psychosis (CBT-p) has been linked with improvements relative to treatment as usual in two separate trials[160,161]. A brief culturally adapted version of CBT-p was also linked with improvements in insight[162,163]. Drake et al[164] found that cognitive therapy which followed cognitive remediation was linked to greater improvements in clinical insight in psychosis. Motivational interviewing[165,166] and mindfulness based interventions[167-169] have also been linked to improved insight relative to treatment as usual.

Other psychological therapies that have also been reported to lead to improvement in insight relative to treatment as usual include integrated psychological therapy[170], a self-management skills program[171], and an adaptation to schizophrenia of an individualized treatment developed to help adults with diabetes accept their medical condition, called guided self-determination[172].

Pharmacological interventions which can affect insight

Pijnenborg et al[173] analyzed the effects of various antipsychotic drugs across trials for first-episode psychosis and observed a general effect of medications in leading to improved insight above gains that could be accounted for by symptom reduction, especially in the first three months of the trials.

Hou et al[174] reported that patients on clozapine were more likely than their peers on other agents to have better insight. Mattia et al[175], using databases from over 14 drug trials, found that second-generation antipsychotic medications were linked with improvements in insight in schizophrenia.

Developments of novel treatments to address insight

Metacognitive training can affect insight

Other treatments developed to target some of the underlying causes of poor insight have also shown promising results. The most broadly researched of these, metacognitive training for schizophrenia patients (MCT), was originally designed as an eight module group intervention, but has also been implemented individually[176].

Consistent with initial work reporting the acceptability and feasibility of MCT, Balzan et al[177] found that this treatment led to improvements in symptoms, cognitive biases and clinical insight relative to treatment as usual in a sample with mild delusions. Gaweda et al[178] reported that MCT led to significant improvements in clinical insight compared with treatment as usual, despite a lack of effects on psychotic symptoms, reasoning bias and theory of mind.

Detailed case work has described how insight improved after four weeks of MCT[179]. However, Briki et al[180] reported that, relative to treatment as usual, MCT led to improvements only at a trend level.

The links of MCT with cognitive insight have been equivocal. Lam et al[181] reported that MCT was linked to greater improvements in self-reflectiveness relative to treatment as usual. Ochoa et al[182] compared the effects of MCT to psychoeducation and found a positive impact on the self-reflection component of cognitive insight along with improvement in several facets related to reasoning style. Ussorio et al[183] compared the effects of MCT on patients with schizophrenia with a longer or shorter duration of untreated psychosis and found that both groups experienced gains in cognitive insight. However, van Oosterhout et al[184] found no effects on cognitive insight in a randomized controlled trial comparing MCT to treatment as usual.

Metacognitively oriented integrative psychotherapy can affect insight

A second intervention developed to target a core cause of poor insight is metacognitive reflection and insight therapy

(MERIT)[185]. This is an individual psychotherapy that seeks to enhance the reflective capacity necessary for adults who have experienced severe mental illness to form a complex and integrated sense of self and others. It targets the more synthetic integrative aspects of metacognition which could form the structures that enable insight.

Detailed case studies of MERIT have reported improvements in clinical insight in people with prolonged and early psychosis[186-189]. Vohs et al[190] recruited patients with first-episode psychosis with poor clinical insight and randomly assigned them to receive six months of MERIT vs. treatment as usual. Treatment completion rate was 80%. There were significant improvements in objective measures of clinical insight in the treatment group as compared to controls, and no evidence of heightened levels of depression.

SUMMARY AND FUTURE DIRECTIONS

Research over the last few years confirms that poor insight is a common and influential element of serious mental illness across cultures and phases of disorder. The path from insight to treatment engagement appears relatively straightforward, with poor insight leading to a greater likelihood of rejecting antipsychotic medication and greater difficulties in forming a therapeutic alliance. From that point, the path is complicated and paradoxically fails to match traditional models of wellness which position acceptance of illness as the first step towards health. Just as poor insight seems associated with higher levels of symptoms and poorer community function, better insight may lead to depression, demoralization and low levels of self-perceived quality of life, especially when coupled with stigma and social disadvantage.

Studies of the roots and correlates of poor insight indicate that this is more than a matter of a single failure in perception or apperception. This research evidence suggests that persons with serious mental illness fail to integrate the complex array of current and past information into a coherent representation allowing for adaptive decisions, as a result of multiple factors. Potential causes of poor insight include experiences which seem incomprehensible, cortical dysfunction, problems with memory, attention and executive function, deficits in social cognition and metacognition, cultural differences, and socio-political factors.

On the positive side, contrary to previous pessimistic appraisals, some existing treatments may lead to improvements in insight. Emerging treatments targeting the metacognitive processes that cause and sustain poor insight may also help persons with serious mental illness to reflect and make personal meaning of experiences of mental illness without the paradoxical negative consequences of good insight, such as depression or suicidality.

Longitudinal and more inclusive studies of insight over the course of schizophrenia are needed to understand the complex interactions between insight, treatment engagement, symptoms, function and well-being over time. Such work has the potential, for example, to explore whether the relationship of insight with treatment and function extends back before the onset of the illness, as suggested by associations of poor insight with duration of untreated psychosis and poorer premorbid social function. Additionally, are there instances in which improvements in self-esteem and stigma occur before and then allow for the development of insight? Research is also needed examining the long-term outcomes of persons with poor insight who drop out of treatment and never return, or those who refuse treatment across their lifetime.

The multifaceted nature of poor insight raises more questions than it answers. Do the neurocognitive, social cognitive and metacognitive correlates of poor insight predate the emergence of the illness? What are the neural pathways which connect abnormalities in brain function with disturbances in higher-order cognition paving the way for poor insight? Should forms of poor insight with biological, social or psychological correlates be distinguished from one another? For instance, intuitively, denial of illness in order to avoid social entrapment and humiliation should be different at many levels from poor insight secondary to deficits in executive function. Even at the symptoms level, is poor insight associated with negative symptoms a different phenomenon from poor insight linked with other symptoms? What are the processes which support a path from brain function to conscious integration of information about changes in internal states, external circumstances, the views of others and the larger trajectory of a life?

Concerning the conceptualization of insight, we began this review noting that the first views of insight cast it as a complex judgment which could vary in its complexity. This view gave way over a half century ago to all-or-nothing models which equated poor insight with anosognosia or a failure to perceive or apprehend a morbid change obvious to others. The complex relationships of insight with behavior, mood and perceived quality of life suggest that the anosognosia model is far from a complete account and may have been partially a wrong turn in the road.

Poor insight is not just the consequence of a failure to notice a problem, grasp a fact or accept a label. It is a failure to make consensually valid sense of complex and potentially traumatic experiences. This view is consistent with studies of first-person experience of psychosis, which demonstrate that insight is most significant to persons with psychosis when it helps them make sense or meaning of life events in a narratized manner[191-194].

The body of work we have reviewed has at least two major implications for the continuing development of interventions to address insight. First, the therapeutic action of these interventions likely goes well beyond providing information about mental illness and rests on the process of helping persons pull together a broad range of potentially deeply painful experiences into a coherent picture of their psychiatric challenges.

Essential here is that the experiences which could potentially be integrated into an adaptive form of awareness of mental illness are strikingly diverse and will vary from person to person. As an illustration, the elements which may inform insight could include historical events (e.g., hospitalization, deaths or major life changes), changes over time in thoughts and emotions (e.g., emergence of positive and negative symptoms, the ability to concentrate or a persistent sense of fear), changes in interpersonal function (e.g., changes in friendships, family relations, housing or romantic partners), changes in instrumental function (e.g., difficulties performing once familiar tasks in vocational settings), as well as internal and external stigma and alterations in social status (e.g., being subject to stigma by mental health professionals, feeling that being mentally ill means that one is less valuable than others, or that one should submit to the control of institutions).

The second implication is that treatments targeting insight must position persons with mental illnesses to make their own sense of their challenges. As this necessarily will also vary among persons, treatments which are flexible and responsive to individual needs are needed. This would be consistent with the rising recovery movement, which holds that wellness is about taking charge of and making sense of one's life and challenges.

REFERENCES

1. Pick A. Über Krankheitsbewusstsein in psychischen Krankheiten. Arch Psychiatr Nervenkr 1882;13:518-81.
2. Lewis A. The psychopathology of insight. Br J Med Psychol 1934;14:332-48.
3. Weinstein EA, Robert LK. Denial of illness: symbolic and physiological aspects. Springfield: Thomas, 1955.
4. Phelps S. Blind to their blindness: a history of the denial of illness. PhD Dissertation, Harvard University, Boston, 2014.
5. Kane JM. Improving patient outcomes in schizophrenia: achieving remission, preventing relapse, and measuring success. J Clin Psychiatry 2013;74: e18.
6. Schennach R, Meyer S, Seemuller F et al. Insight in schizophrenia-course and predictors during the acute treatment phase of patients suffering from a schizophrenia spectrum disorder. Eur Psychiatry 2012;27: 625-33.
7. Wang Y, Xiang YT, Wang CY et al. Insight in Chinese schizophrenia patients: a 12-month follow-up. J Psychiatr Ment Health Nurs 2011;18:751-7.
8. Cuesta MJ, Peralta V, Campos MS et al. Can insight be predicted in first-episode psychosis patients? A longitudinal and hierarchical analysis of predictors in a drug-naïve sample. Schizophr Res 2011;130:148-56.
9. Koren D, Viksman P, Giuliano AJ et al. The nature and evolution of insight in schizophrenia: a multi-informant longitudinal study of first-episode versus chronic patients. Schizophr Res 2013;151:245-51.
10. Osatuke K, Ciesla J, Kasckow JW et al. Insight in schizophrenia: a review of etiological models and supporting research. Compr Psychiatry 2008; 49:70-7.
11. Vohs JL, Lysaker PH, Liffick E et al. Metacognitive capacity as a predictor of insight in first-episode psychosis. J Nerv Ment Dis 2015;203:372-8.
12. Braw Y, Sitman R, Sela T et al. Comparison of insight among schizophrenia and bipolar disorder patients in remission of affective and positive symptoms: analysis and critique. Eur Psychiatry 2012;27:612-8.
13. Lysaker PH, Vohs J, Hillis JD et al. Poor insight into schizophrenia: contributing factors, consequences, and emerging treatment approaches. Expert Rev Neurother 2013;13:785-93.
14. Amador XF, Flaum M, Andreasen NC et al. Awareness of illness in schizophrenia and schizoaffective and mood disorders. Arch Gen Psychiatry 1994;51:826-36.
15. Martin JM, Warman DM, Lysaker PH. Cognitive insight in non-psychiatric individuals and individuals with psychosis: an examination using the Beck Cognitive Insight Scale. Schizophr Res 2010;121:39-45.
16. Riggs SE, Grant PM, Perivoliotis D et al. Assessment of cognitive insight: a qualitative review. Schizophr Bull 2012;38:338-50.
17. Lincoln TM, Lullman E, Rief W. Correlates and long-term consequences of poor insight in patients with schizophrenia. A systematic review. Schizophr Bull 2007;33:1324-42.
18. Misdrahi D, Tessier A, Swendesen J et al. Determination of adherence profiles in schizophrenia using self-reported adherence: results from the FACE-SZ dataset. J Clin Psychiatry 2016;77:1130-6.
19. Lincoln TM, Jung E, Wiesiahn M et al. The impact of negative treatment experiences on persistent refusal of antipsychotics. Compr Psychiatry 2016;70:165-73.
20. Chan KW, Hui LM, Wong HY et al. Medication adherence, knowledge about psychosis, and insight among patients with a schizophrenia-spectrum disorder. J Nerv Ment Dis 2014;202:25-9.
21. Chandra IS, Kumar KL, Reddy MP et al. Attitudes toward medication and reasons for non-compliance in patients with schizophrenia. Indian J Psychol Med 2014;36:294-8.
22. Eticha T, Teklu A, Ali D et al. Factors associated with medication adherence among patients with schizophrenia in Mekelle, Northern Ethiopia. PLoS One 2015;10:1-11.
23. Na E, Yim SJ, Lee JN et al. Relationships among medication adherence, insight, and neurocognition in chronic schizophrenia. Psychiatry Clin Neurosci 2015;69:298-304.
24. Hui CL, Lau WW, Leung CM et al. Clinical and social correlates of duration of untreated psychosis among adult-onset psychosis in Hong Kong Chinese: the JCEP study. Early Interv Psychiatry 2015;9:118-25.
25. O'Donoghue B, Lyne J, Kinsella A et al. Detection and characteristics of individuals with a very long duration of untreated psychosis in an early intervention for psychosis service. Early Interv Psychiatry 2014;8: 332-9.
26. Karthik MS, Warikoo N, Chakrabarti S et al. Attitudes towards antipsychotics among patients with schizophrenia on first- or second-generation medications. Indian J Psychol Med 2014;36:288-93.
27. Samalin L, de Chazeron I, Blanc O et al. Attitudes toward antipsychotic medications as a useful feature in exploring medication non-adherence in schizophrenia. Schizophr Res 2016;178:1-5.
28. Ramachandran AS, Ramanathan R, Praharai SK et al. A cross-sectional, comparative study of insight in schizophrenia and bipolar patients in remission. Indian J Psychol Med 2016;38:207-12.
29. Kako Y, Ito K, Hashimoto N et al. The relationship between insight and subjective experience in schizophrenia. Neuropsychiatr Dis Treat 2014; 10:1415-22.
30. Zhou J, Xiang YT, Li Q et al. Gender differences in attitudes towards antipsychotic medications in patients with schizophrenia. Psychiatry Res 2016;245:276-81.
31. Zhou Y, Rosenheck R, Mohamed S et al. Factors associated with complete discontinuation of medication among patients with schizophrenia in the year after hospital discharge. Psychiatry Res 2017;250:129-35.
32. Czobor P, Van Dorn RA, Citrome L et al. Treatment adherence in schizophrenia: a patient-level meta-analysis of combined CATIE and EUFEST studies. Eur Neuropsychopharmacol 2015;25:1158-66.
33. Abdel Aziz K, Elamin MH, El-Saadouni NM et al. Schizophrenia: impact of psychopathology, faith healers and psycho-education on adherence to medication. Int J Soc Psychiatry 2016;62:719-25.
34. Siu CO, Harvey PD, Agid O et al. Insight and subjective measures of quality of life in chronic schizophrenia. Schizophr Res Cogn 2015;2:127-32.
35. Noordraven EL, Wierdsma AI, Blanken P et al. Depot-medication compliance for patients with psychotic disorders: the importance of illness insight and treatment motivation. Neuropsychiatr Dis Treat 2016;12:269-74.
36. Velligan DI, Sajatovic M, Hatch A et al. Why do psychiatric patients stop antipsychotic medication? A systematic review of reasons for nonadherence to medication in patients with serious mental illness. Patient Prefer Adherence 2017;11:449-68.
37. Lincoln TM, Rief W, Westermann S et al. Who stays, who benefits? Predicting dropout and change in cognitive behaviour therapy for psychosis. Psychiatry Res 2014;216:198-205.
38. Berry K, Gregg L, Lobban F et al. Therapeutic alliance in psychological therapy for people with recent onset psychosis who use cannabis. Compr Psychiatry 2016;67:73-80.

39. Cavelti M, Rusch N, Vauth R. Is living with psychosis demoralizing? Insight, self-stigma, and clinical outcome among people with schizophrenia across 1 year. J Nerv Ment Dis 2014;202:521-9.
40. Ruchlewska A, Kamperman AM, van der Gaag M et al. Working alliance in patients with severe mental illness who need a crisis intervention plan. Community Ment Health J 2016;52:102-8.
41. Bo B, Ottesen OH, Gjestad R et al. Patient satisfaction after acute admission for psychosis. Nord J Psychiatry 2016;70:321-8.
42. Mintz AR, Dobson KS, Romney DM. Insight in schizophrenia: a meta-analysis. Schizophr Res 2003;61:75-88.
43. Johnson S, Sathyaseelan M, Charles H et al. Predictors of disability: a 5-year cohort study of first-episode schizophrenia. Asian J Psychiatry 2014; 9:45-50.
44. Bianchini O, Porcelli S, Nespeca C et al. Effects on antipsychotic drugs on insight in schizophrenia. Psychiatry Res 2014;218:20-4.
45. Kumar A, Sharma P, Das S et al. Insight in psychotic disorder: relation with psychopathology and frontal lobe function. Psychopathology 2014; 47:32-8.
46. Pousa E, Ochoa S, Cobo J et al. A deeper view of insight in schizophrenia: insight dimensions, unawareness and misattribution of particular symptoms and its relation with psychopathological factors. Schizophr Res 2017;189:61-8.
47. Zhou Y, Rosenheck R, Mohamed S et al. Insight in inpatients with schizophrenia: relationship to symptoms and neuropsychological functioning. Schizophr Res 2015;161:376-81.
48. Chan KK. Associations of symptoms, neurocognition, and metacognition with insight in schizophrenia spectrum disorders. Compr Psychiatry 2016;65:63-9.
49. Chang WC, Lau CF, Chan SS et al. Premorbid, clinical and cognitive correlates of primary negative symptoms in first-episode psychosis. Psychiatry Res 2016;242:144-9.
50. Nestsiarovich A, Obyedkov V, Kandratsenka H et al. Disorganization at the stage of schizophrenia clinical outcome: clinical-biological study. Eur Psychiatry 2016;42:44-8.
51. Sanchez-Torres AM, Zarzuela A, Peralta V et al. The association of lifetime insight and cognition in psychosis. Schizophr Res 2015;162:183-8.
52. Yanos PT, Vavshenker B, Pleskach P et al. Insight among people with severe mental illness, co-occurring PTSD and elevated psychotic symptoms: correlates and relationship to treatment participation. Compr Psychiatry 2016;68:172-7.
53. Volavka J, Van Dorn RA, Citrome L et al. Hostility in schizophrenia: an integrated analysis of the combined Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) and the European First Episode Schizophrenia Trial (EUFEST) studies. Eur Psychiatry 2016;31:13-9.
54. Berge D, Mane A, Salgado P et al. Predictors of relapse and functioning in first-episode psychosis: a two-year follow-up study. Psychiatry Serv 2016; 67:227-33.
55. Gumley AI, Schwannauer M, Macbeth A et al. Insight, duration of untreated psychosis and attachment in first-episode psychosis: prospective study of psychiatric recovery over 12-month follow-up. Br J Psychiatry 2014;205: 60-7.
56. Misiak B, Frydecka D, Beszlei JA et al. Effects of antipsychotics on insight in schizophrenia: results from independent samples of first-episode and acute relapsed patients. Int Clin Psychopharmacol 2016;31:185-91.
57. Quee PJ, van der Meer L, Krabbendam L et al. Insight change in psychosis: relationship with neurocognition, social cognition, clinical symptoms, and phase of illness. Acta Psychiatr Scand 2014;129:126-33.
58. Chan SK, Chan KK, Hui CL et al. Correlates of insight with symptomatology and executive function in patients with first-episode schizophrenia-spectrum disorder: a longitudinal perspective. Psychiatry Res 2014;216:177-84.
59. Cobo J, Nieto L, Ochoa S et al. Insight and gender in schizophrenia and other psychoses. Psychiatry Res 2016;243:268-77.
60. Kimhy D, Jobson-Ahmed L, Ben-David S et al. Cognitive insight in individuals at clinical high risk for psychosis. Early Interv Psychiatry 2014;8:130-7.
61. Uchida T, Matsumoto K, Ito F et al. Relationship between cognitive insight and attenuated delusional symptoms in individuals with at-risk mental state. Psychiatry Res 2014;217:20-4.
62. O'Connor JA, Ellett L, Ainakina O et al. Can cognitive insight predict symptom remission in a first episode psychosis cohort? BMC Psychiatry 2017;17:54.
63. Benoit A, Harvey PO, Bherer L et al. Does the Beck Cognitive Insight Scale predict response to cognitive remediation in schizophrenia? Schizophr Res Treatment 2016;2016:6371856.
64. Zhang Q, Li X, Parker GJ et al. Theory of mind correlates with clinical insight but not cognitive insight in patients with schizophrenia. Psychiatry Res 2016;237:188-95.
65. Erol A, Delibas H, Bora O et al. The impact of insight on social functioning in patients with schizophrenia. Int J Soc Psychiatry 2015;61:379-85.
66. Firmin RL, Luther L, Lysaker PH et al. Self-initiated helping behaviors and recovery in severe mental illness: implications for work, volunteerism, and peer support. Psychiatr Rehabil J 2015;38:336-41.
67. Montemagni C, Castagna F, Crivelli B et al. Relative contributions of negative symptoms, insight, and coping strategies to quality of life in stable schizophrenia. Psychiatry Res 2014;220:102-11.
68. Tastet H, Verdoux H, Bouisson J et al. Impact of interpersonal factors on insight in schizophrenia. Schizophr Res 2014;159:527-32.
69. Morandi S, Golay P, Lambert M et al. Community treatment order: identifying the need for more evidence based justification of its use in first episode psychosis patients. Schizophr Res 2016;185:67-72.
70. Drake RJ, Nordentoft M, Haddock G et al. Modeling determinants of medication attitudes and poor adherence in early nonaffective psychosis: implications for intervention. Schizophr Bull 2015;41:584-96.
71. Chong CS, Siu MW, Kwan CH et al. Predictors of functioning in people suffering from first-episode psychosis 1 year into entering early interventions service in Hong Kong. Early Interv Psychiatry (in press).
72. Cannayo D, Minutolo G, Battaglia E et al. Insight and recovery in schizophrenic patients. Int J Psychiatry Clin Pract 2016;20:83-90.
73. Klaas HS, Clemence A, Marion-Veyron R et al. Insight as a social identity process in the evolution of psychosocial functioning in the early phase of psychosis. Psychol Med 2017;47:718-29.
74. Ayesa-Arriola R, Moriñigo JD, David AS et al. Lack of insight 3 years after first-episode psychosis: an unchangeable illness trait determined from first presentation? Schizophr Res 2014;157:271-7.
75. Lysaker PH, Roe D, Yanos PT. Toward understanding the insight paradox: internalized stigma moderates the association between insight and social functioning, hope, and self-esteem among people with schizophrenia spectrum disorders. Schizophr Bull 2007;33:192-9.
76. Onwuameze OE, Uga A, Paradiso S. Longitudinal assessment of clinical risk factors for depression in schizophrenia spectrum disorders. Ann Clin Psychiatry 2016;28:167-74.
77. Misdrahi D, Denard S, Swendsen J et al. Depression in schizophrenia: the influence of the different dimensions of insight. Psychiatry Res 2014; 216:6-12.
78. Rossi A, Amore M, Galderisi S et al. The complex relationship between self-reported 'personal recovery' and clinical recovery in schizophrenia. Schizophr Res (in press).
79. Thomas N, Ribaux D, Phillips LJ. Rumination, depressive symptoms and awareness of illness in schizophrenia. Behav Cogn Psychother 2014;42:1-13.
80. Margariti M, Ploumpidis D, Economou M et al. Quality of life in schizophrenia spectrum disorders: association with insight and psychopathology. Psychiatry Res 2015;225:695-701.
81. Schrank B, Amering M, Hay AG et al. Insight, positive and negative symptoms, hope, depression and self-stigma: a comprehensive model of mutual influences in schizophrenia spectrum disorders. Epidemiol Psychiatr Sci 2014;23:271-9.
82. Belvederi Murri M, Amore M, Calcagno P et al. The "insight paradox" in schizophrenia: magnitude, moderators and mediators of the association between insight and depression. Schizophr Bull 2016;42:1225-33.
83. Lien YJ, Chang HA, Kao YC et al. Insight, self-stigma and psychosocial outcomes in schizophrenia: a structural equation modeling approach. Epidemiol Psychiatr Sci 2016;15:1-10.
84. Valiente C, Provencio M, Espinosa R et al. Insight in paranoia: the role of experiential avoidance and internalized stigma. Schizophr Res 2015;164: 214-20.
85. MacDougall AG, Vandermeer MR, Norman RM. Negative future self as a mediator in the relationship between insight and depression in psychotic disorders. Schizophr Res 2015;165:66-9.
86. Grover S, Sahoo S, Nehra R et al. Relationship of depression with cognitive insight and socio-occupational outcome in patients with schizophrenia. Int J Soc Psychiatry 2017;6:181-94.
87. Belvederi Murri M, Respino M, Innamorati M et al. Is good insight associated with depression among patients with schizophrenia? Systematic review and meta-analysis. Schizophr Res 2015;162:234-47.
88. Buck KD, Roe D, Yanos PT et al. Challenges to assisting with the recovery of personal identity and wellness for persons with serious mental illness: considerations for mental health professionals. Psychosis 2013;5:127-33.

89. Giusti L, Ussorio D, Tosone A et al. Is personal recovery in schizophrenia predicted by low cognitive insight? Community Ment Health J 2015;51: 30-7.
90. Kim JH, Lee S, Han AY et al. Relationship between cognitive insight and subjective quality of life in outpatients with schizophrenia. Neuropsychiatr Dis Treat 2015;11:2041-8.
91. Phalen PL, Viswanadhan K, Lysaker PH et al. The relationship between cognitive insight and quality of life in schizophrenia spectrum disorders: symptom severity as potential moderator. Psychiatry Res 2015;230:839-45.
92. Palmer EC, Gilleen J, David AS. The relationship between cognitive insight and depression in psychosis and schizophrenia: a review and meta-analysis. Schizophr Res 2015;166:261-8.
93. Weintraub MJ, Weismann de Mamani A. Effects of sub-clinical psychosis and cognitive insight on psychological well-being: a structural equation model. Psychiatry Res 2015;226:149-55.
94. Adan A, Capella MD, Prat G et al. Executive functioning in men with schizophrenia and substance use disorders: influence of lifetime suicide attempts. PLoS One 2017;12:1-16.
95. Lopez-Moriñigo JD, Wiffen B, O'Connor J et al. Insight and suicidality in first-episode psychosis: understanding the influence of suicidal history on insight dimensions at first presentation. Early Interv Psychiatry 2014; 8:113-21.
96. Barrett EA, Mork E, Faerden A et al. The development of insight and its relationship with suicidality over one year follow-up in patients with first episode psychosis. Schizophr Res 2015;162:97-102.
97. Massons C, Lopez-Moriñigo JD, Pousa E et al. Insight and suicidality in psychosis: a cross-sectional study. Psychiatry Res 2017;252:147-53.
98. Henriksen MG, Parnas J. Self-disorders and schizophrenia: a phenomenological reappraisal of poor insight and noncompliance. Schizophr Bull 2014;40:542-7.
99. Xavier RM, Vorderstrasse A. Neurobiological basis of insight in schizophrenia: a systematic review. Nurs Res 2016;65:224-37.
100. Asmal L, du Plessis S, Vink M et al. Insight and white matter fractional anisotropy in first-episode psychosis. Schizophr Res 2016;183:88-94.
101. Chen X, Duan M, He H et al. Functional abnormalities of the right posterior insula are related to the altered self-experience in schizophrenia. Psychiatry Res 2016;256:26-32.
102. Gerretsen P, Menon M, Mamo DC et al. Impaired insight into illness and cognitive insight in schizophrenia spectrum disorders: resting state functional connectivity. Schizophr Res 2014;160:43-50.
103. Lee JS, Chun JW, Lee SH et al. Altered neural basis of the reality processing and its relation to cognitive insight in schizophrenia. PLoS One 2015; 10:1-15.
104. Shad MU, Keshayan MS. Neurobiology of insight deficits in schizophrenia: an fMRI study. Schizophr Res 2015;165:220-6.
105. Zhang L, Opmeet EM, Ruhe HG et al. Brain activation during self- and other-reflection in bipolar disorder with a history of psychosis: comparison to schizophrenia. Neuroimage Clin 2015;8:202-9.
106. Larabi DI, Liemburg EJ, Pijnenborg GH et al. Association between prefrontal N-acetylaspartate and insight in psychotic disorders. Schizophr Res 2017;179:112-8.
107. Sapara A, Ftytche DH, Cooke MA et al. Voxel-based magnetic resonance imaging investigation of poor and preserved clinical insight in people with schizophrenia. World J Psychiatry 2016;6:311-21.
108. Emami S, Guimond S, Mallar Chakravarty M et al. Cortical thickness and low insight into symptoms in enduring schizophrenia. Schizophr Res 2016;170:66-72.
109. Parellada M, Pina-Camacho L, Moreno C et al. Insular pathology in young people with high-functioning autism and first-episode psychosis. Psychol Med 2017;47:2472-82.
110. Buchy L, Barbato M, MacMaster FP et al. Cognitive insight is associated with cortical thickness in first-episode psychosis. Schizophr Res 2016; 172:16-22.
111. Curcic-Blake B, van der Meer L, Pijnenborg GH et al. Insight and psychosis: functional and anatomical brain connectivity and self-reflection in schizophrenia. Hum Brain Mapp 2015;36:4859-68.
112. Kuang C, Buchy L, Barbato M et al. A pilot study of cognitive insight and structural covariance in first-episode psychosis. Schizophr Res 2017;179: 91-6.
113. Shenton ME, Whitford TJ, Kubicki M. Structural neuroimaging in schizophrenia: from methods to insights to treatments. Dialogues Clin Neurosci 2010;12:317-32.
114. Bergé DS, Carmona M, Rovira A et al. Gray matter volume deficits and correlation with insight and negative symptoms in first-psychotic-episode subjects. Acta Psychiatr Scand 2011;123:431-9.
115. Buchy L, Hawco C, Joober R et al. Cognitive insight in first-episode schizophrenia: further evidence for a role of the ventrolateral prefrontal cortex. Schizophr Res 2015;166:65-8.
116. Sapara A, Ffytche DH, Birchwood M et al. Preservation and compensation: the functional neuroanatomy of insight and working memory in schizophrenia. Schizophr Res 2014;152:201-9.
117. Sapara A, Ffytche DH, Cooke MA et al. Is it me? Verbal self-monitoring neural network and clinical insight in schizophrenia. Psychiatry Res 2015;234:328-35.
118. Bassitt DP, Neto MRL, de Castro CC et al. Insight and regional brain volumes in schizophrenia. Eur Arch Psychiatry Clin Neurosci 2007;257:58-62.
119. David A, Van Os J, Jones P et al. Insight and psychotic illness. Cross-sectional and longitudinal associations. Br J Psychiatry 1995;167:621-8.
120. Rossell S, Coakes J, Shapleske J et al. Insight: its relationship with cognitive function, brain volume and symptoms in schizophrenia. Psychol Med 2003;33:111-9.
121. Hwang SS, Ahn YM, Kim YS. Neurocognitive functioning as an intermediary variable between psychopathology and insight in schizophrenia. Psychiatry Res 2015;230:792-9.
122. Diez-Martin J, Moreno-Ortega M, Bagney A et al. Differential relationships between set-shifting abilities and dimensions of insight in schizophrenia. Psychopathology 2014;47:86-92.
123. de Vos AE, Pijnenborg GH, Aleman A et al. Implicit and explicit self-related processing in relation to insight in patients with schizophrenia. Cogn Neuropsychiatry 2015;20:311-29.
124. Zaytseva Y, Korsakova N, Gurovich IY et al. Luria revisited: complex motor phenomena in first episode schizophrenia and schizophrenia spectrum disorders. Psychiatry Res 2014;220:145-51.
125. MacDougall AG, McKinnon MC, Herdman KA et al. The relationship between insight and autobiographical memory for emotional events in schizophrenia. Psychiatry Res 2015;226:392-5.
126. Bhagyavathi HD, Mehta UM, Thirthalli J et al. Cascading and combined effects of cognitive deficits and residual symptoms on functional outcome in schizophrenia – a path-analytical approach. Psychiatry Res 2015;229:264-71.
127. Cernis E, Vassos E, Brebion G et al. Schizophrenia patients with high intelligence: a clinically distinct sub-type of schizophrenia? Eur Psychiatry 2015;30:628-32.
128. Gerretsen P, Plitman E, Raiji TK et al. The effects of aging on insight into illness in schizophrenia: a review. Int J Geriatr Psychiatry 2014;29:45-61.
129. Gerretsen P, Voineskos AN, Graff-Guerrero A et al. Insight into illness and cognition in schizophrenia in earlier and later life. J Clin Psychiatry 2017; 150:217-22.
130. Gilleen J, David A, Greenwood K. Self-reflection and set-shifting mediate awareness in cognitively preserved schizophrenia patients. Cogn Neuropsychiatry 2016;21:185-96.
131. Aleman A, Agrawal N, Morgan KD et al. Insight in psychosis and neuropsychological function: meta-analysis. Br J Psychiatry 2006;189:204-12.
132. Nair A, Palmer EC, Aleman A et al. Relationship between cognition, clinical and cognitive insight in psychotic disorders: a review and meta-analysis. Schizophr Res 2014;152:191-200.
133. Kansal V, Patriciu I, Kiang M. Illness insight and neurophysiological error-processing deficits in schizophrenia. Schizophr Res 2014;156:122-7.
134. Poyraz BC, Arikan MK, Poyraz CA et al. Clinical and cognitive insight in patients with acute-phase psychosis: association with treatment and neuropsychological functioning. Nord J Psychiatry 2016;70:528-35.
135. Gonzalez-Blanch C, Alvarez-Jimenez M, Ayesa-Arriola R et al. Differential associations of cognitive insight components with pretreatment characteristics in first-episode psychosis. Psychiatry Res 2014;215:308-13.
136. Ohmuro N, Katsura M, Obara C et al. The relationship between cognitive insight and cognitive performance among individuals with at-risk mental state for developing psychosis. Schizophr Res (in press).
137. Brüne M. Theory of mind in schizophrenia: a review of the literature. Schizophr Bull 2005;31:21-42.
138. Green MF, Leitman DI. Social cognition in schizophrenia. Schizophr Bull 2008;34:670-2.
139. Vohs JL, George S, Leonhardt BL et al. An integrative model of the impairments in insight in schizophrenia: emerging research on causal factors and treatments. Expert Rev Neurother 2016;16:1193-204.

140. Ng R, Fish S, Granholm E. Insight and theory of mind in schizophrenia. Psychiatry Res 2015;225:169-74.
141. Popolo R, Dimaggio G, Luther L et al. Theory of mind in schizophrenia: associations with clinical and cognitive insight controlling for levels of psychopathology. J Nerv Ment Dis 2016;204:240-3.
142. Bora E. Relationship between insight and theory of mind in schizophrenia: a meta-analysis. Schizophr Res (in press).
143. Flavell JH. Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. Am Psychol 1979;34:906.
144. Lysaker PH, Vohs JL, Ballard R et al. Metacognition, self reflection and recovery in schizophrenia: review of the literature. Future Neurol 2013;8: 103-15.
145. Lysaker PH, Vohs J, Minor K et al. Metacognitive deficits in schizophrenia: presence and associations with psychosocial outcomes. J Nerv Ment Dis 2015;203:530-6.
146. Bedford NJ, David AS. Denial of illness in schizophrenia as a disturbance of self-reflection, self-perception and insight. Schizophr Res 2014;152:89-96.
147. Moe AM, Breitborde NJ, Shakeel MK et al. Idea density in the life-stories of people with schizophrenia: associations with narrative qualities and psychiatric symptoms. Schizophr Res 2016;172:201-5.
148. Vidovic D, Brecic P, Vilibic M et al. Insight and self-stigma in patients with schizophrenia. Acta Clin Croat 2016;55:23-8.
149. Singh A, Mattoo SK, Grover S. Stigma and its correlates in patients with schizophrenia attending a general hospital psychiatric unit. Indian J Psychiatry 2016;58:291-300.
150. Drake RE, Whitley R. Recovery and severe mental illness: description and analysis. Can J Psychiatry 2014;59:236-42.
151. Karow A, Naber D, Lambert M et al. EGOFORS initiative: remission as perceived by people with schizophrenia, family members and psychiatrists. Eur Psychiatry 2012;27:426-31.
152. Slade M, Longden E. Empirical evidence about recovery and mental health. BMC Psychiatry 2015;15:1-14.
153. Ehrlich-Ben Or S, Hasson-Ohayon I, Feingold D et al. Meaning in life, insight and self-stigma among people with severe mental illness. Compr Psychiatry 2013;54:195-200.
154. Hasson-Ohayon I, Mashiach-Eizenberg M, Elhasid N et al. Between self-clarity and recovery in schizophrenia: reducing the self-stigma and finding meaning. Compr Psychiatry 2014;55:675-80.
155. Berg AO, Barrett EA, Nerhus M et al. Psychosis: clinical insight and beliefs in immigrants in their first episode. Early Interv Psychiatry (in press).
156. Mohamed S, Rosenheck R, He H et al. Insight and attitudes towards medication among inpatients with chronic schizophrenia in the US and China. Soc Psychiatry Psychiatr Epidemiol 2014;49:1063-70.
157. Zisman-Ilani Y, Hasson-Ohayon I, Lev-Frank I et al. Self stigma, insight, and family burden among Israeli mothers of people with serious mental illness: ethno-national considerations. Transcult Psychiatry 2017;54:423-41.
158. Macgregor A, Norton J, Bortolon C et al. Insight of patients and their parents into schizophrenia: exploring agreement and the influence of parental factors. Psychiatry Res 2015;228:879-86.
159. Raffard S, Bortolon C, Macgregor A et al. Cognitive insight in schizophrenia patients and their biological parents: a pilot study. Schizophr Res 2014;159:471-7.
160. Habib N, Dawood S, Kingdon D et al. Preliminary evaluation of culturally adapted CBT for psychosis (CA-CBTp): findings from developing culturally-sensitive CBT project (DCCP). Behav Cogn Psychother 2015;43:200-8.
161. Li ZJ, Guo ZH, Wang N et al. Cognitive-behavioural therapy for patients with schizophrenia: a multicentre randomized controlled trial in Beijing, China. Psychol Med 2015;45:1893-905.
162. Naeem F, Saeed S, Irfan M et al. Brief culturally adapted CBT for psychosis (CaCBTp): a randomized controlled trial from a low income country. Schizophr Res 2015;164:143-8.
163. Guo ZH, Li ZJ, Ma Y et al. Brief cognitive-behavioural therapy for patients in the community with schizophrenia: randomised controlled trial in Beijing, China. Br J Psychiatry 2017;210:223-9.
164. Drake RJ, Day CJ, Picucci R et al. A naturalistic, randomized, controlled trial combining cognitive remediation with cognitive-behavioural therapy after first-episode non-affective psychosis. Psychol Med 2014;44:1889-99.
165. Chien WT, Muj JH, Cheung EF et al. Effects of motivational interviewing-based adherence therapy for schizophrenia spectrum disorders: a randomized controlled trial. Trials 2015;16:1-14.
166. Chien WT, Muj J, Gray R et al. Adherence therapy versus routine psychiatric care for people with schizophrenia spectrum disorders: a randomized controlled trial. BMC Psychiatry 2016;16:1-14.
167. Chien WT, Thompson DR. Effects of a mindfulness-based psychoeducation programme for Chinese patients with schizophrenia: 2-year follow-up. Br J Psychiatry 2014;205:52-9.
168. Wang LQ, Chien WT, Yip LK et al. A randomized controlled trial of a mindfulness-based intervention program for people with schizophrenia: 6-month follow-up. Neuropsychiatr Dis Treat 2016;12:3097-110.
169. Yilmaz E, Okanli A. Test of mindfulness-based psychosocial skills training to improve insight and functional recovery in schizophrenia. West J Nurs Res (in press).
170. Rakitzi S, Georgila P, Efthimiou K et al. Efficacy and feasibility of the integrated psychological therapy for outpatients with schizophrenia in Greece: final results of a RCT. Psychiatry Res 2016;242:137-43.
171. Zhou B, Gu Y. Effect of self-management training on adherence to medications among community residents with chronic schizophrenia: a single blind randomized controlled trial in Shanghai, China. Shanghai Arch Psychiatry 2014;26:332-8.
172. Jorgensen R, Licht RW, Lysaker PL et al. Effects on cognitive and clinical insight with the use of guided self-determination in outpatients with schizophrenia: a randomized open trial. Eur Psychiatry 2015;30:655-63.
173. Pijnenborg GH, Timmerman ME, Derks EM et al. Differential effects of antipsychotic drugs on insight in first episode schizophrenia: data from the European First-Episode Schizophrenia Trial (EUFEST). Eur Neuropsychopharmacol 2015;25:808-16.
174. Hou CL, Cai MY, Ma XR et al. Clozapine prescription and quality of life in Chinese patients with schizophrenia treated in primary care. Pharmacopsychiatry 2015;48:200-4.
175. Mattila T, Koeter M, Wohlfarth T. The impact of second generation antipsychotics on insight in schizophrenia: results from 14 randomized, placebo controlled trials. Eur Neuropsychopharmacol 2017;27:82-6.
176. Moritz S, Andreou C, Schneider BC, et al. Sowing the seeds of doubt: a narrative review on metacognitive training in schizophrenia. Clin Psychol Rev 2014;34:358-66.
177. Balzan RP, Delfabbro PH, Galletly CA et al. Metacognitive training for patients with schizophrenia: preliminary evidence for a targeted, single-module programme. Aust N Z J Psychiatry 2014;48:1126-36.
178. Gaweda L, Krezolek M, Olbrys J et al. Decreasing self-reported cognitive biases and increasing clinical insight through meta-cognitive training in patients with chronic schizophrenia. J Behav Ther Exp Psychiatry 2015; 48:98-104.
179. Balzan RP, Galletly C. Metacognitive therapy {MCT+) in patients with psychosis not receiving antipsychotic medication: a case study. Front Psychol 2015;6:1-6.
180. Briki M, Monnin J, Haffen E et al. Metacognitive training for schizophrenia: a multicentre randomized controlled trial. Schizophr Res 2014;157: 99-106.
181. Lam KC, Ho CP, Wa JC et al. Metacognitive training (MCT) for schizophrenia improves cognitive insight: a randomized controlled trial in a Chinese sample with schizophrenia spectrum disorders. Behav Res Ther 2015;64:38-42.
182. Ochoa S, Lopez-Carrilero R, Barrigon ML et al. Randomized control trial to assess the efficacy of metacognitive training compared with a psychoeducational group in people with a recent-onset psychosis. Psychol Med 2017;47:1573-84.
183. Ussorio D, Giusti L, Wittekind CE et al. Metacognitive training for young subjects (MCT young version) in the early stages of psychosis: is the duration of untreated psychosis a limiting factor? Psychol Psychother 2016;89:50-65.
184. van Oosterhout B, Krabbendam L, de Boer K et al. Metacognitive group training for schizophrenia spectrum patients with delusions: a randomized controlled trial. Psychol Med 2014;44:3025-35.
185. Lysaker PH, Klion RE. Recovery, meaning-making, and severe mental illness: a comprehensive guide to metacognitive reflection and insight therapy. Abingdon-on-Thames: Routledge (in press).
186. Hamm JA, Firmin RL. Disorganization and individual psychotherapy for schizophrenia: a case report of metacognitive reflection and insight therapy. J Contemp Psychother 2016;46:227-34.
187. Leonhardt BL, Benson K, George S et al. Targeting insight in FEP: a case study of metacognitive reflection insight therapy (MERIT). J Contemp Psychother 2016;46:207-16.
188. Leonhardt BL, Kukla M, Chadoin K et al. Emergence of psychotic content in psychotherapy: a qualitative analysis of content, process, and therapist variables. Psychother Res (in press).

189. Leonhardt BL, Ratliff K, Buck KD. Recovery in FEP: a case study of metacognitive reflection and insight therapy (MERIT). Am J Psychother (in press).
190. Vohs JL, Leonhardt BL, James AV et al. Metacognitive reflection and insight therapy for early psychosis: a preliminary study of a novel integrative psychotherapy. Schizophr Res (in press).
191. Connell M, Schweitzer R, King R. Recovery from first-episode psychosis and recovering self: a qualitative study. Psychiatr Rehabil J 2015;38:359-64.
192. Jacob KS. Insight in psychosis: standards, science, ethics and value judgment. Int J Soc Psychiatry 2017;63:345-51.
193. Ogden LP. "My life as it has value": narrating schizophrenia in later years. Qual Health Res 2014;24:1342-55.
194. Touskova T, Bob P. Consciousness, awareness of insight and neural mechanisms of schizophrenia. Rev Neurosci 2015;26:295-304.

DOI:10.1002/wps.20508

A paradigm shift in psychiatric classification: the Hierarchical Taxonomy Of Psychopathology (HiTOP)

Many have argued that a hierarchical dimensional approach to psychiatric classification would better align the nosology with data on the natural organization of psychopathology[1]. However, such proposals have often been resisted on the grounds that: a) consensus among dimensional models is lacking and b) categorical diagnoses are considered to be essential to clinical decision-making.

The Hierarchical Taxonomy Of Psychopathology (HiTOP) consortium (see https://medicine.stonybrookmedicine.edu/HITOP) was formed by psychiatric nosologists to develop a consensus dimensional classification that is more clinically informative than the traditional diagnostic systems (DSM and ICD).

This group of scientists (now including 69 members) reviewed studies on the structure of psychopathology and developed a consensual model[2]. The resulting system offers to address problems of arbitrary disorder boundaries (consequences of which include subthreshold and not otherwise specified cases) and substantial unreliability of traditional diagnoses, by characterizing psychopathology in terms of dimensions rather than categories.

The system resolves the problem of within-disorder heterogeneity by constructing dimensions on the basis of the observed covariation of symptoms, thus identifying coherent constructs. It deals with comorbidity by identifying higher-order dimensions that reflect associations among lower-order dimensions. This hierarchy summarizes patterns of comorbidity and enables practitioners to study and treat characteristics common to multiple conditions. Importantly, HiTOP encompasses both transient symptoms and stable maladaptive traits.

The HiTOP hierarchy has five levels. It combines symptoms, signs and maladaptive behaviors into tight-knit symptom components (e.g., insomnia) and maladaptive traits (e.g., emotional lability). These, in turn, are combined with closely related components/traits into dimensional syndromes, such as vegetative depression (that includes insomnia, psychomotor retardation, lassitude and appetite loss)[3]. Similar syndromes are combined into subfactors, such as a distress dimension that includes depression, generalized anxiety, post-traumatic stress and some borderline personality traits. Larger constellations of syndromes form broad spectra, such as an internalizing dimension that consists of distress, fear, eating pathology and sexual problems. Finally, spectra can be aggregated into extremely broad super-spectra, such as the general factor of psychopathology that reflects characteristics shared by all mental disorders.

HiTOP organizes psychopathology according to evidence from statistical modeling and validation studies[2], but it is a phenotypic model and does not directly incorporate etiology. Would such an approach perform substantially better than the traditional diagnostic systems? There are two reasons to expect that it will. First, dimensional phenotypes have been found to have greater reliability and stronger associations with validators than categorical diagnoses[4], indicating that dimensional descriptions are more informative. Second, dimensions have been shown to be more useful in clinical research. HiTOP aligns much better than traditional diagnostic systems with the genetic architecture of mental disorders and with the effects of environmental risk factors, such as childhood maltreatment[2,5,6]. HiTOP dimensions can explain nearly all long-term chronicity of psychopathology[7]. HiTOP also far outperforms traditional systems in accounting for functional impairment[3]. Moreover, HiTOP dimensions can help to explain why disorders from different classes respond to the same treatment (e.g., social anxiety disorder to antidepressants)[5]. Indeed, some spectra already have become useful targets for treatment development[8].

Another response to shortcomings of traditional diagnostic systems is the Research Domain Criteria (RDoC) framework, a dimensional classification of basic psychological processes potentially relevant to psychiatric problems. The RDoC initiative aims to develop an etiologically-based nosology, but its scope is largely limited to constructs conserved across species and linked empirically to neural circuitry. Also, the RDoC framework is focused primarily on basic levels of analysis, and its clinical translation lies well in the future. In contrast, HiTOP was designed to be immediately useful in clinical research and practice.

HiTOP can inform the RDoC initiative by identifying key clinical dimensions that need to be studied. Conversely, HiTOP is a descriptive system, and RDoC research can clarify the nature and validity of HiTOP dimensions. It is likely that some RDoC dimensions lack coherent phenotypes and that some HiTOP dimensions have intractable biology, but in areas of convergence these models may ultimately produce a unified nosology, achieving a comprehensive understanding of psychopathology.

Furthermore, HiTOP can help to improve clinical practice immediately. Clinicians often forego a formal diagnostic assessment, as many consider it to have little clinical utility[9]. Initial evidence suggests that dimensional models can be more informative than traditional diagnoses in clinical decision-making[10]. Indeed, dimensional descriptors are indispensable in other areas of medicine (e.g., body mass index, blood pressure, laboratory test results). In psychiatry, dimensional measures have a long history of clinical use (e.g., personality inventories, symptom ratings, intelligence tests, neuropsychological tests).

To date, HiTOP has not been used clinically as a complete system, but it relies heavily on concepts and constructs embedded in widely-used dimensional measures. In fact, available HiTOP-aligned measures (see http://psychology.unt.edu/hitop)

allow practitioners to implement many aspects of the system already.

HiTOP can be used most feasibly in a stepwise manner, beginning with a brief measure of the six spectra. If problems are detected in some spectra, lengthier measures can be administered to characterize dimensions within those domains (while the other domains do not require further assessment). Thus, a HiTOP diagnosis is a patient's profile on relevant dimensions. Although such profiles may include a large number of scales, they are often simpler than traditional manuals, with their hundreds of codes and numerous permutations necessitated by comorbidities[10].

Clinical decisions require cut-offs on dimensions to guide specific actions. The HiTOP consortium aims to develop such cut-offs empirically, and cut-offs based on statistical deviance already exist (e.g., two standard deviations above the mean indicate high severity).

Indeed, HiTOP is a work in progress. Ongoing efforts aim to extend the system to all forms of psychopathology, construct an integrated measure of all HiTOP dimensions, and develop detailed guidance for clinicians using the system. Much more needs to be done, but HiTOP already can be applied in a variety of contexts. At minimum, it provides a framework for conceptualizing research phenotypes and individual patients dimensionally. Ultimately, HiTOP is expected to offer a roadmap for researchers and clinicians that is much more informative than traditional diagnostic systems.

Roman Kotov[1], Robert F. Krueger[2], David Watson[3]
[1]Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA; [2]Department of Psychology, University of Minnesota, Minneapolis, MN, USA; [3]Department of Psychology, University of Notre Dame, South Bend, IN, USA

1. Helzer JE, Kraemer HC, Krueger RF et al (eds). Dimensional approaches in diagnostic classification: refining the research agenda for DSM-V. Arlington: American Psychiatric Publishing, 2009.
2. Kotov R, Krueger RF, Watson D et al. J Abnorm Psychol 2017;126:454-77.
3. Waszczuk MA, Kotov R, Ruggero C et al. J Abnorm Psychol 2017;126:613-34.
4. Markon KE, Chmielewski M, Miller CJ. Psychiatr Bull 2011;137:856-79.
5. Andrews G, Goldberg DP, Krueger RF et al. Psychol Med 2009;39:1993-2000.
6. Keyes KM, Eaton NR, Krueger RF et al. Br J Psychiatry 2012;200:107-15.
7. Vollebergh WA, Iedema J, Bijl RV et al. Arch Gen Psychiatry 2001;58:597-603.
8. Barlow DH, Farchione TJ, Bullis JR et al. JAMA Psychiatry (in press).
9. Taylor D. World Psychiatry 2016;15:224-5.
10. Verheul R. J Pers Disord 2005;19:283-302.

DOI:10.1002/wps.20478

Schizotypy, schizotypic psychopathology and schizophrenia

The term schizotypy refers to a latent personality organization that putatively harbors the liability for schizophrenia and can give rise to a variety of schizophrenia-related phenotypic outcomes[1,2].

This personality organization, which is determined by any number of as-yet-unknown schizophrenia-related genetic influences acting against a background of polygenic assets and liabilities as well as impacts from the environment (e.g., stressors, epigenetic inputs), can manifest itself variously at the phenotypic level, ranging from clinically diagnosable schizophrenia through pathological personality manifestations (e.g., schizotypal, paranoid, avoidant and schizoid personality disorders) to subtle, sub-clinical psychotic-like phenomenology (e.g., perceptual aberrations, magical ideation, referential thinking, interpersonal aversiveness).

Schizotypy may also manifest itself in an imperceptible manner, undetectable by the unaided naked eye, through deviance on endophenotypes that have established valid relations with schizophrenia.

Moreover, schizotypy as a latent construct (personality organization) is centrally embedded in a diathesis-stressor theoretical model that has considerable utility as an organizing framework for the study of schizophrenia, schizophrenia-related psychopathology (e.g., delusional disorder, psychosis not otherwise specified, schizotypal, paranoid and other related personality disorders) as well as putative schizophrenia endophenotypes, a view I have advocated for several decades[3-6].

Note, the term schizotypy is not restricted to describe only those clinical manifestations that are associated with schizotypal personality disorder[2,5,6]. Nor is the term reserved to indicate a methodological preference, e.g. for self-report psychometric assessments. Rather, schizotypy can be assessed using a variety of approaches such as interviews, psychometric inventories, familial risk and/or laboratory measures. Schizotypic persons may indeed display some of the phenomenology associated with schizotypal personality disorder, but they may also show other features[6-8].

There is a long history of describing clinical states bearing the imprint of schizotypy and an implicit connection to schizophrenia liability, including observations by Kraepelin, Bleuler, Rado, Meehl, Gottesman and myself. It has been argued that a clear demarcation in an underlying schizophrenia liability continuum (e.g., a pronounced threshold effect or discontinuity) is required to explain the emergence of schizotypic indicators in psychological functioning. An alternative position regarding schizotypy holds that it is a dimension of normal personality, not necessarily connected to schizophrenia liability, and representing something of a "healthy" personality factor. However, observers of schizophrenia and schizotypic psychopathology, in the main, do not view schizotypy as benign or reflective of healthy psychological adjustment.

Non-psychotic schizotypic states (defined using clinical, laboratory and/or familial risk) have been associated with a wide range of findings, including sustained attention deficits,

working memory deficits, smooth pursuit eye movement dysfunction, schizophrenia-related psychometric deviance on the Minnesota Multiphasic Personality Inventory (MMPI), executive functioning deficits, dysfunctional anti-saccade performance, subtle formal thought disorder, clinical schizotypal and paranoid personality features, schizophrenia-related social cognition deficits, exteroceptive and proprioceptive somatosensory deficits, psychomotor abnormalities, and candidate polymorphisms (e.g., ZNF804A, Val158Met-COMT, neuregulin-1). That schizotypic persons manifest such a panorama of deficits, similar in nature albeit less in degree to those seen in schizophrenia, argues for a connection or common underlying construct for conditions defined phenotypically (i.e., schizotypic subjects vs. schizophrenia-affected subjects).

An area of continued speculation concerns the underlying structure of schizotypy and the precise nature of the variation expressed in that latent construct. Considerable statistical evidence, using a variety of latent structure methods, points to the existence of possible underlying discontinuities or severe threshold effects in schizotypy, and work in this area continues. Such evidence stimulates the caveat, to wit, that the use of continuous measures to assess phenotypic manifestations of schizotypy does not *ipso facto* mean that the underlying (or latent) schizotypy construct is fully quantitative or uniformly graded by degree.

The course and clinical outcome for those designated as harboring schizotypy remains an area of active inquiry. It is entirely conceivable that many individuals possessing schizotypy may traverse the life course escaping psychotic illness as well as other diagnosable schizotypic manifestations. The expectation that some people validly at risk for schizophrenia may never manifest the illness is well established in the reality of discordant monozygotic twins, in which one twin is affected by schizophrenia and the co-twin is not psychotic (perhaps not even diagnosable as having a non-psychotic, but detectable, clinical schizotypic condition like schizotypal or paranoid personality disorder).

Individuals that achieve elevated scores on psychometric measures of schizotypy have been shown to be at increased risk for schizophrenia and schizophrenia-related psychoses later in life, as well as a variety of other related outcomes. Such individuals also display poorer psychosocial functioning, lower rates of marriage, increased use of psychiatric medications, and increased utilization of psychiatric services[2]. It is entirely conceivable that many individuals designated as "prodromal" for schizophrenia, but who do not convert to schizophrenia (which is 60-70% of such subjects), are in fact harboring schizotypy and will, even if not psychotic, show impairments across the life span, perhaps adopting an eccentric or odd manner of personality functioning.

The schizotypy model has helped to adjust the boundaries of schizophrenia phenotype in the DSM-5 (e.g., schizotypal pathology is now included with schizophrenia). Furthermore, illuminating the nature of schizotypy may aid in unraveling the current puzzle of the very low conversion to schizophrenia rates seen in "prodromal" schizophrenia research[9].

Finally, I have argued that the schizotypy framework may be useful in understanding *configurations* (rather than simple additive summation) of genes relevant to schizophrenia variants[2], an idea that is beginning to gain traction. There is no doubt that incorporation of schizotypy indicators into genomic studies of schizophrenia increase their statistical power.

The advantages of a cleaner unit of analysis (the schizotype), free from the effects of medication, institutionalization and neurocognitive decline, are axiomatic. However, the understanding (and misunderstanding) of the schizotypy model as well as alternative approaches to the construct require vigilance, in order to ensure that the approach continues to yield the fruit that it can.

Mark F. Lenzenweger
Department of Psychology, State University of New York at Binghamton, and Department of Psychiatry, Weill Cornell Medical College, New York, NY, USA

1. Meehl PE. J Pers Disord 1990;4:1-99.
2. Lenzenweger MF. Schizotypy and schizophrenia: the view from experimental psychopathology. New York: Guilford Press, 2010.
3. Lenzenweger MF. In: Lenzenweger MF, Dworkin RH (eds). Origins and development of schizophrenia: advances in experimental psychopathology. Washington: American Psychological Association, 1998:93-121.
4. Lenzenweger MF. Curr Dir Psychol Sci 2006;15:162-6.
5. Lenzenweger MF. Schizophr Bull 2015;41(Suppl. 2):S483-91.
6. Lenzenweger MF. In: Blaney PH, Millon T, Krueger R (eds). Oxford textbook of psychopathology, 3rd ed. New York: Oxford University Press, 2015:729-67.
7. Meehl PE. Manual for use with Checklist of Schizotypic Signs. Minneapolis: University of Minnesota, 1964.
8. Kwapil TR, Barrantes-Vidal N. In: Widiger TA (ed). The Oxford handbook of personality disorders. Oxford: Oxford University Press, 2012:437-77.
9. Fusar-Poli P, Borgwardt S, Bechdolf A et al. Arch Gen Psychiatry 2013;70:107-20.

DOI:10.1002/wps.20479

The value of polygenic analyses in psychiatry

The last decade of genetics research in psychiatry (and in other fields) has been dominated by genome-wide association (GWA) studies, in which common variants across the genome are tested for association with a trait or disorder. These studies have shown that polygenicity is the rule, i.e., psychiatric disorders are influenced by many (likely thousands of) genetic variants, each with a small effect[1].

This is best illustrated by the flagship GWA meta-analysis on schizophrenia, which is the first disorder that has achieved the sample size needed to detect the effect sizes that have been dealt by nature's hand. By analysing 37,000 cases and 113,000 controls, 108 associated regions were identified[2]. However, the significant variants together only explained 3.4% on the liability scale for schizophrenia, indicating there are many more vari-

ants involved. This high degree of polygenicity means that everyone harbours risk variants, but those affected likely carry a higher, and possibly unique burden of risk factors, which is fully consistent with the spectrum of clinical presentations.

Still, because of the small effect sizes, the usefulness of the results of GWA analyses has been questioned. In this paper, we show the value of the identification of genetic variants in psychiatric disorders and illustrate how analyses of GWA data have further advanced our knowledge, beyond the identification of associated genetic variants.

One major problem in psychiatry is that there have hardly been any new drugs developed in the last decades[3]. Even though effect sizes are small, significantly associated genetic variants can point to new drug targets, as shown for other diseases[3]. With 108 regions associated and no immediate knowledge about the functional effects of the far majority of the hits, further analyses are necessary, but could lead to new targets.

Functional annotation of genetic variants associated with psychiatric disorders using bioinformatic analyses is an active area of research[4]. This includes analyses that aim to explore which trait-associated genetic variants are also associated with inter-individual variation in gene expression levels, and gene-based analyses investigating which biological pathways are enriched with genes harbouring associated genetic variants[3]. For psychiatric disorders, neuronal, immune and histone pathways are reported to be involved[5], and these analyses will become more informative with new technologies, such as single cell gene expression studies.

GWA data can also be used to increase knowledge on the mechanisms underlying the frequent comorbidity within psychiatric disorders or between psychiatric disorders and other traits. This is interrogated by polygenic analyses, investigating the joint effect of genetic variants[1,4]. Traditionally, to demonstrate a genetic relationship between disorders was difficult, especially for the rarer disorders, because recording of psychiatric diagnoses was needed on large samples of twins or families to demonstrate the increased risk of a second disorder in family members of those affected by a first disorder. However, direct measurement of DNA variants has allowed direct measures of genetic sharing using independently collected case-control samples.

It has become apparent that psychiatric disorders not only share genetic risk with other psychiatric disorders, but also with somatic diseases and traits such as educational attainment[6]. If genetic correlations between disorders and traits are identified, a key question is whether the association reflects shared biological pathways (pleiotropy) or if there is a causal relationship. Using two-step Mendelian randomization, it has been shown that cannabis initiation does result in a small increase in risk to develop schizophrenia, but that schizophrenia leads to a larger increase in risk for cannabis initiation[7]. More insight into directions of effect and causality can direct the development of prevention programs.

Such knowledge is also important for research aiming to develop treatments targeted to children at high risk that their disorder develops into an adult psychiatric disorder, either the same or a different one. A polygenic risk score is an estimate of the cumulative genetic risk of an individual. In schizophrenia research, polygenic risk scores have been found to predict various psychiatric traits during childhood and adolescence, indicating that genetic variants play a role in the transition from internalizing or externalizing symptoms during childhood or adolescence to schizophrenia later in life[8].

These polygenic risk scores cannot be used as diagnostic predictors of psychiatric disease, as risk to psychiatric disorders is only partly explained by genetic risk factors, and, to date, only a small proportion of genetic risk has been identified. Nonetheless, out-of-sample prediction explains about 7% in liability to schizophrenia[2], so those with highest polygenic risk scores have an increased risk approximately equivalent to having a first-degree relative affected.

While this has little clinical utility in the general population, it may have clinical application in the context of prodromal presentation at a mental health clinic. Recently, an individualized risk calculator has been developed that with reasonable accuracy could predict the conversion to psychosis[9]. Predictors included were already existing symptoms and poorer functioning on cognitive tests. Possibly, risk prediction can be improved by adding further variables to the model, including, but not limited to, genetic risk scores[10]. Based on these profiles, individuals could be stratified into high and low risk groups for transition into a severe mental illness[10] and the effects of different treatment programs for these groups could be tested.

Overall, the progress in genetic research has substantially increased our insight into the etiology of psychiatric disorders. Genetic discoveries in schizophrenia have been achieved by large sample sizes, and the current data show that, with larger samples, similar results can be obtained for other disorders. Genotyping technologies are no longer the limiting factor (500,000 DNA variants can be measured for less than $100/person). The limiting factors are availability of large samples with consistently measured clinical symptoms and environmental risk factors. International collaborations, such as the Psychiatric Genomics Consortium (PGC) (www.med.unc.edu/pgc) and the EArly Genetics Lifecourse Epidemiology consortium (EAGLE) (www.wikigenes.org/e/art/e/348.html), and long-term planning are required for cost-effective generation of the data sets needed to deliver on the promise of precision or stratified medicine in psychiatry.

The new genetic discoveries of the last five years are opening previously unknown avenues of research. If ultimately these lead to new treatments, as in other fields of medicine, these treatments could be specific to stratified patient groups.

Christel M. Middeldorp[1-3], Naomi R. Wray[4,5]

[1]Child Health Research Centre, University of Queensland, Brisbane, QLD, Australia; [2]Child and Youth Mental Health Service, Children's Health Queensland Hospital and Health Services, Brisbane, QLD, Australia; [3]Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [4]Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD, Australia; [5]Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia

C.M. Middeldorp acknowledges the Netherlands Organization for Health Research and Development grant "Genetic influences on stability and change in psychopathology from childhood to young adulthood" (ZonMW 912-10-020) and funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 721567. N.R. Wray acknowledges grants from the Australian National Health and Medical Research Council (1078901, 1113400, 1087889).

1. Visscher PM, Wray NR, Zhang Q et al. Am J Hum Genet 2017;101:5-22.
2. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Nature 2014;511:421-7.
3. Breen G, Li Q, Roth BL et al. Nat Neurosci 2016;19:1392-6.
4. Maier RM, Visscher PM, Robinson MR et al. Psychol Med (in press).
5. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Nat Neurosci 2015;18:199-209.
6. Pickrell JK, Berisa T, Liu JZ et al. Nat Genet 2016;48:709-17.
7. Gage SH, Jones HJ, Burgess S et al. Psychol Med 2017;47:971-80.
8. Nivard MG, Gage SH, Hottenga JJ et al. Schizophr Bull 2017;43:1197-207.
9. Cannon TD, Yu C, Addington J et al. Am J Psychiatry 2016;173:980-8.
10. Chatterjee N, Shi J, Garcia-Closas M. Nat Rev Genet 2016;17:392-406.

DOI:10.1002/wps.20480

The promise and challenges of drug repurposing in psychiatry

The term "repurposing" literally means to give a new purpose or use to a drug. Some researchers have sub-classified repurposing into "reformulation", which is the development of a different formulation for the same drug, and "repositioning", which is the process of identifying a new therapeutic use for an already known drug[1]. One may argue that only repositioning is closely aligned with the term repurposing. Therefore, the focus of this paper will be only on the repositioning form of repurposing.

Drug repurposing is viewed as an approach to rediscovering value in "old molecules" and finding new therapeutic uses, particularly in areas with high risk of failure, such as psychiatry. It is considered a cost-effective and de-risked strategy[2]: having already established the safety and tolerability of a compound diminishes the risks of further development.

The importance of repurposing was recently acknowledged by the European Commission, which formed the Commission Expert Group on Safe and Timely Access to Medicines for Patients (STAMP). STAMP aims to recognize the importance of comprehensive investigation of different opportunities that a molecule could bring to patients, with faster development times, and at reduced costs and risk for pharmaceutical companies[3].

The scope and extent of drug repurposing in the central nervous system (CNS) area has been recently reviewed[4]. The authors performed an extensive search of compounds, with the initial and target indication and the type of repurposing strategy (repositioning, reformulation or both). Their study identified 118 source products which were repurposed 203 times, with 80 products repurposed once, 16 products repurposed twice and 22 products repurposed three times or more[4].

Among products repurposed multiple times, over two thirds (68%) came from the CNS area, and half of the new indications (102 cases) were approved[3]. Most of the cases were repositioned (N=171), while only 16 were reformulated, and 16 were reformulated and repositioned at the same time[4]. Among new therapeutic indications, Alzheimer's disease was targeted most often (22 cases), followed by substance dependence (alcohol, opioids, tobacco), bipolar disorder, depression, neuropathy/neuralgia, multiple sclerosis and schizophrenia, with 10 or more cases each[4].

A prototypical example of a repurposed drug in psychiatry is valproic acid/valproate[5]. The anticonvulsant properties of N-dipropylacetic acid (valproic acid) were discovered in 1967 and the drug quickly became widely used in epilepsy, generally in the form of sodium valproate. Antimanic and prophylactic activity in bipolar disorder was only subsequently demonstrated for both valproic acid and sodium valproate, with divalproex (an equimolar combination of valproic acid and sodium valproate) being approved by the US Food and Drug Administration (FDA) in 1995 for this new indication[5].

Traditionally, there have been three major approaches to drug repurposing/repositioning.

One approach is the discovery at the bedside, where a clinician observes/discovers the benefit in a given condition of a compound approved for a different condition. A classic example is bupropion for smoking cessation. Bupropion was first approved by the FDA for the treatment of depression in the 1980s. L. Ferry, at the time Chief of Preventive Medicine at the Loma Linda Veterans Hospital, and her colleagues tried the drug in the mid 1990s in a small group of smokers, with impressive results, as about half were able to quit smoking for at least a year. This led to a series of positive placebo-controlled trials and to the approval of bupropion for smoking cessation in 1997[6].

Another approach involves leveraging the knowledge of the potential benefits of specific pharmacological actions in certain conditions, and identifying compounds initially developed for the treatment of other conditions and sharing similar pharmacological actions. A clear example of this is atomoxetine. This compound, a norepinephrine reuptake inhibitor, was originally developed for the treatment of depression and then abandoned despite good tolerability. T. Spencer, J. Biederman and colleagues, whose group at Massachusetts General Hospital had shown the efficacy in attention deficit disorder of desipramine[7], a tricyclic antidepressant with norepinephrine reuptake inhibition properties, approached the maker of atomoxetine about testing it in this condition, and showed it to be effective[8]. Atomoxetine was subsequently approved by the FDA in December 2002 for the treatment of attention-deficit/hyperactivity disorder.

The third approach to drug repurposing comes from the advances in the understanding of the neurobiology and genetics of psychiatric disorders. The identification of specific neu-

ral pathways associated with certain genetic polymorphisms can lead to the use of approved compounds which have shown to affect those molecular targets. Alternatively, the identification of subtypes associated with specific biomarkers may lead to the exploitation of compounds addressing the specific neurobiological target. An example of this approach is the development of anti-inflammatory compounds for the treatment of the subtype of major depressive disorder associated with chronic inflammation. A recent publication on the antidepressant properties of ixekizumab, approved by the FDA in 2016 for the treatment of moderate-to-severe plaque psoriasis in adult patients who are candidates for systemic therapy or phototherapy, exemplifies this[9].

One of the issues related to the repurposing of older molecules can be the relatively short patent protection, once the drug is approved for the new indication. Some drug companies have used the approach of adding deuteriums (instead of plain hydrogens) to drug structures. Deuterated forms of older compounds can have a more extended patent protection. The first example of this was the approval in August 2017 of deutetrabenazine tablets for the treatment of tardive dyskinesia in adults, nine years after the FDA approval of the older compound (tetrabenazine) to treat chorea associated with Huntington's disease.

In light of the recent and continuous advances in the understanding of the neuroscience of psychiatric disorders, repurposing drugs is likely to yield even greater promise in the future. Having already established the safety and tolerability of a compound diminishes the risk of its further development and, therefore, would allow the use of more cost-effective study designs[10], which require smaller sample sizes and would reduce the cost of the clinical development.

Maurizio Fava
Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA

1. Murteira S, Ghezaiel Z, Karray S et al. J Mark Access Health Policy 2013;1: 21131.
2. Hemphill CS, Sampat BN. J Health Econ 2012;31:327-39.
3. Commission Expert Group on Safe and Timely Access to Medicines for Patients. Repurposing of established medicines/active substances. http://ec.europa.eu/health/files/committee/stamp.
4. Caban A, Pisarczyk K, Kopacz K et al. J Mark Access Health Policy 2017;5: 1299833.
5. Lempérière T. Encéphale 2001;27:365-72.
6. Ferry L, Johnston JA. Int J Clin Pract 2003;57:224-30.
7. Biederman J, Baldessarini RJ, Wright V et al. J Am Acad Child Adolesc Psychiatry 1989;28:777-84.
8. Spencer T, Biederman J, Heiligenstein J et al. J Child Adolesc Psychopharmacol 2001;11:251-65.
9. Griffiths CEM, Fava M, Miller AH et al. Psychother Psychosom 2017;86:260-7.
10. Baer L, Ivanova A. Clin Investig 2013;3:832-3.

DOI:10.1002/wps.20481

Measuring and improving the quality of mental health care: a global perspective

Amy M. Kilbourne[1,2], Kathryn Beck[2], Brigitta Spaeth-Rublee[3], Parashar Ramanuj[4,5], Robert W. O'Brien[1], Naomi Tomoyasu[1], Harold Alan Pincus[6]

[1]Health Services Research and Development Service, Veterans Health Administration, US Department of Veterans Affairs, Washington, DC, USA; [2]Department of Psychiatry, University of Michigan Medical School, Ann Arbor, MI, USA; [3]Department of Behavioral Health Services and Policy Research, New York State Psychiatric Institute, New York, NY, USA; [4]RAND Europe, Cambridge, UK; [5]Royal National Orthopaedic Hospital, Stanmore, UK; [6]Department of Psychiatry and Irving Institute for Clinical and Translational Research, Columbia University and New York-Presbyterian Hospital, New York, NY, USA

Mental disorders are common worldwide, yet the quality of care for these disorders has not increased to the same extent as that for physical conditions. In this paper, we present a framework for promoting quality measurement as a tool for improving quality of mental health care. We identify key barriers to this effort, including lack of standardized information technology-based data sources, limited scientific evidence for mental health quality measures, lack of provider training and support, and cultural barriers to integrating mental health care within general health environments. We describe several innovations that are underway worldwide which can mitigate these barriers. Based on these experiences, we offer several recommendations for improving quality of mental health care. Health care payers and providers will need a portfolio of validated measures of patient-centered outcomes across a spectrum of conditions. Common data elements will have to be developed and embedded within existing electronic health records and other information technology tools. Mental health outcomes will need to be assessed more routinely, and measurement-based care should become part of the overall culture of the mental health care system. Health care systems will need a valid way to stratify quality measures, in order to address potential gaps among subpopulations and identify groups in most need of quality improvement. Much more attention should be devoted to workforce training in and capacity for quality improvement. The field of mental health quality improvement is a team sport, requiring coordination across different providers, involvement of consumer advocates, and leveraging of resources and incentives from health care payers and systems.

Key words: Mental disorders, quality of care, quality measurement, health informatics, electronic health records, patient-centered outcomes, health care systems, health policy

(World Psychiatry 2018;17:30–38)

Mental disorders are responsible worldwide for 32% of years of disability and 13% of disability adjusted life years[1]. In addition, persons with these disorders face increased rates of morbidity from general medical conditions[2-4] and a higher risk of premature mortality[5]. Among persons with mental disorders, disparities in quality and outcomes of care are more pronounced for racial/ethnic minorities[6-8], and those from lower socio-economic status groups[9]. Severe mental illness (e.g., schizophrenia and bipolar disorder) is emerging as a prominent health disparity category, given estimates that persons in this group die 8-25 years younger than the general population[10,11]. Despite the contribution of mental disorders to the global burden of disease, the quality of care for these disorders remains suboptimal, and there are persistent gaps in access to and receipt of mental health services worldwide[12-18].

Quality of care, as described by the Donabedian framework, includes structure, or organization of care, the influence of structure on clinical processes of care as delivered by providers, and ultimately, patient-level health care outcomes [19-21] (see Table 1). This system-level perspective of health care quality (structure, process, outcomes) became the foundation for two US Institute of Medicine's reports: Crossing the Quality Chasm[22] in 2001 and Improving the Quality of Health Care for Mental and Substance-Use Conditions[23] in 2006.

The Crossing the Quality Chasm report highlighted six aims towards quality improvement – safe, effective, patient-centered, timely, efficient, and equitable care – and stated that "quality problems occur typically not because of failure of goodwill, knowledge, effort or resources devoted to health care, but because of fundamental shortcomings in the ways care is organized"[22]. The 2006 report further noted the persistent gaps in quality of mental health care and called for systematic efforts to improve quality in this area[23].

Nonetheless, the overall quality of mental health care has hardly improved since publication of these reports and, in some cases, has worsened over time[24]. In the US, only a third of those in need receive adequate mental health care[25]. The level of mental health quality of care is poor and the rate of improvement is slow compared to general medical conditions[26]. For example, recent data indicate that less than half of patients with publically funded insurance get adequate follow-up after mental health hospitalization[27]. This persistent gap in quality of mental health care is due in part to lack of systematic methods for measuring quality. We cannot improve what we cannot measure.

As health care costs continue to rise and mental disorders become more prevalent worldwide, health care leaders and providers will need valid information on quality of care, in order to: a) identify population needs and make decisions on how to provide the best services, and b) apply effective strategies to improve quality and reduce disparities. This paper describes the current state of quality measurement of mental health care and the challenges it poses to health care systems internationally, and suggests next steps for health care systems around the world to better implement quality mea-

Table 1 Mental health quality measures: key examples

	Description	Examples
Structure	Are adequate personnel, training, facilities, quality improvement infrastructure, information technologies, and policies available for providing care?	Adequate number of components available in assertive community treatment program
		Availability of mental health specialists in primary care practices
		Presence of a mental health care manager
Process	Are evidence-based processes of care delivered?	Percent of patients in mental health program who have documented substance use screening
		Receipt of adequate dose of psychotherapy
		Outpatient follow-up within 7 days after mental health hospitalization discharge
Outcome	Does care improve clinical outcomes?	Functioning (e.g., assessed by WHO-DAS)
		Employment (% patients returning to work)
		Symptoms (e.g., depressive, assessed by PHQ-9)
		Recovery

WHO-DAS – World Health Organization Disability Assessment Scale, PHQ-9 – Patient Health Questionnaire-9

surement and ultimately improve quality of mental health care.

CURRENT STATE OF MEASURING MENTAL HEALTH CARE QUALITY

Worldwide, efforts to standardize mental health care quality measurement are slowly evolving. Measuring and reporting quality of care on a routine basis enables the application of quality improvement at provider, clinic and health system levels, as well as accountability mechanisms that include public reporting and financial penalties and rewards. However, measuring quality of mental health care is challenging worldwide, as it can vary based on the organization of services by country. In general, structure, process and outcome measures have all been employed for accreditation, standard setting, quality improvement and accountability in health care generally and in mental health care. Each have strengths and weaknesses and, ultimately, a balanced portfolio across these categories is needed.

Health care structural components, such as resources (personnel, training, facilities) and policies that support measurement-based care, are fundamental to achieving high quality care. However, while adequate structure measures create the necessary infrastructure for reporting on processes and outcomes and conducting improvement activities, they do not provide sufficient detail as to whether quality services are actually being delivered as intended (fidelity) nor if the outcomes obtained are acceptable.

Ideally, process measures can fill this gap and assess whether evidence-based practices are in fact being implemented. These measures generally involve operationalizing clinical guidelines into specifically defined denominators and numerators, using data that can be reliably obtained from feasibly accessed data sources. However, many widely used mental health process measures lack evidence to be used in mental health quality and outcome improvement. Only a few studies have linked quality of care process measures to improvements in patient functioning and clinical outcomes, calling into question the clinical validity of these measures. Some notable exceptions that have been reported recently show that measures for improved processes of care (e.g., appropriate pharmacotherapy, continuity of care, and psychotherapy use) are associated with reduced mortality [28-31] and reduced symptom severity[32]. Still, even among existing mental health process measures that could be reported, not all have been validated[25,26,33-39].

Outcome measures assess whether the care that a patient receives actually improves his/her symptoms – e.g., improvement or remission in Patient Health Questionnaire-9 (PHQ-9) scores – or functioning. These measures can also assist providers in planning, monitoring and adjusting treatment options (e.g., change in medication, multi-component treatment collaboration). However, in order to address the complexity of mental disorders, mental health outcome measures should not only focus on symptoms and functioning, but also on issues such as quality of life, recovery, and community tenure.

Furthermore, the use of outcome measures for the purpose of evaluating the quality of mental health care requires sophisticated risk adjustment approaches to control for underlying patient risk factors beyond providers' control, such as severity of illness, medical history/health status, socio-demographic factors, in order to minimize "cherry-picking" of the healthiest patients. This, however, may be challenging, due to typically limited available data on psychiatric symptoms, social context and other patient characteristics. Increasingly, there are calls to add patients' experiences to a balanced portfolio of measures, to get their view about a system's structures, the care they have received, as well as self-reported outcomes.

In addition, the mental health service field lacks consistent outcome measures

and tools that are embedded in current information systems and other rapidly changing technologies. Lack of ability for system-wide routine data collection within existing electronic health care systems can ultimately impede continuous quality improvement for patients. To mitigate this challenge, mental health experts are embracing measurement-based care to promote the use of outcome measures on a routine basis.

Measurement-based care is a core component of the chronic care model[40-42], which uses proactive data collection to provide patient-centered care plans. These are delivered by a care manager who also coordinates care between different providers so that it is tailored to the patient's current disposition and self-management preferences. The chronic care model has been shown in multiple randomized trials to improve physical and mental health outcomes across different mental disorders, with little to no added cost[42]. Measurement-based care relies on clinical measures (e.g., PHQ-9, mental health vital signs) as well as systematic, longitudinal and action-oriented care to track, assess and respond to changes in individualized outcomes, such as symptom severity and goal attainment, frequently and over the long term.

Key international examples of measurement-based care include the Improving Access to Psychological Therapies (IAPT) program within the UK National Health Service[43,44], the Dutch Depression Initiative primary mental health collaborative care model[45], and the Australian True-Blue model[46]. Notably, after initial pilot testing and successful evaluation, the IAPT was expanded in the UK for at least 1.5 million adults to access care each year by 2020/21[47], and the Depression Initiative primary mental health collaborative care model was included in the Netherlands into the list of national essential benefits as part of the Health Insurance Act[45]. However, these programs do not reach all patients with mental disorders, and a majority of health care providers do not routinely apply measurement-based care[48,49].

In the US, there are a few notable examples of public and private measurement-based care programs in primary and specialty mental health care settings that are adopted as clinical tools, but to date not widely used for quality measurement. For example, the Sequenced Treatment Alternatives to Relieve Depression (STAR*D)[50], the US Department of Veterans Affairs Behavioral Health Laboratory model[51,52] and the US Department of Defense Behavioral Health Data Portal[53] are all examples of measurement-based care applied to patient populations. In the State of Minnesota, the Depression Improvement Across Minnesota, Offering a New Direction (DIAMOND) initiative implemented measurement-based care to help benchmark quality improvement efforts as part of a bundled payment initiative for depression care management[54].

UNIQUE CHALLENGES TO MENTAL HEALTH CARE QUALITY MEASUREMENT

In the US and worldwide, mental health care quality measurement and measurement-based care have a weak infrastructure in health care systems. This is due to a multitude of barriers specifically related to mental health, that involve limitations in policy and technology as well as limited scientific evidence for mental health quality measures, lack of provider training and support, and cultural barriers to integrating mental health care within general health environments.

The development and application of mental health care quality measures has lagged behind other areas of medicine, in part to lagging policy and technological initiatives. For example, in the US, quality measures are used for chronic medical conditions to set reimbursement through Medicare, the government's public insurance program for elderly individuals (e.g., Value-Based Purchasing Modifier[55]), Medicaid[56] and State Medicaid Reporting Programs[57], and to benchmark care quality in the private sector (e.g., PhysicianCompare.Gov[58], HospitalCompare.Gov[59]). Yet, despite the mental health parity laws passed ten years ago, which stipulate equal coverage for mental health conditions, and the availability of over 500 measures for monitoring quality of mental health care, only 5% of these measures are actually used in the above major quality reporting programs, and only 10% of the measures have been endorsed by the US National Quality Forum[60] (e.g., Value Based Inpatient Psychiatry Quality Reporting Program[61]). Of these available measures, the majority (72%) focus on processes quite distal to outcomes (e.g., screening/assessment)[60] rather than on process measures that indicate treatment adequacy or intensity for mental health care.

On the other hand, there are many important gaps in the evidence base to support mental health quality measurement, especially for outcomes that are most meaningful to consumers, as well as for specific populations such as children. Measures are also lacking for mental health conditions commonly experienced in populations, such as anxiety disorders, and lacking in depth for evidence-based treatments such as psychotherapy. While there is well-established evidence for mental health interventions such as pharmacotherapies, specific manualized psychotherapies (e.g., cognitive behavioral therapy), and team-based interventions (e.g., assertive community treatment), the evidence base for many other psychosocial interventions needs to be strengthened[62]. For evidence-based psychotherapies, quality measures may not fully capture whether they were delivered adequately. Moreover, many providers are able to codify psychosocial interventions in administrative data, but not whether the intervention was delivered with fidelity[23,63].

There is also insufficient attention to the development and implementation of performance measures that reflect patients' views and treatment choices. As a result, few endorsed mental health quality measures assess patient-centered care, notably mental health recovery. The US Substance Abuse and Mental Health Services Administration defines mental health recovery as "a process of change through which individuals improve their

health and wellness, live a self-directed life, and strive to reach their full potential"[64]. Yet, identifying valid recovery measures has been hampered by a lack of consensus about an operational and measurable definition of recovery among providers, the research community and, most importantly, consumers of mental health services. While this is partially inherent to the subjective process of recovery, it has resulted in a large variation in reliability and validity of recovery measures and tools. Beyond the needs for further evidence to support clinical guideline development and a broader array of valid and useful patient reported outcomes, there has been little investment in the development and testing of mental health care quality and recovery measures to assure their validity, utility and comprehensiveness.

Furthermore, the mental health field is far behind other areas of medicine with regard to the implementation of technologies, notably health information technology to capture relevant health information that could support reporting on mental health care quality measures. Despite some incentives to implement electronic health records (e.g., the HITECH Act in the US), there is no specific requirement worldwide to include mental health data in electronic records. Currently, many mental health care quality measures are not linked to existing data sources, which mostly rely on claims data rather than data derived from electronic health records or electronically-reported patient outcomes[26,65]. As a result, these measures cannot be automated to generate meaningful data[60], which in return could support quality measurement and inform routine medical practices and procedures. In addition, mental health providers often use separate electronic medical record systems from their general medical provider counterparts, or do not have access to these systems at all, creating big challenges to engage the mental health field as a whole in quality measurement and improvement of care for patients who often require coordinated services across different sectors.

In some countries with common claims datasets or electronic medical records, mental health care measures have been variably adopted[66,67]. For example, the UK National Health Service has a long tradition of using electronic medical records in primary care for routine quality measurement, most notably through the Quality Outcomes Framework, the largest payment-by-results program in the world. Over the past ten years, the National Health Service has tried to implement a similar outcome-based reimbursement program in mental health care[68]. This would have made routine measurement mandatory for funding. However, the administrative burden involved and the risk of gaming (i.e., biased reporting to improve apparent performance) has led to resistance from the profession[68,69]. The program has now been indefinitely postponed in implementation in favor of smaller areas of work[70]. One of these areas is the above-mentioned IAPT initiative, which embedded routine outcome measurement – using validated tools such as the PHQ-9 and the Clinical Outcomes in Routine Evaluation (CORE) – and could demonstrate good outcomes that have led to further funding into the initiative[71]. In Canada, there has been the adoption of mental health care quality measures in electronic medical records[67]. Still, due to long-standing stigmatization and functional challenges, consumers of mental health services may feel burdened by the data gathering. Overall, integrating health information technology into routine mental health treatment practices is paramount to support measurement-based care for mental health[72,73].

In addition, heterogeneity of provider training and certification requirements within mental health care can also hinder quality measurement implementation. For example, in spite of their extensive involvement in mental health care, less than one third of US social workers receive training in quality measurement and effective clinical practices[74]. Moreover, many of the challenges that providers address with their patients include service needs beyond health care (employment, housing, education, criminal justice and welfare), and quality of care for these services is rarely measured to ensure improved mental health outcomes and recovery. These services often require coordination across different providers, settings, agencies and even sectors, but there is little incentive to improve quality when there are no measures to assess accountability for these services. A notable exception to this has been the US cross-agency priority goal of ending Veteran homelessness, where the US Department of Veterans Affairs began working with other federal, state and local agencies to provide housing vouchers and track outcomes over time[75].

Finally, cultural and administrative differences between physical and mental health providers hinder quality measurement. "Physical" and "mental health" services, in many if not most countries, are often administratively separated at clinical, organizational, policy and financial levels. Mental health care also requires more of a team effort between psychiatrists, social workers, psychologists and case managers, and mental health visits are typically longer, due to the nature of the illnesses.

INNOVATIONS IN MENTAL HEALTH CARE QUALITY MEASUREMENT AND IMPROVEMENT

Several innovations are underway worldwide for measuring and improving quality of mental health care. These initiatives combine advances in technology or measurement-based care with concerted efforts to obtain patient and provider buy-in towards continuous quality measurement and improvement.

International innovations in quality measurement include the World Health Organization (WHO)'s Assessment Instrument for Mental Health Systems[76], and the International Initiative for Mental Health Leadership[77], which provides data on reporting, ability to report, and ascertainment of data across countries.

Table 2 Learning health care system framework for mental health care quality improvement

	Barriers	Leverage opportunities in learning health care systems
Patients	Medical and behavioral health conditions co-occur	Adopt mental health measurement-based care (continuous use of validated outcome assessments that inform changes in treatment decisions)
	The majority of patients are still seen in small primary care practices	Consumer organizations link patients to recovery-oriented services in the region
Providers	The majority of providers lack training in quality improvement and evidence-based practice implementation	Professional organizations mandating training in quality measurement and improvement methods
	Lack of incentives for non-mental health providers to incorporate mental health services where patients are more likely seen (e.g., primary care), and lack of integration with social services	Same-day billing for mental health and physical health care
		Mental health professional organizations adopt common quality measures, guidelines, and improvement strategies
Practices/ Organizations	Limited electronic medical record use in the majority of mental health sites	Standard health information exchanges need to include mental health services
	Lack of effective strategies to scale up and spread evidence-based mental health treatments and models of care	Embed quality improvement experts to help identify, test and scale up treatment models to promote measurement-based care
Purchasers/ National health systems	Primarily fee for service, few bundled payment models	Plan-level mental health care coordination
	Instability in health insurance markets	Value-based reimbursement payment models benchmark on improved quality rather than volume
Population	Stigma	Public reporting of quality measures

In the Netherlands, routine outcome monitoring has been incorporated into health insurance reimbursement mechanisms. This evaluates three aspects of quality – effectiveness of treatment, safety and client satisfaction – through ten measures that are repeated at the start and end of treatment[78]. The initiative stipulates that the indicators are collected centrally and published transparently to stimulate continuous quality improvement.

In Australia, the use of standard outcome measures for all mental health service users was mandated in 2000, and all Australian states have signed agreements to submit routinely collected outcomes and case mix data. The principal outcome measures are the Health of the Nation Outcome Scales (HoNOS) and a quality of life instrument. To be able to implement this initiative on such a large scale required considerable investment in mental health providers, ongoing training and a broad program of engagement[79].

In New Zealand, mental health providers focus on monitoring of key indicators, such as seclusion and restraint minimization, and suicide reduction[80]. In the UK, the National Health Service Benchmarking Network[81] is a collaboration between all mental health provider organizations, which supply data to benchmark their own practice against others. The Benchmarking Network was developed because of the perceived inadequacy of the national data collection system and the lack of feedback on the large amount of data collected. As a ground-up initiative, the Benchmarking Network required a large degree of engagement and dynamic leadership.

In the US, national efforts are underway to identify cross-cutting mental health care quality measures and to determine who "owns" responsibility for improving quality. In the Department of Veterans Affairs, quality measures are set by central leadership for implementation in over 160 medical centers. While quality of mental health care in the Department has been widely documented, regional variations in processes and outcomes of care are common[82-86]. Hence, while regional service directors are ultimately responsible for improving quality, the Department has launched national initiatives to improve quality of care and reduce disparities in mental health care, notably through the implementation of the Uniform Mental Health Services Handbook[87] and the deployment of mental health care managers in primary care settings to promote integrated care. The Department has also sponsored the national implementation of evidence-based psychotherapy for post-traumatic stress disorder[88].

Pay-for-performance (now more often termed "value-based payment"[89]) models are also increasingly being advocated in the US and internationally. These initiatives reward providers for outcomes improvement and are also increasingly becoming used in mental health care[90,91]. Other innovations involve care beyond the clinic walls, including the measurement of recovery-oriented services[92] and incorporation of mobile health to capture outcome data[65,93]. The US Centers for Medicare & Medicaid Services is also deploying initiatives that seek to improve provider use/engagement in evidence-based practices as well as delivery system changes to sustain them. The main focus has been to integrate mental health treatment into primary care, where most patients with mental health symptoms initially present. The Institute for Healthcare Improvement Breakthrough series used business practices to integrate chronic illness care management for depression in primary care settings[94]. There

Figure 1 Multi-stakeholder roadmap for measuring and improving quality of mental health care

also exist other pockets of innovations in integrating mental health into primary care (e.g., the Health Care System Research Network, the Community Mental Health - Cherokee Health System[95]), but few frameworks to scale up and spread.

In the UK, the Commission for Quality and Innovation is implementing pay-for-performance for mental health, in which payments are based on meeting national quality improvement targets[96]. The targets are set locally, but with centrally agreed goals. Nonetheless, inevitable variations in care delivery make the development of quality measures a more difficult process in the mental health field.

Finally, there are emerging efforts to engage multi-stakeholder groups to solicit feedback throughout the entire process of quality measurement development and implementation. While frontline clinicians are often able to provide input for quality measures development, garnering feedback from consumers and their caregivers is also considered essential for buy-in[97]. Byron et al[98] describe a process of engaging stakeholders at all levels of measure development and implementation for Children's Health Insurance Program Reauthorization Act (CHIPRA) quality measures. The Measure Development Plan outlines the planned process, including engaging stakeholders[99]. The National Quality Forum uses a consensus process for review and endorsement of measures, including periods for public comment[100]. Moreover, the Centers for Medicare & Medicaid Services recently convened technical expert panels to help develop, select and maintain measures including clinicians, statisticians, quality improvement experts and methodologists[101].

RECOMMENDATIONS

We offer several recommendations for implementing quality measurement as an ultimate tool for improving quality of mental health care. First, health care payers and providers will need a portfolio of validated measures of patient-centered outcomes across a spectrum of conditions commonly experienced, as well as for special populations, including children/youth[102]. Moreover, valid measures that assess mental health care access are also needed, in order to more comprehensively determine quality of care beyond what happens within the clinical encounter. Measures need to be validated across the Donabedian spectrum (structure, process, outcome).

Second, common data elements should be developed and implemented for diagnoses, clinical measures and mental health "vital signs" and embedded within existing electronic health records and other information technology tools such as smartphones. Other elements that need to be standardized include coding in both electronic health records and administrative datasets for interventions such as medications, psychotherapies (including fidelity measures) and other treatments or care processes. Innovations such as natural language processing, or the automated capture of information from electronic medical records, are already being used to facilitate data capture for information (e.g., homelessness or suicide risk) not readily apparent from claims data.

Third, mental health outcomes will need to be assessed more routinely, and measurement-based care not only needs to be embedded within existing technologies, but should become part of the overall culture of the treatment setting and health care system. Regular outcome assessments have been linked to improvements in service delivery and lower readmission rates[103], whereas infrequent

outcome measurement did little to improve quality[104]. Moreover, routine outcome measurement that was fed back to the clinician and used to make joint treatment decisions with the patient did lead to better quality of life[105]. Quality measures need to be used in health systems that can generate near-real time data on quality in order to promote continuous quality improvement, and need to be monitored for unintentional consequences such as gaming.

Fourth, health systems need to provide investment, leadership and coordination to improve and link data sources in order to measure quality across settings. Systems will need to involve frontline providers and consumers in quality measurement endorsement and design measures that fit the needs of these providers and consumers rather than those of the administrators. Too often systematic quality outcome measurement is driven by a desire to inform policy or reduce expenditure rather than improve treatment decisions for individuals, which may have an adverse effect if staff (who are meant to be collecting the data) perceive it as a distraction with little value. Efforts like the UK Benchmarking Network are a good way of incorporating these perspectives[106].

Finally, health care systems need a valid way to stratify quality measures, in order to address potential gaps among subpopulations and identify groups in most need of quality improvement. A much greater expectancy for workforce training in and capacity for quality improvement is essential. Strategies for quality improvement and accountability need to be adapted, developed, and applied routinely in mental health settings.

In Table 2, we propose a broad multilevel process that outlines barriers to quality measurement and potential facilitators leading to quality improvement[107]. This process, based on the US National Academy of Medicine Learning Health Care System framework, is updated to include "levers" that address organizational barriers experienced in mental health care[108]. Learning health care systems leverage existing data (e.g., electronic health records) to deploy and evaluate innovations and best practices across health care organizations with the goal of improving population health.

CONCLUSIONS

Improving quality of mental health care is a team sport, requiring coordination across different providers, involvement of consumer advocates, and leveraging of resources and incentives from health care payers and systems. Figure 1 offers a roadmap for measuring and improving quality of mental health care. First, patients, providers and health care systems need to provide input on the choice of measures and their implementation. The steps to be taken include establishing an evidence base for quality measures through practice guidelines, operationalizing guidelines into quality measures that have a numerator and denominator based on data easily captured from health care settings, testing quality measures for their reliability and validity (ensuring that they also do not lead to gaming or manipulation), finalizing measures based on endorsement from patients, providers and system leaders as well as professional organizations, adopting the measures for use in routine practice, aligning measures across multiple settings (e.g., primary care, social services), and finally, identifying a group to "own" the measures that will continually monitor and provide strategies to incorporate quality improvement where necessary.

The recommendations for improving quality of mental health services presented here can apply to health care in general. Indeed, mental health has led the way in other health care innovations, including moving care into the community, use of innovative models of integrated care, as well as measures of patient-centered recovery. Moreover, there are lessons learned from mental health services that will inform the rest of health care to adopt a learning health care system. For years, mental health consumers and their family members have advocated for "patient-centered" care and greater focus on the personal goals of the patient, above and beyond receipt of medical services.

The diverse nature of mental health providers also challenges the health care system to take into consideration the perspectives of frontline staff including nurses, social workers, and increasingly peer specialists in owning quality improvement. It is not surprising that many of the quality improvement methods used in mental health care have influenced the growing field of implementation science[109], which is the study of provider behavioral change within the context of organizational constraints. Finally, the growth of value-based payment models that reward health systems and providers on achieving outcomes rather than on volume of services holds great promise for improving the quality of mental health care.

ACKNOWLEDGEMENTS

This paper was supported by the Department of Veterans Affairs, Veterans Health Administration, the Irving Institute, and the National Institutes of Health (R01 MH 099898). Additional funding support was provided by the Commonwealth Fund (grant no. 20141104) and by the National Center for Advancing Translational Sciences, National Institutes of Health (grant no. UL1 TR000040). The views expressed in the paper are those of the authors and do not necessarily represent those of the Department of Veterans Affairs, the National Institutes of Health, or other public entities.

REFERENCES

1. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry 2016;3:171-8.
2. De Hert M, Correll CU, Bobes J et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. World Psychiatry 2011;10:52-77.
3. Goldberg D. The detection and treatment of depression in the physically ill. World Psychiatry 2010;9:16-20.
4. Lin EH, Rutter CM, Katon W et al. Depression and advanced complications of diabetes: a prospective cohort study. Diabetes Care 2010; 33:264-9.
5. Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. JAMA Psychiatry 2015;72:334-41.
6. Carson N, Vesper A, Chen C et al. Quality of follow-up after hospitalization for mental illness among patients from racial-ethnic minority groups. Psychiatr Serv 2014;65:888-6.

7. Cook BL, Zuvekas SH, Carson N et al. Assessing racial/ethnic disparities in treatment across episodes of mental health care. Health Serv Res 2014;49:206-29.
8. Coleman KJ, Stewart C, Waitzfelder BE et al. Racial-ethnic differences in psychiatric diagnoses and treatment across 11 health care systems in the Mental Health Research Network. Psychiatr Serv 2016;67:749-57.
9. Case A, Deaton A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. Proc Natl Acad Sci USA 2015;112:15078-83.
10. Parks J, Svendsen D, Singer P et al. Morbidity and mortality in people with serious mental illness. Alexandria: National Association of State Mental Health Program Directors, 2006.
11. Liu NH, Daumit GL, Dua T et al. Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. World Psychiatry 2017;16:30-40.
12. 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.
13. Esposito E, Wang JL, Adair CE et al. Frequency and adequacy of depression treatment in a Canadian population sample. Can J Psychiatry 2007;52:780-9.
14. Harris MG, Hobbs MJ, Burgess PM et al. Frequency and quality of mental health treatment for affective and anxiety disorders among Australian adults. Med J Aust 2015;202:185-9.
15. Lopes CS, Hellwig N, de Azevedo e Silva G et al. Inequities in access to depression treatment: results of the Brazilian National Health Survey – PNS. Int J Equity Health 2016;15:154.
16. Lu CY, Roughead E. New users of antidepressant medications: first episode duration and predictors of discontinuation. Eur J Clin Pharmacol 2012;68:65-71.
17. 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.
18. 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.
19. Donabedian A. Evaluating the quality of medical care. Milbank Q 2005;83:691-729.
20. Donabedian A. The quality of care: how can it be assessed? JAMA 1988;260:1743-8.
21. Kilbourne AM, Fullerton C, Dausey D et al. A framework for measuring quality and promoting accountability across silos: the case of mental disorders and co-occurring conditions. Qual Saf Health Care 2010;19:113-6.
22. US Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington: National Academies Press, 2001.
23. US Institute of Medicine. Improving the quality of health care for mental and substance-use conditions. Washington: National Academies Press, 2006.
24. Hayes JF, Marston L, Walters K et al. Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000–2014. Br J Psychiatry 2017;211:175-81.
25. Wang PS, Demler O, Kessler RC. Adequacy of treatment for serious mental illness in the United States. Am J Publ Health 2002;92:92-8.
26. Pincus HA, Scholle SH, Spaeth-Rublee B et al. Quality measures for mental health and substance use: gaps, opportunities, and challenges. Health Aff 2016;35:1000-8.
27. US National Committee for Quality Assurance. The state of health care quality, 2016. Washington: National Committee for Quality Assurance, 2016.
28. Harris AHS, Gupta S, Bowe T et al. Predictive validity of two process-of-care quality measures for residential substance use disorder treatment. Addict Sci Clin Pract 2015;10:22.
29. Schmidt EM, Gupta S, Bowe T et al. Predictive validity of outpatient follow-up after detoxification as a quality measure. J Addict Med 2017;11:205-10.
30. Watkins KE, Paddock SM, Hudson TJ et al. Association between process measures and mortality in individuals with opioid use disorders. Drug Alcohol Depend 2017;177:307-14.
31. Watkins KE, Paddock SM, Hudson TJ et al. Association between quality measures and mortality in individuals with co-occurring mental health and substance use disorders. J Subst Abuse Treat 2016;69:1-8.
32. Schmidt EM, Gupta S, Bowe T et al. Predictive validity of a quality measure for intensive substance use disorder treatment. Subst Abus 2017;38:317-23.
33. Bremer RW, Scholle SH, Keyser D et al. Pay for performance in behavioral health. Psychiatr Serv 2008;59:1419-29.
34. Dausey DJ, Pincus HA, Herrell JM. Performance measurement for co-occurring mental health and substance use disorders. Subst Abuse Treat Prev Policy 2009;4:18.
35. Hepner KA, Watkins KE, Farmer CM et al. Quality of care measures for the management of unhealthy alcohol use. J Subst Abuse Treat 2017;76:11-7.
36. Kilbourne AM, Keyser D, Pincus HA. Challenges and opportunities in measuring the quality of mental health care. Can J Psychiatry 2010;55:549-57.
37. Martsolf GR, Osilla KC, Mandel D et al. Assessing the quality and value of psychological health care in civilian health plans: lessons and implications for the Military Health System. Rand Health Q 2016;5:16.
38. Watkins K, Horvitz-Lennon M, Caldarone LB, et al. Developing medical record-based performance indicators to measure the quality of mental healthcare. J Health Qual 2011;33:49-66.
39. Watkins KE, Keyser DJ, Smith B, et al. Transforming mental healthcare in the Veterans Health Administration: a model for measuring performance to improve access, quality, and outcomes. J Health Qual 2010;32:33-42.
40. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Q 1996;74:511-44.
41. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. JAMA 2002;288:1775-9.
42. Woltmann E, Grogan-Kaylor A, Perron B et al. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: systematic review and meta-analysis. Am J Psychiatry 2012;169:790-804.
43. Pincus HA, Jun M, Franx G et al. How can we link general medical and behavioral health care? International models for practice and policy. Psychiatr Serv 2015;66:775-7.
44. Richards DA, Bower P, Pagel C et al. Delivering stepped care: an analysis of implementation in routine practice. Implement Sci 2012;7:3.
45. Goorden M, Huijbregts KML, van Marwijk HWJ et al. Cost-utility of collaborative care for major depressive disorder in primary care in the Netherlands. J Psychosom Res 2015;79:316-23.
46. Morgan MAJ, Coates MJ, Dunbar JA et al. The TrueBlue model of collaborative care using practice nurses as case managers for depression alongside diabetes or heart disease: a randomised trial. BMJ Open 2013;3:e002171.
47. NHS England. Adult mental health: common mental health problems. Implementing the five year forward view for mental health. London: NHS England, 2016.
48. Hatfield D, McCullough L, Frantz SH et al. Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. Clin Psychol Psychother 2010;17:25-32.
49. Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? J Clin Psychiatry 2008;69:1916-9.
50. Trivedi MH, Rush AJ, Wisniewski SR et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 2006;163:28-40.
51. Oslin DW, Ross J, Sayers S et al. Screening, assessment, and management of depression in VA primary care clinics: the Behavioral Health Laboratory. J Gen Intern Med 2006;21:46-50.
52. Tew J, Klaus J, Oslin DW. The Behavioral Health Laboratory: building a stronger foundation for the patient-centered medical home. Fam Syst Health 2010;28:130-45.
53. US Department of Defense. Plan for development of procedure to measure data on mental health care provided by the Department of Defense. Washington: US Department of Defense, 2016.
54. US Institute for Clinical Systems Improvement. The DIAMOND program: treatment for patients with depression in primary care. Bloomington: US Institute for Clinical Systems Improvement, 2016.
55. US Centers for Medicare & Medicaid Services. Value-based payment modifier. Baltimore: US Centers for Medicare & Medicaid Services, 2017.
56. Zivin K, O'Malley A, Bigby J et al. Behavioral health integration in primary care: a review and implications for payment reform. Mathematica Policy Research, 2016.
57. Seibert J, Fields S, Fullerton CA et al. Use of quality measures for Medicaid behavioral health services by state agencies: implications for health care reform. Psychiatr Serv 2015;66:585-91.
58. US Centers for Medicare & Medicaid Services. Physician compare. Baltimore: US Centers for Medicare & Medicaid Services, 2017.
59. US Centers for Medicare & Medicaid Services. Hospital compare. Baltimore: US Centers for Medicare & Medicaid Services, 2017.

60. Patel MM, Brown JD, Croake S, et al. The current state of behavioral health quality measures: where are the gaps? Psychiatr Serv 2015; 66:865-71.
61. US Centers for Medicare & Medicaid Services. Inpatient Psychiatric Facilities Quality Reporting (IPFQR) Program. Baltimore: US Centers for Medicare & Medicaid Services, 2017.
62. US Institute of Medicine. Psychosocial interventions for mental and substance use disorders: a framework for establishing evidence-based standards. Washington: National Academies Press, 2015.
63. Gaynes B, Brown C, Lux LJ et al. Relationship between use of quality measures and improved outcomes in serious mental illness. Rockville: US Agency for Healthcare Research and Quality, 2015.
64. US Substance Abuse and Mental Health Services Administration. Recovery and recovery support. Rockville: US Substance Abuse and Mental Health Services Administration, 2015.
65. Ranallo PA, Kilbourne AM, Whatley AS et al. Behavioral health information technology: from chaos to clarity. Health Aff 2016;35: 1106-13.
66. Druss BG, Dimitropoulos L. Advancing the adoption, integration and testing of technological advancements within existing care systems. Gen Hosp Psychiatry 2013;35:345-8.
67. Riahi S, Fischler I, Stuckey MI et al. The value of electronic medical record implementation in mental health care: a case study. JMIR Med Inform 2017;5:e1.
68. Yeomans D. Clustering in mental health payment by results: a critical summary for the clinician. Adv Psychiatr Treat 2014;20:227-34.
69. UK Royal College of Psychiatrists. Royal College of Psychiatrists' statement on mental health Payment Systems (formerly Payment by Results). London: UK Royal College of Psychiatrists, 2014.
70. Wang R, Shaw I, Middleton H. Delaying the implementation of Payment by Results in mental health: the application of standardisation. Ment Health Rev 2015;20:156-65.
71. McShane M, Mitchell E. Person centred coordinated care: where does the QOF point us? BMJ 2015;350:h2540.
72. Harding KJ, Rush AJ, Arbuckle M et al. Measurement-based care in psychiatric practice: a policy framework for implementation. J Clin Psychiatry 2011;72:1136-43.
73. Fortney JC, Unützer J, Wrenn G et al. A tipping point for measurement-based care. Psychiatr Serv 2017;68:179-88.
74. Insel TR. Translating scientific opportunity into public health impact: a strategic plan for research on mental illness. Arch Gen Psychiatry 2009;66:128-33.
75. US Department of Veterans Affairs. VA is working to end homelessness among veterans. Washington: US Department of Veterans Affairs, 2017.
76. World Health Organization. WHO-AIMS mental health systems in selected low-income and middle-income countries: a WHO-AIMS cross-national analysis. Geneva: World Health Organization, 2009.
77. International Initiative for Mental Health Leadership. IIMHL and IIDL Annual report. Lambton Quay: International Initiative for Mental Health Leadership, 2016.
78. Dutch Association of Mental Health and Addiction Care. Performance indicators mental health care in the Netherlands. www.ggznederland.nl.
79. Burgess P, Coombs T, Clarke A et al. Achievements in mental health outcome measurement in Australia: reflections on progress made by the Australian Mental Health Outcomes and Classification Network (AMHOCN). Int J Ment Health Syst 2012;6:4.
80. New Zealand Mental Health and Addictions KPI Programme. KPI Dashboard for financial years 2013/14 to 2016/17 YTD (Jul-Dec). www.mhakpi.health.nz.
81. UK National Health Service Benchmarking Network. Work programme report 2016/17. Manchester: UK National Health Service, 2017.
82. Hussey PS, Ringel JS, Ahluwalia S, et al. Resources and capabilities of the Department of Veterans Affairs to provide timely and accessible care to veterans. Rand Health Q 2016;5:14.
83. Shields M, Rosenthal M. Quality of inpatient psychiatric care at VA, other government, nonprofit, and for-profit hospitals: a comparison. Psychiatr Serv 2017;68:225-30.
84. O'Hanlon C, Huang C, Sloss E et al. Comparing VA and non-VA quality of care: a systematic review. J Gen Intern Med 2017;32:105-21.
85. Watkins KE, Smith B, Akincigil A et al. The quality of medication treatment for mental disorders in the Department of Veterans Affairs and in private-sector plans. Psychiatr Serv 2016;67:391-6.
86. Hepner KA, Paddock SM, Watkins KE et al. Veterans' perceptions of behavioral health care in the Veterans Health Administration: a national survey. Psychiatr Serv 2014;65:988-96.
87. US Department of Veterans Affairs. Uniform Mental Health Services in VA Medical Centers and Clinics. Washington: US Department of Veterans Affairs, 2008.
88. Sayer NA, Rosen CS, Bernardy NC et al. Context matters: team and organizational factors associated with reach of evidence-based psychotherapies for PTSD in the Veterans Health Administration. Adm Policy Ment Health 2017; 44:904-18.
89. Roland M, Guthrie B. Quality and outcomes framework: what have we learnt? BMJ 2016; 354:i4060.
90. Bao Y, McGuire TG, Chan YF et al. Value-based payment in implementing evidence-based care: the Mental Health Integration Program in Washington state. Am J Manag Care 2017;23:48-53.
91. Unützer J, Chan Y-F, Hafer E et al. Quality improvement with pay-for-performance incentives in integrated behavioral health care. Am J Publ Health 2012;102:e41-5.
92. Pincus HA, Spaeth-Rublee B, Sara G et al. A review of mental health recovery programs in selected industrialized countries. Int J Ment Health Syst 2016;10:73.
93. Price M, Yuen EK, Goetter EM et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21:427-36.
94. US Institute for Health Care Improvement. The Breakthrough Series: IHI's collaborative model for achieving breakthrough improvement. Cambridge: US Institute for Healthcare Improvement, 2003.
95. Agency for Healthcare Research and Quality. Cherokee health systems. www.integrationacademy.ahrq.gov.
96. Ramanuj PP, Breslau J, Strathdee G et al. Carrots and sticks on opposite sides of the Atlantic: integration incentives for people with serious mental illness in England. Psychiatr Serv 2016;68:430-2.
97. US Institute of Medicine. Quality measurement. Psychosocial interventions for mental and substance use disorders: a framework for establishing evidence-based standards. Washington: National Academies Press, 2015.
98. Byron SC, Gardner W, Kleinman LC et al. Developing measures for pediatric quality: methods and experiences of the CHIPRA Pediatric Quality Measures Program grantees. Acad Pediatr 2014;14(Suppl. 5):S27-32.
99. US Centers for Medicare & Medicaid Services. CMS quality measure development plan: supporting the transition to the Merit-based Incentive Payment System (MIPS) and Alternative Payment Models (APMs). Baltimore: Centers for Medicare & Medicaid Services, 2016.
100. US National Quality Forum. Consensus development process. www.qualityforum.org.
101. US Centers for Medicare & Medicaid Services. Technical expert panels. Baltimore: US Centers for Medicare & Medicaid Services, 2017.
102. Hall CL, Moldavsky M, Taylor J et al. Implementation of routine outcome measurement in child and adolescent mental health services in the United Kingdom: a critical perspective. Eur Child Adolesc Psychiatry 2014; 23:239-42.
103. Slade M, McCrone P, Kuipers E et al. Use of standardised outcome measures in adult mental health services: randomised controlled trial. Br J Psychiatry 2006;189:330-6.
104. Ashaye OA, Livingston G, Orrell MW. Does standardized needs assessment improve the outcome of psychiatric day hospital care for older people? A randomized controlled trial. Aging Ment Health 2003;7:195-9.
105. Priebe S, McCabe R, Bullenkamp J et al. The impact of routine outcome measurement on treatment processes in community mental health care: approach and methods of the MECCA study. Epidemiol Psychiatr Soc 2002; 11:198-205.
106. NHS Scotland. Mental health project final report. Edinburgh: Scottish Government, 2008.
107. Pincus HA, Hough L, Houtsinger JK et al. Emerging models of depression care: multilevel ('6 P') strategies. Int J Methods Psychiatr Res 2003;12:54-63.
108. US Institute of Medicine. Best care at lower cost: the path to continuously learning health care in America. Washington: National Academies Press, 2013.
109. Chambers DA, Feero WG, Khoury MJ. Convergence of implementation science, precision medicine, and the learning health care system: a new model for biomedical research. JAMA 2016;315:1941-2.

DOI:10.1002/wps.20482

Order of operations in using expanded measurement to promote treatment quality improvement

Kilbourne et al[1] make a compelling case that measurement is required for mental health treatment quality improvement. While I agree that a "balanced portfolio" of structure, process and outcome measures will *ultimately* be needed for this purpose, I believe that the formidable barriers to quality improvement detailed by the authors make it critical to focus initially on outcome measurement.

To support this suggestion, I note that Kilbourne et al's assertion that low mental health treatment quality "is due in part to lack of systematic methods for measuring quality" is not entirely true. Such methods exist for measuring outcomes, but administrative burden is often cited as the reason for not using them. This is not justified. Primary care doctors routinely monitor hypertension and diabetes by measuring blood pressure and blood glucose. It would be less burdensome to assess patient self-reported anxiety, depression and other symptom constellations with existing validated self-report symptom scales. As noted by Kilbourne et al, patient engagement, quality of life and clinical outcomes all increase significantly when this is done. And, importantly, there would be no need for complex infrastructure, as this kind of ongoing outcome monitoring can be done using paper and pencil charting.

Why, then, do mental health clinicians not engage in systematic outcome monitoring? The overview of barriers presented by Kilbourne et al mostly apply to process measures (which make up most measures in mandated quality reporting programs). Process measures are difficult to use for quality assurance, especially for non-pharmacologic treatments, because the subtleties of treatment quality cannot be captured in structured measures or administrative data. Clinicians consequently lack enthusiasm for such measures. The situation is somewhat different in the case of pharmacologic treatments, as it is generally possible to abstract from claims data information on prescribing patterns (although often not on whether prescriptions are filled) and frequency of visits. But even here it is of great value also to track self-reported patient symptoms, as such data can help evaluate clinician decisions about medication titration, switching and augmentation.

The only barriers discussed by Kilbourne et al that apply to outcome assessment are "administrative burden", the possibility that "consumers of mental health services might feel burdened by the data gathering", and the fact that few mental health treatment providers have received training in quality measurement. These are all barriers that can be overcome. With regard to administrative burden, I noted above that primary care physicians routinely engage in monitoring activities that are more administratively burdensome. The notion that patients might feel burdened is similarly unconvincing, given that primary care patients routinely put up with much more onerous tests. In addition, perceived burden would presumably be reduced if clinicians informed patients that systematic symptom tracking leads to improved treatment outcomes, and if checklist responses were used as starting points for more nuanced discussions with patients about recent symptoms. Nor is the fact that many mental health treatment providers have never received training in quality measurement a major barrier. In the case of cognitive behavioural therapy, ongoing patient outcome monitoring is a standard part of the clinical process[2]. It should not be difficult to convince other clinicians to engage in similar outcome monitoring if relatively simple technological approaches were available to facilitate patient reporting and generation of benchmarked symptom tracking curves. Easy-to-use products to do this are beginning to become available (e.g., http://www.mobiletherapy.com/; http://www.selfecho.com/).

I believe a more important barrier to outcome monitoring is one not mentioned by Kilbourne et al: that outcome assessments could be aggregated by payers and used to impose financial constraints (e.g., pay-for-performance) on treatment providers. Although risk adjustment methods exist to address the problem of adverse selection[3], uncertainties inherent in this process create reluctance to be held responsible for outcomes on the part of both clinicians and health plans. Yet we know from case studies, most notably the New York State Coronary Artery Bypass Surgery Report-Card System[4], that mandated risk-adjusted outcome assessment can be highly effective in improving patient outcomes even in the absence of publically available measures of structure or process. This occurs because unrestricted public reporting of risk-adjusted comparative outcome scores leads to market pressures that encourage quality improvements.

Importantly, measures of structure and process are centrally involved in guiding these quality improvements, but these measures tend to be part of something referred to by management consultants as "research to know" (i.e., for internal use to help treatment organizations detect problems they need to fix to improve patient outcomes), whereas risk-adjusted outcome measures are "research to show" (i.e., for public disclosure as ultimate measures of treatment quality). It is critical to appreciate the necessary order of operations and distinct roles of these different kinds of measurement. Publically reported risk-adjusted outcome measures need to come first to create market pressures for quality improvement. Internal structure and process measures come next and are implemented by treatment organizations to improve quality of care in an effort to change scores on publically-available outcome measures.

Based on these considerations, I believe that it is critical to focus initially on

risk-adjusted outcome measures. The obvious way to do this is for governments or other organized payers to mandate outcome assessments, carry out centralized risk adjustment, and report results publically at as low a level of aggregation as possible. Kilbourne et al mention initiatives along these lines in Australia and the Netherlands. The Access to Psychological Therapies program in England is another example[5]. The International Consortium of Health Outcomes Measurement is attempting to develop standard sets of outcome measures for this purpose across many different medical conditions, and to create an implementation network to help facilitate data collection and risk adjustment[6].

When governments and other organized payers are reluctant to embrace outcome assessments, I would suggest that organized consumer groups might fill the gap by creating an electronic system that allows patients to provide information about their conditions and relevant background at the beginning of treatment on a patient-report website, and to complete ongoing symptom tracking surveys on the same site. The baseline data would be used by the system developers for risk adjustment, and the tracking data would be used to generate patient-specific treatment response curves that could be made available to clinicians for monitoring treatment response. Patients would have an incentive to participate to provide information known to promote treatment quality improvement. Once such a system is in place, risk-adjusted treatment quality profiles could be generated to create the market pressures needed to encourage providers to engage in quality improvement initiatives.

Once such a system becomes the norm, the "balanced portfolio" of structure, process and outcome measures called for by Kilbourne et al will evolve naturally, with payers using outcome assessments for pay-for-performance, and provider groups using structure and process measures for continuing quality improvement. But order of operations is important. The process needs to begin with risk-adjusted outcome assessments to create market pressures for quality improvement and to provide objective standards for quality assurance, with structure and process measures used primarily by provider organizations for internal purposes to improve patient outcomes.

Ronald C. Kessler
Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

1. Kilbourne AM, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
2. Dobson KS. Handbook of cognitive-behavioral therapies, 3rd ed. New York: Guilford, 2010.
3. Centers for Medicare & Medicaid Services. HHS-operated risk adjustment methodology meeting discussion paper, 2016. https://www.cms.gov.
4. Hannan EL, Cozzens K, King SB et al. J Am Coll Cardiol 2012;59:2309-16.
5. Clark DM. In: McHugh KR, Barlow DH (eds). Dissemination and implementation of evidence-based psychological interventions. New York: Oxford University Press, 2012:61-77.
6. Foley M. Dr Christina Åkerman Interview. http://www.learninghealthcareproject.org.

DOI:10.1002/wps.20483

Improving the quality of global mental health care requires universal agreement on minimum national investment

Kilbourne et al[1] provide an informative review of current theory and approaches to the measurement of quality of mental health services from a number of higher income countries across the world. A welcome emphasis is given to social outcomes, and the authors note that quality of life, personal recovery and community tenure are as relevant as more traditional outcomes such as symptoms and functioning.

The authors also acknowledge that any outcomes framework needs to take into account variables such as morbidity and socioeconomic factors, to avoid "cherry picking" and gaming by providers. This is an important point and one that is as relevant to social outcomes as "clinical" outcomes, but more difficult to adjust for. Quality of life is a notoriously slippery concept which has a complex relationship with relative expectations of what a "good" quality of life comprises[2]. Similarly, personal recovery is, by definition, a subjective concept and it is no surprise that the development of valid recovery measures has been hampered by a lack of consensus amongst providers, researchers and service users. This is unlikely to be solved by further investment in tool development.

Apart from the problems of actually measuring relevant outcomes in mental health systems, a major issue with many existing measurement based schemes is that they only focus on the simpler parts of the system. The three examples from mental health systems given by Kilbourne et al (the UK's Improving Access to Psychological Therapy, the Dutch Depression Initiative and the Australian TrueBlue model) are all primary care facing models that aim to address common mental disorders. These services deliver specific, time limited, evidence based interventions and are ideally suited to straightforward monitoring of their structures, processes and clinical outcomes. This has led to greater investment and their being embedded into national service models in the UK and the Netherlands.

However, people with more complex mental health problems tend to require multiple interventions from multiple services, often spanning statutory health, social care and non-governmental organization providers. The problem of identifying standard, universal metrics and measures that can capture the impact of these complex arrangements in order to assess whether "quality" is being delivered has, unsurprisingly, proved insoluble, not least because

social outcomes are often more relevant to this group than clinical outcomes such as symptoms.

The abandonment of the outcome based reimbursement system for mental health in the UK probably had more to do with this issue than with administrative burden or risk of gaming. The main clinician rated outcome measure that was under consideration, the Health of the National Outcome Scale (HoNOS)[3], is one of the most widely used mental health outcome assessment tools worldwide, but there are concerns about its appropriateness and sensitivity to change for people with longer term and more complex mental health problems. As such, it cannot reliably indicate whether a service is providing effective care and should be reimbursed.

In Australia, universal routine outcome data (including HoNOS) have been collected systematically for around 20 years, but this has not stopped the gradual disinvestment in statutory mental health services for those with the most complex needs, and concerns are now being raised about the quality of care provided by other sectors for this group[4].

A bigger issue, something of an elephant in the room, is that there is not such good evidence that improving quality of care actually leads to better clinical outcomes, particularly when we consider longer term, complex conditions. Evaluation of the impact of the national Quality and Outcomes Framework for diabetes care in the UK found no clear association with improved clinical outcomes over the three years before and after its introduction[5].

Nevertheless, it would clearly be counterproductive not to attempt to understand how to organize services to be as safe, effective and efficient as possible. The difficulty in identifying robust universal measures for mental health services that can do this may explain why, as Kilbourne et al point out, most "outcome" measures are actually process measures. In complex systems such as these, it is much easier to describe what you are doing than to assess whether it has had an impact. Perhaps New Zealand has adopted the most pragmatic approach: to focus on monitoring key indicators that can be agreed on as universal markers of basic service quality, such as the minimization of seclusion and restraint, and suicide reduction[6].

Indeed, the increased support for "pay-for-performance" or "activity" rather than "payment for results" models probably reflects a growing acceptance that there is no simple way to assess outcomes in most mental health services. Consequently, comparative benchmarking that uses various process metrics has become increasingly popular in England and Wales through the voluntary National Health Service benchmarking network. However, this can only work within a publicly funded system where sharing data does not potentially threaten an organization through competitive market forces.

Finally, the biggest issue (an even larger elephant) is resourcing. Across the world, most countries lack even basic mental health care. The nuances of different approaches to quality assessment in higher income countries pale into insignificance when considering the appalling consequences of this. Globally, most people with serious mental health problems are in long-term institutions, often living in unacceptable, inhumane conditions[7]. Taylor Salisbury et al[8] recently showed that, across Europe, the proportion of the national health budget spent on mental health was positively correlated with the quality of the country's longer term facilities.

It seems that adoption of a universal national minimum percentage investment in mental health care should be the first crucial step in any global quality improvement initiative.

Helen Killaspy
Rehabilitation Psychiatry, Division of Psychiatry, University College London, London, UK

1. Kilbourne A, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
2. Carr A, Gibson B, Robinson P. BMJ 2001;322:1240.
3. Wing J, Beevor A, Curtis R et al. Br J Psychiatry 1998;172:11-8.
4. Morgan V, Waterreus A, Carr V et al. Aust N Z J Psychiatry 2016;51:124-40.
5. Calvert M, Shankar A, McManus R et al. BMJ 2009;338:b1870.
6. New Zealand Mental Health and Addictions KPI Programme. Dashboard for financial years 2013/14 to 2016/17 YTD (Jul-Dec). www.mhakpi.health.nz.
7. Saxena S, Thornicroft G, Knapp M et al. Lancet 2007;370:878-89.
8. Taylor Salisbury T, Killaspy H, King M. Br J Psychiatry 2017;211:45-9.

DOI:10.1002/wps.20484

Exploiting routine data for international benchmarking of quality in mental health care

The paper by Kilbourne et al[1] provides an extensive overview of the challenges in assessing quality of mental health care. Service users, informal carers, policy makers and the general public increasingly demand that mental health systems provide good "value for money", and thus the need for validated, meaningful and purposeful data on quality of mental health care is growing. As outlined by the authors, many countries have taken actions to identify, define, collect and analyze such data.

In parallel with national activities, there is a growing interest for international benchmarking of mental health systems to inform national policies. The challenges in standardizing measurements become even larger when comparing mental health systems in different countries, due to differences in those systems and, in many cases, absence of common international definitions. Common indicator definitions and standardized data collection procedures are prerequisites for meaningful benchmarking between countries.

In spite of the above-mentioned challenges, international benchmarking is

an important moving force to foster development of mental health services in countries. For feasibility reasons, such benchmarking endeavours have to rely on routinely collected data, which tend to be dominated by hospital care data. So far, international comparisons of mental health systems have relied on existing administrative databases. Today, the increasing use of centralized repositories of electronic medical records presents a largely unchartered area of new possibilities for collecting and assessing quality of care data. It needs to be considered that data mining of electronic medical records will always require special attention to data security and confidentiality.

Electronic medical records can be extended to include patient-reported outcomes and thus integrate the patients' views. In the future, artificial intelligence may be applied to the analysis of data available from central repositories, contributing to the learning health care system. Kilbourne et al[1] touch upon these possibilities, but a clear vision and detailed roadmap is needed to outline the research and policy actions that will enable us to use clinician- and patient-derived information from repositories of electronic health records in order to assess and improve mental health care quality. A first step is to facilitate widespread adoption of patient-reported outcome data collection into electronic health record systems[2].

In the European Union, the need for comparable data and standardized definitions to enable comparisons of national mental health system performance has been identified, and processes to harmonize indicators and data collection have been initiated[3,4]. Indeed, international comparisons based on disparate and non-harmonized data may cause more confusion than clarification[5]. In an effort to foster comparability of data, most European countries have joined the Health Care Quality Indicators project, led by the Organisation for Economic Co-operation and Development. The project's key areas include mental health care indicators but, due to variation in health care systems, so far only four quality indicators have been implemented and are reported annually; two of these relate to suicides during or after a hospitalization, and two are based on excess mortality in schizophrenia and bipolar disorder[6].

Re-hospitalization within 30 days after discharge from a psychiatric ward has been suggested as a further indicator for international comparisons of quality in mental health care, but uncertainties regarding the meanings of this indicator remain: is it an indicator of poor hospital care and premature discharge, or does it reflect insufficient community services and lack of continuity of care? Unplanned re-hospitalizations are often disruptive for the patient and constitute a strain on limited health care resources. However, in some mental health systems, planned re-hospitalizations are an integral part of individual treatment plans, making differences in readmission rates difficult to interpret.

The recent CEPHOS-LINK (Comparative Effectiveness Research on Psychiatric Hospitalization by Record Linkage of Large Administrative Data Sets) project compared re-hospitalization and its predictors in six different European countries, based on retrospective cohort studies with data from country-specific large electronic health care registries. The study showed a clear interaction of case-mix and country with readmission rates, even after harmonizing the national datasets[7].

The European Commission-funded BRIDGEHEALTH (BRidging Information and Data Generation for Evidence-Based Health Policy and Research) project recently recommended to establish a European Research Infrastructure Consortium on health information for collection of comparable information on health system performance. In response to this recommendation, European Union member states are currently setting up a joint action for health information as a first step towards common governance, collection and analysis of health and health care information. Rendering large national health care databases and electronic health records repositories interoperable and thus comparable across countries is essential for comparing outcomes and processes in health service utilization in different countries.

The concluding recommendations by Kilbourne et al[1] outline the way forward. However, like in all team sports, international rules are needed even for friendly tournaments. Involvement in the development of international standards for assessment of quality in care and adherence to internationally agreed standards will enable us to make meaningful comparisons of mental health systems. This is of special importance in times of electronic records, which open up new possibilities for assessment of quality of care.

It is not yet clear to which extent the above-mentioned collaborative effort of the European Commission and the member states will cover harmonization and collection of relevant mental health systems data for comparison and benchmarking. From a European perspective, a joint mental health observatory is urgently needed to lead development and implementation of monitoring of mental health service provision. Such a centre would enable development and worldwide dissemination of indicators that reflect the European mental health care values of universality, access to good quality care, equity, and solidarity[8].

Kristian Wahlbeck
Finnish Association for Mental Health, Helsinki, Finland

1. Kilbourne AM, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
2. Ahmed S, Ware P, Gardner W et al. J Clin Epidemiol 2017;89:160-7.
3. Wahlbeck K. Eur Psychiatry 2007;22:I-III.
4. Wahlbeck K. Epidemiol Psychiatr Sci 2011;20:15-8.
5. Ekholm O, Bronnum-Hansen H. Scand J Publ Health 2009;37:661-3.
6. Organisation for Economic Co-operation and Development. Health at a glance 2015: OECD indicators. Paris: OECD Publishing, 2015.
7. National Institute for Health and Welfare, Finland. www.cephos-link.org.
8. Council of the European Union. Official Journal of the European Union 2006;C 146:1-3.

DOI:10.1002/wps.20485

Increasing equity in access to mental health care: a critical first step in improving service quality

As Kilbourne et al[1] describe, measuring quality of mental health care serves as an important step towards reducing service inequities. However, quality measurement that predominantly focuses on treatment outcomes overlooks individuals with mental health needs who cannot access value-based treatments. Untreated mental health and substance use disorders are associated with premature mortality, productivity loss, high rates of disability, and increased risk for chronic disease. Thus, ensuring equitable access within a value-based framework is needed to not only close existing treatment gaps but also to improve patient outcomes.

The degree of inequity in access to mental health care varies among countries with different models of health care system and welfare regimes. Findings from a study of seventeen low-, middle- and high-income countries revealed low mental health care utilization despite documented high need: in each country, at least two-thirds of individuals with common mental disorders went untreated[2]. Twelve-month service utilization rates also tended to be lower in less developed countries and to align with the percentage of gross domestic product spent on health care[3]. Furthermore, members of socially disadvantaged groups such as ethnic/racial minorities and low-income patients have lower mental health service utilization compared with members of advantaged groups[4].

Inequities in access to mental health care can arise due to myriad reasons, including: eligibility criteria to enter programs (e.g., receiving a specific required diagnosis); lack of linguistic capacity; policies that discriminate based on legal status (e.g., refugees, immigrants, racial/ethnic minorities); lack of information regarding where and how to obtain care; and logistical, psychological and economic barriers (e.g., transportation, childcare, beliefs about self-sufficiency, stigma-related concerns, concerns about privacy, long waiting times for services, high costs, or inflexible work schedules). To adequately document mental health care inequities, measures and procedures to evaluate access must be consistently and globally implemented across mental health care systems. But you cannot evaluate what you have not measured, and unmet need is typically absent from conventional administrative or service data.

Methods of measuring access might focus on one's potential ability to access care, including the package of benefits included under mental health coverage and the availability of appropriate and effective service providers within reasonable geographic proximity. Additionally, access measures should incorporate obstacles that arise once someone has decided to enter care, such as insufficient choice of providers, low doses of services, and ineffective or low-quality services. In a value-based framework, both horizontal equity (understood as the provision of equal care for equal needs) and vertical equity (understood as different treatments for people with different needs or preferences) must be considered when striving for equitable access.

Conducting national surveys of household and institutionalized individuals every five to ten years might offer a benchmark of those who need care and where they are located. Results of such investigations can also inform testable hypotheses about why some individuals do not receive services, including explanations related to specific preferences for care. Additionally, quick assessments obtained through computerized adaptive testing might simplify diagnostic evaluation and assure linguistic diversity, as well as afford more attention to differential item functioning (i.e., the extent to which an item measures different abilities for members of different groups), so that providers can adequately operate across cultures, diverse populations, and languages[5]. Measuring both barriers and facilitators to mental health and substance use treatment access through geographic mapping can also provide a more comprehensive picture of specific areas in need of immediate intervention. Finally, overlapping measures of need with administrative service use data can facilitate allocation of resources, adjustments of risk for inclusion of underserved populations, and payment incentives for providers to reach those with unmet behavioral health needs.

But, before these methods can be widely adopted, a shift in the purpose of medical records (i.e., from being used mostly for billing to being used predominantly to monitor access, quality, and patient's service preferences) must come into vogue. Stratification by need level (e.g., those with comorbid conditions), age, race/ethnicity, income, sexual orientation, gender, urbanicity, or linguistic subgroups can assist in isolating where there are inequities and who or what is responsible for them. However, it will likely be more difficult to make the health care system accountable for collecting these data, given the inadequate budgets and resources granted by ministries of health for mental health care[6].

Despite having several potential methods to measure access to mental health and substance use care, systems may not utilize these methods in a meaningful way if they are not incentivized to do so. Although reporting requirements tied to provider accreditation or funding vary across oversight agencies, states and countries, they typically focus on service outcomes of those *in care* rather than outcomes of those *eligible for services*. Recommended performance metrics that include access to care have been proposed by several relevant organizations but, without some form of mandated accountability, health care administrators do not reliably collect or report this information[7].

More research should focus on strategies to make service administrators and policy makers responsive to reducing access inequities and incentivized to develop

leadership, implementation plans, and resources to ensure prompt action. As the field stands now, recommended paths seem to converge on care that maximizes value for patients rather than volume and profitability of services[8]. Transitioning to value-based care delivery should force health care systems to focus more on potential and existing patients, as well as those patients' preferences and needs.

Many creative solutions have been offered to increase access, including integrating behavioral health services into primary or community-based care, augmenting the workforce through task-shifting (e.g., utilizing community health workers or peer navigators to provide some services), imparting training and supervision to novel providers via the Internet, or delivering services to people where they live (e.g., via minute clinics, medical vans, or telemental health services) rather than expecting people to travel long distances to access services[9,10].

But it is surprising that, given all we know about how to expand access, re-engineering service delivery seems sluggish. Why is access not a priority given existing rates of untreated mental health and substance use disorders? Have we not made a compelling case to policy makers, the general population, or health care system leaders? Do we need more evidence than the opioid epidemic, the massive incarceration of people with mental health conditions, or the suicide pandemic?

Now, more than ever, maximizing patient outcomes will require reaching out to patients earlier in their illness trajectory, helping them recognize mental health needs, and making them co-leaders in their care. It might necessitate psychoeducation dissemination campaigns, home visits, and continuous communication to understand what patients prefer as high-value health care delivery. It will also entail measuring mental health outcomes that matter to the patient rather than mental health outcomes related to symptoms, even when those patients do not always come to care.

We have an ethical obligation to make our communities healthier, with a universal approach to treatment rather than treatment for the very few. Like Martin Luther King Jr. said, "of all forms of inequity, injustice in health care is the most shocking and inhumane".

Margarita Alegría[1,2], Ora Nakash[1,3], Amanda NeMoyer[1,4]
[1]Disparities Research Unit, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA; [2]Departments of Medicine and Psychiatry, Harvard Medical School, Boston, MA, USA; [3]Baruch Ivcher School of Psychology, Interdisciplinary Center Herzliya, Herzliya, Israel; [4]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

The authors are supported by the US National Institute on Aging (grant no. R01AG046149), the US National Institute on Minority Health and Health Disparities (grant no. R01MD009719) and the US National Institute of Mental Health (award no. T32MH019733). The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of any funding institutions.

1. Kilbourne AM, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
2. Wang PS, Aguilar-Gaxiola S, Alonso J et al. Lancet 2007;370:841-50.
3. Kakuma R, Minas H, van Ginneken N et al. Lancet 2011;378:1654-63.
4. Jimenez DE, Cook B, Bartels SJ et al. J Am Geriatr Soc 2013;61:18-25.
5. Gibbons RD, Weiss DJ, Kupfer DJ et al. Psychiatr Serv 2008;59:361-8.
6. World Health Organization. Mental health atlas 2011. Geneva: World Health Organization, 2011.
7. Essock SM, Olfson M, Hogan MF. Int Rev Psychiatry 2015;27:296-305.
8. Porter ME, Lee TH. Harvard Bus Rev 2013;91:1-19.
9. Fairburn CG, Patel V. Am J Psychiatry 2014;171:495-8.
10. Lund C, Tomlinson M, Patel V. Br J Psychiatry 2016;208(Suppl. 56):s1-3.

DOI:10.1002/wps.20486

Mental health quality improvement goes global

The field of quality assessment in health care traces its origins back more than 50 years to A. Donabedian's seminal article outlining a framework for understanding quality of care from a health systems perspective[1].

Quality improvement in mental health is a younger enterprise. A 2006 report from the US Institute of Medicine outlined the challenges in assessing and improving quality of care for mental disorders, including lack of standardized approaches for diagnosing mental health and substance use disorders; limited evidence base supporting current quality measures, and fragmentation and lack of information technology infrastructure to measure quality. Nonetheless, the report argued that quality improvement principles could, and should, be applied to efforts to improve mental health care[2].

Since the publication of that report over a decade ago, innovations in health technology, the growth of large integrated health systems, and movement of mental health into the mainstream of health care have helped spur a rapid growth of mental health quality improvement in the US[3].

Kilbourne et al's important paper[4] provides an update on the state of mental health care quality improvement worldwide. The authors describe initiatives to measure and improve care within the US and other developed countries, and make recommendations to better incorporate them into routine practice. The paper provides a valuable framework for understanding quality improvement from an international perspective.

There are tensions inherent in mental health care quality improvement that become particularly evident in efforts to consider this process from a global perspective. Historically, performance measurement systems in the US and other developed countries have been built on existing administrative datasets[5]. The elements of these datasets vary across countries and health systems – for instance, fee-for-service systems typically aggregate billable claims, whereas countries whose systems focus on inpatient care primarily collect data from these settings.

Differences in structure and financing of mental health systems may shape the availability of data and constrain the col-

lection of uniform measures across countries. They may also change the relative priorities placed on measures across domains of care. For instance, inpatient readmission has become a focus of incentive payment systems in the US Medicare program, where cost reduction and shifting to outpatient care have been central policy goals[6].

Low- and middle-income countries may face particular challenges in measuring quality. These countries commonly lack an information technology infrastructure to measure and track quality. Quality assessment in these countries typically relies heavily on counts of facilities or providers per capita[7]. The World Health Organization uses the treatment gap – the proportion of those with a mental disorder diagnosis who do not receive any care – as a common metric of unmet need across countries[8]. This measure has been important from a policy and advocacy perspective. However, it also has limited ability to capture more nuanced views of quality of care among individuals once they have begun treatment.

The field of mental health care quality assessment is currently moving from an emphasis on measuring delivery of care towards assessing clinical outcomes. This is an important development, since improving outcomes is a central goal of any health system. However, it is important to keep in mind that, more than with other health conditions, mental disorders can only be understood in their individual and social contexts. This is particularly the case for functional outcomes such as quality of life, employment and recovery. There is a need to further develop the science of mental health functioning, and to understand how to adapt and interpret these measures of function across countries and cultures.

The most important function of quality assessment is in using these measures to improve care. Moving from quality assessment to care improvement requires attention to both desired goals and potential unintended consequences. D. Campbell famously noted that, once a quantitative indicator is used for decision making, there is the potential for the measure and the underlying processes to become corrupted[9]. Pay-for-performance and public reporting of outcome measures can create incentives to cherry-pick patients who are more likely to improve. They can also lead providers to "teach to the test" and neglect domains of care that are difficult to measure. These potential challenges are universal, but may take different forms across countries based on incentive structures and organization of health systems.

We should approach quality improvement like we do with any other intervention in medicine – as a powerful tool that carries both potential benefits and risks. We need to develop and test the best practices for improving quality of care across different systems. A variety of social science disciplines can help inform these efforts. Behavioral economics can provide guidance on how to structure monetary and non-monetary incentives to change mental health provider and patient behaviors. Implementation science can provide insights into how to disseminate mental health interventions such as quality improvement within complex systems of care. Medical anthropology can provide a greater understanding of how to understand and improve mental health function across countries and cultures.

These challenges and opportunities are not unique to mental health care quality improvement. In the final interview prior to his death, Donabedian highlighted the importance of focusing on patients' experiences rather than simply on the technical aspects of care in quality improvement efforts: "The view of quality that is taken in the hospital is really limited to technical competence and, more recently, to superficial attention to the interpersonal process. The role of the doctor is to actively make sure that the patient arrives at a decision that is a reasonable one for him or her. . . it is the ethical dimension of individuals that is essential to a system's success"[10].

This emphasis on the patient should be the central organizing principle in future efforts to measure and improve mental health care across the globe.

Benjamin G. Druss
Rollins School of Public Health, Atlanta, GA, USA

1. Donabedian A. Milbank Mem Fund Q 1966; 44(Suppl. 3):166-206.
2. US Institute of Medicine. Improving the quality of health care for mental and substance-use conditions. Washington: National Academies Press, 2006.
3. Pincus HA, Scholle SH, Spaeth-Rublee B et al. Health Aff 2016;35:1000-8.
4. Kilbourne AM, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
5. Lauriks S, Buster MC, de Wit MA et al. BMC Publ Health 2012;12:214.
6. Berenson RA, Paulus RA, Kalman NS. N Engl J Med 2012;366:1364-6.
7. Saxena S, Lora A, van Ommeren M et al. Psychiatr Serv 2007;58:816-21.
8. Kohn R, Saxena S, Levav I et al. Bull World Health Organ 2004;82:858-66.
9. Campbell D. Eval Program Plann 1976;2:67-90.
10. Donabedian A. Health Aff 2001;20:137-41.

DOI:10.1002/wps.20487

Why measuring quality of mental health care is still an unmet challenge and how to meet it

Quality improvement programs aim to reduce clinical decision variability due to lack of knowledge or subjectivity as well as the gap between guidelines and real world practices. In psychiatry, the attempt to disseminate these initiatives has generated two opposite reactions: on the one hand, some researchers have been fascinated with the idea of moulding clinical practice based exclusively on research findings, considering all deviations from scientific evidence as an inappropriate action; on the other, several clinicians have regarded the entire process of quality improvement as biased by reductionism, and as an expression of disdain of the value of their own experience.

Substantial gaps between evidence and

practice, and non-attainment of optimal health outcomes, are well documented in the whole field of medicine, even in high-income countries and across disciplines. A seminal review of studies on quality of primary care conducted in UK, Australia and New Zealand in 2001 concluded that "in almost all studies the process of care did not reach the standards set out in national guidelines or set by the researchers themselves"[1]. In the US, a more recent analysis of 48 state and regional measure sets found that they included more than 500 different measures, only 20% of which were used by more than one program; a study of 29 private health plans identified approximately 550 distinct measures, which had little overlap with the measures used by public programs[2]. In the UK, the translation of guidelines into legal standards of care in the field of surgery and anesthesiology has been found not to have a sufficient endorsement by the relevant professional bodies[3]. In the Netherlands, a study performed in two different regions proved that several professionals and patients experienced barriers in guideline implementation in the field of obstetrics and gynecology[4].

In psychiatry, quality of care programs have to face even more significant problems in terms of feasibility: heterogeneity of services in the various countries or regions; high professional burden; uneven provision of resources; inequality in the access to care; lack of support or rewards for services or professionals undertaking initiatives based on quality improvement; factors beyond the control of clinicians that interfere with treatment outcomes; the necessity to integrate interventions provided in different settings[5].

Evidence collected from non-research clinicians and clinical consensus may play a particularly salient role in this regard. Trying to understand the reasons for current discrepancies between routine practice and guideline recommendations, and promoting a shared culture of quality that pays appropriate respect to the clinicians' experience[6], are fundamental ingredients for the success and dissemination of quality of care initiatives[7].

Routine outcome monitoring, performed at the local level, including both objective and subjective variables, has been thought to have the potential to facilitate quality assurance programs. So far, however, there have been inconsistent findings even from studies performed in similar contexts, and systematic reviews have been inconclusive on the regular feedback and use of those data for improving patient outcomes or management[8].

Any effort to implement quality indicators should not overlook a number of ethical and theoretical questions. What is the role of ethical values in shaping practices? We know that the mechanisms involved in the collection of evidence are rather complex and sometimes include economic interests. The concepts themselves of "effectiveness" and "quality of care" might differ markedly between clinicians and administrators. Furthermore, some aspects of treatment are more easily measurable and likely to be standardizable (e.g., drug treatments) than others (psychosocial treatments or psychodynamic therapies), with important consequences on the amount of evidence that is available. What course of action should a clinician take when confronted with different levels of evidence? What are the limits of evidence? What should be done in clinical situations for which there is no scientific evidence?

On the other hand, we should not forget that providing evidence-based treatments and quality measurements is indeed an ethical imperative. It means not only respecting patients' rights and mental health workers' competences, but also professionals' right and duty to be updated on and to appropriately discuss emerging practices.

Furthermore, in order to be effective, quality assessment programs should be solidly rooted in the frame of the specific context, which involves: a) gathering information on the core values and norms underlying ideologies and assumptions within an organization; b) identifying the level of commitment and motivation for improvement of professionals and support by leaders; c) clarifying the extent to which the implementation of the programs is likely to actually improve the quality of care provided rather than impairing the process (for example, by subtracting resources dedicated to the direct care of patients in favor of bureaucratic practices); d) involving service users in the identification of emergent service needs[9]; e) remaining vigilant to the risk that local adaptation alters a given recommendation to better adjust it to real practice.

This latter issue has become quite pressing: adherence to specific programmatic standards – frequently referred to as fidelity of implementation – is necessary to produce the expected outcomes. However, this issue too is not a simple one. It is much easier to manualize interventions in research settings than in the real world. Some evidence suggests that the best outcomes are obtained by therapists who demonstrate flexibility by adapting interventions to their individual clients' needs[10]. The degree to which interventions should involve flexible modification and integration to be effective when applied in everyday clinical work remains a matter of debate.

In conclusion, all actors involved in mental health planning, organization and routine care processes are called upon to work towards obviating the risk that, due to negligence or a lack of coordination or consideration of the local context, even the most advanced mental health care models miss the opportunity to significantly improve routine care practices.

Mirella Ruggeri
Section of Psychiatry, Department of Neuroscience, Biomedicine and Movement, University of Verona, Verona, Italy

1. Seddon ME, Marshall MN, Campbell SM et al. Qual Health Care 2001;10:152-8.
2. Cassel CD, Conway PH, Delbanco SF et al. N Engl J Med 2014;371:2145-7.
3. Fearnley RA, Bell MDD, Bodenham AR. Br J Anesth 2012;4:557-61.
4. Van Boogaard E, Hermens RP, Leschot NJ et al. Acta Obstet Gynecol Scand 2011;2:186-91.
5. Kirchner JE, Woodward EN, Smith JL et al. Prim Care Companion CNS Disord 2016;18(6).
6. Ruggeri M, Bonetto C, Lasalvia A et al. Schizophr Bull 2015;41:1192-203.
7. Ruggeri M, Lora A, Semisa D et al. Epidemiol Psychiatr Sci 2008;17:358-68.
8. Kendrick T, El-Gohary M, Stuart B et al. Cochrane Database Syst Rev 2016;7:CD011119.
9. Furnival J, Boaden R, Walshe C. Int J Qual Health Care 2017;29:604-11.
10. Cohen DJ, Crabtree BF, Etz RS et al. Am J Prev Med 2008;35(Suppl. 5):S381-9.

DOI:10.1002/wps.20488

Improving quality of mental health care in low-resource settings: lessons from PRIME

Kilbourne et al[1] present a useful framework for measuring and improving the quality of mental health care. They identify several barriers to this undertaking, highlight examples of innovations that can overcome these barriers in several countries, and offer recommendations for improving the quality of mental health care. The vast majority of the examples that are cited are from high-income country settings. It is worth reflecting on whether similar challenges exist in low- and middle-income countries (LMIC), and indeed what solutions may be found in these diverse low-resource settings.

Several key challenges can be identified for the improvement of the quality of mental health care in LMIC settings. First is the pervasive reality of limited resources, which has a major bearing not only on treatment coverage, but also quality. As just one structural indicator, there are 1.4 and 4.8 mental health workers per 100,000 population, respectively, in the African and South East Asian World Health Organization (WHO) regions, compared to 43.5 in the European region[2]. A second challenge is the lack of standardized service quality monitoring tools in LMIC, although instruments like the WHO Assessment Instrument for Mental health Systems (WHO-AIMS)[3] and the WHO QualityRights tool[4] are assisting with this. Third is the weak health system environment, including problems of suboptimal and at times dysfunctional general health management information systems. And fourth are the diverse cultural environments and pathways to care, which make assessment of processes and outcomes of care highly challenging.

In the Programme for Improving Mental health carE (PRIME), we have faced all of these challenges in various forms, while attempting to integrate mental health care into diverse low-resource primary care systems, and improve quality of care. PRIME is a research programme consortium working in Ethiopia, India, Nepal, South Africa and Uganda[5]. Our aim is to address the following "how" questions: how can we deliver evidence-based psychosocial and pharmacological interventions in a manner that is integrated into existing primary care systems and sensitive to local cultural needs; how can we ensure high-quality care while utilizing a task-sharing approach, employing general primary care workers to deliver mental health care; and how can we ensure continuous quality improvement in challenging low-resource settings?

Commencing in 2011, we worked closely with Ministry of Health partners to establish one district demonstration site in each country. Our early engagement with partners in each district site entailed the use of Theory of Change methods, to collaboratively map out hypothesized causal pathways from entry into the system to achieving the desired patient and population level outcomes[6]. In many respects the Theory of Change maps that we developed were underpinned by Donabedian's "structure, process, outcome" framework, so central to Kilbourne et al's paper.

In addition to facilitating local stakeholder partnerships, the Theory of Change approach enabled the PRIME country teams to identify a set of structure, process and outcome indicators. The indicators were then integrated into four main study designs to assess the implementation and impact of the PRIME mental health care plans in each district[7]. These studies included repeat community surveys to assess changes in population treatment coverage over a 3-year period; repeat facility detection surveys to assess improvements in the capacity of primary health care workers to identify depression and alcohol use disorders; cohort studies of individuals living with psychosis, epilepsy, depression and alcohol use disorders, to assess improvements in individual level clinical symptoms, functioning and economic circumstances (in some countries including nested randomized controlled trials); and finally case studies to assess structural measures such as medication supply, human resources, facilities and process measures such as numbers of patients treated and referred. The findings of each of these studies are currently being analyzed.

The process of implementing these mental health care plans has highlighted a number of context-specific quality of care challenges. In response to these, we have developed several local quality improvement initiatives, which are ongoing. These include: specific measures to improve detection of depression and reduce drop-out from care in people with psychosis in Ethiopia; improving screening for depression, pharmacological management and health management information systems indicators in India; improving individual patient follow-up in Nepal; facilitating the transition of primary care clinics to chronic disease management and patient-centred care in South Africa; and building the capacity of records staff, health workers and facility managers to collect and use health management information systems data for mental health care in Uganda. PRIME country teams have simultaneously enrolled in online quality improvement courses hosted by the Institute for Healthcare Improvement, to build their own capacity to develop quality improvement measures and interventions.

We have also developed or adapted several tools to assist in the improvement of the quality of care. One example is the Enhancing Assessment of Common Therapeutic Factors (ENACT) scale in Nepal, which is used by mental health specialist supervisors to routinely assess clinical competence of non-specialist health workers in the delivery of mental health care[8]. Another is the adaptation and use of the Institute for Healthcare Improvement's Plan Do Study Act (PDSA) cycle in South Africa.

Although the LMIC settings in which PRIME is working are very diverse, and differ substantially from the high-income

country settings referred to in Kilbourne et al's paper, there are several principles which the authors highlight that are equally relevant to these diverse settings.

First is the importance of using indicators that include structure, process and outcome variables to develop an integrated means of assessing quality of care. As mentioned earlier, the Donabedian framework of structure, process and outcome underpins the PRIME district mental health care plans and the Theory of Change approach that we used. This was essential to map out the steps in the pathway from engagement in the district sites to impact at the patient, system and population level. The challenge remains one of ensuring that routine health management information systems data can be used to monitor services, without the presence of a research infrastructure, and this is one of our key areas of work as PRIME winds up its programme in 2019.

A second principle is the importance of developing common metrics. Although PRIME countries differ substantially, we have been able to apply cross-country study designs and quality improvement measures which have common elements and approaches[9]. A third principle is the importance of health systems investing in routine quality improvement measures. These were not present in any of the sites when we began our work in 2011.

Our hope is that by including and piloting these measures in our diverse LMIC settings we will be able to demonstrate both feasibility and impact. We therefore join Kilbourne et al in making the case for further investment in integrated quality improvement measures for mental health care, particularly in low-resource settings.

Crick Lund
Department of Psychiatry and Mental Health, Alan J. Flisher Centre for Public Mental Health, University of Cape Town, South Africa; and Department of Population and Health Services Research, Centre for Global Mental Health, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

This paper is an output from the PRIME Research Programme Consortium, funded by the UK Department of International Development for the benefit of developing countries. The views expressed are not necessarily those of the UK government.

1. Kilbourne AM, Beck K, Spaeth-Rublee B et al. World Psychiatry 2018;17:30-8.
2. World Health Organization. Mental health atlas 2014. Geneva: World Health Organization, 2015.
3. World Health Organization. WHO Assessment Instrument for Mental Health Systems (WHO-AIMS) Version 2.1. Geneva: World Health Organization, 2005.
4. World Health Organization. WHO QualityRights tool kit: assessing and improving quality and human rights in mental health and social care facilities. Geneva: World Health Organization, 2012.
5. Lund C, Tomlinson M, De Silva M et al. PLoS Med 2012;9:e1001359.
6. Breuer E, De Silva MJ, Shidaye R et al. Br J Psychiatry 2016;208(Suppl. 56):s55-62.
7. De Silva MJ, Rathod SD, Hanlon C et al. Br J Psychiatry 2016;208(Suppl. 56):s63-70.
8. Kohrt BA, Ramaiya MK, Rai S et al. Global Mental Health 2015;2:e23.
9. Hanlon C, Fekadu A, Jordans MJD et al. Br J Psychiatry 2016;208(Suppl. 56):s47-54.

DOI:10.1002/wps.20489

What causes psychosis? An umbrella review of risk and protective factors

Joaquim Radua[1-3], Valentina Ramella-Cravaro[1,4], John P.A. Ioannidis[5-8], Abraham Reichenberg[9-12], Nacharin Phiphopthatsanee[1], Taha Amir[1], Hyi Yenn Thoo[1], Dominic Oliver[1], Cathy Davies[1], Craig Morgan[9,13], Philip McGuire[9,13], Robin M. Murray[9,13], Paolo Fusar-Poli[1,13,14]

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

Psychosis is a heterogeneous psychiatric condition for which a multitude of risk and protective factors have been suggested. This umbrella review aimed to classify the strength of evidence for the associations between each factor and psychotic disorders whilst controlling for several biases. The Web of Knowledge database was searched to identify systematic reviews and meta-analyses of observational studies which examined associations between socio-demographic, parental, perinatal, later factors or antecedents and psychotic disorders, and which included a comparison group of healthy controls, published from 1965 to January 31, 2017. The literature search and data extraction followed PRISMA and MOOSE guidelines. The association between each factor and ICD or DSM diagnoses of non-organic psychotic disorders was graded into convincing, highly suggestive, suggestive, weak, or non-significant according to a standardized classification based on: number of psychotic cases, random-effects p value, largest study 95% confidence interval, heterogeneity between studies, 95% prediction interval, small study effect, and excess significance bias. In order to assess evidence for temporality of association, we also conducted sensitivity analyses restricted to data from prospective studies. Fifty-five meta-analyses or systematic reviews were included in the umbrella review, corresponding to 683 individual studies and 170 putative risk or protective factors for psychotic disorders. Only the ultra-high-risk state for psychosis (odds ratio, OR=9.32, 95% CI: 4.91-17.72) and Black-Caribbean ethnicity in England (OR=4.87, 95% CI: 3.96-6.00) showed convincing evidence of association. Six factors were highly suggestive (ethnic minority in low ethnic density area, second generation immigrants, trait anhedonia, premorbid IQ, minor physical anomalies, and olfactory identification ability), and nine were suggestive (urbanicity, ethnic minority in high ethnic density area, first generation immigrants, North-African immigrants in Europe, winter/spring season of birth in Northern hemisphere, childhood social withdrawal, childhood trauma, Toxoplasma gondii IgG, and non-right handedness). When only prospective studies were considered, the evidence was convincing for ultra-high-risk state and suggestive for urbanicity only. In summary, this umbrella review found several factors to be associated with psychotic disorders with different levels of evidence. These risk or protective factors represent a starting point for further etiopathological research and for the improvement of the prediction of psychosis.

Key words: Schizophrenia, psychosis, risk, environment, socio-demographic factors, parental factors, perinatal factors, antecedents, ultra-high-risk state for psychosis, Black-Caribbean ethnicity, urbanicity

(World Psychiatry 2018;17:49–66)

Psychotic disorders like schizophrenia are among the world's leading causes of disability[1]. They have a mean incidence of 31.7 per 100,000 person-years in England[2] and a 12-month prevalence of 1.1% among the US population[3]. Despite many decades of research, the etiology of these disorders remains undetermined[4].

The model that has received most empirical support suggests that the etiology of psychotic disorders, schizophrenia for example, involves direct genetic and environmental risk factors along with their interaction[5,6]. In reality, some of the risk factors that have been associated with psychotic disorders – such as family history of mental illness – include both a genetic and an environmental component, and hence a distinction between genetic and environmental risk factors may be spurious.

With this in mind, in this study we adopted a pragmatic approach and used the term "non-purely genetic factors" to define socio-demographic, parental, perinatal, later factors and antecedents[7-9] that may increase (risk factors) or decrease (protective factors) the likelihood of developing psychotic disorders. The clinical importance of investigating these factors is threefold. First, they could potentially be used to advance the prediction of psychosis in populations at risk of developing the disorder[10,11]. Second, some, albeit not all, of these factors may be potentially modifiable by preventive interventions[4]. Third, they could inform outreach campaigns targeting the general public to raise awareness of risk factors for psychosis and to promote mental health.

Numerous studies investigating the association between potential risk or protective factors and psychotic disorders have been published. The body of literature in this area is substantial, presumably due to the severe societal burden that is associated with these disorders and thus the urgent need to understand their causes. However, to date, for all of those factors, there is no conclusive evidence with respect to both the association itself and its direction (i.e., risk or protective), because published findings have often been conflicting.

Furthermore, some of these results have been found to be affected by several types of biases[12,13]. These are particularly

relevant to this area of research because experimental support for etiology, in the sense of randomized allocation to the above-mentioned exposures[13], is naturally lacking, and most evidence is based on observational studies. Finally, previously published literature did not generate clear hierarchies of evidence across those factors, rendering the overall interpretation of the findings particularly complex. In fact, until recently there were no stringent evaluation criteria by which to hierarchically stratify the robustness of the evidence whilst at the same time controlling for the presence of biases.

Umbrella reviews can overcome these problems by assessing the level of the evidence provided by systematic reviews and meta-analyses[14] for each risk or protective factor, through strict criteria that probe a standard list of potential biases. These criteria have been extensively validated in various areas of medicine, such as neurology, oncology, nutrition medicine, internal medicine, psychiatry, paediatrics, dermatology and neurosurgery[15-33]. In the current study, we applied the umbrella review approach to the published evidence on risk or protective factors for psychotic disorders.

Our umbrella review advances knowledge in the field of psychosis etiology, providing the first state-of-the-art classification based on the robustness of associations between putative risk or protective factors and psychotic disorders, controlling at the same time for several biases. The use of classification criteria for levels of evidence can help overcome some of the ambiguity experienced by clinicians and researchers when confronted with conflicting meta-analyses[34] on complex topics and trying to base their decisions on them. Furthermore, our analysis will hopefully promote further etiological clinical research in psychosis, support the refinement of risk prediction in at-risk populations, and inform future preventive strategies.

METHODS

The protocol of the study was registered on PROSPERO 2016: CRD42016054101.

Search strategy and selection criteria

An umbrella review (i.e., a systematic collection and assessment of multiple systematic reviews and meta-analyses published on a specific research topic)[35] was conducted. We searched the Web of Knowledge database (incorporating Web of Science and MEDLINE) to identify systematic reviews or meta-analyses of observational studies that examined the association between a number of factors and psychotic disorders, published from 1965 to January 31, 2017. The search strategy used the keywords ("systematic review" OR "meta-analysis") and ("psychosis" OR "schizophrenia"). We then conducted a manual search of the reference lists of the retrieved articles.

Articles were initially screened on the basis of title and abstract reading. The full texts of potentially eligible articles were then independently scrutinized by two investigators (PFP, VRC), with no language restrictions. We selected systematic reviews or meta-analyses of individual observational studies (case-control, cohort, cross-sectional and ecological studies) that examined the association between socio-demographic, parental, perinatal, later factors or antecedents and any non-organic psychotic disorder as defined by any edition of the ICD or the DSM, including a comparison group of non-psychotic healthy controls, and reporting enough data to perform the analyses.

When data were incomplete, the corresponding author was contacted and invited to send additional information. When two articles presented overlapping datasets on the same factor, only the article with the largest dataset was retained for the main analysis. However, if the overlap was minimal, the articles were used conjointly, counting overlapping individual studies once only[36-43]. Moreover, we also excluded articles that did not report quantitative data, and articles with an outcome other than the onset of an established psychotic disorder, such as those related to relapse, remission or treatment response.

Articles that investigated pure genetic markers of psychotic disorders were excluded, because they have been examined extensively elsewhere[44,45]. Articles that investigated the association between biomarkers and psychotic disorders were not included, because this would have required specific methodological approaches and separate analyses. However, some putative biomarkers have been defined as antecedents (e.g., premorbid IQ[38,39], minor physical anomalies[46], non-right handedness[47], dermatoglyphic abnormalities[48] and neurological soft signs[49]) or perinatal factors (vitamin D[50]), and the relevant articles were therefore included.

The same inclusion/exclusion criteria were checked for each individual study comprised in every eligible meta-analysis or systematic review. The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) recommendations[51] and the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines[52] were followed.

Data extraction

Data extraction was performed independently by at least two investigators. Any existing discrepancies were resolved in consensus meetings with two of the authors (VRC, PFP). Factors were extracted as defined in the corresponding meta-analysis or systematic review. We did not combine similar factors if they were considered and analyzed separately by meta-analyses/systematic reviews[53]. Similarly, we did not split factors into subgroups if they were considered as a whole[54]. When a meta-analysis or systematic review reported both pooled results and results divided according to subgroups, pooled results were preferred, since they had a larger sample size.

Such a conservative approach was adopted to minimize the chance of introducing risk or protective factors that had not been defined by the corresponding articles, and that may have been too heterogeneous to allow meaningful interpretation. This approach also minimized the risk of artificially inflating

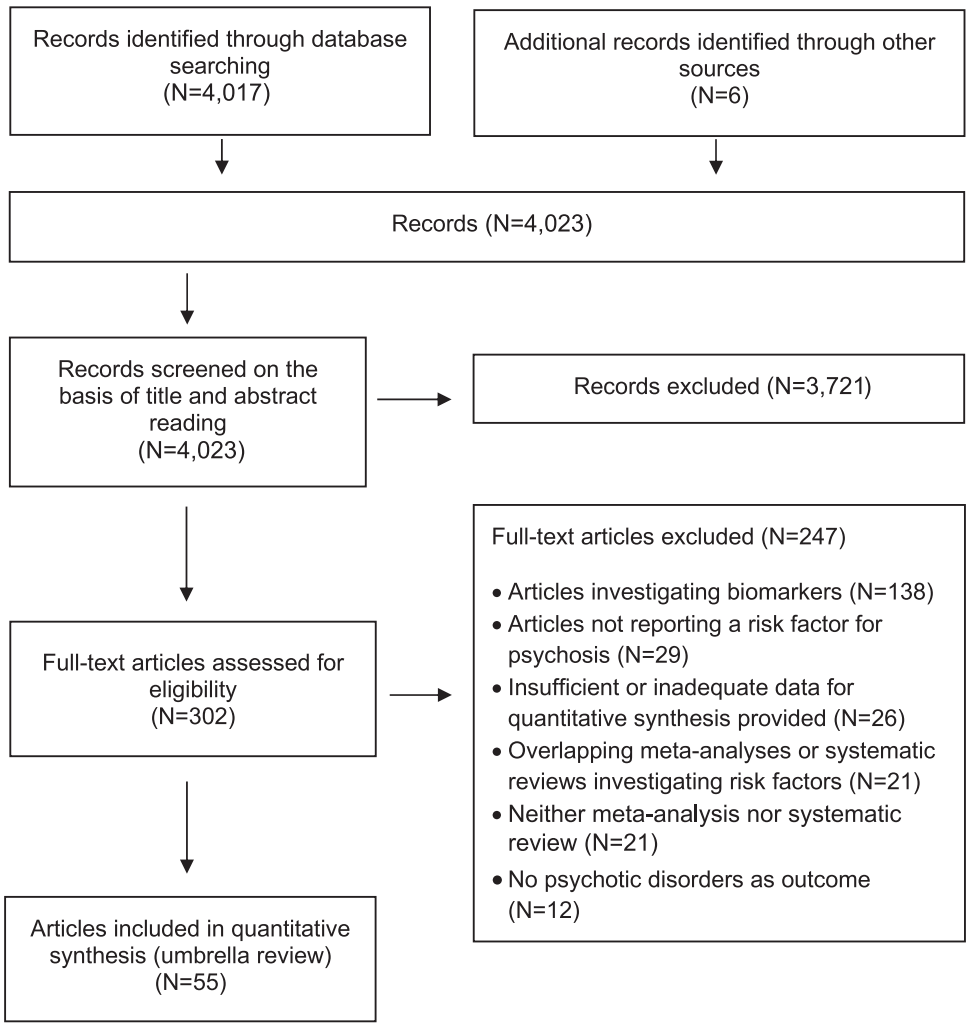

Figure 1 PRISMA flow chart

the sample size and therefore biasing the hierarchical classification of the evidence. The exception was for analyses not based on individual-level data, for which we specifically created new risk factors[55,56], as detailed in the statistical analysis section.

For descriptive purposes, risk/protective factors for psychotic disorders were clustered as previously suggested: socio-demographic and parental factors, perinatal factors, later factors (i.e., factors intervening in the post-perinatal period) and antecedents[7-9]. In line with previous definitions[7,8], antecedents were conceptualized as premorbid deviations in functioning and developmental milestones that could indicate an early expression of the disorder or active risk-modifying mechanisms and processes involved in psychosis onset. Risk factors, instead, would indicate a passive exposure to environmental agents that could play a role in the development of psychosis. One could argue that this distinction remains arbitrary, since the exact timing and mechanisms involved in the etiology of psychotic disorders remain to be elucidated, but this issue is beyond the aim of this review.

Several variables were recorded: the type of factor studied, the first author of the paper, the year of publication, the type of psychotic diagnosis, and the measure of association between the factor and psychotic disorders (preferably unadjusted), with the corresponding 95% confidence interval (CI) and the sample size (where available). If studies contained several types of control groups, data from healthy controls were used. When data were reported only in graphic form, they were digitally measured and extracted using WebPlotDigitizer[57]. The methodological quality of included studies was assessed using the validated AMSTAR (A Measurement Tool to Assess Systematic Reviews) instrument[58-60].

Statistical analysis

This umbrella review is composed of a number of meta-analyses of the included articles conducted separately with a series of scripts in R[61]. The effect size measures of the associa-

Table 1 Characteristics of meta-analyses and systematic reviews studying the association between psychotic disorders and socio-demographic and parental factors

Study	Factors examined	k	Diagnosis	AMSTAR index
Bosqui et al[71]	Ethnic minority in high ethnic density area; ethnic minority in low ethnic density area	5, 5	FEP, SZ, NAP, AP	9/11
Bourque et al[53]	First generation immigrants; second generation immigrants	12, 9	SZ, NAP, AP	10/11
Torrey et al[72]	Paternal age >35 years; paternal age >45 years; paternal age >55 years	8, 7, 4	SZ, NAP, AP	3/11
Kinney et al[55]	Disadvantaged group; latitude	2, 29	SZ	2/11
Kirkbride et al[2]	Age/gender; African ethnicity; Asian ethnicity; mixed ethnicity; other white ethnicity (all examined only in England)	9, 4, 4, 2, 3	FEP, SZ, NAP, AP	11/11
Kwok[73]	Low paternal socio-economic status	9	FEP, SZ, NAP	6/11
O'Donoghue et al[74]	Neighbourhood level social deprivation	3	FEP	8/11
Rasic et al[75]	Parental severe mental illness	9	FEP, SZ, NAP, AP	6/11
Saha et al[56]	Gross national income per capita	88	SZ	9/11
Tortelli et al[76]	Black-Caribbean ethnicity in England	7	FEP, SZ, NAP, AP	10/11
van der Ven et al[77]	North African immigrants in Europe	5	NAP	9/11
Vassos et al[78]	Urbanicity	8	SZ, NAP	6/11

k – number of studies for each factor, FEP – first episode of psychosis, SZ – schizophrenia, NAP – non-affective psychosis other than schizophrenia, AP – affective psychosis, AMSTAR – A Measurement Tool to Assess Systematic Reviews

tion between each factor and psychotic disorders were: incidence rate ratio (IRR), odds ratio (OR), risk ratio (RR), and standardized mean difference (Hedges' g) for continuous measures. Primarily, the effect size measure and its CI were used.

Since authors usually round off the measures, the first step was to "unround" them by estimating a more exact measure and CI, in which the (logarithm of the) lower and upper bounds were symmetrical around the (logarithm of the) measure. Subsequently, the variance was calculated from the standard formula for the CI. If two or more studies shared the non-exposed sample, the size of this sample was divided equally between these studies. This approach minimized the dependence produced by the sharing of the non-exposed sample, whilst allowing estimation of heterogeneity across the exposed samples[62].

Some factors required special adjustments, such as: a) the transformation of measures other than OR into OR in factors where effect size was reported using different types of measures (to have a single outcome measure), and b) the combination of effect sizes for left and right nostrils in olfaction studies[63], conservatively assuming a weak to moderate correlation (r=0.3)[64]. Ultimately, we used the "metainc" (IRR), "metabin" (RR and OR), or "metacont" (Hedges' g) functions in the R "meta" package[65] to calculate the meta-analytic effect size and its p value, the CI, and the heterogeneity (summarized with the I^2 statistic and the p value associated with the Q value). Resulting statistics were also used to calculate the prediction interval[66].

A few specific adjustments were also adopted for age and gender, where IRRs were available for schizophrenia and affective psychosis separately, and stratified by 5 or 10-year age ranges and gender[2,67]. We combined schizophrenia and affective psychoses and then meta-analyzed the IRR of each 10-year age range (vs. other ages), and the IRR of males (vs. females, globally and within each 10-year age range). Since age and gender were considered as basic factors and excluded by previous reviews on psychosis[8,9] and by umbrella reviews on other neuropsychiatric conditions[23,25,27], these analyses were considered exploratory.

Alternative analyses were also conducted for latitude[55] and gross national income per capita (GNI)[56], when the prevalence rates were provided in a series of locations. Specifically, the incidence in each location was (logistic) regressed by the latitude or GNI, obtaining the OR of 10° increase in latitude or 10,000 USD increase in GNI. These results were also considered exploratory because they are based on ecological analyses rather than individual-level data, and were traditionally excluded from previous umbrella reviews of risk factors[23,25,27].

Complementary analyses included: a) an Egger test to assess small-study effects that lead to potential reporting or publication bias[68]; b) a test of excess significance bias[69] as described below, and c) an OR equivalent. The test of excess significance bias consisted of a binomial test to compare the observed vs. the expected number of studies yielding statistically significant results. This expected number was calculated as the sum of the statistical power of the studies, which was estimated using the standard t-test formulas for Hedges' g, and random simulations for OR, RR and IRR. Specifically, the statistical power of study A was estimated as the proportion of times in which a

Table 2 Characteristics of meta-analyses and systematic reviews studying the association between psychotic disorders and perinatal factors

Study	Factors examined	k	Diagnosis	AMSTAR index
Cai et al[40]	Gestational influenza	6	SZ, NAP, AP	10/11
Cannon et al[37]	Anaemia in pregnancy; antepartum haemorrhage; asphyxia; baby detained in hospital; birth weight <2000 g; birth weight <2500 g; birth weight <2500 g + prematurity; breech delivery; caesarean section; cephalopelvic disproportion; congenital malformations; diabetes in pregnancy; emergency caesarean section; forceps/vacuum delivery; gestational age <37 weeks; gestational age >42 weeks; induction of labour; low Apgar score; non-vertex presentation; placental abruption; preeclampsia; Rhesus incompatibility; small birth length; being small for gestational age; small head circumference; smoking in pregnancy; threatened premature delivery; urinary infection in pregnancy; uterine atony	2, 4, 3, 2, 2, 4, 3, 3, 5, 2, 3, 2, 3, 6, 4, 3, 2, 2, 5, 2, 5, 2, 3, 5, 2, 1, 2, 2, 3	SZ	6/11
Christesen et al[50]	Neonatal vitamin D (<19.7 vs. 40.5-50.9 nmol/L); neonatal vitamin D (19.7-30.9 vs. 40.5-50.9 nmol/L); neonatal vitamin D (30.9-40.4 vs. 40.5-50.9 nmol/L); neonatal vitamin D (>50.9 vs. 40.5-50.9 nmol/L)	1, 1, 1, 1	SZ	6/11
Davies et al[79]	Winter/spring season of birth in Northern hemisphere	7	SZ	6/11
Geddes & Lawrie[81]	Obstetric complications	10	SZ	6/11
Geddes et al[36]	Antepartum haemorrhage; birth weight <2500 g; caesarean section; congenital malformations; cord complications; forceps delivery; gestational age <37 weeks; incubator or resuscitation; labour >24 hours; non-vertex presentation; preeclampsia; Rhesus incompatibility; rubella or syphilis; twin birth	9, 9, 9, 7, 9, 9, 3, 8, 4, 9, 9, 7, 9, 9	SZ	4/11
McGrath & Welham[80]	Winter/spring season of birth in Southern hemisphere	7	SZ	9/11
Selten et al[106]	Maternal stress during pregnancy	4	SZ	5/11
Selten & Termoshuizen[41]	Gestational influenza	7	SZ, AP	7/11
Van Lieshout et al[82]	Pre-pregnancy and pregnancy maternal obesity	4	SZ, NAP	10/11

k – number of studies for each factor, FEP – first episode of psychosis, SZ – schizophrenia, NAP – non-affective psychosis other than schizophrenia, AP – affective psychosis, AMSTAR – A Measurement Tool to Assess Systematic Reviews

simulated study using binomial or Poisson random cases was considered "statistically significant"; the simulated studies had the same mean incidence and person-times or sample sizes as study A (using the full sample sizes in case of sharing a sample), and the same effect size as the largest study in the meta-analysis.

Small-study effects and excess significance bias were claimed at one-sided p values <0.05, as in previous studies[27]. In order to easily compare meta-analyses using different outcome measures, OR equivalents were provided for the above measures. Given the low incidence of psychotic disorders, RR was assumed to be equivalent to OR, after checking that the difference between an OR and a RR of the same data was negligible. IRR was assumed to be equivalent to RR, and Hedges' g was converted to OR using a standard formula[70].

IRR, OR and RR greater than 1 or Hedges' g greater than 0 indicated that the factor was associated with an increased likelihood of psychotic disorders. IRR, OR and RR lower than 1 or Hedges' g lower than 0 indicated that the factor was associated with a reduced likelihood of psychotic disorders, i.e. it was protective.

The levels of evidence of the associations between putative risk (or protective) factors and psychotic disorders were classified in accordance with previous umbrella reviews[23,25,27]: convincing (class I) when number of cases >1000, $p < 10^{-6}$, $I^2 < 50\%$, 95% prediction interval excluding the null, no small-study effects, and no excess significance bias; highly suggestive (class II) when number of cases >1000, $p < 10^{-6}$, largest study with a statistically significant effect, and class I criteria not met; suggestive (class III) when number of cases >1000, $p<10^{-3}$, and class I-II criteria not met; weak (class IV) when $p < 0.05$ and class I-III criteria not met; non-significant when $p > 0.05$.

Finally, a sensitivity analysis was conducted for the factors classified as class I-III by using only prospective studies (as defined in each meta-analysis/systematic review or, when this was not provided, as defined by each individual study). Prospective studies allow one to address the temporality of the

Table 3 Characteristics of meta-analyses and systematic reviews studying the association between psychotic disorders and later factors

Study	Factors examined	k	Diagnosis	AMSTAR index
Arias et al[83]	BK virus; Borna disease virus; Chlamydia psittaci; Chlamydia trachomatis; cytomegalovirus; Epstein-Barr virus; human endogenous retrovirus; human endogenous retrovirus type k115; human endogenous retrovirus type W; human herpes virus 2; influenza; JC virus; Toxocara spp; varicella zoster virus	1, 8, 2, 2, 8, 3, 4, 1, 4, 5, 2, 1, 1, 4	SZ	8/11
Attademo et al[84]	Benzene; carbon monoxide; nitrogen dioxide; nitrogen oxides; tetrachloroethylene; traffic	1, 1, 1, 1, 1, 1	SZ	2/11
Beards et al[85]	Adult life events	6	FEP, SZ, NAP, AP	8/11
Clancy et al[86]	Epilepsy	1	SZ	6/11
Cunningham et al[87]	Bullying	1	SZ, NAP	7/11
De Sousa et al[88]	Parental communication deviance	4	SZ	6/11
Gurillo et al[89]	Tobacco use	5	FEP, SZ, NAP, AP	11/11
Gutierrez-Fernandez et al[90]	Chlamydia pneumoniae; human herpes virus 1; human herpes virus 6	3, 11, 3	SZ, NAP	8/11
Khandaker et al[91]	Central nervous system infection during childhood	2	SZ, NAP	10/11
Linszen et al[92]	Hearing impairment	5	SZ	8/11
Marconi et al[93]	Heavy cannabis use	2	FEP, SZ, NAP	7/11
Molloy et al[94]	Traumatic brain injury	8	SZ	7/11
Sutterland et al[95]	Toxoplasma gondii IgG; Toxoplasma gondii IgM	40, 15	FEP, SZ	9/11
Varese et al[54]	Childhood trauma	20	FEP, SZ, NAP, AP	10/11

k – number of studies for each factor, FEP – first episode of psychosis, SZ – schizophrenia, NAP – non-affective psychosis other than schizophrenia, AP – affective psychosis, AMSTAR – A Measurement Tool to Assess Systematic Reviews, Ig – immunoglobulin

association, thus dealing with the problem of reverse causation that may affect, for example, case-control studies[23,25,27].

RESULTS

Database

Overall, 4,023 records were searched, 302 were screened and 55 articles were eligible[2,36-43,46-50,53-56,63,71-106] (see Figure 1). The eligible articles were published between 1995 (when meta-analyses in this field first became available)[107] and 2017. All of the studies utilized a healthy control group except one, investigating the ultra-high-risk state[98]. This latter study used as controls help-seeking individuals undergoing an ultra-high-risk assessment but not meeting the relevant criteria. The mental health status of this control group was not well defined.

Overall, the 55 eligible meta-analyses or systematic reviews, including 683 individual studies, reported on 170 putative risk/protective factors of psychotic disorders (Tables 1-4). For paternal socio-economic status[73], neighbourhood-level social deprivation[74], pre-pregnancy and pregnancy maternal obesity[82], neonatal levels of vitamin D[50], polluting agents (benzene, carbon monoxide, nitrogen dioxide, nitrogen oxides, tetrachloroethylene, traffic)[84] and the ultra-high-risk state[98], the studies did not provide a quantitative synthesis of individual findings, but reported adequate data to allow meta-analyses.

Summary of associations

The number of cases was greater than 1,000 for 48 factors (28.2%). One hundred three of the 170 analyzed factors (60.6%) presented a statistically significant effect ($p<0.05$) under the random-effects model, but only 39 (22.9%) reached $p<10^{-6}$. Fifty-three factors (31.2%) presented a large heterogeneity ($I^2>50\%$), while for 28 (16.5%) the 95% prediction interval did not include the null. Additionally, the evidence for small-study effects and excess significance bias was noted for 16 (9.4%) and 17 (10.0%) factors, respectively.

Classification of level of evidence of associations between socio-demographic and parental, perinatal, later factors or antecedents and psychotic disorders

Among the 170 factors, one socio-demographic factor (Black-Caribbean ethnicity in England: OR=4.87, 95% CI: 3.96-6.00) and

Table 4 Characteristics of meta-analyses and systematic reviews studying the association between psychotic disorders and antecedents

Study	Factors examined	k	Diagnosis	AMSTAR index
Dickson et al[96]	Motor function pre-onset of psychosis; poor academic achievement pre-onset of psychosis; poor mathematic academic achievement pre-onset of psychosis	4, 4, 3	FEP, SZ, NAP	7/11
Filatova et al[97]	Delay in grabbing object; delay in holding head up; delay in sitting unsupported; delay in standing unsupported; delay in walking unsupported	3, 3, 4, 4, 5	SZ, NAP	9/11
Fusar-Poli et al[98]	Ultra-high-risk state for psychosis	9	FEP	9/11
Golembo-Smith et al[48]	ATD angle; fingertip pattern asymmetry; fluctuating asymmetry A-B ridge count; fluctuating asymmetry finger ridge count; total A-B ridge count; total finger ridge count	5, 4, 3, 3, 13, 13	SZ	6/11
Hirnstein & Hugdahl[47]	Non-right handedness	41	SZ, NAP	5/11
Khandaker et al[39]	Premorbid IQ	5	SZ, NAP	8/11
Kaymaz et al[99]	Psychotic-like experiences	4	FEP, SZ, NAP, AP	10/11
Koning et al[100]	Dyskinesia in antipsychotic-naïve schizophrenic patients; parkinsonism in antipsychotic-naïve schizophrenic patients	5, 3	FEP, SZ	5/11
Matheson et al[43]	Childhood social withdrawal	5	SZ, NAP	8/11
Moberg et al[63]	Olfactory detection ability; olfactory identification ability; olfactory discrimination ability; olfactory memory ability; olfactory hedonics ability (pleasant odours); olfactory hedonics ability (unpleasant odours); olfactory hedonics ability (unspecified odours)	18, 51, 8, 2, 9, 8, 7	SZ, NAP	9/11
Neelam et al[49]	Neurological soft signs	7	SZ, NAP	8/11
Ohi et al[101]	Cooperativeness; harm avoidance; novelty seeking; persistence; reward dependence; self-directedness; self-transcendence	7, 7, 7, 7, 7, 7, 7	SZ	4/11
Ohi et al[102]	Agreeableness; conscientiousness; extraversion; neuroticism; openness	6, 7, 8, 8, 7	SZ	6/11
Potvin & Marchand[103]	Hypoalgesia	9	SZ	5/11
Tarbox & Pogue-Geile[42]	Childhood antisocial and externalizing behaviour; childhood social withdrawal and internalizing behaviour	2, 6	SZ, NAP	3/11
Ward et al[104]	Extracranial size	7	SZ, NAP	3/11
Woodberry et al[38]	Premorbid IQ	11	SZ	7/11
Xu et al[46]	Minor physical anomalies	14	SZ	4/11
Yan et al[105]	Trait anhedonia	44	SZ, NAP	6/11

k – number of studies for each factor, FEP – first episode of psychosis, SZ – schizophrenia, NAP – non-affective psychosis other than schizophrenia, AP – affective psychosis, AMSTAR – A Measurement Tool to Assess Systematic Reviews, ATD angle – dermatoglyphic feature that compares the length of the hand to the width

one antecedent (ultra-high-risk state: OR=9.32, 95% CI: 4.91-17.72) presented a convincing level of association (class I: >1000 cases, $p < 10^{-6}$, no evidence of small-study effects or excess significance bias, 95% prediction interval not including the null, and no large heterogeneity).

For six factors there was highly suggestive evidence for association (class II: >1000 cases, $p<10^{-6}$, largest study with a statistically significant effect, and class I criteria not met). These were two socio-demographic and parental factors (ethnic minority in low ethnic density area: OR=3.71; and second generation immigrants: OR=1.68); none of the perinatal and later factors; and four antecedents (minor physical anomalies: OR=5.30; trait anhedonia: OR=4.41; olfactory identification ability: OR=0.19; and premorbid IQ: OR=0.47).

There was suggestive evidence for association (class III) for nine further factors: four socio-demographic and parental factors (North-African immigrants in Europe: OR=2.22; urbanicity: OR=2.19; ethnic minority in high ethnic density area: OR=2.11; and first generation immigrants: OR=2.10); one perinatal factor (winter/spring season of birth in Northern hemisphere: OR=1.04); two later factors (childhood trauma: OR=2.87; and Toxoplasma gondii IgG: OR=1.82); and two antecedents (childhood social withdrawal: OR=2.91; and non-right handedness: OR=1.58). There was either weak (class IV)

Table 5 Level of evidence for the association of socio-demographic and parental factors and psychotic disorders

			Features used for classification of level of evidence							
Factor	k	Random-effects measure, ES (95% CI)	N	p random-effects	I^2 (p)	PI (95% CI)	SSE/ESB	LS	eOR	CE
Black-Caribbean ethnicity in England[76]	9	IRR, 4.87 (3.96-6.00)	3,446	2.8×10^{-50}	38% (0.12)	2.95-8.03	No/No	Yes	4.87	I
Ethnic minority in low ethnic density area[71]	5	IRR, 3.71 (2.47-5.58)	1,328	3.1×10^{-10}	70% (0.09)	0.95-14.43	Yes/No	Yes	3.71	II
Second generation immigrants[53]	26	IRR, 1.68 (1.42-1.92)	28,753	7.6×10^{-10}	77% (<0.001)	0.92-3.06	No/No	Yes	1.68	II
North African immigrants in Europe[77]	12	IRR, 2.22 (1.58-3.12)	2,577	4.2×10^{-6}	65% (0.001)	0.77-6.41	No/NA	Yes	2.22	III
Urbanicity[78]	8	OR, 2.19 (1.55-3.09)	45,791	8.9×10^{-5}	99% (<0.001)	0.62-7.77	No/No	Yes	2.19	III
Ethnic minority in high ethnic density area[71]	5	IRR, 2.11 (1.39-3.20)	1,328	4.3×10^{-4}	58% (0.04)	0.57-7.81	No/No	Yes	2.11	III
First generation immigrants[53]	42	IRR, 2.10 (1.72-2.56)	25,063	1.9×10^{-13}	89% (<0.001)	0.75-5.89	No/Yes	No	2.10	III
Parental severe mental illness[75]	9	RR, 5.94 (2.99-11.79)	90	3.5×10^{-7}	0% (0.85)	2.60-13.59	No/No	Yes	5.94	IV
Black African ethnicity in England[2]	4	IRR, 4.72 (3.30-6.77)	452	2.3×10^{-17}	49% (0.12)	1.25-17.82	No/NA	Yes	4.72	IV
Asian ethnicity in England[2]	6	IRR, 2.83 (1.59-5.05)	613	4.2×10^{-4}	55% (0.05)	0.53-15.00	No/Yes	No	2.83	IV
Other white ethnicity in England[2]	3	IRR, 2.62 (1.35-5.10)	274	0.004	87% (<0.001)	0.93-21.88	No/NA	Yes	2.62	IV
Paternal age >45 years[72]	4	OR, 2.36 (1.35-4.11)	392	0.003	0% (0.66)	0.69-8.01	No/Yes	No	2.36	IV
Disadvantaged vs. advantaged groups[55]	3	RR, 2.27 (1.21-4.27)	532	0.010	69% (0.04)	0-2016.72	No/No	Yes	2.27	IV
Mixed ethnicity in England[2]	3	IRR, 2.19 (1.08-4.44)	330	0.030	0% (0.41)	0.02-14.53	No/NA	No	2.19	IV
Low paternal socio-economic status[73]	9	OR, 1.30 (1.02-1.65)	15,922	0.032	94% (<0.001)	0.58-2.90	No/No	Yes	1.30	IV
Paternal age >35 years[72]	9	OR, 1.22 (1.06-1.41)	2,181	0.007	30% (0.18)	0.89-1.67	No/Yes	No	1.22	IV
Neighbourhood level social deprivation[74]	3	OR, 1.64 (0.83-3.23)	5,560	0.156	88% (<0.001)	0-5961.52	No/No	No	1.64	ns
Paternal age >55 years[72]	7	OR, 1.21 (0.82-1.78)	57	0.341	47% (0.07)	0.45-3.22	No/No	No	1.21	ns

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small-study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, IRR – incidence rate ratio, OR – odds ratio, RR – relative risk, NA – not assessable, ns – not significant

or no evidence of association with psychotic disorders for all other factors (see Tables 5-8).

Exploratory analyses

Results of the exploratory analyses on the association between age and gender strata (total of 23 strata) showed a main effect for male gender (males vs. females: IRR=1.34, 95% CI: 1.05-1.71, class IV). There was also a main effect for 15-35 year-old age (25-34 year-old vs. other: IRR=1.45, 95% CI: 1.29-1.63, class II; 20-29 year-old vs. other: IRR=2.43, 95% CI: 1.58-3.74, class IV; 15-24 year-old vs. other: IRR=1.46, 95% CI: 1.14-1.87, class IV). Age older than 35 was found to be a protective factor (60-69 year-old vs. other: IRR=0.26, 95% CI: 0.14-0.51, class IV; 55-64 year-old vs. other: IRR=0.30, 95% CI: 0.17-0.51, class IV; 50-59 year-old vs. other: IRR=0.50, 95% CI: 0.27-0.93, class IV; 40-49 year-old vs. other: IRR=0.54, 95% CI: 0.35-0.83, class IV; 35-44 year-old vs. other: IRR=0.80, 95% CI: 0.70-0.93, class IV).

There was also weak (class IV) association between psychotic disorders and male gender for 15-40 year-old age (male vs. female within 20-29 year-old: IRR=2.19, 95% CI: 1.69-2.84; male vs. female within 15-24 year-old: IRR=1.98, 95% CI: 1.62-2.41; male vs. female within 30-39 year-old: IRR=1.72, 95% CI: 1.22-2.41; male vs. female within 25-34 year-old: IRR=1.60, 95% CI: 1.26-2.03). The other ten strata were all not associated with psychotic disorders.

Additional exploratory analyses on latitude (per 10°)[55] and GNI per capita (per 10,000 USD)[56] found significant associations, with ORs of 1.22 and 0.80, respectively. Although these factors include >1000 patients and have a p<0.001, it was not possible to apply the classification of the evidence.

Classification of level of evidence of associations between socio-demographic and parental, perinatal, later factors or antecedents and psychotic disorders after sensitivity analysis

A sensitivity analysis was not possible for four of the associations categorized as class I-III in the overall analysis (winter/spring season of birth in Northern hemisphere, olfactory iden-

Table 6 Level of evidence for the association of perinatal factors and psychotic disorders

			Features used for classification of level of evidence							
Factor	**k**	**Random-effects measure, ES (95% CI)**	**N**	**p random-effects**	**I^2 (p)**	**PI (95% CI)**	**SSE/ESB**	**LS**	**eOR**	**CE**
Winter/spring season of birth in Northern hemisphere[79]	27	OR, 1.04 (1.02-1.06)	115,010	2.1×10^{-6}	0% (0.99)	1.02-1.06	No/No	Yes	1.04	III
Diabetes in pregnancy[37]	2	OR, 10.12 (1.84-55.72)	243	0.008	0% (0.69)	NA	NA/NA	No	10.12	IV
Emergency caesarean section[37]	3	OR, 3.36 (1.48-7.63)	825	0.004	0% (0.92)	0.02-685.69	No/No	No	3.36	IV
Birth weight <2000 g[37]	2	OR, 2.46 (1.11-5.46)	507	0.027	0% (0.85)	NA	NA/NA	Yes	2.46	IV
Congenital malformations[36,37]	10	OR, 2.31 (1.29-4.13)	1,080	0.005	0% (0.99)	1.16-4.57	No/Yes	Yes	2.31	IV
Incubator or resuscitation[36]	8	OR, 2.12 (1.29-3.47)	438	0.003	0% (0.85)	1.14-3.92	No/No	No	2.12	IV
Neonatal vitamin D (<19.7 vs. 40.5-50.9 nmol/L)[50]	1	RR, 2.11 (1.28-3.49)	424	0.004	NA (1.00)	NA	NA/NA	Yes	2.11	IV
Neonatal vitamin D (30.9-40.4 vs. 40.5-50.9 nmol/L)[50]	1	RR, 2.10 (1.30-3.40)	424	0.003	NA (1.00)	NA	NA/NA	Yes	2.10	IV
Threatened premature delivery[37]	2	OR, 2.05 (1.02-4.10)	314	0.043	0% (0.56)	NA	NA/NA	Yes	2.05	IV
Neonatal vitamin D (19.7-30.9 vs. 40.5-50.9 nmol/L)[50]	1	RR, 2.02 (1.27-3.19)	424	0.003	NA (1.00)	NA	NA/NA	Yes	2.02	IV
Pre-pregnancy and pregnancy maternal obesity[82]	4	OR, 1.99 (1.26-3.14)	305	0.003	27% (0.24)	0.47-8.50	No/No	No	1.99	IV
Uterine atony[37]	3	OR, 1.93 (1.35-2.76)	836	3.3×10^{-4}	0% (0.37)	0.19-19.78	No/No	Yes	1.93	IV
Obstetric complications[81]	10	OR, 1.84 (1.25-2.70)	373	0.002	25% (0.21)	0.80-4.22	Yes/Yes	No	1.84	IV
Neonatal vitamin D (>50.9 vs. 40.5-50.9 nmol/L)[50]	1	RR, 1.71 (1.04-2.80)	424	0.033	NA (1.00)	NA	NA/NA	Yes	1.71	IV
Antepartum haemorrhage[36,37]	14	OR, 1.63 (1.12-2.38)	1,489	0.011	6% (0.38)	0.92-2.89	No/No	No	1.63	IV
Birth weight <2500 g[36,37]	13	OR, 1.57 (1.20-2.07)	1,815	0.001	0% (0.45)	1.16-2.14	No/Yes	Yes	1.57	IV
Small head circumference[37]	2	OR, 1.41 (1.00-1.97)	762	0.048	0% (0.58)	NA	NA/NA	No	1.41	IV
Placental abruption[37]	2	OR, 4.54 (0.32-64.63)	314	0.264	72% (0.05)	NA	NA/NA	No	4.54	ns
Rhesus incompatibility[36,37]	9	OR, 1.96 (0.88-4.33)	1,097	0.098	0% (0.98)	0.75-5.11	No/NA	No	1.96	ns
Asphyxia[37]	3	OR, 1.95 (0.77-4.97)	1,122	0.160	76% (0.01)	0-108727	No/No	Yes	1.95	ns
Forceps delivery[36]	9	OR, 1.67 (0.90-3.08)	554	0.103	42% (0.08)	0.34-8.15	Yes/No	Yes	1.67	ns
Rubella or syphilis[36]	9	OR, 1.64 (0.47-5.71)	567	0.435	0% (0.099)	0.37-7.39	No/No	No	1.64	ns
Twin birth[36]	9	OR, 1.53 (0.79-2.97)	558	0.208	0% (0.45)	0.69-3.40	Yes/No	No	1.53	ns
Gestational age <37 weeks[36,37]	7	OR, 1.35 (0.99-1.84)	1,502	0.057	0% (0.66)	0.90-2.03	Yes/No	No	1.35	ns
Being small for gestational age[37]	5	OR, 1.34 (0.82-2.19)	1,436	0.240	58% (0.04)	0.28-6.41	No/No	Yes	1.34	ns
Smoking in pregnancy[37]	1	OR, 1.29 (0.72-2.31)	76	0.393	NA (1.00)	NA	NA/NA	No	1.29	ns
Birth weight <2500 g and prematurity[37]	4	OR, 1.25 (0.52-3.00)	959	0.610	65% (0.03)	0.03-46.31	No/No	Yes	1.25	ns
Anaemia in pregnancy[37]	3	OR, 1.22 (0.46-3.28)	528	0.688	56% (0.10)	0- 41770	No/No	No	1.22	ns
Maternal stress during pregnancy[106]	5	RR, 1.16 (0.94-1.43)	4,412	0.166	71% (0.01)	0.60-2.25	No/No	No	1.16	ns
Low Apgar score[37]	2	OR, 1.13 (0.69-1.84)	405	0.622	0% (0.67)	NA	NA/NA	No	1.13	ns
Preeclampsia[36,37]	15	OR, 1.07 (0.78-1.46)	2,277	0.690	22% (0.20)	0.53-2.15	No/No	Yes	1.07	ns
Forceps/vacuum delivery[37]	7	OR, 1.07 (0.81-1.42)	1,888	0.643	34% (0.16)	0.55-2.09	No/No	Yes	1.07	ns
Cord complications[36]	9	OR, 1.06 (0.47-2.39)	549	0.894	0% (0.54)	0.40-2.83	No/No	No	1.06	ns
Small birth length[37]	3	OR, 1.05 (0.86-1.30)	929	0.619	0% (0.91)	0.28-4.03	No/No	No	1.05	ns
Baby detained in hospital[37]	3	OR, 1.04 (0.59-1.86)	976	0.883	76% (0.01)	0-903.90	No/No	Yes	1.04	ns
Winter/spring season of birth in Southern hemisphere[80]	7	OR, 1.03 (0.88-1.19)	15,023	0.738	16% (0.30)	0.77-1.37	No/NA	No	1.03	ns
Influenza during pregnancy[40,41]	14	OR, 0.99 (0.91-1.08)	7,620	0.867	46% (0.03)	0.79-1.24	No/No	No	0.99	ns

Table 6 Level of evidence for the association of perinatal factors and psychotic disorders *(continued)*

			Features used for classification of level of evidence							
Factor	**k**	**Random-effects measure, ES (95% CI)**	**N**	**p random-effects**	**I^2 (p)**	**PI (95% CI)**	**SSE/ESB**	**LS**	**eOR**	**CE**
Non-vertex presentation[36,37]	15	OR, 0.99 (0.75-1.31)	2,272	0.953	6% (0.38)	0.65-1.51	No/No	No	0.99	ns
Gestational age >42 weeks[37]	3	OR, 0.97 (0.48-1.95)	1,193	0.933	42% (0.18)	0-1000	No/No	No	0.97	ns
Caesarean section[36,37]	15	OR, 0.95 (0.71-1.28)	1,920	0.734	0% (0.46)	0.68-1.32	No/No	No	0.95	ns
Breech delivery[37]	3	OR, 0.95 (0.49-1.84)	470	0.879	0% (0.78)	0.01-68.26	No/No	No	0.95	ns
Urinary infection in pregnancy[37]	3	OR, 0.90 (0.44-1.84)	690	0.776	29% (0.24)	0-498.73	No/No	No	0.90	ns
Induction of labor[37]	3	OR, 0.82 (0.53-1.28)	528	0.387	24% (0.26)	0.02-35.30	No/No	No	0.82	ns
Cephalopelvic disproportion[37]	2	OR, 0.60 (0.18-1.99)	243	0.407	0% (0.48)	NA	NA/NA	No	0.60	ns
Labour >24 hours[36]	4	OR, 0.84 (0.39-1.78)	266	0.643	20% (0.28)	0.09-8.11	No/No	No	0.84	ns

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small-study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, IRR – incidence rate ratio, OR – odds ratio, RR – relative risk, NA – not assessable, ns – not significant

tification ability, trait anhedonia and minor physical anomalies), because they did not include any prospective studies.

Within class I factors, only ultra-high-risk state maintained the same level of evidence, whereas Black-Caribbean ethnicity in England downgraded to a weak (class IV) level of evidence. Equally, all other available class II and III factors were downgraded either to a weak (ethnic minority in low ethnic density area, North-African immigrants in Europe, childhood trauma, ethnic minority in high ethnic density area, childhood social withdrawal, first and second generation immigrants, Toxoplasma gondii IgG, and premorbid IQ) or a non-significant (non-right handedness) level of evidence, except urbanicity, that remained a class III risk factor (Table 9).

DISCUSSION

To our knowledge, this is the first umbrella review of risk and protective factors for psychotic disorders that includes a robust hierarchical classification of the published evidence. Overall, 55 meta-analyses or systematic reviews, with a total of 683 individual studies and 170 socio-demographic and parental, perinatal, later factors or antecedents of psychotic disorders, were included. There was convincing evidence (class I) for only two factors, which were the ultra-high-risk state for psychosis and Black-Caribbean ethnicity in England. However, six other factors were characterized by highly suggestive evidence (class II), and another nine by suggestive evidence (class III). Sensitivity analyses that limited data to prospective studies indicated that ultra-high-risk state and urbanicity showed the largest evidence of association (class I and class III, respectively) with psychotic disorders.

Overall, our umbrella review indicates that, although a large number of risk factors for psychotic disorders have been evaluated in multiple studies, reviews and meta-analyses, the number of those that have suggestive or stronger support is far more limited. This is consistent with previous findings about the etiology of other neuropsychiatric conditions where umbrella reviews have been performed, such as dementia, Parkinson's disease, amyotrophic lateral sclerosis, multiple sclerosis and bipolar disorder[16,23,25-27].

Although the past two decades have clearly shown that the ultra-high-risk state is substantially associated with an increased risk of psychosis[11,108,109], this result should be interpreted with caution. Firstly, this state is the closest antecedent of psychosis by definition, with onset of the disorder occurring from within a few months of ultra-high-risk diagnosis[110]. Indeed, some ultra-high-risk individuals already present with severe symptoms, including short-lived psychotic episodes[111,112], affective symptoms[113] and impaired functioning[114]. Secondly, the ultra-high-risk state is intrinsically heterogeneous[10,115], including different subgroups[115] and varying diagnostic operationalizations[116]. Furthermore, from an epidemiological perspective, it is a spurious condition, characterized by the accumulation of a number of risk factors[117] which enrich the risk in an uncontrolled manner[118-122].

Ethnic minority status and urbanicity may better represent true risk factors, contributing to the development of psychotic disorders through increased socio-environmental adversities[123]. In fact, the effect of both factors on the risk of developing psychotic disorders may be explained (mediated) by environmental exposures at an individual level, such as substance use, social isolation, social defeat, social fragmentation, and discrimination[124]. Interestingly, many of these exposures appear to share a common factor of social stress and defeat[125,126], and have been – mostly indirectly – associated with various neurobiological sequelae of potential relevance to psychotic disorders[127], such as alterations in the hypothalamic-pituitary-adrenal axis[128,129], inflammation[130], altered brain functioning[131,132], reduced brain volumes[133], and neurochemical dysfunctions[126,134,135]. However, studies to directly assess the correlations between these factors

Table 7 Level of evidence for the association of later factors and psychotic disorders

Factor	k	Random-effects measure, ES (95% CI)	Features used for classification of level of evidence						eOR	CE
			N	p random-effects	I^2 (p)	PI (95% CI)	SSE/ESB	LS		
Childhood trauma[54]	20	OR, 2.87 (2.07-3.98)	2,363	2.5×10^{-14}	77% (<0.001)	0.75-11.01	No/Yes	No	2.87	III
Toxoplasma gondii IgG[95]	42	OR, 1.82 (1.51-2.18)	8,796	2.1×10^{-10}	78% (<0.001)	0.68-4.88	Yes/Yes	No	1.82	III
Toxocara spp[83]	1	OR, 41.61 (9.71-178.32)	98	5.1×10^{-7}	NA (1.00)	NA	NA/NA	Yes	41.61	IV
Chlamydia psittaci[83]	2	OR, 29.05 (8.91-94.69)	82	2.2×10^{-8}	0% (0.71)	NA	NA/NA	Yes	29.05	IV
Human endogenous retrovirus type W[83]	5	OR, 19.78 (6.50-60.22)	256	1.4×10^{-7}	33% (0.20)	1.05-372.34	No/No	Yes	19.78	IV
Parental communication deviance[88]	4	g, 1.35 (0.97-1.73)	74	2.3×10^{-12}	0% (0.41)	0.51-2.19	No/No	No	11.55	IV
Chlamydia pneumoniae[90]	3	OR, 6.02 (2.86-12.66)	116	2.1×10^{-6}	0% (0.57)	0.05-745.30	No/No	Yes	6.02	IV
Traffic[84]	1	RR, 5.55 (1.63-18.87)	29	0.006	NA (<0.001)	NA	NA/NA	Yes	5.55	IV
Adult life events[85]	6	OR, 5.34 (3.84-7.43)	317	2.1×10^{-23}	3% (0.39)	3.22-8.87	No/No	Yes	5.34	IV
Heavy cannabis use[93]	2	OR, 5.17 (3.64-7.36)	748	6.3×10^{-20}	42% (0.18)	NA	NA/NA	Yes	5.17	IV
Benzene[84]	1	RR, 3.20 (1.01-10.12)	29	0.048	NA (1.00)	NA	NA/NA	Yes	3.20	IV
Tobacco use[89]	6	RR, 2.19 (1.36-3.53)	8,488	0.001	99% (<0.001)	0.38-12.50	No/No	Yes	2.19	IV
Borna disease virus[83]	21	OR, 1.94 (1.30-2.91)	1,919	0.001	36% (0.05)	0.65-5.81	No/No	Yes	1.94	IV
Traumatic brain injury[94]	8	OR, 1.49 (1.09-2.05)	9,653	0.013	78% (<0.001)	0.57-3.89	Yes/No	No	1.49	IV
Human herpes virus 2[83]	5	OR, 1.44 (1.14-1.81)	901	0.002	0% (0.97)	0.99-2.09	No/No	Yes	1.44	IV
Chlamydia trachomatis[83]	2	OR, 4.39 (0.03-587.92)	82	0.554	85% (0.01)	NA	NA/NA	No	4.39	ns
Human endogenous retrovirus[83]	4	OR, 3.64 (0.72-18.37)	128	0.117	36% (0.19)	0.01-1019	No/No	Yes	3.64	ns
Tetrachloroethylene[84]	1	RR, 3.41 (0.48-24.24)	4	0.219	NA (1.00)	NA	NA/NA	No	3.41	ns
Carbon monoxide[84]	1	RR, 3.07 (0.96-9.82)	29	0.059	NA (1.00)	NA	NA/NA	No	3.07	ns
Epilepsy[86]	1	OR, 3.06 (0.31-29.95)	4	0.337	NA (1.00)	NA	NA/NA	No	3.06	ns
Nitrogen oxides[84]	1	RR, 2.02 (0.74-5.53)	29	0.171	NA (1.00)	NA	NA/NA	No	2.02	ns
Central nervous system infection during childhood[91]	2	RR, 1.99 (0.31-12.78)	2,369	0.466	80% (0.02)	NA	NA/NA	No	1.99	ns
Epstein-Barr virus[83]	3	OR, 1.98 (0.23-16.85)	55	0.532	0% (0.81)	0-2121495	No/No	No	1.98	ns
Nitrogen dioxide[84]	1	RR, 1.91 (0.70-5.19)	29	0.205	NA (1.00)	NA	NA/NA	No	1.91	ns
Hearing impairment[92]	5	OR, 1.64 (0.85-3.15)	597	0.141	76% (0.002)	0.18-15.17	No/No	Yes	1.64	ns
Toxoplasma gondii IgM[95]	15	OR, 1.24 (0.97-1.59)	2,867	0.083	2% (0.43)	0.91-1.70	No/No	No	1.24	ns
Human herpes virus 1[90]	11	OR, 1.24 (0.98-1.58)	1,117	0.074	5% (0.39)	0.87-1.78	No/No	No	1.24	ns
Cytomegalovirus[83]	8	OR, 1.20 (0.65-2.20)	171	0.558	0% (1.00)	0.56-2.56	No/No	No	1.20	ns
Varicella zoster virus[83]	4	OR, 1.17 (0.16-8.58)	69	0.878	0% (0.99)	0.01-92.93	No/No	No	1.17	ns
BK virus[83]	1	OR, 1.05 (0.02-55.41)	20	0.979	NA (1.00)	NA	NA/NA	No	1.05	ns
JC virus[83]	1	OR, 1.05 (0.02-55.41)	20	0.979	NA (1.00)	NA	NA/NA	No	1.05	ns
Human endogenous retrovirus type k115[83]	1	OR, 0.89 (0.43-1.84)	178	0.753	NA (1.00)	NA	NA/NA	No	0.89	ns
Influenza[83]	2	OR, 0.87 (0.05-15.48)	33	0.925	0% (0.92)	NA	NA/NA	No	0.87	ns
Human T-lymphotropic virus 1[83]	2	OR, 0.57 (0.20-1.62)	209	0.294	0% (0.87)	NA	NA/NA	No	0.57	ns
Bullying[87]	1	OR, 0.38 (0.13-1.10)	30	0.075	NA (1.00)	NA	NA/NA	No	0.38	ns
Human herpes virus 6[90]	3	OR, 0.34 (0.05-2.42)	55	0.284	0% (0.71)	0-106440	No/No	No	0.34	ns

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small-study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, IRR – incidence rate ratio, OR – odds ratio, RR – relative risk, Ig – immunoglobulin, NA – not assessable, ns – not significant

Table 8 Level of evidence for the association of antecedents and psychotic disorders

Factor	k	Random-effects measure, ES (95%CI)	Features used for classification of level of evidence							
			N	p random-effects	I^2 (p)	PI (95% CI)	SSE/ESB	LS	eOR	CE
Ultra-high-risk state for psychosis[98]	9	RR, 9.32 (4.91 to 17.72)	1,226	9.5×10^{-12}	0% (0.91)	4.30 to 20.24	No/No	No	9.32	I
Minor physical anomalies[46]	14	g, 0.92 (0.61 to 1.23)	1,212	5.8×10^{-9}	91% (<0.001)	−0.34 to 2.18	No/Yes	Yes	5.30	II
Trait anhedonia[105]	44	g, 0.82 (0.72 to 0.92)	1,601	9.2×10^{-57}	43% (0.002)	0.37 to 1.27	No/Yes	Yes	4.41	II
Olfactory identification ability[63]	55	g, −0.91 (–1.05 to −0.78)	1,703	4.0×10^{-41}	67% (<0.001)	−1.72 to −0.10	Yes/Yes	Yes	0.19	II
Premorbid IQ[38,39]	16	g, −0.42 (−0.52 to −0.33)	4,459	1.1×10^{-18}	73% (<0.001)	−0.70 to −0.14	No/No	Yes	0.47	II
Childhood social withdrawal[42,43]	15	g, 0.59 (0.33 to 0.85)	1,810	6.4×10^{-6}	93% (<0.001)	−0.44 to 1.62	No/No	Yes	2.91	III
Non-right handedness[47]	41	OR, 1.58 (1.35 to 1.86)	2,652	2.0×10^{-8}	21% (0.12)	0.99 to 2.54	No/No	No	1.58	III
Neurological soft signs[49]	8	g, 1.83 (1.28 to 2.38)	564	7.7×10^{-11}	93% (<0.001)	−0.15 to 3.81	Yes/No	Yes	27.59	IV
Neuroticism[102]	8	g, 1.20 (0.88 to 1.52)	430	2.7×10^{-13}	73% (<0.001)	0.18 to.21	No/No	Yes	8.76	IV
Harm avoidance[101]	7	g, 0.98 (0.78 to 1.18)	384	4.5×10^{-21}	48% (0.07)	0.43 to 1.53	No/No	Yes	5.92	IV
Parkinsonism in antipsychotic-naïve schizophrenic patients[100]	3	OR, 5.33 (1.75 to 16.23)	84	0.003	0% (0.81)	0 to 7310	No/No	Yes	5.33	IV
Psychotic like experiences[99]	4	RR, 3.84 (2.55 to 5.79)	118	1.2×10^{-10}	0% (0.65)	1.56 to 9.45	No/No	No	3.84	IV
Dyskinesia in antipsychotic-naïve schizophrenic patients[100]	5	OR, 3.59 (1.53 to 8.42)	189	0.003	0% (0.75)	0.90 to 14.32	No/No	Yes	3.59	IV
Self-transcendence[101]	7	g, 0.61 (0.48 to 0.75)	384	7.8×10^{-19}	0% (0.67)	0.43 to 0.79	No/No	Yes	3.03	IV
Antisocial and externalizing behaviour[42]	3	g, 0.48 (0.22 to 0.74)	68	3.1×10^{-4}	36% (0.20)	−1.97 to 2.93	No/No	Yes	2.39	IV
Delay in walking unsupported[97]	5	g, 0.48 (0.27 to 0.68)	368	4.3×10^{-6}	81% (<0.001)	−0.27 to 1.22	Yes/NA	Yes	2.37	IV
Hypoalgesia[103]	9	g, 0.46 (0.13 to 0.79)	204	0.006	64% (0.005)	−0.57 to 1.49	No/No	No	2.31	IV
Extracranial size[104]	7	g, 0.27 (0.05 to 0.50)	192	0.018	15% (0.31)	−0.15 to 0.70	No/No	Yes	1.64	IV
Delay in standing unsupported[97]	4	g, 0.25 (0.12 to 0.39)	307	2.6×10^{-4}	48% (0.12)	−0.26 to 0.76	Yes/NA	No	1.58	IV
Delay in sitting unsupported[97]	4	g, 0.19 (0.05 to 0.33)	386	0.006	48% (0.12)	−0.33 to 0.70	Yes/NA	No	1.41	IV
Delay in holding head up[97]	3	g, 0.13 (0.01 to 0.24)	352	0.029	0% (0.91)	−0.61 to 0.86	Yes/NA	No	1.26	IV
Olfactory memory ability[63]	2	g, −1.62 (–2.24 to −1.01)	67	2.0×10^{-7}	56% (0.13)	NA	NA/NA	Yes	0.05	IV
Self-directedness[101]	7	g, −0.96 (–1.10 to −0.82)	384	7.7×10^{-42}	0% (0.75)	−1.14 to −0.78	No/No	Yes	0.17	IV
Extraversion[102]	8	g, −0.90 (–1.05 to −0.75)	430	3.6×10^{-32}	5% (0.38)	−1.13 to −0.67	No/No	Yes	0.20	IV
Olfactory discrimination ability[63]	8	g, −0.88 (–1.16 to −0.60)	226	4.1×10^{-10}	45% (0.07)	−1.61 to −0.15	No/No	Yes	0.20	IV
Olfactory hedonics ability (pleasant odours)[63]	10	g, −0.76 (−0.99 to −0.54)	298	2.5×10^{-11}	38% (0.10)	−1.34 to −0.19	No/No	Yes	0.25	IV
Conscientiousness[102]	7	g, −0.68 (−0.92 to −0.44)	399	2.2×10^{-8}	51% (0.05)	−1.33 to −0.04	No/No	Yes	0.29	IV
Olfactory detection ability[63]	18	g, −0.63 (−0.94 to −0.32)	498	5.9×10^{-5}	80% (<0.001)	−1.92 to 0.66	Yes/Yes	No	0.32	IV
Motor function pre-onset of psychosis[96]	4	g, −0.56 (−0.73 to −0.38)	152	4.1×10^{-10}	0% (0.60)	−0.94 to −0.17	No/No	Yes	0.36	IV
Olfactory hedonics ability (unspecified odours)[63]	7	g, −0.51 (−0.78 to −0.24)	142	2.1×10^{-4}	21% (0.26)	−1.06 to 0.05	No/No	No	0.40	IV
Agreeableness[102]	6	g, −0.47 (−0.88 to −0.07)	375	0.022	81% (<0.001)	−1.82 to 0.88	No/No	Yes	0.42	IV
Cooperativeness[101]	7	g, −0.47 (−0.60 to −0.33)	384	7.9×10^{-12}	0% (0.88)	−0.64 to −0.29	Yes/Yes	Yes	0.43	IV
Reward dependence[101]	7	g, −0.43 (−0.56 to −0.30)	384	2.7×10^{-10}	0% (0.43)	−0.61 to −0.26	No/No	Yes	0.46	IV
Openness[102]	7	g, −0.40 (−0.67 to −0.13)	399	0.003	62% (0.01)	−1.18 to 0.38	No/Yes	No	0.49	IV
Olfactory hedonics ability (unpleasant odours)[63]	9	g, −0.35 (−0.53 to −0.17)	244	1.3×10^{-4}	0% (0.79)	−0.57 to −0.13	No/No	No	0.53	IV
Persistence[101]	7	g, −0.24 (−0.39 to −0.08)	384	0.003	22% (0.26)	−0.56 to 0.09	No/No	No	0.65	IV
Total A-B ridge count[48]	13	g, −0.15 (−0.28 to −0.02)	979	0.027	46% (0.35)	−0.53 to 0.23	No/No	No	0.76	IV

Table 8 Level of evidence for the association of antecedents and psychotic disorders *(continued)*

Factor	k	Random-effects measure, ES (95%CI)	Features used for classification of level of evidence							
			N	p random-effects	I^2 (p)	PI (95% CI)	SSE/ESB	LS	eOR	CE
Fluctuating asymmetry A-B ridge count[48]	4	g, 0.74 (−0.65 to 2.13)	241	0.295	98% (<0.001)	−6.00 to 7.49	No/Yes	Yes	3.84	ns
Fluctuating asymmetry finger ridge count[48]	4	g, 0.31 (−0.50 to 1.12)	233	0.448	94% (<0.001)	−3.54 to 4.17	No/No	Yes	1.76	ns
Fingertip pattern asymmetry[48]	5	g, 0.25 (−0.08 to 0.59)	249	0.138	66% (0.02)	−0.85 to 1.35	No/No	Yes	1.58	ns
Poor general academic achievement pre-onset of psychosis[96]	4	g, 0.20 (−0.12 to 0.51)	1,007	0.219	93% (<0.001)	−1.25 to 1.65	No/No	Yes	1.43	ns
ATD angle[48]	5	g, 0.16 (−0.02 to 0.34)	261	0.083	0% (0.54)	−0.13 to 0.46	No/No	No	1.34	ns
Poor mathematic academic achievement pre-onset of psychosis[96]	3	g, 0.11 (−0.24 to 0.47)	136	0.527	63% (0.06)	−3.77 to 3.99	No/No	Yes	1.23	ns
Delay in grabbing object[97]	3	g, 0.05 (−0.07 to 0.17)	351	0.440	14% (0.31)	−0.90 to 1.00	Yes/NA	No	1.09	ns
Novelty seeking[101]	7	g, −0.31 (−0.68 to 0.05)	384	0.092	85% (<0.001)	−1.56 to 0.93	No/No	No	0.57	ns
Total finger ridge count[48]	13	g, −0.12 (−0.29 to 0.04)	935	0.149	65% (0.001)	−0.69 to 0.44	No/Yes	No	0.80	ns

k – number of samples for each factor, ES – effect size, N – number of cases, PI – prediction interval, CI – confidence interval, SSE – small study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CE – class of evidence, IRR – incidence rate ratio, OR – odds ratio, RR – relative risk, ATD angle – dermatoglyphic feature that compares the length of the hand to the width, NA – not assessable, ns – not significant

(e.g., urbanicity) and neurobiological alterations in psychotic disorders have only just started to emerge[133,136]. Until the exact mechanisms that lead to an increased risk of psychosis are determined, the requirement for biological or psychological plausibility for these factors cannot be fully met. Importantly, future research is required to clarify the contextual specifics of ethnic minority status and urbanicity, because their effects may also be modulated by geographical location or predominant population factors, rather than having universal value.

Several other factors beyond the ultra-high-risk state, ethnic minority status, and urbanicity provided a highly suggestive or a suggestive level of evidence of association with psychotic disorders, mostly confirming the role that perinatal factors (winter/spring season of birth in Northern hemisphere) or later factors/antecedents (childhood trauma and childhood social withdrawal, Toxoplasma gondii IgG, minor physical anomalies, trait anhedonia, low olfactory identification ability, low premorbid IQ, and non-right handedness) might have in psychosis onset. At the same time, a number of the explored factors showed only weak evidence of association with psychotic disorders. Some of these factors, such as heavy cannabis use and obstetric complications, were expected to have stronger evidence. However, weak findings in these areas may simply indicate that there are not yet enough data. Our umbrella review also identified only a few putative protective factors, indicating that the vast majority of available studies have focused on the adverse or negative end of several factors. Future research is required to actively seek unstudied protective factors that are not reciprocal to risk factors, such as specific characteristics of the individual, family or wider environment that improve the likelihood of positive outcomes[137].

This study has several conceptual implications. On an etiopathological level, our findings corroborate the notion that psychotic disorders can be related to adversities in an individual's social milieu, whereby environmental exposures during critical developmental periods impact brain, neurocognition, affect, and social cognition[13,138]. It is also apparent that most of these factors are likely not specific to psychosis, but also associated with other mental disorders[139]. From a transdiagnostic perspective, the current study can provide a benchmark for comparing the magnitude of association of these factors with other non-psychotic mental disorders. On a risk prediction level, these results may substantially advance our ability to prognosticate the onset of psychosis in populations at risk, paralleling the recent advancements observed in genetics.

In this latter area, the availability of robust meta-analytical evidence of associations between genetic loci and psychotic disorders – provided by the genome-wide association study (GWAS) meta-analysis conducted by the Schizophrenia Working Group of the Psychiatric Genomics Consortium[44] – has ultimately led to the development of polygenic risk scores to assess the *en masse* genetic effect of several loci[140]. Polygenic risk scores have been used to predict case-control status at the time of a first-episode psychosis, explaining approximately 9% of the variance[140]. The small proportion of variance explained indicates that the use of polygenic risk scores in clinical routine would be unwarranted[44] without first boosting them with other non-purely genetic factors.

Table 9 Sensitivity analysis for the associations of socio-demographic and parental, perinatal, later factors, antecedents and psychotic disorders within individual prospective studies of class I-III factors

				Features used for classification of level of evidence							
Factor	CE	k	Random-effects measure, ES (95% CI)	N > 1000	p random-effects	I^2 (p)	PI (95% CI)	SSE/ESB	LS	eOR	CES
Ultra-high-risk state for psychosis[98]	I	9	RR, 9.32 (4.91 to 17.72)	Yes	9.5×10^{-12}	0% (0.91)	4.30 to 20.24	No/No	No	9.32	I
Urbanicity[78]	III	8	OR, 2.19 (1.55 to 3.09)	Yes	8.9×10^{-6}	99% (<0.001)	0.62 to 7.77	No/No	Yes	2.19	III
Black-Caribbean ethnicity in England[76]	I	7	IRR, 5.54 (4.50 to 6.82)	No	4.9×10^{-59}	0% (0.48)	4.22 to 7.27	No/No	Yes	5.54	IV
Ethnic minority in low ethnic density area[71]	II	3	IRR, 4.27 (1.89 to 9.68)	No	4.9×10^{-4}	82% (0.004)	0 to 75335	Yes/No	Yes	4.27	IV
North African immigrants in Europe[77]	III	8	IRR, 3.20 (2.36 to 4.35)	No	1.0×10^{-13}	21% (0.27)	1.73 to 5.94	No/NA	Yes	3.20	IV
Childhood trauma[54]	III	4	OR, 2.52 (1.27 to 5.02)	Yes	0.009	71% (0.016)	0.14 to 46.01	No/Yes	No	2.52	IV
Ethnic minority in high ethnic density area[71]	III	3	IRR, 2.51 (1.10 to 5.71)	No	0.028	70% (0.037)	0 to 28153	No/No	Yes	2.51	IV
Childhood social withdrawal[42,43]	III	11	g, 0.43 (0.14 to 0.71)	Yes	0.003	94% (<0.001)	−0.63 to 1.48	No/No	Yes	2.16	IV
First generation immigrants[53]	III	12	IRR, 1.83 (1.40 to 2.38)	No	9.6×10^{-6}	0% (0.82)	1.35 to 2.47	No/Yes	No	1.83	IV
Second generation immigrants[53]	II	10	IRR, 1.45 (1.05 to 2.01)	Yes	0.023	76% (<0.001)	0.54 to 3.95	Yes/No	No	1.45	IV
Toxoplasma gondii IgG[95]	III	7	OR, 1.28 (1.06 to 1.55)	Yes	0.012	22% (0.26)	0.86 to 1.91	Yes/No	No	1.28	IV
Premorbid IQ[38,39]	III	9	g, −0.43 (−0.64 to −0.22)	No	5.2×10^{-5}	62% (0.007)	−1.04 to 0.18	No/No	Yes	0.46	IV
Non-right handedness[47]	III	1	OR, 1.83 (0.62 to 5.39)	No	0.273	NA	NA	NA/NA	No	1.83	ns
Minor physical anomalies[46]	II	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
Olfactory identification ability[63]	II	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
Trait anhedonia[105]	II	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC
Winter/spring season of birth in Northern hemisphere[79]	III	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC

CE – class of evidence, k – number of samples for each factor within prospective studies, ES – effect size, CI – confidence interval, N – number of cases, PI – prediction interval, SSE – small study effect, ESB – excess significance bias, LS – largest study with significant effect, eOR – equivalent odds ratio, CES – class of evidence after sensitivity analysis, RR – relative risk, OR – odds ratio, IRR – incidence rate ratio, NA – not assessable, Ig – immunoglobulin, ns – non-significant, NC – not calculable (no prospective studies to be analyzed)

To date, the integration of multiple non-purely genetic factors into "polyrisk" scores has been limited by the lack of established and robust *a priori* knowledge on their association with psychotic disorders[139]. The current umbrella review attempted to fill this knowledge gap by providing the most robust estimates of association (ORs) between several non-purely genetic risk (or protective) factors and psychotic disorders. Assessing these predictive factors may offer some logistic advantages over more complex measurements that are based on cognitive, imaging, central or peripheral measures. Simple demographic factors have already been used to develop an individualized risk estimation tool to predict psychosis onset in at-risk individuals in clinical practice[141].

Recently, geneticists have advocated the development of polyrisk scores encompassing socio-demographic, parental, perinatal, later factors, antecedents, and genetic risk profiling[139,142]. Such an approach may ultimately reveal new, clinically useful predictors, because even the risk factors that we found to be weakly associated with psychotic disorders may eventually contribute to the predictive accuracy of the model, as previously observed for genetic associations[44]. The current umbrella review lays the groundwork for testing the predictive accuracy of integrated polyrisk scores in independent samples[139].

Finally, on a pragmatic level, the current stratification of evidence can be used by clinicians, policy makers and regulatory bodies to inform and strategically target outreach campaigns, to promote the prevention of mental disorders in the youth population, and to raise awareness of risk factors for psychotic disorders.

This study also has some limitations. First, association is not necessarily causation. Reverse causation is a particular concern[13], and thus establishing the temporality of the association is critical. It is possible that some of the later factors and antecedents are actually characteristics of psychotic disorders themselves or secondary to their appearance. To specifically address these problems and the effect of temporality, we con-

ducted a sensitivity analysis restricted to individual studies with prospective designs.

A second limitation is that the umbrella review approach may favour the selection of more commonly and readily studied factors, since they are more likely to be included in a meta-analysis. We cannot exclude the possibility that some promising factors, despite having sufficient data, do not have a corresponding eligible meta-analysis, such as mood and anxiety disorders[143-145], personality disorders[146], attachment[147], alcohol and psychoactive substances[148-151], sleep dysfunction[152], homelessness[153] or pervasive developmental disorders[154]. However, this possibility is becoming less likely in the current era, with meta-analyses being performed massively, to the point that for several topics multiple meta-analyses are available[155,156]. In any case, for most putative risk or protective factors that are difficult to study (or uncommonly studied), the current grade of evidence is unlikely to be remarkable, given the limited data.

A third limitation is that the definition of healthy control groups employed by each meta-analysis/systematic review or, when this was not provided, by each individual study, may not be entirely accurate. Moreover, some of the factors included in this umbrella review may be better conceptualized as risk markers, because they may be the result of different interacting risk factors. The ultra-high-risk state[98], ethnicity[76] and immigration status[53,77] are prototypical examples of risk markers, and their limitations have already been addressed above.

Another caution is that the categories of socio-demographic and parental, perinatal, later factors, and antecedents[7-9] were used only for descriptive purposes. As noted in the Methods section, these categories may actually overlap to some extent. Finally, the relevance of epigenetic risk factors, and the interaction between environmental and genetic factors in psychotic disorders, remains to be elucidated[157].

In conclusion, we found several factors to be associated with psychotic disorders at different levels of evidence. These factors represent a starting point of knowledge that can be used to advance etiological research and improve the prediction of psychosis.

ACKNOWLEDGEMENTS

This study was supported in part by the National Institute for Health Research Biomedical Research Centre at South London and Maudsley National Health Service Foundation Trust and King's College London, by a 2017 Medical Research Council Confidence in Concept grant to P. Fusar-Poli, and by Instituto de Salud Carlos III and FEDER grants (CP14/00041 and PI14/00292) to J. Radua. The funders had no influence on the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The views expressed are those of the authors and not necessarily those of the funders. J. Radua and V. Ramella-Cravaro contributed equally to this work.

REFERENCES

1. Global Burden of Disease Study Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;386:743-800.
2. Kirkbride JB, Errazuriz A, Croudace TJ et al. Incidence of schizophrenia and other psychoses in England, 1950–2009: a systematic review and meta-analyses. PLoS One 2012;7:e31660.
3. Cloutier M, Aigbogun MS, Guerin A et al. The economic burden of schizophrenia in the United States in 2013. J Clin Psychiatry 2016;77:764-71.
4. Fusar-Poli P, McGorry PD, Kane J. Improving outcomes of first-episode psychosis: an overview. World Psychiatry 2017;16:251-65.
5. van Os J, Rutten BPF, Poulton R. Gene-environment interactions in schizophrenia: review of epidemiological findings and future directions. Schizophr Bull 2008;34:1066-82.
6. Howes OD, Murray RM. Schizophrenia: an integrated sociodevelopmental-cognitive model. Lancet 2014;383:1677-87.
7. Matheson SL, Shepherd AM, Laurens KR et al. A systematic meta-review grading the evidence for non-genetic risk factors and putative antecedents of schizophrenia. Schizophr Res 2011;133:133-42.
8. Laurens KR, Luo L, Matheson SL et al. Common or distinct pathways to psychosis? A systematic review of evidence from prospective studies for developmental risk factors and antecedents of the schizophrenia spectrum disorders and affective psychoses. BMC Psychiatry 2015;15:205.
9. Dean K MR. Environmental risk factors for psychosis. Dialogues Clin Neurosci 2005;7:69-80.
10. Fusar-Poli P. The Clinical High-Risk State for Psychosis (CHR-P), Version II. Schizophr Bull 2017;43:44-7.
11. 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.
12. Ioannidis JP, Munafo MR, Fusar-Poli P et al. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. Trends Cogn Sci 2014;18:235-41.
13. van Os J, Kenis G, Rutten BPF. The environment and schizophrenia. Nature 2010;468:203-12.
14. Aromataris E, Fernandez R, Godfrey CM et al. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Health 2015;13:132-40.
15. Dinu M, Pagliai G, Casini A et al. Mediterranean diet and multiple health outcomes: an umbrella review of meta-analyses of observational studies and randomised trials. Eur J Clin Nutr (in press).
16. Bortolato B, Kohler CA, Evangelou E et al. Systematic assessment of environmental risk factors for bipolar disorder: an umbrella review of systematic reviews and meta-analyses. Bipolar Disord 2017;19:84-96.
17. Kyrgiou M, Kalliala I, Markozannes G et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. BMJ 2017;356:j477.
18. Dragioti E, Karathanos V, Gerdle B et al. Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. Acta Psychiatr Scand 2017;136:236-46.
19. Markozannes G, Tzoulaki I, Karli D et al. Diet, body size, physical activity and risk of prostate cancer: an umbrella review of the evidence. Eur J Cancer 2016;69:61-9.
20. Belbasis L, Savvidou MD, Kanu C et al. Birth weight in relation to health and disease in later life: an umbrella review of systematic reviews and meta-analyses. BMC Med 2016;14:147.
21. Belbasis L, Stefanaki I, Stratigos AJ et al. Non-genetic risk factors for cutaneous melanoma and keratinocyte skin cancers: an umbrella review of meta-analyses. J Dermatol Sci 2016;84:330-9.
22. Papageorgiou PN, Deschner J, Papageorgiou SN. Effectiveness and adverse effects of deep brain stimulation: umbrella review of meta-analyses. J Neurol Surg A Cent Eur Neurosurg 2017;78:180-90.
23. Bellou V, Belbasis L, Tzoulaki I et al. Systematic evaluation of the associations between environmental risk factors and dementia: an umbrella review of systematic reviews and meta-analyses. Alzheimers Dement 2017;13:406-18.
24. Carvalho AF, Kohler CA, Fernandes BS et al. Bias in emerging biomarkers for bipolar disorder. Psychol Med 2016;46:2287-97.
25. Bellou V, Belbasis L, Tzoulaki I et al. Environmental risk factors and Parkinson's disease: an umbrella review of meta-analyses. Parkinsonism Relat Disord 2016;23:1-9.
26. Belbasis L, Bellou V, Evangelou E. Environmental risk factors and amyotrophic lateral sclerosis: an umbrella review and critical assessment of current evidence from systematic reviews and meta-analyses of observational studies. Neuroepidemiology 2016;46:96-105.
27. Belbasis L, Bellou V, Evangelou E et al. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. Lancet Neurol 2015;14:263-73.

28. Tsilidis KK, Kasimis JC, Lopez DS et al. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. BMJ 2015; 350:g7607.
29. Theodoratou E, Tzoulaki I, Zgaga L et al. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. BMJ 2014;348:g2035.
30. Tonelli AR, Zein J, Adams J et al. Effects of interventions on survival in acute respiratory distress syndrome: an umbrella review of 159 published randomized trials and 29 meta-analyses. Intensive Care Med 2014;40: 769-87.
31. Ioannidis JP, Zhou Y, Chang CQ et al. Potential increased risk of cancer from commonly used medications: an umbrella review of meta-analyses. Ann Oncol 2014;25:16-23.
32. Doufas AG, Panagiotou OA, Ioannidis JP. Concordance of sleep and pain outcomes of diverse interventions: an umbrella review. PLoS One 2012;7: e40891.
33. Contopoulos-Ioannidis DG, Ioannidis JP. Claims for improved survival from systemic corticosteroids in diverse conditions: an umbrella review. Eur J Clin Invest 2012;42:233-44.
34. Fleischhacker WW. A meta view on meta-analyses. JAMA Psychiatry 2017; 74:684-5.
35. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. CMAJ 2009;181:488-93.
36. Geddes JR, Verdoux H, Takei N et al. Schizophrenia and complications of pregnancy and labor: an individual patient data meta-analysis. Schizophr Bull 1999;25:413-23.
37. Cannon M, Jones PB, Murray RM. Obstetric complications and schizophrenia: historical and meta-analytic review. Am J Psychiatry 2002;159: 1080-92.
38. Woodberry KA, Giuliano AJ, Seidman LJ. Premorbid IQ in schizophrenia: a meta-analytic review. Am J Psychiatry 2008;165:579-87.
39. Khandaker GM, Barnett JH, White IR et al. A quantitative meta-analysis of population-based studies of premorbid intelligence and schizophrenia. Schizophr Res 2011;132:220-7.
40. Cai L, Wan CL, He L et al. Gestational influenza increases the risk of psychosis in adults. Med Chem 2015;11:676-82.
41. Selten J-P, Termorshuizen F. The serological evidence for maternal influenza as risk factor for psychosis in offspring is insufficient: critical review and meta-analysis. Schizophr Res 2017;183:2-9.
42. Tarbox SI, Pogue-Geile MF. Development of social functioning in preschizophrenia children and adolescents: a systematic review. Psychol Bull 2008;134:561-83.
43. Matheson SL, Vijayan H, Dickson H et al. Systematic meta-analysis of childhood social withdrawal in schizophrenia, and comparison with data from at-risk children aged 9-14 years. J Psychiatr Res 2013;47:1061-8.
44. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014;511:421-7.
45. O'Dushlaine C, Rossin L, Lee PH et al. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci 2015;18:199-209.
46. Xu T, Chan RCK, Compton MT. Minor physical anomalies in patients with schizophrenia, unaffected first-degree relatives, and healthy controls: a meta-analysis. PLoS One 2011;6:e24129.
47. Hirnstein M, Hugdahl K. Excess of non-right-handedness in schizophrenia: meta-analysis of gender effects and potential biases in handedness assessment. Br J Psychiatry 2014;205:260-7.
48. Golembo-Smith S, Walder DJ, Daly MP et al. The presentation of dermatoglyphic abnormalities in schizophrenia: a meta-analytic review. Schizophr Res 2012;142:1-11.
49. Neelam K, Garg D, Marshall M. A systematic review and meta-analysis of neurological soft signs in relatives of people with schizophrenia. BMC Psychiatry 2011;11:139.
50. Christesen HT, Elvander C, Lamont RF et al. The impact of vitamin D in pregnancy on extraskeletal health in children: a systematic review. Acta Obstet Gynecol Scand 2012;91:1368-80.
51. Moher D, Liberati A, Tetzlaff J et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006-12.
52. Stroup DF, Berlin JA, Morton SC et al. Meta-analysis of observational studies in epidemiology - a proposal for reporting. JAMA 2000;283:2008-12.
53. Bourque F, van der Ven E, Malla A. A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. Psychol Med 2011;41:897-910.
54. Varese F, Smeets F, Drukker M et al. Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective- and cross-sectional cohort studies. Schizophr Bull 2012;38:661-71.
55. Kinney DK, Teixeira P, Hsu D et al. Relation of schizophrenia prevalence to latitude, climate, fish consumption, infant mortality, and skin color: a role for prenatal vitamin D deficiency and infections? Schizophr Bull 2009;35:582-95.
56. Saha S, Chant D, Welham J et al. A systematic review of the prevalence of schizophrenia. PLoS Med 2005;2:413-33.
57. Rohatgi A. WebPlotDigitizer 3.10. arohatgi.info/WebPlotDigitizer/.
58. Shea BJ, Grimshaw JM, Wells GA et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
59. Shea BJ, Bouter LM, Peterson J et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One 2007;2: e1350.
60. Pieper D, Mathes T, Eikermann M. Can AMSTAR also be applied to systematic reviews of non-randomized studies? BMC Res Notes 2014;7:609.
61. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2011.
62. Higgins JPT, Green S (eds). Cochrane handbook for systematic reviews of interventions version 5.1.0. The Cochrane Collaboration, 2011.
63. Moberg P, Kamath V, Marchetto D et al. Meta-analysis of olfactory function in schizophrenia, first-degree family members, and youths at-risk for psychosis. Schizophr Bull 2014;40:50-9.
64. Rubia K, Alegria AA, Cubillo AI et al. Effects of stimulants on brain function in attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. Biol Psychiatry 2014;76:616-28.
65. Schwarzer G. Meta: an R package for meta-analysis. R News 2007;7:40-5.
66. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc A Stat 2009;172:137-59.
67. Jackson D, Kirkbride J, Croudace T et al. Meta-analytic approaches to determine gender differences in the age-incidence characteristics of schizophrenia and related psychoses. Int J Methods Psychiatr Res 2013; 22:36-45.
68. Egger M, Smith GD, Schneider M et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
69. Ioannidis JPA, Trikalinos TA. An exploratory test for an excess of significant findings. Clin Trials 2007;4:245-53.
70. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 2000;19:3127-31.
71. Bosqui TJ, Hoy K, Shannon C. A systematic review and meta-analysis of the ethnic density effect in psychotic disorders. Soc Psychiatry Psychiatr Epidemiol 2014;49:519-29.
72. Torrey EF, Buka S, Cannon TD et al. Paternal age as a risk factor for schizophrenia: how important is it? Schizophr Res 2009;114:1-5.
73. Kwok W. Is there evidence that social class at birth increases risk of psychosis? A systematic review. Int J Soc Psychiatry 2014;60:801-8.
74. O'Donoghue B, Roche E, Lane A. Neighbourhood level social deprivation and the risk of psychotic disorders: a systematic review. Soc Psychiatry Psychiatr Epidemiol 2016;51:941-50.
75. Rasic D, Hajek T, Alda M et al. Risk of mental illness in offspring of parents with schizophrenia, bipolar disorder, and major depressive disorder: a meta-analysis of family high-risk studies. Schizophr Bull 2014;40:28-38.
76. Tortelli A, Errazuriz A, Croudace T et al. Schizophrenia and other psychotic disorders in Caribbean-born migrants and their descendants in England: systematic review and meta-analysis of incidence rates, 1950–2013. Soc Psychiatry Psychiatr Epidemiol 2015;50:1039-55.
77. van der Ven E, Veling W, Tortelli A et al. Evidence of an excessive gender gap in the risk of psychotic disorder among North African immigrants in Europe: a systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2016;51:1603-13.
78. Vassos E, Pedersen CB, Murray RM et al. Meta-analysis of the association of urbanicity with schizophrenia. Schizophr Bull 2012;38:1118-23.
79. Davies G, Welham J, Chant D et al. A systematic review and meta-analysis of Northern Hemisphere season of birth studies in schizophrenia. Schizophr Bull 2003;29:587-93.
80. McGrath JJ, Welham JL. Season of birth and schizophrenia: a systematic review and meta-analysis of data from the Southern Hemisphere. Schizophr Res 1999;35:237-42.

81. Geddes JR, Lawrie SM. Obstetric complications and schizophrenia: a meta-analysis. Br J Psychiatry 1995;167:786-93.
82. Van Lieshout RJ, Taylor VH, Boyle MH. Pre-pregnancy and pregnancy obesity and neurodevelopmental outcomes in offspring: a systematic review. Obes Rev 2011;12:e548-59.
83. Arias I, Sorlozano A, Villegas E et al. Infectious agents associated with schizophrenia: a meta-analysis. Schizophr Res 2012;136:128-36.
84. Attademo L, Bernardini F, Garinella R et al. Environmental pollution and risk of psychotic disorders: a review of the science to date. Schizophr Res 2017;181:55-9.
85. Beards S, Gayer-Anderson C, Borges S et al. Life events and psychosis: a review and meta-analysis. Schizophr Bull 2013;39:740-7.
86. Clancy MJ, Clarke MC, Connor DJ et al. The prevalence of psychosis in epilepsy; a systematic review and meta-analysis. BMC Psychiatry 2014; 14:75.
87. Cunningham T, Hoy K, Shannon C. Does childhood bullying lead to the development of psychotic symptoms? A meta-analysis and review of prospective studies. Psychosis 2015;8:48-59.
88. de Sousa P, Varese F, Sellwood W et al. Parental communication and psychosis: a meta-analysis. Schizophr Bull 2014;40:756-68.
89. Gurillo P, Jauhar S, Murray RM et al. Does tobacco use cause psychosis? Systematic review and meta-analysis. Lancet Psychiatry 2015;2:718-25.
90. Gutierrez-Fernandez J, de Dios Luna del Castillo J, Mananes-Gonzalez S et al. Different presence of Chlamydia pneumoniae, herpes simplex virus type 1, human herpes virus 6, and Toxoplasma gondii in schizophrenia: meta-analysis and analytical study. Neuropsychiatr Dis Treat 2015;11: 843-52.
91. Khandaker GM, Zimbron J, Dalman C et al. Childhood infection and adult schizophrenia: a meta-analysis of population-based studies. Schizophr Res 2012;139:161-8.
92. Linszen MMJ, Brouwer RM, Heringa SM et al. Increased risk of psychosis in patients with hearing impairment: review and meta-analyses. Neurosci Biobehav Rev 2016;62:1-20.
93. 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.
94. Molloy C, Conroy RM, Cotter DR et al. Is traumatic brain injury a risk factor for schizophrenia? A meta-analysis of case-controlled population-based studies. Schizophr Bull 2011;37:1104-10.
95. Sutterland AL, Fond G, Kuin A et al. Beyond the association. Toxoplasma gondii in schizophrenia, bipolar disorder, and addiction: systematic review and meta-analysis. Acta Psychiatr Scand 2015;132:161-79.
96. Dickson H, Laurens KR, Cullen AE et al. Meta-analyses of cognitive and motor function in youth aged 16 years and younger who subsequently develop schizophrenia. Psychol Med 2012;42:743-55.
97. Filatova S, Koivumaa-Honkanen H, Hirvonen N et al. Early motor developmental milestones and schizophrenia: a systematic review and meta-analysis. Schizophr Res 2017;188:13-20.
98. 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.
99. Kaymaz N, Drukker M, Lieb R et al. Do subthreshold psychotic experiences predict clinical outcomes in unselected non-help-seeking population-based samples? A systematic review and meta-analysis, enriched with new results. Psychol Med 2012;42:2239-53.
100. Koning JPF, Tenback DE, van Os J et al. Dyskinesia and Parkinsonism in antipsychotic-naive patients with schizophrenia, first-degree relatives and healthy controls: a meta-analysis. Schizophr Bull 2010;36:723-31.
101. Ohi K, Hashimoto R, Yasuda Y et al. Personality traits and schizophrenia: evidence from a case-control study and meta-analysis. Psychiatry Res 2012;198:7-11.
102. Ohi K, Shimada T, Nitta Y et al. The Five-Factor Model personality traits in schizophrenia: a meta-analysis. Psychiatry Res 2016;240:34-41.
103. Potvin S, Marchand S. Hypoalgesia in schizophrenia is independent of antipsychotic drugs: a systematic quantitative review of experimental studies. Pain 2008;138:70-8.
104. Ward KE, Friedman L, Wise A et al. Meta-analysis of brain and cranial size in schizophrenia. Schizophr Res 1996;22:197-213.
105. Yan C, Cao Y, Zhang Y et al. Trait and state positive emotional experience in schizophrenia: a meta-analysis. PLoS One 2012;7:e40672.
106. Selten JP, Cantor-Graae E, Nahon D et al. No relationship between risk of schizophrenia and prenatal exposure to stress during the Six-Day War or Yom Kippur War in Israel. Schizophr Res 2003;63:131-5.
107. O'Rourke K. An historical perspective on meta-analysis: dealing quantitatively with varying study results. J Roy Soc Med 2007;100:579-82.
108. 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.
109. Fusar-Poli P, Rutigliano G, Stahl D et al. Long-term validity of the At Risk Mental State (ARMS) for predicting psychotic and non-psychotic mental disorders. Eur Psychiatry 2017;42:49-54.
110. Kempton MJ, Bonoldi I, Valmaggia L et al. Speed of psychosis progression in people at ultra-high clinical risk: a complementary meta-analysis. JAMA Psychiatry 2015;72:622-3.
111. Fusar-Poli P, Cappucciati M, De Micheli A et al. Diagnostic and prognostic significance of brief limited intermittent psychotic symptoms (BLIPS) in individuals at ultra high risk. Schizophr Bull 2017;43:48-56.
112. Fusar-Poli P, Cappucciati M, Bonoldi I et al. Prognosis of brief psychotic episodes: a meta-analysis. JAMA Psychiatry 2016;73:211-20.
113. Falkenberg I, Valmaggia L, Byrnes M et al. Why are help-seeking subjects at ultra-high risk for psychosis help-seeking? Psychiatry Res 2015;228: 808-15.
114. Fusar-Poli P, Rocchetti M, Sardella A et al. Disorder, not just state of risk: meta-analysis of functioning and quality of life in people at high risk of psychosis. Br J Psychiatry 2015;207:198-206.
115. Fusar-Poli P, Cappucciati M, Borgwardt S et al. Heterogeneity of psychosis risk within individuals at clinical high risk: a meta-analytical stratification. JAMA Psychiatry 2016;73:113-20.
116. Fusar-Poli P, Cappucciati M, Rutigliano G et al. Towards a standard psychometric diagnostic interview for subjects at ultra high risk of psychosis: CAARMS versus SIPS. Psychiatry J 2016;2016:7146341.
117. Fusar-Poli P, Tantardini M, De Simone S et al. Deconstructing vulnerability for psychosis: meta-analysis of environmental risk factors for psychosis in subjects at ultra high-risk. Eur Psychiatry 2017;40:65-75.
118. Fusar-Poli P, Rutigliano G, Stahl D et al. Deconstructing pretest risk enrichment to optimize prediction of psychosis in individuals at clinical high risk. JAMA Psychiatry 2016;73:1260-7.
119. Fusar-Poli P, Schultze-Lutter F, Cappucciati M et al. The dark side of the moon: meta-analytical impact of recruitment strategies on risk enrichment in the clinical high risk state for psychosis. Schizophr Bull 2016;42: 732-43.
120. Fusar-Poli P, Schultze-Lutter F, Addington J. Intensive community outreach for those at ultra high risk of psychosis: dilution, not solution. Lancet Psychiatry 2016;3:18.
121. Fusar-Poli P, Palombini E, Davies C et al. Why transition risk to psychosis is not declining at the OASIS ultra high risk service: the hidden role of stable pretest risk enrichment. Schizophr Res (in press).
122. Fusar-Poli P. Why ultra high risk criteria for psychosis prediction do not work well outside clinical samples and what to do about it. World Psychiatry 2017;16:212-3.
123. Kirkbride J, Hameed Y, Ankireddypalli G et al. The epidemiology of first-episode psychosis in early intervention in psychosis services: findings from the Social Epidemiology of Psychoses in East Anglia [SEPEA] Study. Am J Psychiatry 2017;174:143-53.
124. Heinz A, Deserno L, Reininghaus U. Urbanicity, social adversity and psychosis. World Psychiatry 2013;12:187-97.
125. Selten J-P, van Os J, Cantor-Graae E. The social defeat hypothesis of schizophrenia: issues of measurement and reverse causality. World Psychiatry 2016;15:294-5.
126. Mizrahi R. Social stress and psychosis risk: common neurochemical substrates? Neuropsychopharmacology 2016;41:666-74.
127. Holtzman CW, Trotman HD, Goulding SM et al. Stress and neurodevelopmental processes in the emergence of psychosis. Neuroscience 2013;249: 172-91.
128. Steinheuser V, Ackermann K, Schonfeld P et al. Stress and the city: impact of urban upbringing on the (re)activity of the hypothalamus-pituitary-adrenal axis. Psychosom Med 2014;76:678-85.
129. Mewes R, Reich H, Skoluda N et al. Elevated hair cortisol concentrations in recently fled asylum seekers in comparison to permanently settled immigrants and non-immigrants. Transl Psychiatry 2017;7:e1051.
130. Powell ND, Sloan EK, Bailey MT et al. Social stress up-regulates inflammatory gene expression in the leukocyte transcriptome via beta-adrenergic induction of myelopoiesis. Proc Natl Acad Sci USA 2013;110:16574-9.
131. Akdeniz C, Tost H, Streit F et al. Neuroimaging evidence for a role of neural social stress processing in ethnic minority-associated environmental risk. JAMA Psychiatry 2014;71:672-80.

132. Lederbogen F, Kirsch P, Haddad L et al. City living and urban upbringing affect neural social stress processing in humans. Nature 2011;474:498-501.
133. Haddad L, Schafer A, Streit F et al. Brain structure correlates of urban upbringing, an environmental risk factor for schizophrenia. Schizophr Bull 2015;41:115-22.
134. Selten JP, Booij J, Buwalda B et al. Biological mechanisms whereby social exclusion may contribute to the etiology of psychosis: a narrative review. Schizophr Bull (in press).
135. Egerton A, Howes OD, Houle S et al. Elevated striatal dopamine function in immigrants and their children: a risk mechanism for psychosis. Schizophr Bull 2017;43:293-301.
136. Frissen A, van Os J, Habets P et al. No evidence of association between childhood urban environment and cortical thinning in psychotic disorder. PLoS One 2017;12:e0166651.
137. Keskinen E, Marttila R, Koivumaa-Honkanen H et al. Search for protective factors for psychosis - a population-based sample with special interest in unaffected individuals with parental psychosis. Early Interv Psychiatry (in press).
138. Millan MJ, Andrieux A, Bartzokis G et al. Altering the course of schizophrenia: progress and perspectives. Nat Rev Drug Discov 2016;15:485-515.
139. Uher R, Zwicker A. Etiology in psychiatry: embracing the reality of polygene-environmental causation of mental illness. World Psychiatry 2017; 16:121-9.
140. Vassos E, Di Forti M, Coleman J et al. An examination of polygenic score risk prediction in individuals with first episode psychosis. Biol Psychiatry 2016;81:470-7.
141. Fusar-Poli P, Rutigliano G, Stahl D et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. JAMA Psychiatry 2017;74:493-500.
142. Iyegbe C, Campbell D, Butler A et al. The emerging molecular architecture of schizophrenia, polygenic risk scores and the clinical implications for GxE research. Soc Psychiatry Psychiatr Epidemiol 2014;49:169-82.
143. Hartley S, Barrowclough C, Haddock G. Anxiety and depression in psychosis: a systematic review of associations with positive psychotic symptoms. Acta Psychiatr Scand 2013;128:327-46.
144. Mishara AL, Fusar-Poli P. The phenomenology and neurobiology of delusion formation during psychosis onset: Jaspers, Truman symptoms, and aberrant salience. Schizophr Bull 2013;39:278-86.
145. Achim AM, Maziade M, Raymond E et al. How prevalent are anxiety disorders in schizophrenia? A meta-analysis and critical review on a significant association. Schizophr Bull 2011;37:811-21.
146. Newton-Howes G, Tyrer P, North B et al. The prevalence of personality disorder in schizophrenia and psychotic disorders: systematic review of rates and explanatory modelling. Psychol Med 2008;38:1075-82.
147. Gumley AI, Taylor HEF, Schwannauer M et al. A systematic review of attachment and psychosis: measurement, construct validity and outcomes. Acta Psychiatr Scand 2014;129:257-74.
148. Large M, Sharma S, Compton MT et al. Cannabis use and earlier onset of psychosis. Arch Gen Psychiatry 2011;68:555-61.
149. Murray RM, Quigley H, Quattrone D et al. Traditional marijuana, high-potency cannabis and synthetic cannabinoids: increasing risk for psychosis. World Psychiatry 2016;15:195-204.
150. Bramness JG, Rognli EB. Psychosis induced by amphetamines. Curr Opin Psychiatry 2016;29:236-41.
151. Roncero C, Ros-Cucurull E, Daigre C et al. Prevalence and risk factors of psychotic symptoms in cocaine-dependent patients. Actas Esp Psiquiatr 2012;40:187-97.
152. Reeve S, Sheaves B, Freeman D. The role of sleep dysfunction in the occurrence of delusions and hallucinations: a systematic review. Clin Psychol Rev 2015;42:96-115.
153. Folsom D, Jeste DV. Schizophrenia in homeless persons: a systematic review of the literature. Acta Psychiatr Scand 2002;105:404-13.
154. Padgett FE, Miltsiou E, Tiffin PA. The co-occurrence of nonaffective psychosis and the pervasive developmental disorders: a systematic review. J Intellect Dev Disabil 2010;35:187-98.
155. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016;94:485-514.
156. Siontis KC, Hernandez-Boussard T, Ioannidis JP. Overlapping meta-analyses on the same topic: survey of published studies. BMJ 2013;347: f4501.
157. Millan MJ. An epigenetic framework for neurodevelopmental disorders: from pathogenesis to potential therapy. Neuropharmacology 2013;68:2-82.

DOI:10.1002/wps.20490

Prediction of psychosis across protocols and risk cohorts using automated language analysis

Cheryl M. Corcoran[1,2], Facundo Carrillo[3,4], Diego Fernández-Slezak[3,4], Gillinder Bedi[2,5,6], Casimir Klim[2,5], Daniel C. Javitt[2,5], Carrie E. Bearden[7], Guillermo A. Cecchi[8]

[1]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [2]New York State Psychiatric Institute, New York, NY, USA; [3]Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina; [4]Instituto de Investigación en Ciencias de la Computación, Universidad de Buenos Aires, Buenos Aires, Argentina; [5]Department of Psychiatry, Columbia University Medical Center, New York, NY, USA; [6]Centre for Youth Mental Health, University of Melbourne, and Orygen National Centre of Excellence in Youth Mental Health, Melbourne, Australia; [7]Department of Psychiatry and Biobehavioral Sciences and Psychology, University of California Los Angeles; Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA; [8]Computational Biology Center - Neuroscience, IBM T.J. Watson Research Center, Ossining, NY, USA

Language and speech are the primary source of data for psychiatrists to diagnose and treat mental disorders. In psychosis, the very structure of language can be disturbed, including semantic coherence (e.g., derailment and tangentiality) and syntactic complexity (e.g., concreteness). Subtle disturbances in language are evident in schizophrenia even prior to first psychosis onset, during prodromal stages. Using computer-based natural language processing analyses, we previously showed that, among English-speaking clinical (e.g., ultra) high-risk youths, baseline reduction in semantic coherence (the flow of meaning in speech) and in syntactic complexity could predict subsequent psychosis onset with high accuracy. Herein, we aimed to cross-validate these automated linguistic analytic methods in a second larger risk cohort, also English-speaking, and to discriminate speech in psychosis from normal speech. We identified an automated machine-learning speech classifier – comprising decreased semantic coherence, greater variance in that coherence, and reduced usage of possessive pronouns – that had an 83% accuracy in predicting psychosis onset (intra-protocol), a cross-validated accuracy of 79% of psychosis onset prediction in the original risk cohort (cross-protocol), and a 72% accuracy in discriminating the speech of recent-onset psychosis patients from that of healthy individuals. The classifier was highly correlated with previously identified manual linguistic predictors. Our findings support the utility and validity of automated natural language processing methods to characterize disturbances in semantics and syntax across stages of psychotic disorder. The next steps will be to apply these methods in larger risk cohorts to further test reproducibility, also in languages other than English, and identify sources of variability. This technology has the potential to improve prediction of psychosis outcome among at-risk youths and identify linguistic targets for remediation and preventive intervention. More broadly, automated linguistic analysis can be a powerful tool for diagnosis and treatment across neuropsychiatry.

Key words: Automated language analysis, prediction of psychosis, semantic coherence, syntactic complexity, high-risk youths, machine learning

(World Psychiatry 2018;17:67–75)

Language offers a privileged view into the mind: it is the basis by which we infer others' thought processes, such that disorganized language is considered to reflect disorder in thought. Language disturbance is prevalent in schizophrenia and is related to functional disability, given that an individual needs to think and speak clearly in order to maintain friends and a job[1]. In schizophrenia, the speaker "violates the syntactical and semantic conventions which govern language usage", yielding reduction in syntactic complexity (concrete speech, poverty of content) and loss of semantic coherence, e.g. the disruption in flow of meaning in language (derailment, tangentiality)[2]. This language disturbance is an early core feature of schizophrenia, evident in subtle form prior to initial psychosis onset, in cohorts of both familial[3] and clinical[4-7] high-risk youths, as assessed using clinical ratings.

Beyond clinical ratings, there has been an effort to characterize early subtle language disturbances in clinical high-risk (CHR) individuals using linguistic analysis, with the aim of improving prediction. Bearden et al[8] applied manually coded linguistic analyses to brief speech transcripts in a CHR cohort, finding that both semantic features (illogical thinking) and reduction in syntactic complexity (poverty of speech) predicted psychosis onset with an accuracy of 71%, as compared with 35% accuracy for clinical ratings. Psychosis onset was also predicted by reduced referential cohesion, such that the use of pronouns and comparatives ("this" or "that") frequently did not clearly indicate who or what was previously described.

While this manual linguistic approach appears to be superior to clinical ratings in psychosis prediction, it depends on predefined measures that may not capture other subtle language features. Therefore, we have used automated natural language processing methods to analyze speech in CHR cohorts. These are probabilistic linguistic analyses based on the computer's acquisition of vocabulary (semantics) and learning of grammar (syntax) through machine-learning algorithms trained on very large bodies of text, enabled by exponential increases in computing power, and the flood of text that arrived with the Internet.

For semantics, a common approach is latent semantic analysis, in which a word's meaning is learned based on its co-occurrence with other words, inspired by theories of vocabulary acquisition[9,10]. In this analysis, each word is assigned a multi-dimensional semantic vector, such that the cosine between word-vectors represents the semantic similarity between words. Grouping of successive word-vectors can be used to estimate the semantic coherence of a narrative.

Latent semantic analysis has been applied to speech in schizophrenia, finding an association of decreased semantic coherence with clinical ratings of thought disorder and functional impairment, and with abnormal task-related activation in language circuits[11,12].

Table 1 Demographic features of the two samples

	UCLA site				NYC site	
	CHR+ (N=19)	CHR– (N=40)	CTR (N=21)	FEP (N=16)	CHR+ (N=5)	CHR– (N=29)
Age at baseline (years, mean±SD)	17.3 ± 3.7	16.4 ± 3.0	18.0 ± 2.8	15.8 ± 1.7[a]	22.2 ± 3.4	21.2 ± 3.6
Gender (% male)	89.5	55.0[b]	61.9[b]	68.7	80.0	65.5
Ethnicity (% Caucasian)	63.1	50.0	66.7	62.5	40.0	37.9
Parental socio-economic status (Hollingshead index, mean±SD)	4.4 ± 2.1[a]	4.4 ± 1.7[a]	5.7 ± 1.4	4.9 ± 1.8	NA	NA

Significant differences at $p<0.05$ level: [a]vs. CTR, [b]vs. CHR+

UCLA – University of California Los Angeles, NYC – New York City, CHR+ – clinical high-risk subjects who converted to psychosis during follow-up, CHR– – clinical high-risk subjects who did not convert to psychosis during follow-up, CTR – healthy controls, FEP – subjects with first-episode psychosis, NA – not available

For syntax, part-of-speech tagging is used to determine sentence length and rates of usage of different parts of speech[13,14].

In an earlier proof-of-principle study in a narrative-based protocol with a small CHR cohort, we used both latent semantic analysis and part-of-speech tagging, with machine learning, to identify a classifier of psychosis that comprised minimum semantic coherence, shortened sentence length, and a decrease in the use of determiner pronouns (e.g., "that" or "which") to introduce dependent clauses[15]. These three features were correlated with but outperformed clinical ratings in prediction of psychosis.

In the present study, we applied the same automated natural language processing approach with machine learning, including latent semantic analysis and part-of-speech tagging, to the larger CHR prompt-based protocol speech dataset that Bearden et al previously analyzed using manually coded linguistic methods[8].

We hypothesized that a classifier trained with the larger prompt-based protocol dataset[8] would be highly accurate (~80%) in predicting psychosis onset when tested intra-protocol as well as when retested in the narrative-based protocol[15] (cross-protocol). We also hypothesized that the automated and manual linguistic features derived from the training dataset would be correlated with one another.

We further tested the ability of the classifier to discriminate speech in adolescents with recent-onset psychosis from normal speech, as a putative early illness marker.

METHODS

Participants

Participants at the University of California Los Angeles (UCLA) site included 59 CHR individuals. They were defined by meeting criteria for one of three prodromal syndrome categories, as assessed by the Structured Interview for Prodromal Syndromes/Scale of Prodromal Symptoms (SIPS/SOPS)[16]: a) attenuated positive symptoms, b) brief intermittent psychotic symptoms, or c) a substantial drop in social/role functioning in conjunction with a schizotypal personality disorder diagnosis or a first-degree relative with a psychotic disorder. Of these subjects, 19 developed a psychotic disorder within two years ("converters", CHR+) and 40 did not (CHR–). Transition to psychosis was determined using the SIPS/SOPS "presence of psychosis" criteria. Transcripts from UCLA were also available for 16 recent-onset psychosis patients and 21 healthy individuals similar in demographics, recruited from local schools and the community.

Participants at the New York City (NYC) site included 34 CHR individuals, defined by meeting the above SIPS/SOPS criteria. Of these subjects, five developed psychosis within 2.5 years (CHR+) according to SIPS/SOPS criteria, and 29 did not (CHR–).

The demographic features of the two samples are presented in Table 1. The institutional review boards at New York State Psychiatric Institute/Columbia University and UCLA approved the study, and informed consent was obtained from all participants (parental consent with assent for minors).

Speech assay

UCLA (prompt-based protocol dataset)

Speech was elicited using Caplan's "Story Game", in which participants retell and then answer questions about a story they hear ("what do you like about it?"; "is it true?"), and then construct and tell a new story[17]. Speech samples were transcribed and de-identified, which means that proper nouns such as names were substituted.

Manual linguistic analyses included administration of the Kiddie Formal Thought Disorder Rating Scale (K-FTDS) and the Caplan modification of the Halliday and Hassan approach to analysis of cohesion[17]. The K-FTDS scores included frequency counts of illogical thinking, loose associations, and poverty of content. Cohesion categories included referential (pronomial, demonstrative and comparative – "this", "that"), conjunction

("and", "but", "because") and unclear/ambiguous[17]. This dataset was used to analyze intra-protocol prediction accuracy.

NYC (narrative-based protocol dataset)

Open-ended narrative interviews of about one hour were obtained by interviewers trained by an expert in qualitative research methods. Prompts queried impact of life changes experienced, and expectations for the future[18]. This dataset was used to study cross-protocol prediction accuracy.

Speech analyses

Speech pre-processing

The speech transcripts were pre-processed and prepared for computer-based analyses. We used the Natural Language Toolkit, which is an open source program available on the Internet (NLTK; http://www.nltk.org). First, punctuation (e.g., commas, periods) was discarded, words were tokenized (identified as parts of speech), and then each transcript was parsed into phrases, using rules of grammar in English. Words were then converted to the roots from which they are inflected, or lemmatized, using the NLTK WordNet lemmatizer.

The resulting pre-processed speech data yielded for each transcript a series of lemmatized words, maintaining the original order in which they were spoken, without punctuation and in lower case.

Latent semantic analysis

Latent semantic analysis[9,10] was used to convert each transcript from a series of words into a series of semantic vectors, maintaining the original order of the transcribed text. In this analysis, a high-dimensional semantic vector is assigned to each word in the lexicon based on its co-occurrence with other words in a very large corpus of text, specifically the Touchstone Applied Science Associates (TASA) corpus, a collection of educational materials.

Automated analysis provides a construction of meaning in language that resembles what the human mind does, i.e. to learn the meaning of words in terms of prior experience with those words in different contexts. The computer "learns" the meaning of words computationally, by scanning a very large corpus of text and determining the frequency of co-occurrence of each word with every other word in the lexicon. Words that co-occur more frequently are considered to have greater semantic similarity (e.g., "cat"/"dog" vs. "cat"/"pencil"), and the direction of their vectors will be more aligned. Aggregates of words (e.g., sentences) have semantic vectors that are the sum of semantic vectors for all the words they contain. Semantic coherence between words, or between aggregates (e.g., successive sentences), can be indexed by calculating the cosine between successive semantic vectors (from −1.0 for incoherence to 1.0 for coherence).

As the narrative-based protocol in NYC was open-ended, yielding mean uninterrupted responses of 130 words for CHR– and 182 words for CHR+, there had been sufficient free speech for analysis of semantic coherence at the sentence level in our prior study[15]. However, the prompt-based study at UCLA[8] led to much briefer responses (mean uninterrupted response <20 words; insufficient number of sentences for analysis), such that a k-level measure of semantic coherence was used instead, which computes word-to-word variability at "k" inter-word distances, with k ranging from 5 to 8[19]. As in our prior study[15], we calculated typical statistical measures for each of the k-level measures of coherence, such as mean, standard deviation, minimum, maximum, and 90th percentile (less sensitive to outliers than the maximum), also "normalized" or adjusted for sentence length.

Part-of-speech tagging analyses

Just as each word in every transcript was assigned a semantic vector, each word was also tagged in respect to its grammatical function, using the POS-Tag procedures in the open-access Natural Language Toolkit (www.nltk.org) in reference to a hand-tagged corpus called the Penn Treebank[13]. For example, the sentence "The dog is near the fence" would be tagged as ("The", "DT"), ("dog", "NN"), ("is", "VBZ"), ("near", "IN"), ("the", "DT"), ("fence", "NN"), where DT is the tag for determiners, NN for nouns, VBZ for verbs, and IN for prepositions.

The Penn Treebank has thirty-six part-of-speech tags, which include types of nouns, verbs, adjectives, adverbs, determiners, prepositions and pronouns. For each transcript, we calculated the frequency of use for each grammatical function.

Machine learning classification

The machine learning algorithm classifies speech by whether it is characteristic of individuals who will develop psychosis, as opposed to those who will not. It does this by learning the underlying patterns in a subset of transcripts and then in an iterative fashion, predicting the classification (psychosis or no) in new transcripts not used during the learning phase.

The machine learning analysis was circumscribed to the eleven speech variables that were significantly different between CHR+ and CHR– in the UCLA cohort (nine semantic coherence features and two syntactic elements – frequencies of comparative adjectives and possessive pronouns), plus three variables that predicted psychosis in our prior study[15], including WH-family ("which", "what", "whom") determiners, pronouns and adjectives. The list of these fourteen features used for analyses is provided in Table 2. Each transcript had a vector comprised of these fourteen variables.

We then performed singular value decomposition (which is a type of factor analysis based on linear algebra) on the fourteen features in these transcript vectors, adding the UCLA healthy control sample data to have a better understanding of the intrinsic structure of the speech data. We chose the top four factors that best discriminated transcripts from CHR+ vs. CHR–. A logistic regression model was then trained on the four

factors to classify CHR+ vs. CHR–, using an iteration of learning on a subset and prediction in left-out samples.

Cross-site validation

The same fourteen features were extracted from the NYC data, and aligned to the UCLA features using a simple global coordinate "Procrustean" transformation[20,21], similar to spatial registration in brain imaging[22], that includes scaling (in size), rotation and translation in Euclidean space. This minimized the difference in covariance of the two datasets, while maintaining the relative position among data points.

We further implemented a convex hull embedding method used in our prior study[15] to create a three-dimensional space (the top three factors) to model the accuracy of the classifier derived from the UCLA cohort in discriminating CHR+ from CHR– in the transformed NYC cohort. A convex hull of a set of points is the minimal convex polyhedron that contains them.

Correlations of text features with demographics, clinical ratings and manual features

We tested whether the fourteen identified text features were associated with age, gender, ethnicity (Caucasian/non-Caucasian) and parental socio-economic status[23]. We then assessed whether these text features were correlated with clinical ratings or with the three manually-coded linguistic measures (illogical thought, poverty of content and referential cohesion) that predicted psychosis onset in the UCLA cohort in the earlier study[8]. We calculated the canonical correlation between automated and manual text variables, which is the correlation between two sets of variables obtained from the same individuals.

Utility of the classifier in discriminating psychosis from normal speech

As an independent validation, we determined the accuracy of the CHR speech classifier in discriminating speech from the 21 healthy volunteers and 16 recent-onset psychosis patients ascertained at UCLA, who were also administered the same prompt-based protocol to elicit speech samples. The idea was that healthy controls should have a speech similar to that of CHR–, while recent-onset psychosis patients should have a speech similar to CHR+.

RESULTS

Machine learning classification

Of the four factors in the machine learning classifier, the first three highlighted semantic features, respectively weighted for maximum semantic coherence, variance in semantic coherence, and minimum semantic coherence, while the fourth factor was weighted for frequency of use of possessive pronouns (Figure 1).

Table 2 Syntactic and semantic features used for predictive modeling

Description	Example
a. Adjective, comparative	"braver", "closer", "cuter"
b. Possessive pronoun	"her", "his", "mine", "my", "our", "ours", "their", "your"
c. WH-determiner	"that", "which", "what"
d. WH-pronoun	"that", "what", "which", "who", "whom"
e. WH-adverb	"how", "however", "whenever", "why"
f. Minimum coherence at 5-level, normalized	
g. Minimum coherence at 5-level	
h. 90th percentile coherence at 5-level	
i. Maximum coherence at 6-level	
j. Mean coherence at 7-level, normalized	
k. Standard deviation coherence at 7-level, normalized	
l. 90th percentile at 7-level	
m. Standard deviation coherence at 7-level	
n. 90th percentile at 8-level	

A k-level measure of semantic coherence was used, which computes word-to-word variability at "k" inter-word distances, with k ranging from 5 to 8

The accuracy of the ensemble of these four factors in classifying psychosis outcome in the UCLA cohort was 83% using the logistic regression classifier. The post-hoc analysis yielded an area under the curve (AUC) of 0.87 in the receiver operating characteristic (ROC) curve (Figure 2).

So, a classifier comprising decreased semantic coherence, greater variance in that coherence, and reduced usage of possessive pronouns ("her", "his", "mine", "my", "our", "ours", "their", "your") was highly accurate in predicting subsequent psychosis onset.

Cross-site validation

When this UCLA machine-learning classifier was applied to the original NYC speech data, after Procrustean transformation[20,21,24], it significantly discriminated CHR with respect to psychosis onset ($p<0.05$ upon label randomization), with a true negative ratio of 0.82 (24/29) and a true positive ratio of 0.60 (3/5), that is, an overall accuracy of 0.79. With logistic regression, the UCLA classifier yielded an AUC of 0.72 for the transformed NYC cohort speech data (Figure 2).

In order to compare with our previous study[15], we created a three-dimensional projection of data using the top three factors

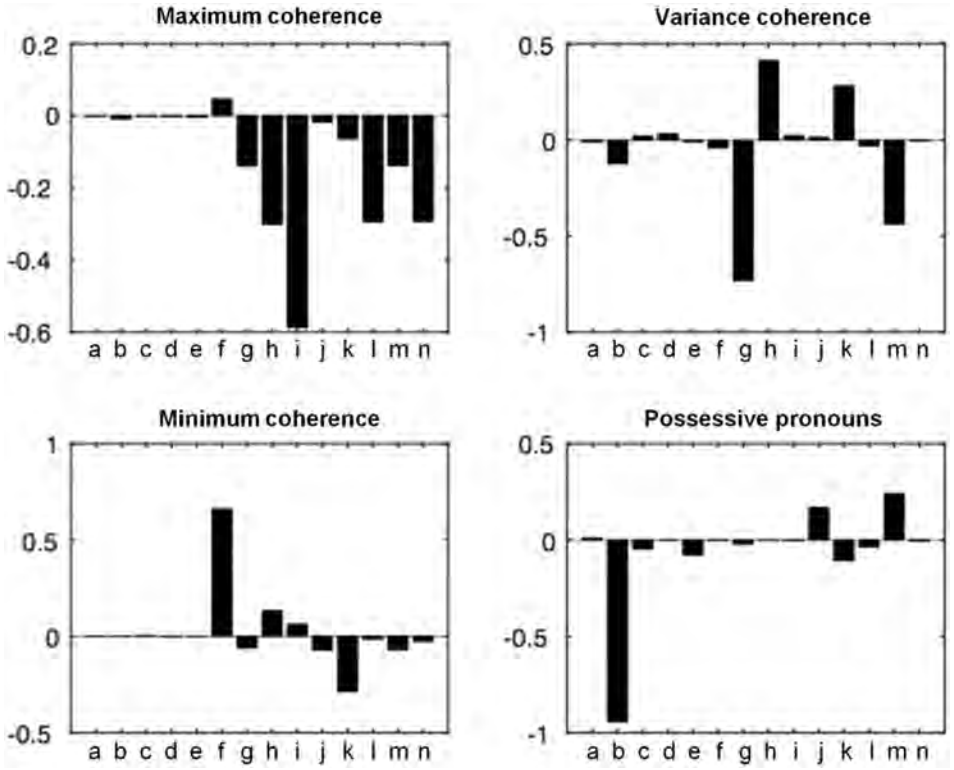

Figure 1 The four-factor University of California Los Angeles (UCLA) machine learning classifier of psychosis outcome. Factors are aggregates of weighted syntactic (a-e) and semantic coherence (f-n) features, as listed in Table 2. The first three factors are weighted toward semantic features (maximum, variance and minimum), and the fourth factor is weighted toward a syntactic feature (possessive pronouns). Y axes show factor weights.

identified from the UCLA CHR speech dataset. This yielded convex hulls that excluded 11 of 19 CHR+ in the UCLA cohort (i.e., 8/19 false negatives) (Figure 3A), indicating that the logistic regression classifier (with all four factors) was more accurate. Using the same three factors from the UCLA classifier, the convex hull of CHR– in NYC excluded three of five CHR+ (Figure 3B). Of note, there was substantial overlap in the convex hulls of CHR– individuals for both the UCLA and NYC speech datasets (Figure 3C).

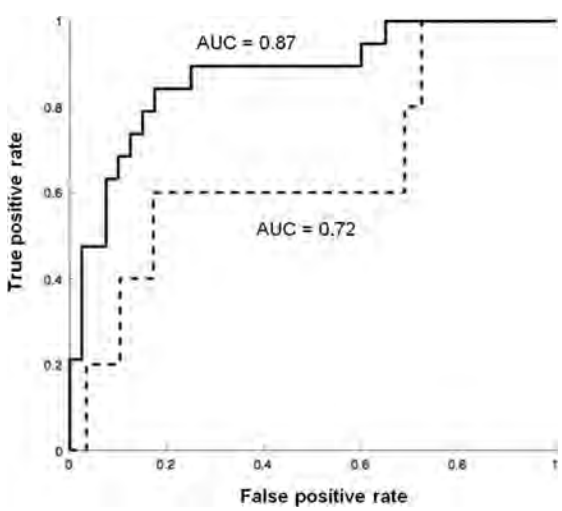

Figure 2 Receiver operating characteristics (ROC) for the University of California Los Angeles (UCLA) clinical high-risk (CHR) classifier of psychosis outcome as applied to the UCLA dataset (solid line) and to the realigned New York City (NYC) dataset (dotted line). AUC – area under the curve.

Correlations with demographics, clinical ratings and manual linguistic features

Among demographic features, age was significantly associated with three of the semantic coherence variables, specifically the 90% order variables for 5-level ($p=0.002$), 7-level ($p=0.01$) and 8-level ($p=0.004$), suggesting increasing semantic coherence with age. By contrast, there were no associations of automated text variables with gender, ethnicity, or parental socioeconomic status[23].

There was no significant association between automated analysis text features and SIPS/SOPS clinical ratings (total positive and total negative). However, the canonical correlation between the fourteen text features identified here, and the three manual linguistic features (illogical thought content, poverty of content and referential cohesion) that predicted psychosis onset in the earlier study[8], was large and highly significant, with $r=0.71$, $p<10^{-6}$.

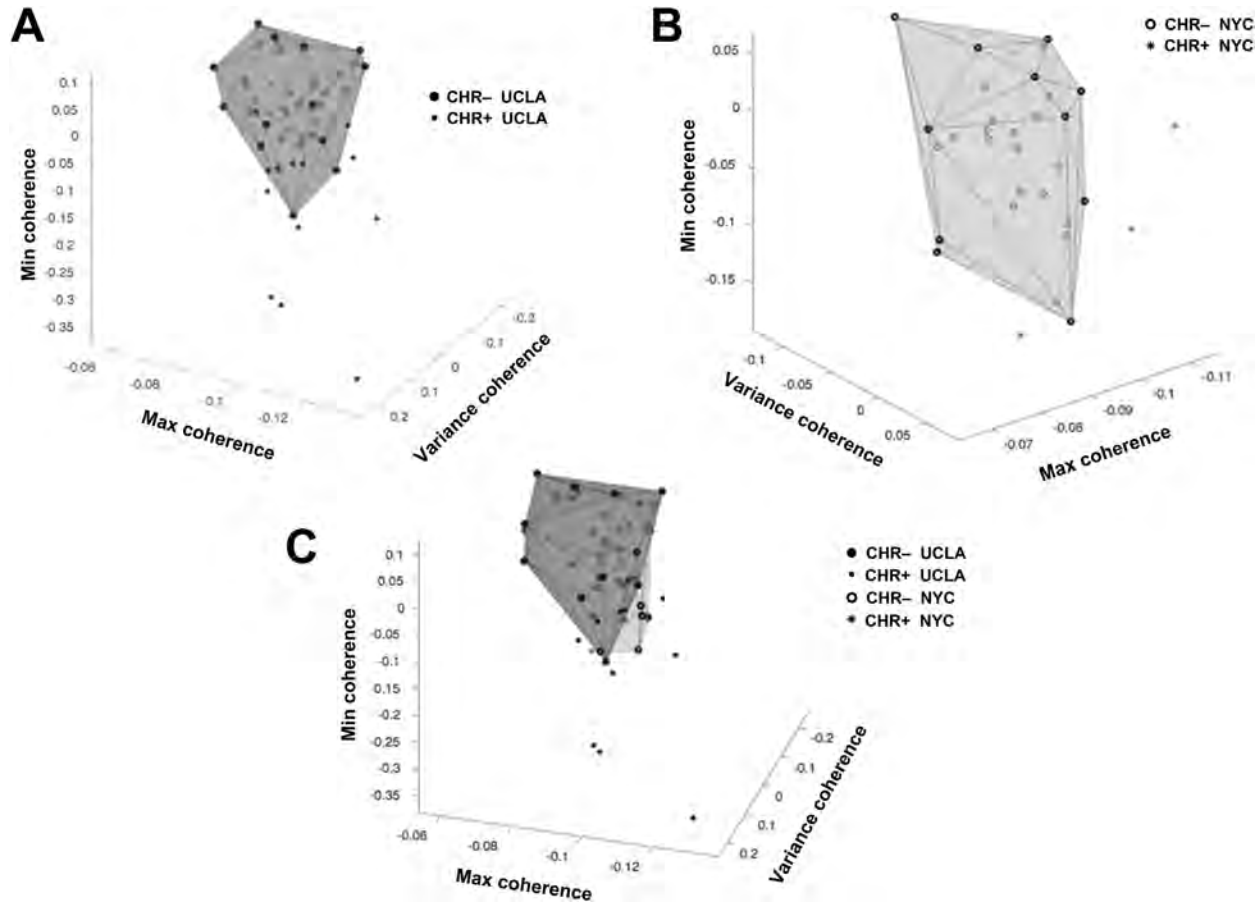

Figure 3 Projection of the top three factors for the University of California Los Angeles (UCLA) and New York City (NYC) clinical high-risk (CHR) cohorts. These factors were weighted for semantic coherence features. A. Convex hull of non-converters (CHR–) in UCLA, with 11 of 19 converters (CHR+) outside of the hull. B. Convex hull of CHR– in NYC, with 3 of 5 CHR+ outside the hull. C. Data in A and B (all CHR) shown together to demonstrate extent of overlap in language properties.

Utility of the classifier in discriminating psychosis from normal speech

A 72% accuracy was obtained with the logistic regression classifier when applied to the speech dataset of healthy controls and recent-onset psychosis patients at UCLA.

Singular value decomposition three-factor representation excluded 11 of 16 recent-onset psychosis patients from the convex hull defined by the data points of healthy volunteers, yielding a true positive rate of 0.69 (Figure 4A). There was spatial overlap between the convex hulls that contained healthy controls and CHR– individuals (Figure 4B).

DISCUSSION

Using automated natural language processing methods with machine learning to analyze speech in a CHR cohort, we generated a classifier comprising decreased semantic coherence, greater variance in that coherence, and reduced usage of possessive pronouns which was highly accurate in predicting subsequent psychosis onset.

This classifier had an intra-protocol accuracy of 83% in the training dataset, and a cross-protocol accuracy of 79% when applied to transcripts from a second independent CHR cohort (test dataset)[15], demonstrating significant transfer of predictability, despite disparate methods of speech elicitation[8,15]. Further, this same classifier discriminated the speech of recent-onset psychosis patients from that of healthy individuals with 72% accuracy, suggesting that its discriminatory power was relatively robust across illness stages, as has been found for clinical ratings of thought disorder[1,6]. Finally, the predictive automated and manual linguistic features were highly correlated in the cohort, providing evidence of concurrent validity.

It has long been observed that language in schizophrenia is characterized by a disturbance in semantic coherence, with Kraepelin describing *Sprachverwirrtheit* (e.g., confused speech)[25], and Bleuler highlighting a "loosening of associations" in language as a primary feature of schizophrenia[26]. Later, Andreasen operationalized decreased semantic coherence as positive thought dis-

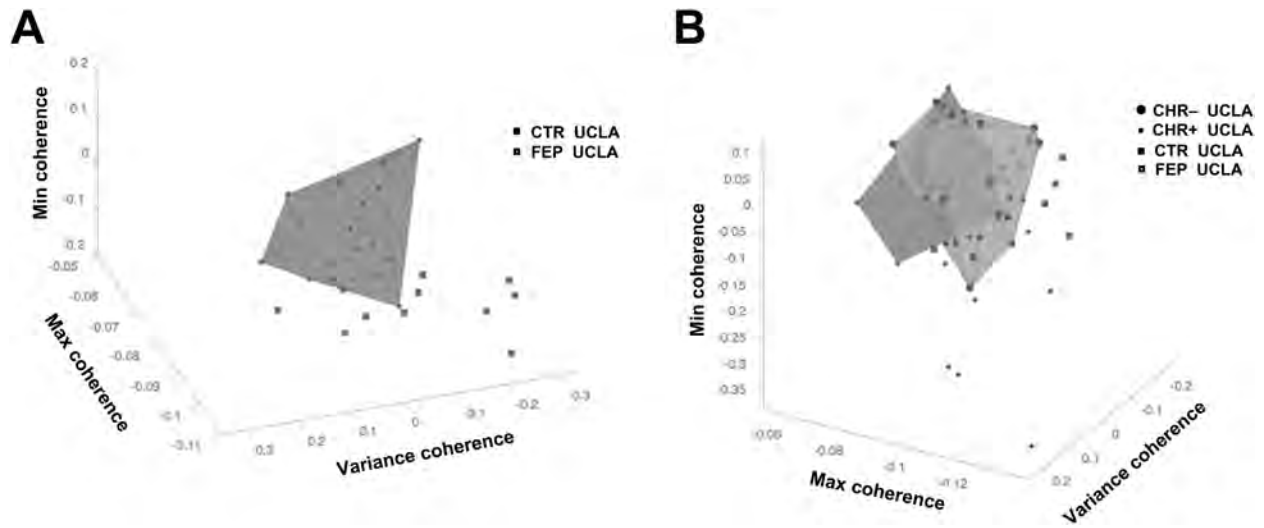

Figure 4 Projection of the top three factors for University of California Los Angeles (UCLA) first-episode psychosis (FEP) patients and healthy controls (CTR). A. Convex hull of healthy controls (CTR) with 11 of 16 FEP patients outside the hull. B. Overlap of convex hulls for FEP vs. CTR, and converters (CHR+) vs. non-converters (CHR–).

order[27]. Hoffman applied manual discourse analysis to transcribed speech from schizophrenia patients, finding a reduction in semantic coherence[28], a finding replicated later using computer-assisted discourse analysis[29].

It has only been in the last decade that natural language processing linguistic corpus-based analyses, specifically latent semantic analysis, have been applied to language production in schizophrenia, finding decreases in semantic coherence that correlate with clinical ratings, functional impairment, and task-related activation in language circuits[11,12]. Now, in the two CHR studies to date, latent semantic analysis with machine learning has shown decreased semantic coherence to predict subsequent psychosis onset.

Disturbance in syntax is also well-documented in schizophrenia. Errors of pronomial reference in schizophrenia speech were described three decades ago by Hoffman[30], a finding since replicated by other investigators using word classification/count strategies[29,31]. In the present study, using part-of-speech tagging, we identified decreased use of possessive pronouns as prognostic for psychosis onset, accounting for most of the weight of the fourth factor in the classifier. This is consistent with prior manual linguistic analysis in this same cohort, which identified decreased referential cohesion as predictive of psychosis[8], such that the use of pronouns and comparatives ("this" or "that") frequently did not clearly indicate who or what was previously described.

More commonly found in schizophrenia speech is a reduction in syntactic complexity[27,32], typically operationalized as shorter sentences, and most evident when open-ended narrative is elicited[12,30,31,33]. In our prior small natural language processing study[15], we found two measures of syntactic complexity – shorter sentences and reduced use of determiner pronouns that introduce dependent clauses – to be both predictive of psychosis and highly correlated with negative symptoms. In the present study, the failure of sentence length to predict psychosis in the training dataset may be a consequence of the brief and structured responses that were elicited (<20 mean words per response)[12], as compared with prior studies (>120 mean words/response[15], ~800 words/response[12] and >10 sentences/response[30]).

In both of our CHR studies, we have created convex hull classifications in which speech datapoints for non-converters (CHR–) were inside the hull, while those with emergent psychosis (CHR+) were outside. A similar convex hull was generated for healthy controls using the CHR classifier, with recent-onset psychosis patients largely outside the hull. Together, these findings suggest that pre-psychotic and psychotic language is deviant from a constrained hull of relatively normal language in respect to semantics and syntax.

As yet, this normal pattern of language, as characterized by automated natural language processing methods, remains poorly understood, including in a developmental context, as both semantic and syntactic complexity increase in adolescence and young adulthood[34]. Of note, the premise that processes underlying normal language production and comprehension are relatively homogeneous is supported by a body of work by Hasson, showing alignment of brain activation time courses across normal individuals (intersubject coherence) during both listening and speaking[35].

Our finding of strong correlations between automated and manual linguistic variables provides evidence of concurrent validity for the natural language processing approach. Automated natural language processing methods are far more rapid and less expensive than manual linguistic approaches, and

can be more readily adapted for research and ultimately in the clinic.

Beyond language semantic analysis and part-of-speech tagging, speech and language can also be evaluated in respect to speech graphs[36], prosody, pragmatics, metaphoricity[37], and for discourse or conversations among interlocutors. Automated natural language processing analyses have also been used to characterize other disturbances in behavior, including intoxication from drugs of abuse[38] and Parkinson's disease[39], such that this technology holds promise for medicine more broadly. Finally, automated approaches can be extended to other behavior, such as facial expressions of emotion[40]. Overall, automated speech analysis is a powerful but inexpensive technology that can be used in psychiatry for diagnosis, prognosis and estimates of treatment response.

The main limitations in the present study include sample size, and remaining gaps in our knowledge in respect to what is normal across development for automated linguistic variables, and how normal and deviant language can be mapped to underlying neural circuits. Further, different methods of speech elicitation were used in the two cohorts, such that sentence-level coherence could not be estimated for the training dataset due to brevity of responses, requiring the use of "k-level" methods to characterize semantic coherence, and an alignment transformation of data for cross-protocol validation. In ongoing studies, we are using open-ended interviews to elicit free natural speech for analysis, so that we can measure semantic coherence at the sentence level, and better capture measures of syntactic complexity.

Overall, we demonstrate the utility and validity of using automated natural language processing methods to characterize subtle disturbances in semantics and syntax across stages of psychotic disorder. This technology has the potential to improve prediction of psychosis outcome among adolescents and young adults at clinical high risk, and may have broader implications for medical research and practice at large.

ACKNOWLEDGEMENTS

This research was supported by the US National Institute of Mental Health (R01 MH 107558; R03 MH 108933 02), the New York State Office of Mental Health, and a NARSAD/BBRF Young Investigator Award and Miller Family Term Chair to C.E. Bearden. These funding sources had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and the decision to submit the manuscript for publication.

REFERENCES

1. Roche E, Creed L, MacMahon D et al. The epidemiology and associated phenomenology of formal thought disorder: a systematic review. Schizophr Bull 2015;41:951-62.
2. Andreasen NC, Grove WM. Thought, language, and communication in schizophrenia: diagnosis and prognosis. Schizophr Bull 1986;12:348-59.
3. Gooding DC, Ott SL, Roberts SA et al. Thought disorder in mid-childhood as a predictor of adulthood diagnostic outcome: findings from the New York High-Risk Project. Psychol Med 2013;43:1003-12.
4. 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.
5. Addington J, Liu L, Buchy L et al. North American Prodrome Longitudinal Study (NAPLS 2): the prodromal symptoms. J Nerv Ment Dis 2015;203:328-35.
6. DeVylder JE, Muchomba FM, Gill KE et al. Symptom trajectories and psychosis onset in a clinical high-risk cohort: the relevance of subthreshold thought disorder. Schizophr Res 2014;159:278-83.
7. Cornblatt BA, Carrion 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.
8. Bearden CE, Wu KN, Caplan R et al. Thought disorder and communication deviance as predictors of outcome in youth at clinical high risk for psychosis. J Am Acad Child Adolesc Psychiatry 2011;50:669-80.
9. Landauer TK, Dumais ST. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol Rev 1997;104:211-40.
10. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. Discourse Process 1998;25:259-84.
11. Elvevag B, Foltz PW, Weinberger DR et al. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. Schizophr Res 2007;93:304-16.
12. Elvevag B, Foltz PW, Rosenstein M et al. An automated method to analyze language use in patients with schizophrenia and their first-degree relatives. J Neurolinguistics 2010;23:270-84.
13. Santorini B. Part-of-speech tagging guidelines for the Penn Treebank Project. 3rd revision. Philadelphia: Department of Computer and Information Science, University of Pennsylvania, 1990.
14. Bird S. Natural language processing and linguistic fieldwork. Comput Linguist 2009;35:469-74.
15. Bedi G, Carrillo F, Cecchi GA et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. NPJ Schizophr 2015;1:15030.
16. Miller TJ, McGlashan TH, Rosen JL et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. Schizophr Bull 2003;29:703-15.
17. Caplan R, Guthrie D, Fish B et al. The Kiddie Formal Thought Disorder Rating Scale: clinical assessment, reliability, and validity. J Am Acad Child Adolesc Psychiatry 1989;28:408-16.
18. Ben-David S, Birnbaum ML, Eilenberg ME et al. The subjective experience of youths at clinically high risk of psychosis: a qualitative study. Psychiatr Serv 2014;65:1499-50.
19. Mandera P, Keuleers E, Brysbaert M. How useful are corpus-based methods for extrapolating psycholinguistic variables? Q J Exp Psychol 2015;68:1623-42.
20. Shönemann P. A generalized solution of the orthogonal procrustes problem. Psychometrika 1966;31:1-10.
21. Haxby JV, Guntupalli JS, Connolly AC et al. A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron 2011;72:404-16.
22. Ashburner J, Firston K. Rigid body registration. In: Penny W, Friston K, Ashburner J et al (eds). Statistical parametric mapping: the analysis of functional brain images. Cambridge: Academic Press, 2007:49-62.
23. Mollica RF, Milic M. Social class and psychiatric practice: a revision of the Hollingshead and Redlich model. Am J Psychiatry 1986;143:12-7.
24. Jorge-Botana G, Olmos R, Luzon JM. Word maturity indices with latent semantic analysis: why, when, and where is Procrustes rotation applied? Wiley Interdiscip Rev Cogn Sci (in press).
25. Kraepelin E. Psychiatrie. Ein Lehrbuch für Studierende und Ärzte. Leipzig: Barth, 1899.
26. Bleuler E. Dementia Praecox oder Gruppe der Schizophrenien. Leipzig: Deuticke, 1911.
27. Andreasen NC. Thought, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. Arch Gen Psychiatry 1979;36:1315-21.
28. Hoffman RE, Stopek S, Andreasen NC. A comparative study of manic vs schizophrenic speech disorganization. Arch Gen Psychiatry 1986;43:831-8.
29. Noel-Jorand MC, Reinert M, Giudicelli S et al. A new approach to discourse analysis in psychiatry, applied to a schizophrenic patient's speech. Schizophr Res 1997;25:183-98.
30. Hoffman RE, Hogben GL, Smith H et al. Message disruptions during syntactic processing in schizophrenia. J Commun Disord 1985;18:183-202.

31. Buck B, Penn DL. Lexical characteristics of emotional narratives in schizophrenia: relationships with symptoms, functioning, and social cognition. J Nerv Ment Dis 2015;203:702-8.
32. Kuperberg GR. Language in schizophrenia Part 2: What can psycholinguistics bring to the study of schizophrenia... and vice versa? Lang Linguist Compass 2010;4:590-604.
33. Andreasen NC. Thought, language, and communication disorders. II. Diagnostic significance. Arch Gen Psychiatry 1979;36:1325-30.
34. Nippold MA, Ward-Lonergan JM, Fanning JL. Persuasive writing in children, adolescents, and adults: a study of syntactic, semantic, and pragmatic development. Lang Speech Hear Serv Sch 2005;36:125-38.
35. Silbert LJ, Honey CJ, Simony E et al. Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. Proc Natl Acad Sci USA 2014;111:E4687-96.
36. Mota NB, Vasconcelos NA, Lemos N et al. Speech graphs provide a quantitative measure of thought disorder in psychosis. PLoS One 2012;7:e34928.
37. Gutierrez ED, Shuotva E, Marghetis T et al. Literal and metaphorical senses in compositional distributional semantic models. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, 2016:183-93.
38. Bedi G, Cecchi GA, Slezak DF et al. A window into the intoxicated mind? Speech as an index of psychoactive drug effects. Neuropsychopharmacology 2014;39:2340-8.
39. Garcia AM, Carrillo F, Orozco-Arroyave JR et al. How language flows when movements don't: an automated analysis of spontaneous discourse in Parkinson's disease. Brain Lang 2016;162:19-28.
40. Baker JT, Pennant L, Baltrušaitis T et al. Toward expert systems in mental health assessment: a computational approach to the face and voice in dyadic patient-doctor interactions. iproc 2016;2:e44.

DOI:10.1002/wps.20491

Income inequality and depression: a systematic review and meta-analysis of the association and a scoping review of mechanisms

Vikram Patel[1], Jonathan K. Burns[2], Monisha Dhingra[3], Leslie Tarver[4], Brandon A. Kohrt[5], Crick Lund[6,7]

[1]Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA; [2]Institute of Health Research, University of Exeter, Exeter, UK; [3]Ashoka University, Sonepat, Rai, Haryana, India; [4]Department of Psychiatry, Massachusetts General Hospital and McLean Hospital, Harvard Medical School, Boston, MA, USA; [5]Department of Psychiatry and Behavioral Sciences, George Washington University, Washington, DC, USA; [6]Health Service and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [7]Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa

Most countries have witnessed a dramatic increase of income inequality in the past three decades. This paper addresses the question of whether income inequality is associated with the population prevalence of depression and, if so, the potential mechanisms and pathways which may explain this association. Our systematic review included 26 studies, mostly from high-income countries. Nearly two-thirds of all studies and five out of six longitudinal studies reported a statistically significant positive relationship between income inequality and risk of depression; only one study reported a statistically significant negative relationship. Twelve studies were included in a meta-analysis with dichotomized inequality groupings. The pooled risk ratio was 1.19 (95% CI: 1.07-1.31), demonstrating greater risk of depression in populations with higher income inequality relative to populations with lower inequality. Multiple studies reported subgroup effects, including greater impacts of income inequality among women and low-income populations. We propose an ecological framework, with mechanisms operating at the national level (the neo-material hypothesis), neighbourhood level (the social capital and the social comparison hypotheses) and individual level (psychological stress and social defeat hypotheses) to explain this association. We conclude that policy makers should actively promote actions to reduce income inequality, such as progressive taxation policies and a basic universal income. Mental health professionals should champion such policies, as well as promote the delivery of interventions which target the pathways and proximal determinants, such as building life skills in adolescents and provision of psychological therapies and packages of care with demonstrated effectiveness for settings of poverty and high income inequality.

Key words: Income inequality, depression, neo-material hypothesis, social capital, social comparison, psychological stress, social defeat, low-income populations

(World Psychiatry 2018;17:76–89)

The unequal distribution of income and wealth has been growing steadily over the past three decades to astonishing levels, fuelled by the wide adoption of neo-liberal economic policies and globalization. In 2016, while the bottom half of the global population collectively owned less than one percent of total wealth, the wealthiest top 10 percent owned 89 percent of all global assets[1].

The growth of income and wealth inequality has been observed in countries at all levels of socio-economic development. In the US, one of the richest countries in the world, the top 10 percent of the population now average nearly nine times as much income as the bottom 90 percent[2]. In India, an exemplar of a low- or middle-income country (LMIC), the richest 1% owned nearly 60% of the total wealth of the country in 2016.

However, there is a three-fold variation in the range of levels of inequality among countries, with the most equal countries mostly clustered in Western Europe and the most unequal countries comprising LMIC and the US. These variations at the country level, as well as at sub-national levels (i.e., provinces or states) allow the exploration of the association between income inequality and a variety of social outcomes, notably health.

There is a robust body of evidence linking inequality and health outcomes, ranging from infant mortality and life expectancy to obesity. A compelling case for a causal relationship between inequality and a number of negative health outcomes has been recently presented[3].

Not surprisingly, there is also evidence linking income inequality with mental health outcomes. A significant positive relationship has been reported between the incidence rate of schizophrenia and country-level Gini coefficient, a widely used measure of the distribution of income or wealth in a population. A possible mechanism proposed for this association was that inequality impacts negatively on social cohesion and capital, and increases chronic stress, placing individuals at a heightened risk of schizophrenia[4].

A review of studies on the association of income inequality and a range of mental health related outcomes reported heterogeneous findings, with about one third of studies observing a positive association between income inequality and the prevalence or incidence of mental health problems, one third observing mixed results for different subgroups, and one third observing no association[5]. Depression was one of the mental health outcomes considered in studies showing a positive association with income inequality.

Although potential mechanisms that underlie the observed association between income inequality and health have been proposed[6], little is known about the mechanisms involved in the case of depression. Hence, there is a need for a systematic review focusing on this association which also sets out to identify potential mechanisms and develops a conceptual framework that can further our understanding and set an agenda for future research in the field.

The present study sought to advance the scientific inquiry of the association between income inequality and mental health in three specific ways. First, we systematically identified and descriptively synthesized the most updated literature on depression and income inequality, with a focus on study characteristics and potential differential impact by gender and level

of poverty. Second, we quantitatively assessed the strength of the association of income inequality and depression prevalence through a meta-analysis. Finally, we conducted a scoping review of the literature to explore the potential mechanisms, and developed a theoretical framework for this association.

By focusing on one mental health outcome (depression), we hoped to provide a more in-depth analysis of potential mechanisms than has previously been possible. Our ultimate goal was to elaborate the implications of this body of evidence on policies which influence the distribution of income and wealth, and identify specific gaps in our knowledge which deserve further research investment.

METHODS

Systematic review and meta-analysis

Search strategy

The search strategy was guided by our protocol (PROSPERO registration: CRD42017072721), which is available on request. In brief, we searched PubMed/Medline, EBSCO and PsycINFO databases. The search string used was "(depress* OR mental) AND (inequal* OR Gini)". The electronic databases were searched for titles or abstracts containing these terms in all published articles between January 1, 1990 and July 31, 2017. The search was limited to studies published in English and involving human subjects. The reference lists of all included studies were hand-searched for additional relevant reports or key terms. If new key terms were identified (new term included: "mood"), additional searches of the above databases were conducted and relevant papers were added until no further publications were found.

We included all studies providing primary quantitative data with a measure of depression or depressive symptoms as an outcome and any measure of income inequality at any geographical scale. Exclusion criteria were: unpublished data of any form including conference proceedings, case reports, dissertations; qualitative studies; and publications reporting duplicate data from the same population (in such cases, the report with the larger sample size was included).

All titles and abstracts identified in the search were screened to exclude those that were obviously irrelevant based on the above exclusion criteria. Full-text versions were obtained for all abstracts remaining after screening. Obtained full-text articles were read and those not satisfying inclusion criteria were subsequently removed. The remaining articles were included in the systematic review (see Figure 1).

Analyses

Study data were extracted onto a customized sheet. Quality assessments were independently performed by using the Systematic Appraisal of Quality in Observational Research (SAQOR) tool, that comprises six domains (each containing two to five questions): sample, control/comparison group, exposure/outcome measurements, follow-up, confounders, and reporting of data[7].

The SAQOR has been adapted for use in cross-cultural psychiatric epidemiology studies[8]. In the current study, two domains were omitted (control/comparison group and follow-up), as they were not applicable to any of the papers identified. A summary quality assessment was made by a single rater (JKB) for each of the four domains, and then an overall summary grade was determined based on adequacy in the four domains. The overall quality of each study was graded as high, moderate or low.

The meta-analysis was conducted using Cochrane Review Manager (RevMan) version 5.3[9]. Data were extracted from studies to calculate risk ratios for the association of income inequality and depression. Studies that included data on depression event rates stratified by income inequality were included in the meta-analysis.

Compared to risk ratios, odds ratios exaggerate effect sizes, with the distortion most pronounced for outcomes with prevalence greater than 10%[10,11]. Because depression prevalence may be greater than 10% in a population, risk is more accurately estimated with risk ratios. To calculate risk ratios, income inequality for each study was categorized as binary outcomes (higher vs. lower income inequality in a given population).

When income inequality was categorized into three or more groups, we re-categorized the groups as follows: in studies with three groups, the lowest income inequality group was used as the reference and the medium and high inequality groups were collapsed; studies with four inequality groups were re-categorized grouping the two lower and two higher inequality strata; for quintiles, the two low inequality quintiles were pooled as the reference group to compare with the three high inequality quintiles, which were collapsed into one stratum; finally, in studies with more than five inequality groups, these were re-categorized into two strata of roughly equal sample size.

Unadjusted prevalence rates of depression were used whenever available, given the lack of consistency across studies in variables used for adjusting outcomes. When unadjusted prevalence rates were not available, demographics and other characteristics used for adjustment are reported. A random effects meta-analysis was conducted because of the heterogeneity in design, populations and outcome measures.

We conducted sensitivity analyses using the leave-one-out approach to test the impact of excluding single studies contributing a disproportionately large effect. A forest plot of the risk ratios with summary statistics (pooled effect sizes) was completed using RevMan. Heterogeneity among trials was calculated using the I^2 measure of inconsistency.

Scoping review of mechanisms

We searched the Introduction and Discussion sections of studies included in the systematic review, to identify authors' hypothesized mechanisms of the relationship between inequal-

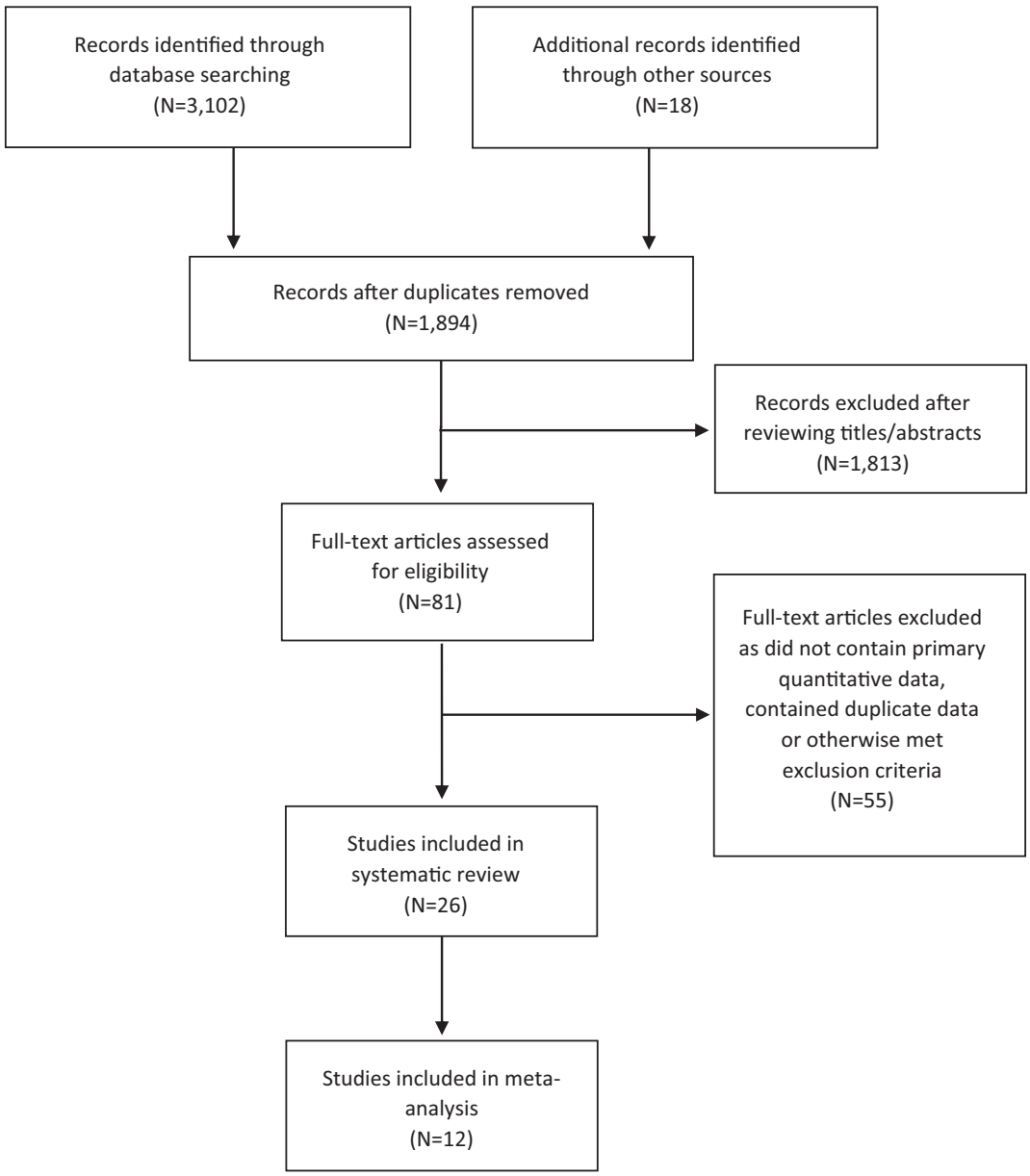

Figure 1 PRISMA flow diagram

ity and depression. We subsequently considered hypothesized mechanisms in a recent review regarding the causal links between inequality and health[3]. We then compiled a list of hypothesized mechanisms based on their plausibility, specifically the extent to which the purported mechanisms were supported by the data reported in the included studies. We sought to improve the overall coherence by grouping different hypothesized mechanisms into conceptual categories.

Finally, we supplemented these findings with an analysis of the variability in findings from the studies included in our systematic review. This led to consideration of a number of other factors that might inform the hypothesized mechanisms, namely the geographical unit of analysis, level of national development of the study country, effects of income inequality on low- vs. high-income groups, cultural variations across countries, broader political and historical context, life course or developmental stage factors, as well as gender and methodological considerations.

RESULTS

Systematic review and meta-analysis

Searches of the listed databases using the search string as well as hand-searching reference lists identified 1,894 potential

articles. After screening titles or abstracts, 1,813 were removed as they were irrelevant or clearly did not contain primary data. Full-text reports were retrieved for 81 articles, which were assessed for eligibility against inclusion and exclusion criteria. Of these, 55 were removed as they did not provide primary quantitative data on the relationship between income inequality and depression, contained duplicate data or otherwise met exclusion criteria. Thus, 26 studies were included in the systematic review (Table 1). The selection process for included studies is illustrated in Figure 1.

Study characteristics

The majority (N=18) of the 26 studies testing associations between income inequality and depression came from high-income countries, with 15 studies reported from the US. In terms of geographical scale, four were conducted at the country level, 14 at regional level (state, county, district, municipality), and eight at local area or neighbourhood level.

A number of studies were conducted in specific populations: five in older persons only[19,22,27,30,35]; four in adolescents only[7,23,26,28]; one in students aged 17–30 years[25]; and one in low-income nursing assistants[21].

The most common measure of depression, used in ten studies, was the Centre for Epidemiologic Studies Depression Scale (CES-D), while four studies used the Composite International Diagnostic Interview (CIDI), two the Patient Health Questionnaire (PHQ), and two the Alcohol Use Disorder and Associated Disabilities Interview Schedule - IV (AUDADIS-IV). Each of the remaining eight studies utilized a different instrument. Godoy et al[17] investigated 655 adults in villages within the Bolivian Amazon and found a positive relationship between village-level Gini coefficient and experiences of "sadness" over the last seven days.

Income inequality was most commonly measured using the Gini coefficient (21 studies), with the remainder using a ratio measure (e.g., 20%:20% ratio; P90/P10 ratio). Notably, all country-level studies utilized the Gini coefficient, while ratio measures were more commonly used in local area-level studies (three out of eight) than in regional-level studies (two out of 14).

Associations between income inequality and depression

Nearly two-thirds (N=16, 61.5%) of studies found a significant positive relationship between income inequality and risk of depression, while another three (11.5%) reported a positive relationship that was significant in bivariate but not multivariate regression analysis. Six studies (23.1%) found no significant relationship, while only one (3.8%) reported a negative relationship between income inequality and risk of depression (Table 1).

Nineteen studies did not stratify their analysis by absolute income. Out of the seven studies that stratified analyses by absolute income, two showed a significant effect of income inequality on depression only in low-income participants, and two demonstrated that the effect size was the strongest in low-income individuals. Studies documenting greater effects in low-income participants were conducted at either the regional (N=2) or the local level (N=2). The three studies reporting no absolute income effect were conducted at either the regional (N=2) or the country level (N=1).

Five studies stratified their analyses by gender. Of these, three found an association between income inequality and depression in females only[18,23,24], one detected no gender effect[4], and one found an association in men only in the bivariate analysis[29].

Although none of the studies stratified the analyses by age group, several were conducted exclusively in adolescent or older adult populations, and some interesting observations can be made here. Of the four studies in adolescents only, three found a significant association between income inequality and depression (two in the regression[23,26] and one in bivariate analysis only[28]). Of the five studies in older adults only, three found an association between income inequality and depression (two in regression[19,22] and one in bivariate analysis only[27]), one found no association[30], and one found a negative association[35].

Participant ethnicity was reported in eight studies, with only three conducted in specific ethnic populations: in nearly 9,000 Hispanic adults aged 60 and older in Mexico[30]; in nearly 6,500 Black and Hispanic adolescents in the US[7]; and in Tsimane villagers in the Bolivian Amazon[17]. Notably, two of these latter studies did not find a relationship between income inequality and depression. Of the five studies that stratified their analysis by ethnicity, only one found an ethnicity effect, with the relationship between income inequality and depression most pronounced in middle-class Blacks in a population representative panel in South Africa[4].

Of the 26 studies, only six were longitudinal, allowing for temporal analyses. Of these, five reported a significant positive relationship between income inequality and depression [4,16,17,21,24]; and one reported no association[31]. All except two studies had large sample sizes of over 1,000 participants (ranging from 1,355[12] to 293,405[15]).

Meta-analysis

Twelve studies were included in the meta-analysis, based on availability of event rates of depression to calculate risk ratios. Quality ratings of the included studies using SAQOR ranged from high to moderate (see Table 2). The pool of studies included six US studies, three multi-country studies, one UK study, one Brazil study, and one South Africa study. Two of the US studies were limited to older adults. One study only included women[18]. One multi-country study limited the sample to university students[25].

Four studies employed three strata of inequality[13,14,18,35]; one study employed four strata[12] and two studies employed

Table 1 Papers included in the systematic review of the association between income inequality and depression

Study	Sample	Study design	Country	Geographical unit of analysis	Inequality measure	Inequality range (Gini)	Depression measure	Absolute income effect	Gender effect
Positive association (higher risk of depression in populations with higher income inequality)									
Ahern & Galea[12]	1,355 adults aged 18 and over	Cross-sectional	US	Local (neighbourhood)	Gini	0.37-0.51	BSI-D	Low income only	NA
Burns et al[4]	25,936 adults aged 15 and over	Longitudinal panel	South Africa	District	P90/P10 ratio	0.46-0.68	CES-D	Low income only	No effect
Chiavegatto Filho et al[13]	3,542 adults aged 18 and over	Cross-sectional	Brazil	Municipality	Gini	0.18-0.34 (means for 1st and 3rd tertiles)	CIDI	NA	NA
Cifuentes at al[14]	251,158 adults	Cross-sectional	65 countries	Country level	Gini	0.25-0.74	DSM-IV and DIPS	High HDI countries only	NA
Fan et al[15]	293,405 adults aged 18 and over	Cross-sectional	US	State	Gini	0.40-0.54	PHQ	NA	NA
Fiscella & Franks[16]	6,913 adults aged 25-74	Longitudinal	US	Local	Aggregate income earned by the poorer 50% of population divided by total aggregate income in the community	0.18-0.37 (ratio range)	GWB-D	NA	NA
Godoy et al[17]	655 adults aged 16 and over	Longitudinal panel	Bolivia	Local (village)	Gini	0.71 ± 0.08 (mean±SD)	Feelings of "sadness" during past week	NA	NA
Kahn et al[18]	8,060 women with children aged 26-48 months	Cross-sectional	US	State	Gini	0.415-0.430 (cut-offs for 1st and 3rd tertiles)	CES-D	Most pronounced in low-income participants	Included only women

Table 1 Papers included in the systematic review of the association between income inequality and depression *(continued)*

Study	Sample	Study design	Country	Geographical unit of analysis	Inequality measure	Inequality range (Gini)	Depression measure	Absolute income effect	Gender effect
Ladin et al[19]	22,777 adults aged 55 and over	Cross-sectional	10 European countries	Country level	Gini	0.25-0.36	Euro-D Scale	NA	NA
Messias[20]	235,067 adults	Cross-sectional	US	State	Gini	0.410-0.495	PHQ	NA	NA
Muntaner et al[21]	241 low-income nursing assistants	Longitudinal	US	County	Gini	0.31-0.48	RCES-D	NA	NA
Muramatsu[22]	6,640 adults aged 70 and over	Cross-sectional	US	County	Gini	Not reported	CES-D	No effect	NA
Pabayo et al[23]	1,614 adolescents aged 14-19	Cross-sectional	US	Local (census tract)	Gini	0.33-0.65	MDS	NA	Effect only in girls
Pabayo et al[24]	34,653 adults aged 18 and over	Longitudinal	US	State	Gini	0.42-0.45 (cut-offs for 1st and 5th quintiles)	AUDADIS-IV	No effect	Effect only in women
Steptoe et al[25]	17,348 students aged 17-30	Cross-sectional	23 countries	Country level	Gini	0.20-0.59	BDI	NA	NA
Vihjalmsdottir et al[26]	5.958 adolescents aged 15-16	Cross-sectional	Iceland	Local (neighbourhood)	20%:20% ratio	4.47-39.90 (ratio range)	SCL-90 (12 depression items)	NA	NA
Equivocal association (positive in bivariate analysis but not in regression)									
Choi et al[27]	34,994 adults aged 50 and over	Cross-sectional	US	County	Gini	0.33-0.60	CES-D	NA	NA
Goodman et al[28]	13,235 adolescents mean age 16	Cross-sectional	US	Local (school)	Proportion of total income held by the lower half of the population	19.7-40.5 (ratio range)	CES-D	NA	NA
Henderson et al[29]	42,862 adults aged 18 and over	Cross-sectional	US	State	Gini	0.38-0.50	AUDADIS-IV	NA	Effect only in men in bivariate analysis

Table 1 Papers included in the systematic review of the association between income inequality and depression *(continued)*

Study	Sample	Study design	Country	Geographical unit of analysis	Inequality measure	Inequality range (Gini)	Depression measure	Absolute income effect	Gender effect
No association									
Fernández-Niño et al[30]	8,874 adults aged 60 and over	Cross-sectional	Mexico	State and municipal	Gini	Not reported	CES-D	NA	NA
Adjaye-Gbewonyo et al[31]	9,664 adults	Longitudinal panel	South Africa	District	Gini	0.46-0.68	CES-D	NA	NA
McLaughlin et al[7]	6,483 adolescents aged 13-17	Cross-sectional	US	Local (census tract)	Gini	0.59-0.65 (cut-offs for 1st and 4th quartiles)	CIDI (modified)	NA	NA
Sturm & Gresenz[32]	9,585 adults	Cross-sectional	US	Municipality	Gini	0.38-0.54	CIDI	NA	NA
Rai et al[33]	187,496 adults aged 18 and over	Cross-sectional	53 countries	Country level	Gini	0.25-0.74	CIDI	No effect	NA
Zimmerman et al[34]	4,817 adults aged 40-45	Cross-sectional	US	County	County-level percentage of households with income over $150,000 annually	Not reported	CES-D	NA	NA
Negative association (lower risk of depression in populations with higher income inequality)									
Marshall et al[35]	10,644 adults aged 50 and over	Cross-sectional	UK	Local (neighbourhood)	Gini	Not reported	CES-D	Most salient in low-income people	NA

BSI-D – Brief Symptom Inventory Depression Scale, CES-D – Center for Epidemiologic Studies - Depression, CIDI – Composite International Diagnostic Interview, DIPS – Diagnosis Item Properties Study, HDI – human development index, PHQ – Patient Health Questionnaire, GWB-D – General Well-Being Schedule - Depression subscale, RCES-D – CES-D Revised, AUDADIS-IV – Alcohol Use Disorder and Associated Disabilities Interview Schedule - IV, MDS – Modified Depression Scale, BDI – Beck Depression Inventory, SCL-90 – Symptom Checklist - 90, NA – not available

Table 2 Quality assessment of papers included in meta-analysis (SAQOR tool)

Paper	Sample	Exposure/outcome measures	Distorting influences	Reporting of data	Overall quality
Adjaye-Gbewonyo et al[31]/Burns et al[4]	Adequate	Adequate	Adequate	Adequate	High
Ahern & Galea[12]	Adequate	Adequate	Adequate	Adequate	High
Ladin et al[19]	Adequate	Adequate	Unclear	Unclear	Moderate
Chiavegatto Filho et al[13]	Adequate	Adequate	Adequate	Adequate	High
Kahn et al[18]	Adequate	Adequate	Adequate	Unclear	Moderate
Choi et al[27]	Adequate	Adequate	Adequate	Adequate	High
Fan et al[15]	Adequate	Adequate	Adequate	Adequate	High
Henderson et al[29]	Adequate	Adequate	Adequate	Adequate	High
Cifuentes et al[14]	Adequate	Adequate	Adequate	Unclear	Moderate
Sturm & Gresenz[32]	Adequate	Adequate	Adequate	Inadequate	Moderate
Steptoe et al[25]	Adequate	Adequate	Adequate	Unclear	Moderate
Marshall et al[35]	Unclear	Adequate	Adequate	Adequate	Moderate

SAQOR – Systematic Appraisal of Quality in Observational Research

quintiles[15,29]. All were re-categorized as dichotomous as described earlier. Ladin et al[19] divided the sample of ten European countries into five high inequality versus five low inequality countries. We followed a similar procedure for Steptoe et al's[25] study of 23 countries, by creating a group of 11 low inequality countries and 12 high inequality countries.

For the South Africa data, we extracted information from the two studies that employed the South Africa National Income Dynamics Study[4,31]. Data were available from Burns et al[4] for the depression prevalence by municipality. Adjaye-Gbewonyo et al[31] calculated Gini coefficients for each municipality based on the 2011 census. We integrated depression prevalence data and Gini coefficient data by municipality and split the dataset into approximate halves around a Gini coefficient of 0.75.

Unadjusted data were used for all studies when available. Unadjusted data were not presented in the study by Cifuentes et al[14], and rates were adjusted for age, gender and marital status. Fan et al[15] only presented adjusted prevalence figures: rates were adjusted for gender, age, race/ethnicity, marital status, education, household income, and chronic medical conditions.

Based on the 12 studies with dichotomized inequality groupings, the pooled risk ratio was 1.19 (95% CI: 1.07-1.31), demonstrating greater risk of depression in populations with higher income inequality relative to those with lower income inequality (see Figure 2). The heterogeneity was very high, $I^2=98\%$, which is likely due in part to the diversity of sample designs, populations, measures used, and adjustments and weighting in analyses. In all sensitivity analyses, the pooled risk ratio was significant for higher income inequality associated with increased risk of depression ($p<0.05$).

Multiple studies conducted moderator analyses by stratifying the samples by gender, absolute income, country economic status, and ethnicity/race. Because of the limited number of studies with outcomes that could be dichotomized by depression and income inequality, we did not create subpools of studies or employ meta-regression to assess these potential moderators.

Scoping review of mechanisms

Based on the results of the systematic review, a number of potential mechanisms of the inequality-depression relationship may be hypothesized, operating at different ecological levels, from the individual to the neighbourhood to the regional or national levels.

At the individual level, the effects of income inequality on general health are likely to be primarily mediated through psychological stress[3]. This may be regarded as the final mechanism mediating the effects of income inequality on depression in a range of pathways.

At the neighbourhood levels, two mechanisms are hypothesized. The first is the social comparison or status anxiety hypotheses[36], which argue that comparing oneself to those who are better off in a highly unequal context creates feelings of social defeat or status anxiety[4,37]. In a similar vein, Walker et al[38] hypothesized feelings of withdrawal and shame experienced by those in lower social positions. The second neighbourhood mechanism is the social capital hypothesis, which argues that income inequality erodes social capital, including two key components: cognitive social capital (especially social trust)[26] and structural social capital (the organizational and structural arrangements which facilitate social interactions and build social trust and cooperation, for example through group membership)[39].

Study	High income inequality Events	Total	Low income inequality Events	Total	Weight	Risk ratio M-H, random, 95% CI
Ahern & Galea[12]	160	982	34	373	4.50%	1.79 (1.26-2.54)
Ladin et al[19]	2,917	9,288	2,327	12,195	9.30%	1.65 (1.57-1.73)
Chiavegatto Filho et al[13]	317	2,831	53	711	5.60%	1.50 (1.14-1.99)
Kahn et al[18]	1,384	6,938	155	1,071	7.80%	1.38 (1.18-1.61)
Adjaye-Gbewonyo et al[31]/ Burns et al[4]	2,731	9,717	2,416	11,448	9.30%	1.33 (1.27-1.40)
Choi et al[27]	5,867	33,334	4,232	31,348	9.30%	1.30 (1.26-1.35)
Fan et al[15]	18,352	176,043	10,254	117,362	9.40%	1.19 (1.17-1.22)
Henderson et al[29]	8,043	30,689	3,094	12,044	9.30%	1.02 (0.98-1.06)
Cifuentes et al[14]	14,936	234,360	2,922	46,203	9.30%	1.01 (0.97-1.05)
Sturm & Gresenz[32]	424	4,420	369	3,815	8.20%	0.99 (0.87-1.13)
Steptoe et al[25]	1,952	9,749	1,685	7,599	9.20%	0.90 (0.85-0.96)
Marshall et al[35]	1,064	6,386	788	4,258	8.90%	0.90 (0.83-0.98)
Total (95% CI)		524,737		248,427	100.00%	1.19 (1.07-1.31)
Total events	58,147		28,329			
Heterogeneity: tau² = 0.03; chi² = 513.55, df = 11 (p<0.00001); I² = 98%						

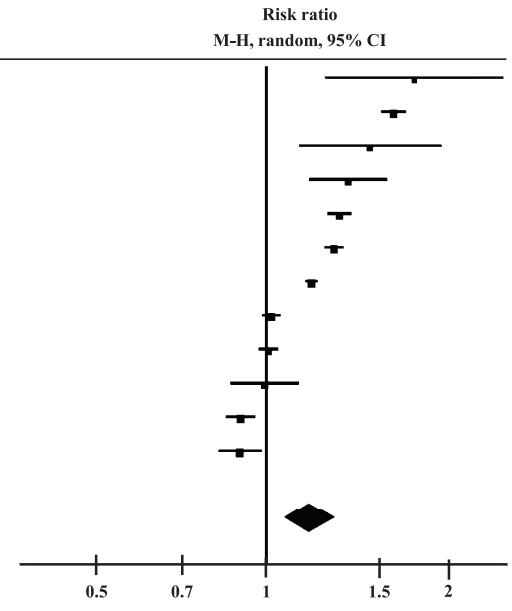

Figure 2 Forest plot of the association between income inequality and depression. M-H – Mantel-Haenszel estimate, total – pooled risk ratio

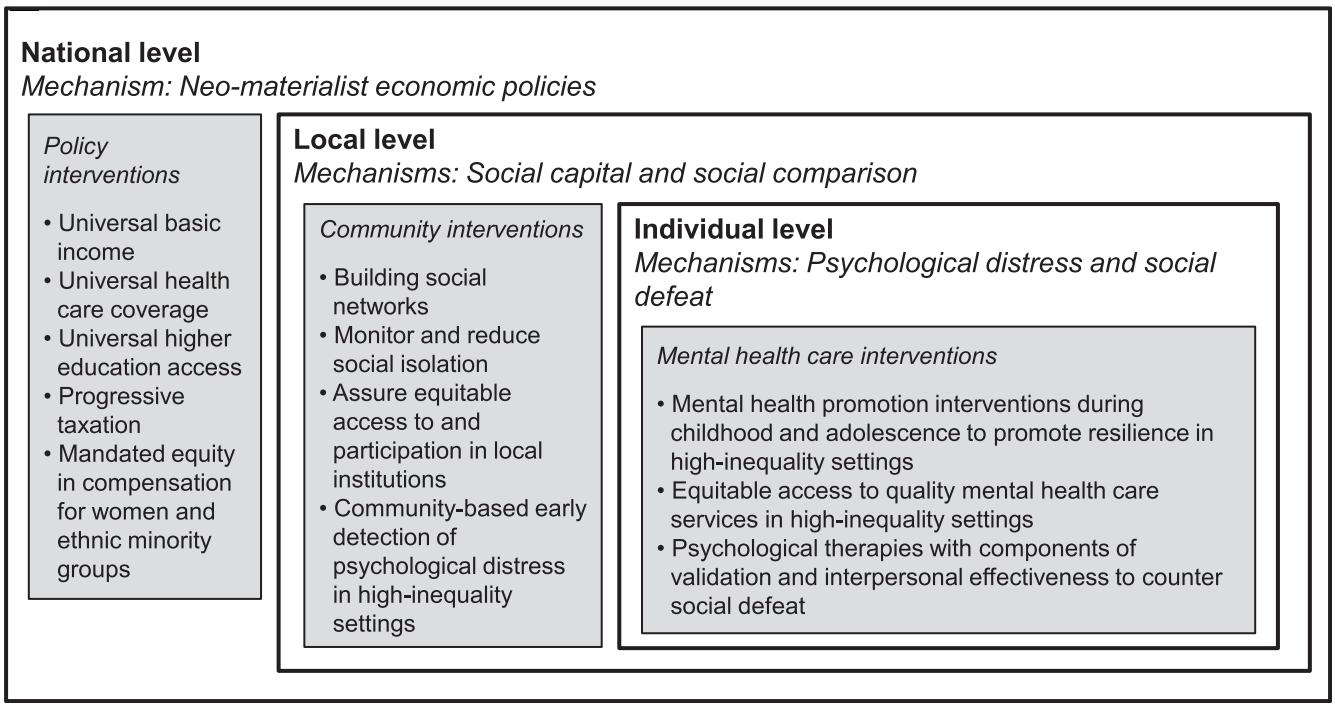

Figure 3 Mechanisms of income inequality on depression by ecological levels and recommendations for interventions

Social capital is critical, because it facilitates social integration (a dynamic process by which members of a social group participate in dialogue or collaborate to achieve a shared social goal). Income inequality therefore undermines social capital and social integration, promoting social isolation, alienation and loneliness. It also undermines perceptions of fairness (a component of trust)[37]. Ichida et al[40] confirmed the social capital hypothesis in Japan, showing that social capital (measured as social trust) mediated the effect of inequality on self-rated health. This is supported by Durkheim's theory of social integration and social regulation[41], the failure of which he linked to suicide.

Perceptions of fairness and trust are also consistent with Merton's *anomie disjunction* between society's goals and normative structures governing the means to attain that goal[42]. This is more exaggerated in societies with higher levels of inequality, where the means of attaining upward social mobility are severely constrained, and therefore there is a disjunction between society's goals or aspirations (for example of acquiring wealth) and the means to attain that goal, which are not accessible to those who are lower on the socio-economic hierarchy.

Both the above neighbourhood mechanisms may be more pronounced at certain developmental stages, in particular in adolescence, when social trust and group membership are being established, and when most mental health problems emerge. For example, social status was associated with depression among adolescents whose parents had lowest levels of education[7]. In addition, social comparison may be amplified by other group identities, for example ethnicity or gender.

At the national or regional levels, the neo-material hypothesis proposes that greater income inequality coexists with a wide range of material deprivations which are relevant to health[43]. These include lack of investment in housing, education and public transport as well as pollution control, healthy food availability and accessibility of health care. Thus, greater inequality leads to worse physical health (for example due to less public spending on health care in more unequal societies), leading in turn to an increased rate of depression.

This hypothesis was supported by Muramatsu[22], who found that the association between inequality and depression was stronger among those with more illnesses. However, it is worth noting Zimmerman et al's opposite finding that more unequal states did not in fact spend less on health care[34]. Also, Fone et al[36] argue that it is unlikely that the neo-material hypothesis would apply at small area level (such as neighbourhoods), as resource allocation decisions for major services are not typically made within these areas.

For all of these potential mechanisms, it is important to consider a range of other factors that may moderate the relationship between income inequality and depression, reflected in the available studies. The first is the geographical unit of analysis. Of the six studies that found no association, five conducted analysis at the district level, and national level effects appeared to be more marked in the studies included in this review. According to Ahern and Galea[12], this is likely to be at least partially influenced by the nature of the area demarcation. For example, if a neighbourhood includes strong contrasts of high- versus low-income groups, the effect of income in-

equality is likely to be more pronounced at that neighbourhood level. But frequently neighbourhoods involve homogeneous demarcations, and the effect may then be less pronounced.

In a similar vein, Fone et al[36] found that in Wales income deprivation was more important than income inequality for common mental disorders at the local neighbourhood level, but that the effect of income inequality became more evident at larger regional levels. Furthermore, Chen and Crawford[44] reported that, when comparing US counties and states, the income inequality/health relationship was more evident at the state than the county level, (although this was true for health insurance as an outcome but not for self-reported health). Thus, it may be possible to argue that different mechanisms operate at different geographical levels or units of analysis.

A second important consideration is the level of national development, for example as measured using the human development index (HDI). One study showed a possible interaction with country HDI level, namely that the inequality/depression association was more evident in higher HDI countries[14]. Income inequality may matter in high-income countries with low levels of poverty, but not in low- or middle-income countries with high levels of poverty, where the effects of material poverty and absolute income may be more significant.

A third consideration is the effect of income inequality on low- vs. high-income groups. Within countries, the effect of inequality on depression appears to be more pronounced among low-income groups[12]. This is consistent with the hypothesized role of upward social mobility, the constraints of which are more likely to be experienced by low-income groups. The hypothesis that inequality is deleterious for high-income groups too is proposed by other authors[3]. Kawachi et al[45] argue that the wealthy in highly unequal societies cannot escape the "pathologies of poverty", including crime, violence and exposure to some infectious diseases.

A fourth consideration is cultural variation across countries. Although this may be difficult to test empirically, Steptoe et al[25] considered the results in a multi-country study with respect to cultural variation along the axis of individualism and collectivism. The likelihood of high levels of depressive symptoms was lower in more individualistic cultures, with 26% reduction in the odds of elevated symptoms with every unit change in individualism-collectivism score.

A fifth consideration is the broader political and historical context within which depression and inequality are measured. For example, in post-apartheid South Africa, there have been expectations of rapid social improvements, and there is clear evidence of improvements for some people, but for those who remain in poverty there is a sense of frustration, alienation, disappointment and anger, manifest in frequent service delivery protests[4]. This may well exaggerate the effects of income inequality on depression.

A sixth consideration is life course or developmental stage. According to one study, childhood social class is more predictive of self-rated health than adult social class[16]. Prevalence of depression varies substantially across the life course[46], and early exposure to inequality may well affect later mental health. Most of the studies included in this review lack a life course or developmental framework, even when the effect of inequality on specific age groups was examined, for example in the case of adolescent depression.

A seventh consideration is gender. In at least one study[23], the effect of inequality on depression was found for adolescent girls but not for boys. This was confirmed by Hiilamo[47] in a study in Finland, which explored changes in municipality-level relative poverty and antidepressant prescriptions from 1995 to 2010, and found a positive association for young adult females.

A final consideration is the methods employed by the studies themselves. For example, contrary to the finding that the association between inequality and depression was less evident in more local, homogeneous populations, Fiscella and Franks[16] did find a positive association at local level. This finding may be attributable to the study design, which employed longitudinal, multi-level methods and collected baseline data on county income inequality, individual income, age, gender, self-rated health, level of depressive symptoms, and severity of biomedical morbidity.

DISCUSSION

In this paper we present, to our knowledge, the most comprehensive review of the evidence on the relationship between income inequality and depression. Despite the relatively small evidence base (especially from LMIC) and methodological limitations of the available studies, we report a compelling quantitative association between income inequality and depression. Even though the absolute effect size was relatively small (risk ratio of 1.19), the translation of this risk to population mental health is likely to be very large.

Further, we note that the primary outcome of the studies we included was a categorical outcome of "case-level" depression. This is a crude indicator of population mental health, and the associations between income inequality and mood are likely to be greater when the latter is treated as a continuous dimension, which could capture dose-effects of the degree of inequality on the distribution of affective symptoms.

If our findings are indicative of a causal relationship, then we should expect worse mental health globally in the years ahead, as income inequality is continuing to increase in most countries, making the United Nations (UN) Sustainable Development Goal targets for mental health[48] even harder to achieve. This is especially likely to be the case for disadvantaged or vulnerable groups in the population that already bear a disproportionate burden of mental health problems, such as women, adolescents, older adults and low-income groups.

The heterogeneity of the findings of studies across populations and over time is not surprising, given the complexity of

likely mechanisms and pathways, and their moderation by a range of contextual factors which we have attempted to delineate. These mechanisms operate at different ecological levels, but the final pathways are, as with any mental health problem, uniquely individual, moderated by a range of distal and proximal determinants.

Although we do strongly endorse the need to "unpack" these mechanisms through carefully designed studies, such research is likely to be complex, time-consuming and costly. Thus, we propose that the evidence which already exists is sufficient to take pre-emptive action to halt the potentially damaging effects of income inequality on the mental health of populations.

Implications for reducing the global burden of depression

Our ecological framework offers indications for the kinds of interventions which hold promise (see Figure 3). Obviously, at the national or regional level, economic policies which promote the fair distribution of income, for example through a universal basic income and progressive taxation, are potentially the most tractable[49]. Additionally, promoting social policies that reduce gender inequities which systematically disadvantage women, and income inequities, such as universal health coverage and expanding opportunities for educational attainment, can reduce the impact of the neo-material effect on low-income populations.

In addition to structural interventions, the mechanisms we propose suggest attractive opportunities for proximal interventions to mitigate the adverse personal consequences of living in unequal societies. The Disease Control Priorities project[50] has recommended a series of interventions for the prevention, treatment and care of mental health problems, most of which can be delivered through community and routine health care platforms, using task-sharing by non-specialist providers. Particularly relevant examples would include interventions in early life through adolescence to build resilience (for example, parenting interventions and life skills interventions), as well as promoting early detection and self-help for mood and anxiety disorders (for example, through improving access to empirically supported digital apps, especially with guidance)[51].

A recent systematic review has demonstrated the effectiveness of psychological therapies delivered by non-specialists in low-resource settings[52]. Such therapies may be modified when delivered in the context of high inequality through a focus on mechanisms related to cognitive comparisons leading to social defeat and worthlessness. For example, interventions that focus on demoralization[53,54] may be especially important in highly inequitable societies and communities. Third-wave psychological therapies that include components of self-validation may also counter social defeat and worthlessness associated with depression and suicidality[55]. These therapies are currently being adapted for delivery in settings of extreme poverty[56].

Interventions that harness the power of social networking sites to build social capital also show promise at mobilizing specific subgroups and reducing the risk of social isolation. Pilot programs in Mexico and South Africa have shown encouraging results at reducing levels of anxiety, depression and feelings of social isolation in adolescents and pregnant women with HIV/AIDS[57-60]. Marshall et al[35] report that social interactions and networks among subgroups in mixed-income neighbourhoods cushion the impact of income inequality on depression.

Other research points to the role of social interactions, cultural biases and belief systems in maintaining and perpetuating conditions for income inequality[61,62]. Thus, it is important that we develop interventions that target social and cultural aspects of inequalities (for example, designing institutional platforms such as schools and health institutions) to enhance social capital, and all interventions must be guided by a strong emphasis on equitable coverage. This is consistent with a shift from cultural competency to "structural competency", which emphasizes the need for mental health providers to be knowledgeable of context and resources of their patients and actively draw upon resources to mitigate social and structural determinants of mental illness[63].

Limitations of the study

There are limitations to our study which should be noted. First, publication bias, namely a propensity for journals to publish positive findings, may overestimate the strength or consistency of the association between inequality and depression. Second, there was a heterogeneity of outcome measures for depression, with some studies not utilizing validated assessment instruments, and a diversity in sample size and sampling strategies, all of which impact depression prevalence estimates[64]. Third, the majority of studies failed to stratify their samples by important socio-demographic factors such as gender, age and absolute income, limiting our ability to explore in greater depth the controversial question of whether the negative effects of income inequality are evenly distributed across the population or if certain vulnerable groups are particularly affected[65].

Regarding the meta-analysis, we were unable to use unadjusted data across all studies, and it is likely that the studies that adjusted inequality by outcomes reflect aspects of the association differently than unadjusted studies. In addition, the inequality cut-off for each study was different, based on the relative levels of inequality within the sample. For example, inequality levels within the South Africa dataset were high on average compared to European nations. Therefore, our findings are reflective of regional and national relative income inequality rather than the effect of absolute inequality (e.g., dividing all samples at one specific Gini coefficient cut-off, which would have been arbitrary, given that inequality is, by definition, a relative measure).

The meta-analysis also demonstrated high heterogeneity. As the pool of studies examining income inequality and men-

tal health grows, it will be possible to perform more subgroup analyses with studies that employ comparable designs and samples in order to reduce heterogeneity.

Implications for research

Future research should aim to unpack the mechanisms underlying the association between inequality and depression, in particular to explain the heterogeneity of findings across contexts. This research should involve prospective studies in diverse countries, in particular in a range of LMIC which are witnessing rapid socio-economic changes, such as the BRICS nations. Notably, Brazil and South Africa, which both have high levels of inequality, showed comparable high effects of income inequality on depression in our meta-analysis (risk ratios were 1.38 and 1.33 respectively).

Future studies should include the effects of changes in income inequality (at different geographical levels and population subgroups of analysis) over time, with embedded assessments of hypothesized individual and area-level mechanisms; and evaluation of the effects of interventions addressing the proposed pathways. Additionally, further exploration of the studies with equivocal findings, such as countries with high levels of income inequality which did not show an increased prevalence of depression, should also be conducted, to understand possible structural differences, policies or socio-cultural factors that mitigate this effect.

It is important to methodologically take note of the historical, political and cultural forces that may shape the association between income inequality and depression in developing countries. Modelling contextually grounded forces can shed greater light on the precise mechanisms that may be operating in these contexts.

CONCLUSIONS

Mental health professionals, regardless of their political persuasion, should carefully assess the evidence presented in this review to shape their position with respect to the ideologically contentious issue of income inequality.

They should ally with other stakeholders in government and civil society who are arguing for a fairer, more equitable distribution of income, as this is a major social determinant of poor mental health, while also drawing attention to the need for greater investments in proven individual interventions for the prevention and treatment of depression.

ACKNOWLEDGEMENTS

V. Patel is supported by a Wellcome Trust Principal Research Fellowship and by the US National Institute of Mental Health (NIMH) (U19MH113211). B.A. Kohrt is supported by the US NIMH (grants K01MH104310 and R21MH111280). V. Patel and C. Lund are supported by UK Aid, as part of the PRogramme for Improving Mental health carE (PRIME). The views expressed in this paper are not necessarily those of the funders.

REFERENCES

1. Davies J, Lluberas R, Shorrocks A. Credit Suisse Global Wealth Databook 2016. Zurich: Credit Suisse Research Institute, 2016.
2. Inequality.org. Income inequality in the United States. https://inequality.org/facts/income-inequality/.
3. Pickett KE, Wilkinson RG. Income inequality and health: a causal review. Soc Sci Med 2015;128:316-26.
4. Burns JK, Tomita A, Lund C. Income inequality widens the existing income-related disparity in depression risk in post-apartheid South Africa: evidence from a nationally representative panel study. Health Place 2017; 45:10-6.
5. Ribeiro WS, Bauer A, Andrade MC et al. Income inequality and mental illness-related morbidity and resilience: a systematic review and meta-analysis. Lancet Psychiatry 2017;4:554-62.
6. Wilkinson R, Pickett K. Inequality and mental illness. Lancet Psychiatry 2017;4:512-3.
7. McLaughlin KA, Costello EJ, Leblanc W et al. Socioeconomic status and adolescent mental disorders. Am J Publ Health 2012;102:1742-50.
8. Kohrt BA, Rasmussen A, Kaiser BN et al. Cultural concepts of distress and psychiatric disorders: literature review and research recommendations for global mental health epidemiology. Int J Epidemiol 2013;43:365-406.
9. Cochrane Community. Review Manager (RevMan). https://gradepro.org/.
10. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ 1997;315:1533-7.
11. Viera AJ. Odds ratios and risk ratios: what's the difference and why does it matter? South Med J 2008;101:730-4.
12. Ahern J, Galea S. Social context and depression after a disaster: the role of income inequality. J Epidemiol Community Health 2006;60:766-70.
13. Chiavegatto Filho AD, Kawachi I, Wang YP et al. Does income inequality get under the skin? A multilevel analysis of depression, anxiety and mental disorders in Sao Paulo, Brazil. J Epidemiol Community Health 2013;67: 966-72.
14. Cifuentes M, Sembajwe G, Tak S et al. The association of major depressive episodes with income inequality and the human development index. Soc Sci Med 2008;67:529-39.
15. Fan AZ, Strasser S, Zhang XY et al. State-level socioeconomic factors are associated with current depression among US adults in 2006 and 2008. J Publ Health Epidemiol 2011;3:462-70.
16. Fiscella K, Franks P. Individual income, income inequality, health, and mortality: what are the relationships? Health Serv Res 2000;35:307.
17. Godoy RA, Reyes-García V, McDade T et al. Does village inequality in modern income harm the psyche? Anger, fear, sadness, and alcohol consumption in a pre-industrial society. Soc Sci Med 2006;63:359-72.
18. Kahn RS, Wise PH, Kennedy BP et al. State income inequality, household income, and maternal mental and physical health: cross sectional national survey. BMJ 2000;321:1311.
19. Ladin K, Daniels N, Kawachi I. Exploring the relationship between absolute and relative position and late-life depression: evidence from 10 European countries. Gerontologist 2010;50:48-59.
20. Messias E. Income inequality, illiteracy rate, and life expectancy in Brazil. Am J Publ Health 2003;93:1294-6.
21. Muntaner C, Li Y, Xue X et al. County level socioeconomic position, work organization and depression disorder: a repeated measures cross-classified multilevel analysis of low-income nursing home workers. Health Place 2006;12:688-700.
22. Muramatsu N. County-level income inequality and depression among older Americans. Health Serv Res 2003;38:1863-84.
23. Pabayo R, Dunn EC, Gilman SE et al. Income inequality within urban settings and depressive symptoms among adolescents. J Epidemiol Community Health 2016;70:997-1003.
24. Pabayo R, Kawachi I, Gilman SE. Income inequality among American states and the incidence of major depression. J Epidemiol Community Health 2014;68:110-5.
25. Steptoe A, Tsuda A, Tanaka Y. Depressive symptoms, socio-economic background, sense of control, and cultural factors in university students from 23 countries. Int J Behav Med 2007;14:97-107.
26. Vilhjalmsdottir A, Gardarsdottir RB, Bernburg JG et al. Neighborhood income inequality, social capital and emotional distress among adolescents: a population-based study. J Adolesc 2016;51:92-102.
27. Choi H, Burgard S, Elo IT et al. Are older adults living in more equal counties healthier than older adults living in more unequal counties? A propensity score matching approach. Soc Sci Med 2015;141:82-90.

28. Goodman E, Huang B, Wade TJ et al. A multilevel analysis of the relation of socioeconomic status to adolescent depressive symptoms: does school context matter? J Pediatr 2003;143:451-6.
29. Henderson C, Liu X, Roux AV et al. The effects of US state income inequality and alcohol policies on symptoms of depression and alcohol dependence. Soc Sci Med 2004;58:565-75.
30. Fernández-Niño JA, Manrique-Espinoza BS, Bojorquez-Chapela I et al. Income inequality, socioeconomic deprivation and depressive symptoms among older adults in Mexico. PLoS One 2014;9:e108127.
31. Adjaye-Gbewonyo K, Avendano M, Subramanian SV et al. Income inequality and depressive symptoms in South Africa: a longitudinal analysis of the National Income Dynamics Study. Health Place 2016;42:37-46.
32. Sturm R, Gresenz CR. Relations of income inequality and family income to chronic medical conditions and mental health disorders: national survey. BMJ 2002;324:20.
33. Rai D, Zitko P, Jones K et al. Country- and individual-level socioeconomic determinants of depression: multilevel cross-national comparison. Br J Psychiatry 2013;202:195-203.
34. Zimmerman FJ, Bell JF. Income inequality and physical and mental health: testing associations consistent with proposed causal pathways. J Epidemiol Community Health 2006;60:513-21.
35. Marshall A, Jivraj S, Nazroo J et al. Does the level of wealth inequality within an area influence the prevalence of depression amongst older people? Health Place 2014;27:194-204.
36. Fone D, Greene G, Farewell D et al. Common mental disorders, neighbourhood income inequality and income deprivation: small-area multilevel analysis. Br J Psychiatry 2013;202:286-93.
37. Buttrick NR, Oishi S. The psychological consequences of income inequality. Soc Pers Psychol Compass 2017;11:e12304.
38. Walker R, Kyomuhendo GB, Chase E et al. Poverty in global perspective: is shame a common denominator? J Soc Policy 2013;42:215-33.
39. Kawachi I, Kennedy BP, Glass R. Social capital and self-rated health: a contextual analysis. Am J Publ Health 1999;89:1187-93.
40. Ichida Y, Kondo K, Hirai H et al. Social capital, income inequality and self-rated health in Chita peninsula, Japan: a multilevel analysis of older people in 25 communities. Soc Sci Med 2009;69:489-99.
41. Durkheim É. Suicide: a study in sociology. New York: Free Press, 1966.
42. Merton RK. Social theory and social structure. New York: Simon and Schuster, 1968.
43. Lynch JW, Smith GD, Kaplan GA et al. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. BMJ 2000;320:1200.
44. Chen Z, Crawford CA. The role of geographic scale in testing the income inequality hypothesis as an explanation of health disparities. Soc Sci Med 2012;75:1022-31.
45. Kawachi I, Berkman L. Social cohesion, social capital, and health. In: Berkman LF, Kawachi I (eds). Social epidemiology. New York: Oxford University Press, 2000:174-90.
46. 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.
47. Hiilamo H. Is income inequality 'toxic for mental health'? An ecological study on municipal level risk factors for depression. PLoS One 2014;9: e92775.
48. United Nations. Transforming our world: the 2030 Agenda for Sustainable Development. New York: United Nations, 2015.
49. Piketty T. Capital in the twenty-first century. Cambridge: Belknap Press, 2014.
50. Patel V, Chisholm D, Dua T et al (eds). Mental, neurological, and substance use disorders: disease control priorities, 3rd ed. Washington: World Bank Publications, 2016.
51. Naslund JA, Aschbrenner KA, Araya R et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017;4: 486-500.
52. Singla DR, Kohrt BA, Murray LK et al. Psychological treatments for the world: lessons from low-and middle-income countries. Ann Rev Clin Psychol 2017;13:149-81.
53. Noordhof A, Kamphuis JH, Sellbom M et al. Change in self-reported personality during major depressive disorder treatment: a reanalysis of treatment studies from a demoralization perspective. Personal Disord (in press).
54. Griffith JL. Hope modules: brief psychotherapeutic interventions to counter demoralization from daily stressors of chronic illness. Acad Psychiatry (in press).
55. Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. Behav Ther 2004;35:639-65.
56. Ramaiya MK, Fiorillo D, Regmi U et al. A cultural adaptation of dialectical behavior therapy in Nepal. Cogn Behav Pract 2017;24:428-44.
57. McClure CM, McFarland M, Legins KE. Commentary: innovations in programming for HIV among adolescents: towards an AIDS-free generation. JAIDS 2014;66:S224-7.
58. The SHM Foundation. Zumbido health. http://shmfoundation.org/.
59. The SHM Foundation. Project Kopano. https://shmfoundation.org/.
60. The SHM Foundation. Project Khulama. https://shmfoundation.org/.
61. Bowles S, Durlauf SN, Hoff K (eds). Poverty traps. Princeton: Princeton University Press, 2006.
62. Hoff KR, Pandey P. Belief systems and durable inequalities: an experimental investigation of Indian caste. Washington: World Bank Publications, 2004.
63. Metzl JM, Hansen H. Structural competency: theorizing a new medical engagement with stigma and inequality. Soc Sci Med 2014;103:126-33.
64. 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.
65. Kawachi I, Subramanian SV, Almeida-Filho N. A glossary for health inequalities. J Epidemiol Community Health 2002;56:647-52.

DOI:10.1002/wps.20492

Psychotherapies for depression in low- and middle-income countries: a meta-analysis

Pim Cuijpers[1,2], Eirini Karyotaki[1,2], Mirjam Reijnders[1,2], Marianna Purgato[3], Corrado Barbui[3]

[1]Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [2]Amsterdam School of Public Health, Amsterdam, The Netherlands; [3]WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation, Department of Neuroscience, Biomedicine and Movement Sciences, Section of Psychiatry, University of Verona, Verona, Italy

Most psychotherapies for depression have been developed in high-income Western countries of North America, Europe and Australia. A growing number of randomized trials have examined the effects of these treatments in non-Western countries. We conducted a meta-analysis of these studies to examine whether these psychotherapies are effective and to compare their effects between studies from Western and non-Western countries. We conducted systematic searches in bibliographical databases and included 253 randomized controlled trials, of which 32 were conducted in non-Western countries. The effects of psychotherapies in non-Western countries were large (g=1.10; 95% CI: 0.91-1.30), with high heterogeneity (I^2=90; 95% CI: 87-92). After adjustment for publication bias, the effect size dropped to g=0.73 (95% CI: 0.51-0.96). Subgroup analyses did not indicate that adaptation to the local situation was associated with the effect size. Comparisons with the studies in Western countries showed that the effects of the therapies were significantly larger in non-Western countries, also after adjusting for characteristics of the participants, the treatments and the studies. These larger effect sizes in non-Western countries may reflect true differences indicating that therapies are indeed more effective; or may be explained by the care-as-usual control conditions in non-Western countries, often indicating that no care was available; or may be the result of the relative low quality of many trials in the field. This study suggests that psychotherapies that were developed in Western countries may or may not be more effective in non-Western countries, but they are probably no less effective and can therefore also be used in these latter countries.

Key words: Depression, psychotherapy, low- and middle-income countries, care-as-usual, meta-analysis

(World Psychiatry 2018;17:90–101)

Depression and other common mental disorders are highly prevalent, with almost one in five people worldwide affected[1,2]. They have a considerable impact on the lives of patients and their families, and are associated with huge economic and societal costs[3]. The disability associated with these disorders results in a loss of more than one million healthy life years, which makes mental disorders the leading cause of years lived with disability worldwide[4]. The economic costs, in terms of production losses and health and social care expenditures, have been estimated at US$2.5 trillion in 2010 worldwide[5-7], and these costs are expected to grow to US$6.0 trillion by 2030[8].

Several evidence-based pharmacotherapies and psychotherapies are available for depression. However, most people with a depressive disorder do not receive treatment, especially in low- and middle-income countries, where only between 7 and 21% of patients are treated[5]. If patients get treatment, this typically consists of pharmacotherapy, while the majority of patients prefer psychotherapies[9].

Several psychotherapies, such as cognitive behavior therapy, interpersonal psychotherapy, problem-solving and behavioral activation, have been developed for the treatment of depression[10]. Since the 1970s, several hundreds of randomized trials have shown that these interventions are effective[11-14], although their effects are modest and have been overestimated because of the low quality of many trials[15] and publication bias[16,17]. The effects of psychotherapies have been found to be comparable to those of pharmacotherapy[18], and probably last longer[19].

Most psychotherapies have been developed in high-income Western countries in North America, Europe and Australia, and the vast majority of the more than 450 randomized trials which have examined their effects[20] have been conducted in those countries. It is therefore not well known whether these therapies are also effective in low- and middle-income countries.

In recent years, a growing number of randomized trials have examined the effects of psychotherapies for depression in countries outside of North America, Europe and Australia. The goal of the present meta-analysis is to examine whether these psychotherapies are also effective in non-Western countries and to compare their effects with those in Western countries. This also gives the opportunity to examine whether the effects of psychotherapies are associated with the income of the country and the region where the trial was conducted.

METHODS

Identification and selection of studies

We used an existing database of studies on psychotherapies for depression. This database has been described in detail elsewhere[20], has been used in a series of earlier published meta-analyses[21], and is continuously updated. For this database we searched four major bibliographical sources (PubMed, PsycINFO, Embase and the Cochrane Library) by combining terms (both index terms and text words) indicative of depression and

psychotherapies, with filters for randomized controlled trials. We also checked the references of earlier meta-analyses.

Because this database was not developed specifically to include studies from non-Western countries, we examined the list produced by the Effective Practice and Organization of Care (EPOC) Group (a Cochrane review group), which contains a collection of databases, websites and journals relevant to low- and middle-income countries. We selected databases that were freely available, could be searched in English, and had a working web address. The following databases were searched with adapted search strings: the International Initiative for Impact Evaluation (3ie); the British Library for Development Studies; the Eldis; the World Health Organization (WHO)'s Global Index Medicus; the Latin-American and Caribbean System on Health Sciences Information (LILACS); the Indice Bibliográfico Español de Ciencias de la Salud (IBECS); the AfricaBib; the IndMed; the KoreaMed; and African Journals Online. The search was made in November 2016.

All records were screened by two independent researchers and all papers that could possibly meet inclusion criteria according to one of the researchers were retrieved as full text. The decision to include or exclude a study was also done by the two independent researchers, and disagreements were solved through discussion.

We included papers reporting on a randomized trial in which a psychotherapy for adult depression was compared with a control group (waiting list, care-as-usual, placebo, other inactive treatment) in a non-Western country (not located in North America, Europe or Australia).

Depression could be established by a diagnostic interview or a score above a cut-off on a self-report scale. Psychotherapies were defined as interventions with a primary focus on language-based communication between a patient and a therapist, or as bibliotherapy supported by a therapist[22]. The therapies could be delivered individually, in groups, or as guided self-help by professionals or para-professionals. Comorbid mental or somatic disorders were not used as an exclusion criterion. Studies on inpatients were excluded. We also excluded maintenance studies aimed at people who had already recovered or partly recovered after an earlier treatment.

In addition to the main analyses of the studies conducted in non-Western countries, we also compared treatment effect sizes in the trials conducted in non-Western countries with those conducted in Western countries. For this comparison, we selected from our database trials on psychotherapies for depression that were conducted in Western countries and in which psychotherapy was compared with a control condition, with the same inclusion and exclusion criteria as for the studies in non-Western countries.

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[23]. This tool assesses possible sources of bias in randomized trials, including the adequate generation of allocation sequence; the concealment of allocation to conditions; the prevention of knowledge of the allocated intervention (masking 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). Assessment of the quality of the included studies was conducted by two independent researchers, and disagreements were solved through discussion.

We also coded participant characteristics (depressive disorder or scoring high on a self-rating scale; recruitment method; target group); characteristics of the psychotherapies (treatment format; number of sessions); and general characteristics of the studies (type of control group; country where the study was conducted).

We rated whether the intervention was adapted to the local setting and population. We considered an intervention not adapted when the authors did not mention adaptation and when the procedures described were comparable to those found in therapies developed in Western countries. An intervention was considered as adapted when it was explicitly mentioned that it was adapted to the local situation. We considered an intervention also as "adapted" when it was developed in a non-Western country and was based on models or theories from non-Western countries. We also considered an intervention as "not adapted" when Western manuals were just translated into the national language.

In order to examine whether the effects of psychotherapy were associated with the per capita income, we recorded the gross national income (GNI) based on purchasing power parity (PPP) per capita in international dollars for each of the countries where a trial was conducted, using data from the World Bank (http://data.worldbank.org). We categorized the countries into low-, lower-middle, upper-middle and high-income ones according to the definition of the World Bank. We also used the six World Bank regions to categorize where the studies were conducted (East Asia and Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, South Asia, and Sub-Saharan Africa).

Primary outcome

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[24]. Effect sizes were calculated 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[25]. If means and standard deviations were not reported, we used the procedures of the Comprehensive Meta-Analysis software (see

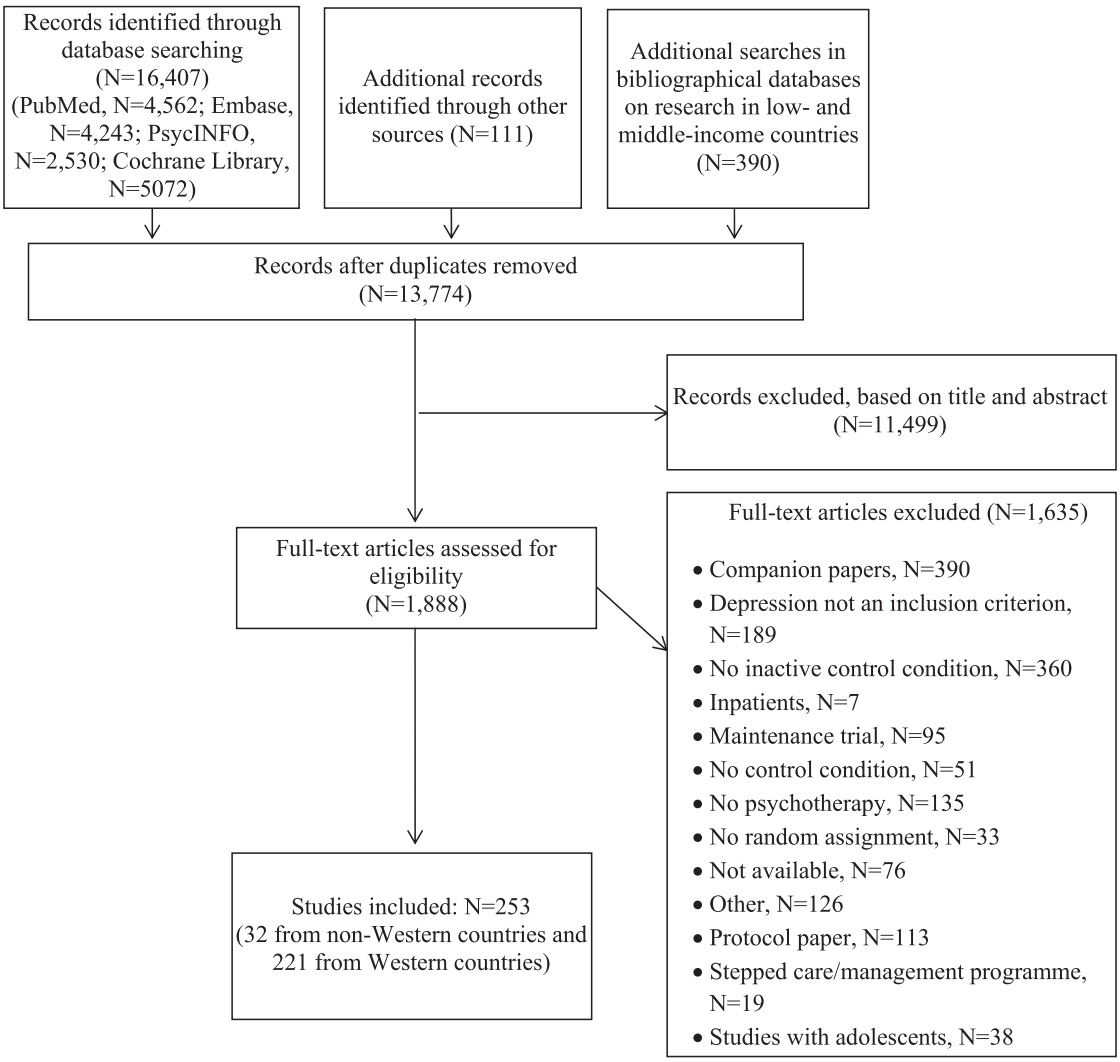

Figure 1 PRISMA flow chart

below) to calculate the effect size using dichotomous outcomes; and if these were not available either, we used other statistics (such as t or p value) to calculate the effect size.

In order to calculate effect sizes, we used all measures examining depressive symptoms, such as the Beck Depression Inventory (BDI-I or BDI-II)[26,27] or the Hamilton Rating Scale for Depression (HAMD-17)[28].

Meta-analyses

To calculate pooled mean effect sizes, we used the computer program Comprehensive Meta-Analysis (version 3.3070). Because we expected considerable heterogeneity among the studies, we employed a random effects pooling model in all analyses.

Numbers-needed-to-be-treated (NNT) were calculated using the formulae provided by Furukawa[29], 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 in psychotherapies for depression)[30]. As a test of homogeneity of effect sizes, we calculated the I^2 statistic, which is an indicator of heterogeneity in percentages. 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[31]. We calculated 95% confidence intervals (CIs) around I^2 using the non-central chi-squared-based approach within the heterogi module for Stata[32,33]. We conducted sensitivity analyses excluding potential outliers. These were defined as studies in which the 95% CI of the effect size did not overlap with the 95% CI of the pooled effect size.

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 effect size, as indicated by a z value and an associated p value. Multivariate meta-regression anal-

Table 1 Selected characteristics of randomized trials comparing psychotherapies for adult depression to control groups in non-Western countries

Study	Conditions	N patients	Format	Cultural adaptation	N sessions	Country	Risk of bias*
Bolton et al[35]	IPT	139	Group	Adapted	16	Uganda	+ – SR +
	CAU	145					
Chan et al[36]	CBT	17	Individual	Not adapted	10	China	– + + –
	MBCT	17	Individual	Non-Western	10		
	WL	16					
Chan et al[37]	Other	14	Individual	Non-Western	5	China	+ – SR +
	CAU	12					
Chen et al[38]	SUP	30	Group	Not adapted	4	Taiwan	– – SR –
	CAU	30					
Chiang et al[39]	CBT	30	Group	Not adapted	12	Taiwan	+ + + –
	CAU	32					
Cho et al[40]	CBT	12	Individual	Not adapted	9	Korea	– – SR –
	CAU	10					
Duarte et al[41]	CBT	41	Group	Not adapted	12	Brazil	– + SR –
	CAU	44					
Faramarzi et al[42]	CBT	29	Group	Not adapted	10	Iran	– – SR –
	CAU	30					
Furukawa et al[43]	CBT	58	Individual	Adapted	8	Japan	+ + SR +
	WL	60					
García-Peña et al[44]	CBT	41	Group	Not adapted	12	Mexico	+ – SR –
	CAU	40					
Hamdan-Mansour et al[45]	CBT	44	Group	Adapted	10	Jordan	– + SR –
	CAU	36					
Hou et al[46]	CBT	104	Individual	Not adapted	19	China	– – SR –
	CAU	109					
Huang et al[47]	CBT	31	Group	Not adapted	12	Taiwan	– – SR –
	CAU	30					
Jiang et al[48]	Other	257	Individual	Not adapted	–	China	+ – SR –
	CAU	514					
Leung et al[49]	CBT	47	Group	Adapted	6	China	– – SR +
	CAU	50					
Liu et al[50]	CBT	27	Guided self-help	Not adapted	10	Taiwan	– – SR –
	WL	25					
Mukhtar[51]	CBT	58	Group	Adapted	8	Malaysia	– – SR –
	WL	55					
Naeem et al[52]	CBT	94	Guided self-help	Adapted	7	Pakistan	+ – SR –
	CAU	89					
Nakimuli-Mpungu et al[53]	SUP	57	Group	Adapted	8	Uganda	+ + SR +
	Other	52					
Ng et al[54]	Other	14	Individual	Not adapted	5	Singapore	– – SR –
	CAU	12					
Ngai et al[55]	CBT	197	Other	Adapted	5	China	+ + SR +
	CAU	200					

Table 1 Selected characteristics of randomized trials comparing psychotherapies for adult depression to control groups in non-Western countries *(continued)*

Study	Conditions	N patients	Format	Cultural adaptation	N sessions	Country	Risk of bias*
Omidi et al[56]	CBT	30	Group	Not adapted	8	Iran	– – SR –
	MBCT	30	Group	Not adapted	8		
	CAU	30					
Petersen et al[57]	IPT	17	Group	Adapted	8	South Africa	+ – SR –
	CAU	17					
Qiu et al[58]	CBT	31	Group	Not adapted	10	China	+ + + +
	WL	31					
Rahman et al[59]	CBT	418	Individual	Adapted	16	Pakistan	+ + + –
	Other	400					
Songprakun & McCann[60]	CBT	26	Guided self-help	Not adapted	8	Thailand	+ + + –
	CAU	28					
Sreevani et al[61]	Other	15	Group	Non-Western	4	India	+ – SR –
	CAU	15					
Teichman et al[62]	CMT	15	Individual	Not adapted	13	Israel	– – SR –
	CBT	15	Individual	Not adapted	13		
	WL	15					
Vitriol et al[63]	DYN	44	Individual	Not adapted	12	Chile	– – + +
	CAU	43					
Wong[64]	CBT	48	Group	Adapted	10	China	– + SR +
	WL	40					
Wong[65]	CBT	163	Group	Adapted	10	China	– + SR –
	WL	159					
Zu et al[66]	CBT	12	Individual	Not adapted	20	China	+ – + –
	CAU	16					

CAU – care as usual, CBT – cognitive behavior therapy, CMT – cognitive marital therapy, CT – cognitive therapy, DR – psychodrama, DYN – psychodynamic therapy, IPT – interpersonal psychotherapy, MBCT – mindfulness based cognitive therapy, SUP – non-directive supportive therapy
*A positive (+) or negative (−) sign is given for four quality criteria: allocation sequence, concealment of allocation to conditions, blinding of assessors, and intention-to-treat analyses; SR indicates that only self-report measures (and no assessor) were used

yses, with the effect size as the dependent variable, were conducted through Comprehensive Meta-Analysis.

We tested for publication bias by inspecting the funnel plot on primary outcome measures and by Duval and Tweedie's trim and fill procedure[34], which yields an estimate of the effect size after the publication bias has been taken into account (as implemented in Comprehensive Meta-Analysis). 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 studies

After examining a total of 16,908 abstracts (13,774 after removal of duplicates), we retrieved 1,888 full-text papers for further consideration. We excluded 1,635 of the retrieved papers. The PRISMA flow chart describing the inclusion process, with the reasons for exclusion, is presented in Figure 1.

A total of 32 studies conducted in non-Western countries (with 35 comparisons between a psychotherapy and a control condition) met inclusion criteria for this meta-analysis (Table 1). Another 221 studies (with 297 comparisons between a treatment and a control group) on psychotherapies in Western countries were included (for the comparison of effect sizes in Western versus non-Western countries). This makes a total of 253 studies that were included in the analyses.

Characteristics of included studies

In the 32 included studies conducted in Non-Western countries, a total of 4,607 patients participated (2,222 for therapy conditions and 2,385 for control conditions). Participants were

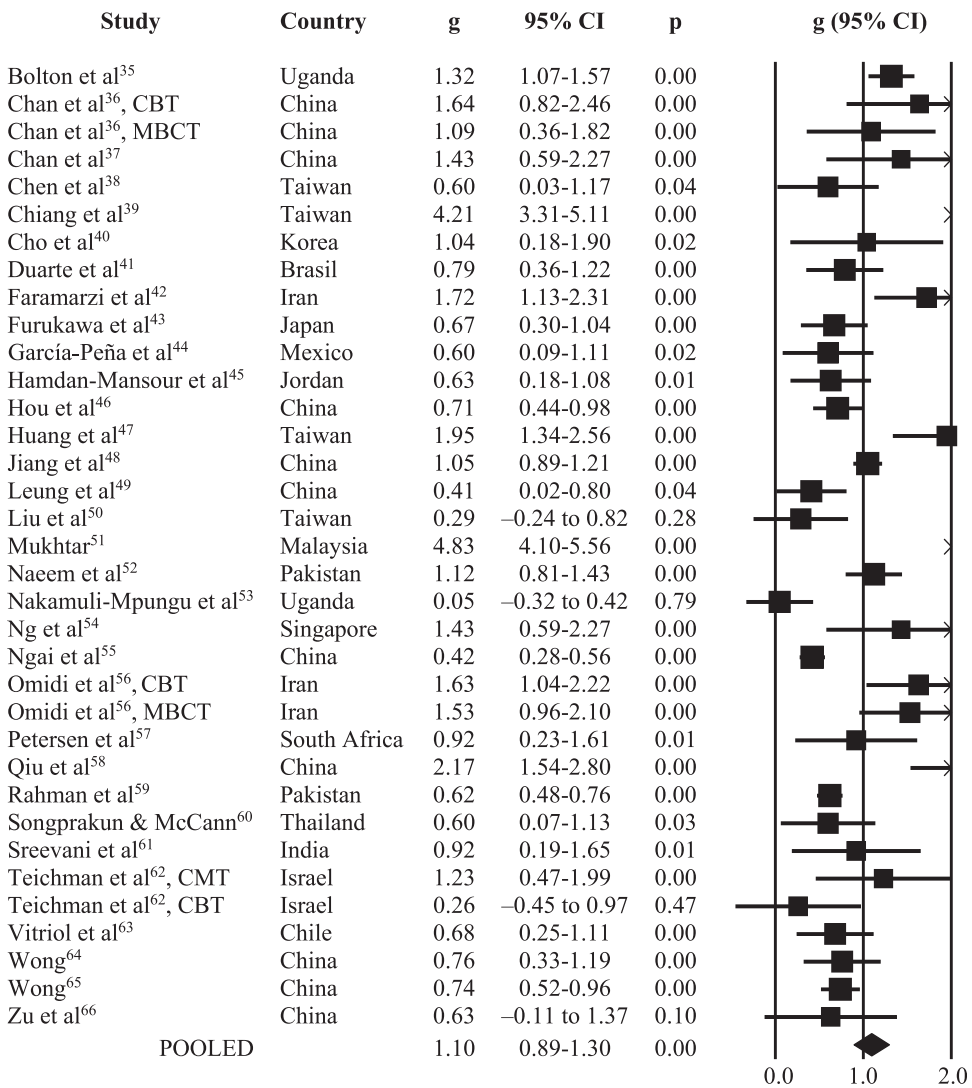

Figure 2 Forest plot of effect sizes from randomized controlled trials of psychotherapies for depression in Non-Western countries. CBT – cognitive behavior therapy, MBCT – mindfulness based cognitive therapy, DR – psychodrama, CMT – cognitive marital therapy

recruited through announcements in local newspapers and other media (four studies), referrals from health services (11 studies), or other strategies such as screening at general medical services (17 studies).

In 25 of the 35 comparisons between a treatment and a control condition, cognitive behavior therapy was used as the intervention. Two studies used interpersonal psychotherapy, one used psychodynamic therapy, one used non-directive supportive therapy, and the remaining six used another type of treatment. Of these treatments, 12 were culturally adapted, 20 were not culturally adapted, and three were non-Western treatments. Eighteen comparisons used a group treatment format, 13 utilized individual treatment, and three used a guided self-help treatment format. The number of treatment sessions ranged from four to 20. Eight studies used a waiting list as control group, 22 studies used care-as-usual, and two used another control group.

Nineteen studies were conducted in East Asia, three in South Asia, three in Latin America and the Caribbean, four in the Middle East and North Africa, and three in Sub-Saharan Africa. The gross national income of the countries ranged from low/low-medium (250 US$) to high (54,580 US$).

Effects of psychotherapies in non-Western countries

The overall effect in the 35 comparisons between psychotherapies and control groups was g=1.10 (95% CI: 0.91-1.30), which corresponds with a NNT of 2.51. Heterogeneity was very high (I^2=90; 95% CI: 87-92). Effect sizes and 95% confidence intervals of each study are presented in the forest plot in Figure 2. The results of these main analyses are presented in Table 2.

Considering only the outcomes measured with the HAM-D-17, the mean effect size was g=1.38 (95% CI: 0.66-2.09; n=7;

Table 2 Psychotherapies for adult depression in non-Western countries compared with control conditions

		N	g	95% CI	I^2	95% CI	p	NNT
All comparisons		35	1.10	0.91-1.30	90	87-92		2.51
One effect size per study (highest only)		32	1.11	0.90-1.32	90	88-92		2.49
One effect size per study (lowest only)		32	1.06	0.85-1.27	90	88-92		2.62
Outliers excluded		26	0.95	0.82-1.08	55	23-70		2.95
Extreme positive outliers excluded		32	0.87	0.73-1.06	78	69-83		3.26
Only HAM-D		7	1.38	0.66-2.09	93	89-95		1.99
Only BDI-I		9	1.33	0.54-2.12	93	90-95		2.06
Only BDI-II		7	1.37	0.76-1.97	91	85-94		2.01
Adjusted for publication bias (9 imputed)		44	0.73	0.51-0.96	93	92-94		3.98
Subgroup analyses								
Region	East Asia	17	0.83	0.64-1.02	77	61-84	0.55	3.44
	Middle East and North Africa	6	1.17	0.69-1.65	74	18-87		2.35
	South Asia	3	0.86	0.47-1.25	77	0-91		3.30
	Other	6	0.73	0.30-1.16	85	64-91		3.98
Income level of country	High	8	0.86	0.48-1.23	71	24-84	0.95	3.30
	Upper middle	18	0.89	0.71-1.08	77	63-84		3.18
	Low/lower middle	6	0.83	0.44-1.22	88	76-93		3.44
Risk of bias	0-1 (high)	10	1.20	0.84-1.56	73	42-84	<0.001	2.29
	2-3	16	0.87	0.70-1.03	61	22-76		3.26
	4 (low)	6	0.51	0.34-0.69	60	0-82		6.01
Control group	Care as usual	22	0.97	0.78-1.16	80	71-86	0.02	2.88
	Waiting list/other	10	0.65	0.45-0.85	61	0-79		4.55
Target group	Adults	15	0.95	0.74-1.16	65	32-79	0.16	2.95
	Perinatal depression	7	0.67	0.44-0.90	84	67-91		4.39
	Other	10	0.97	0.60-1.35	80	60-88		2.88
Diagnosis	Depressive disorder	21	0.91	0.74-1.09	74	57-82	0.48	3.02
	Cut-off on scale	11	0.80	0.53-1.07	84	72-89		3.58
Adaptation	Yes	14	0.74	0.56-0.93	80	65-87	0.06	3.92
	No	18	0.99	0.78-1.19	68	42-79		2.82
Type of therapy	CBT	22	0.85	0.69-1.01	75	60-82	0.71	3.35
	Other	10	0.91	0.63-1.19	76	50-86		3.10
Format of therapy	Individual	12	0.89	0.68-1.10	63	17-79	0.28	3.18
	Group	15	0.94	0.69-1.20	81	68-87		2.99
	Other	5	0.64	0.35-0.94	78	27-89		4.63

BDI – Beck Depression Inventory, HAM-D – Hamilton Rating Scale for Depression, NNT – numbers-needed-to-be-treated, CBT – cognitive behavior therapy

NNT=1.99; I^2=93; 95% CI: 89-95). For the BDI-I, it was g=1.33 (95% CI: 0.54-2.12; n=9; NNT=2.06; I^2=93; 95% CI: 90-95); for the BDI-II, it was g=1.37 (95% CI: 0.76-1.97; n=7; NNT=2.01; I^2=91; 95% CI: 85-94).

Nine studies were potential outliers[39,47,49-51,53,55,58,59]. After exclusion of these studies, the effects dropped to g=0.95 (95% CI: 0.82-1.08; NNT=2.95). Heterogeneity was still moderate (I^2=55; 95% CI: 23-70). There were three potential outliers with extremely high effect sizes (g>2.0)[39,51,58]. The pooled effect size after excluding these extreme outliers was g=0.87 (95% CI: 0.73-1.06; I^2=78; 95% CI: 69-83).

In this meta-analysis, we included three studies in which two psychological treatments were compared with the same control group. This means that multiple comparisons were included in the same analysis, which may have resulted in an artificial reduction of heterogeneity and may have affected the pooled effect size. We examined the possible effects of this by conducting an analysis in which we included only one effect

Table 3 Psychotherapies for adult depression in Western and non-Western countries compared with control conditions

		N	g	95% CI	I^2	95% CI	p	NNT
	Western	291	0.60	0.55-0.64	59	53-64	<0.001	4.99
	Non-Western	32	0.87	0.73-1.02	78	69-83		3.26
Region	North America	165	0.67	0.59-0.74	61	53-67	<0.001	4.39
	Europe	107	0.51	0.45-0.57	47	32-58		6.01
	Australia	19	0.62	0.38-0.85	74	56-82		4.80
	East Asia	17	0.83	0.64-1.02	77	61-84		3.44
	Middle East and North Africa	6	1.17	0.69-1.65	74	18-87		2.35
	South Asia	3	0.86	0.47-1.25	77	0-91		3.30
	Other	6	0.73	0.30-1.16	85	64-91		3.98
Income level of country	High	297	0.60	0.55-0.65	59	54-64	0.002	4.99
	Upper middle	20	0.92	0.74-1.11	76	61-83		3.06
	Low/lower middle	6	0.83	0.44-1.22	88	76-93		3.44
Income level of country	High, Western	289	0.59	0.55-0.64	58	53-63	0.003	5.08
	High, non-Western	8	0.86	0.48-1.23	71	24-84		3.30
	Upper middle[a]	18	0.93	0.73-1.12	78	64-85		3.02
	Low/lower middle	6	0.83	0.44-1.22	88	76-93		3.44

[a]Turkey was excluded from this analysis because it is a Western country but also an upper middle-income one
NNT – Numbers-needed-to-be-treated

size per study. First, we included only the comparisons with the largest effect size from these studies and then we included only the smallest effect sizes. As can be seen from Table 2, the resulting effect sizes were almost the same as in the overall analyses. Heterogeneity was still very high in these analyses.

Visual inspection of the funnel plot, as well as Duval and Tweedie's trim and fill procedure, pointed at considerable publication bias. After adjustment for publication bias, the mean effect size was reduced from g=1.10 to g=0.73 (95% CI: 0.51-0.96; number of missed studies: 9). Egger's test also pointed at significant asymmetry of the funnel plot (p=0.004; intercept: 2.42; 95% CI: 0.65-4.20).

In the subgroup analyses, excluding the extreme outliers, we found that the risk of bias was significantly associated with the effect size (p<0.001). The six comparisons with the lowest risk of bias (no risk of bias for any of the four items of the assessment tool) had an effect size of g=0.51 (95% CI: 0.34-0.69; NNT=6.01) compared to g=1.20 (95% CI: 0.84-1.56; NNT=2.29) in the studies with the highest risk of bias.

We also found that the type of control group was significantly associated with the effect size, with care-as-usual control groups resulting in higher effect sizes than waiting list and other control groups (p=0.02).

None of the other subgroup analyses resulted in significant differences between subgroups, and that included the region (East Asia, Middle East and North Africa, South Asia, other), the level of income of the country (high, upper middle, low/lower middle), and whether or not the treatment was adapted to the local situation.

We conducted a series of bivariate meta-regression analyses. In these analyses, we found no indication that the effect size was significantly associated with the GNI (coefficient: 0.00; 95% CI: −0.00 to 0.00; p=0.56), the number of treatment sessions (coefficient: 0.00; 95% CI: −0.04 to 0.04; p=1.00), and year of publication (coefficient: 0.00; 95% CI: −0.03 to 0.04; p=0.84).

Comparison between the effects of psychotherapy in Western vs. non-Western countries

We considered the 32 comparisons from non-Western countries vs. the 291 comparisons from Western countries (Table 3; extreme outliers with g>2.0 were excluded from these analyses). We found that the effect size in Western countries (g=0.60; 95% CI: 0.55-0.64; I^2=59; 95% CI: 53-64; NNT=4.99) was significantly lower than in non-Western countries (p<0.001).

We also examined the effect sizes in the different regions and found that they differed significant across regions (p<0.001), with the lowest effect sizes in North America, Europe and Australia, and the highest in East Asia, South Asia and the Middle East and North Africa. We also found a significant difference across countries with different incomes, with the highest effect sizes in low- and middle-income countries.

Table 4 Standardized regression coefficients of characteristics of studies on psychotherapies for depression in Western and non-Western countries (full multivariate meta-regression analyses, excluding extreme outliers)

		Coeff	SE	p	Coeff	SE	p	Coeff	SE	p
Western vs. non-Western countries		0.26	0.08	<0.001						
Region	North America				Ref					
	Europe				−0.02	0.06	0.83			
	Australia				0.08	0.10	0.44			
	East Asia				0.17	0.11	0.11			
	Middle East and North Africa				0.44	0.18	0.02			
	South Asia				0.44	0.20	0.03			
	Other				0.25	0.16	0.11			
Income level of country	High							Ref		
	Low/lower middle							0.43	0.15	0.004
	Upper middle							0.31	0.10	0.002
Diagnosis vs. cut-off		−0.02	0.05	0.63	−0.01	0.05	0.88	−0.01	0.05	0.83
Target group	Unselected adults	Ref			Ref			Ref		
	Older adults	−0.05	0.07	0.52	−0.04	0.08	0.56	−0.04	0.07	0.55
	Women with PPD	−0.04	0.08	0.65	−0.04	0.08	0.61	−0.04	0.08	0.58
	General medical disease	0.04	0.07	0.57	0.04	0.07	0.60	0.04	0.07	0.53
	Other	0.05	0.07	0.45	0.03	0.07	0.64	0.06	0.07	0.38
Type of therapy	CBT	Ref			Ref			Ref		
	IPT	−0.08	0.09	0.39	−0.07	0.09	0.44	−0.09	0.09	0.33
	PST	−0.03	0.10	0.75	−0.02	0.10	0.84	−0.03	0.09	0.73
	Supportive	0.03	0.11	0.81	0.05	0.11	0.67	0.05	0.11	0.65
	Other	0.02	0.06	0.75	0.03	0.06	0.64	0.02	0.06	0.72
Format of therapy	Individual	Ref			Ref			Ref		
	Group	−0.10	0.06	0.08	−0.10	0.06	0.07	−0.12	0.06	0.03
	Guided self-help	0.05	0.07	0.53	0.04	0.08	0.57	0.03	0.07	0.67
	Other/mixed	−0.17	0.10	0.09	−0.15	0.10	0.13	−0.18	0.10	0.07
Number of sessions (continuous)		−0.00	0.01	0.68	0.00	0.01	0.54	0.00	0.01	0.65
Risk of bias (continuous)		−0.12	0.02	<0.001	−0.12	0.02	<0.001	−0.12	0.02	<0.001
Control group	Waiting list	Ref			Ref			Ref		
	Care as usual	−0.09	0.06	0.14	−0.10	0.06	0.13	−0.11	0.06	0.08
	Other	−0.21	0.07	<0.01	−0.23	0.07	<0.001	−0.23	0.07	<0.001
Intercept		1.01	0.10	<0.001	1.00	0.10	<0.001	1.03	0.10	<0.001
R^2 analog		0.36			0.36			0.38		

Coeff – regression coefficient, Ref – reference group, PPD – post-partum depression, CBT – cognitive behavior therapy, IPT – interpersonal psychotherapy, PST – problem solving therapy

In addition, we conducted a separate subgroup analysis in which we separated high-income countries into Western and non-Western countries (Table 3). We found that the eight studies in high-income, non-Western countries resulted in an effect size of g=0.86 (95% CI: 0.48-1.23; NNT=3.30; I^2=71; 95% CI: 24-84) compared to g=0.59 in Western countries (Table 2). A direct comparison between high-income countries in Western and non-Western countries did not indicate a significant difference (p=0.17), but this may have been related to the small number of studies from high-income non-Western countries.

We conducted a series of multivariate meta-regression analyses with the effect size as dependent variable (Table 4). In the first analysis, we included a dummy variable indicating whether the study was conducted in a Western or non-Western country, and also included other variables of the participants (a diagnosis of depression versus scoring above a cut-off on a self-report scale; the target group), the therapies (type,

Table 5 Standardized regression coefficients of characteristics of studies on psychotherapies for depression in Western and non-Western countries (parsimonious multivariate meta-regression analyses)

		Coeff	SE	p	Coeff	SE	p	Coeff	SE	p
Western vs. non-Western countries		0.23	0.07	<0.001						
Region	North America				Ref					
	Europe				−0.01	0.06	0.91			
	Australia				0.08	0.10	0.42			
	East Asia				0.13	0.10	0.21			
	Middle East and North Africa				0.43	0.17	0.01			
	South Asia				0.40	0.19	0.04			
	Other				0.22	0.15	0.15			
Income level of country	High							Ref		
	Low/lower middle							0.36	0.14	0.01
	Upper middle							0.24	0.09	0.01
Risk of bias (continuous)		−0.10	0.02	<0.001	−0.10	0.02	<0.001	−0.11	0.02	<0.001
Control group	Waiting list	Ref			Ref			Ref		
	Care as usual	−0.12	0.05	0.02	−0.12	0.05	0.02	−0.13	0.05	0.02
	Other	−0.23	0.06	<0.001	−0.25	0.06	<0.001	−0.25	0.06	<0.001
Intercept		0.98	0.05	<0.001	0.98	0.06	<0.001	0.99	0.05	<0.001
R^2 analog		0.37			0.37			0.38		

Coeff – regression coefficient, Ref – reference group

treatment format, number of sessions) and characteristics of the studies (type of control group and risk of bias). Whether the study was conducted in a Western or non-Western country remained a significant predictor of the effect size after adjusting for all other characteristics of the participants, interventions and studies ($p<0.001$).

In the second meta-regression analysis we used the same predictors, except that the dummy variable indicating that the study was conducted in a Western vs. a non-Western country was removed, and instead we added the variable indicating the region where the study was conducted. We found that studies conducted in the Middle East and North Africa, and in South Asia had significantly higher effect sizes than the reference group (studies from the United States).

In the third meta-regression analysis, we included the income of the country as predictor, and we found that both studies conducted in upper middle- ($p=0.002$) and in low/lower middle-income countries ($p=0.004$) had significantly higher effect sizes than those in high-income countries, while adjusting for all other variables.

We did not include the dummies indicating Western versus non-Western countries, the regions and the income level in one analysis, because the overlap across these variables was considered too large.

To avoid overfit of the meta-regression models, we repeated the above three meta-regression analyses with a (manual) stepwise backward elimination of the least significant predictor until only significant predictors remained in the model. The results of these parsimonious analyses are presented in Table 5. As can be seen, in all three models, risk of bias and type of control group remained significant, as well as the dummies indicating Western vs. non-Western countries, the regions and the income level.

DISCUSSION

Our study documents that psychotherapies for depression that have been developed in Western countries are also effective in non-Western countries. We even found indications that these therapies may be more effective in non-Western than in Western countries. This finding remained significant in multivariate meta-regression analyses in which we controlled for characteristics of the participants, the interventions and the studies.

We classified these studies in different ways, one in which we simply differentiated between Western and non-Western countries, one in which we categorized the countries into the major regions of the world according to the World Bank, and one in which we classified the countries according to their income (high, upper middle and low/lower middle). We found that the studies in non-Western countries had better outcomes than those from Western countries; that the effect sizes were

especially high in the Middle East and North-Africa, and in South Asia (although the lack of statistical significance for other regions may be caused by lack of power) and that studies in upper middle- and low/lower middle-income countries resulted in significantly higher effect sizes than studies in high-income countries.

It is not clear why the studies in non-Western countries had better outcomes. It is possible that these therapies simply work better in (some) non-Western countries, but it is not clear why that would be the case. Another explanation could be that most studies in non-Western countries had care-as-usual control groups, and that care-as-usual in these cases simply means to get no treatment at all, while in Western countries care-as-usual implies that patients have access to several treatments, like regular care provided by general practitioners or specialized mental health services. Another explanation could be that the quality of the studies conducted in non-Western countries was not optimal. Risk of bias was low in only 6 of the 32 included comparisons, and these studies with low risk of bias had considerably lower effect sizes than those with higher risk, very comparable to the ones found in Western countries.

We did not find indications that a specific adaptation of the treatment to the context where the therapy was conducted was associated with better outcomes. This finding should be considered with caution, because the description of the intervention was very brief in most papers, so that it cannot be excluded that the interventions were still adapted although this was not mentioned in the paper.

These findings do suggest that psychotherapies developed in Western countries can be implemented in non-Western countries when sufficient resources are available and without culturally adapting them. It has been argued recently that an investment in mental health care in low- and middle-income countries has considerable economic support[5]. Because we found no indication that the effects are associated with the treatment format, it would be possible to introduce low intensity interventions as a first line treatment, because these are easier and cheaper to implement than high intensity ones.

This study has several limitations that have to be taken into account when interpreting the results. One important limitation is that we may have missed studies because our searches mainly focused on Western databases, while studies published in other languages were not directly accessible. That implies that our results may be distorted because of bias in the selection of studies. Another limitation is that the quality of most of the included studies was not optimal, and only a handful of them had a high quality. Furthermore, these high-quality studies found considerably smaller effect sizes than the others, suggesting that the true effects are probably smaller than we found. However, after adjustment for study quality, studies in non-Western countries were still had better outcomes than those in Western countries. Another limitation is that most studies in non-Western countries were conducted in a selected sample of countries in Asia, and only few in Africa and Latin America.

Despite these limitations, this study suggests that psychotherapies developed in Western countries may or not be more effective in non-Western countries, but are probably no less effective and can therefore also be used in these latter countries, regardless of their income level.

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. Kessler RC, Berglund P, Demler O et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005;62:593-602.
3. Smit F, Comijs H, Schoevers R et al. Target groups for the prevention of late-life anxiety. Br J Psychiatry 2007;190:428-34.
4. 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.
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. Gustavsson A, Svensson M, Jacobi F et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011;21:718-79.
7. Hu T-W. Perspectives: an international review of the national cost estimates of mental illness, 1990-2003. J Ment Health Policy Econ 2006;9:3-13.
8. Bloom DE, Cafiero E, Jané-Llopis E et al. The global economic burden of noncommunicable diseases. Geneva: World Economic Forum, 2012.
9. 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.
10. Cuijpers P. Personalized treatment for functional outcome in depression. Medicographia 2014;36:476-81.
11. 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.
12. Cuijpers P, Geraedts AS, van Oppen P et al. Interpersonal psychotherapy for depression: a meta-analysis. Am J Psychiatry 2011;168:581-92.
13. Cuijpers P, van Straten A, Warmerdam L. Problem solving therapies for depression: a meta-analysis. Eur Psychiatry 2007;22:9-15.
14. Ekers D, Webster L, Van Straten A et al. Behavioural activation for depression; an update of meta-analysis of effectiveness and sub group analysis. PLoS One 2014;9:e100100.
15. 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.
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. Cuijpers P, Cristea IA, Karyotaki E et al. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry 2016;15:245-58.
18. 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.
19. Karyotaki E, Smit Y, Beurs DP et al. The long-term efficacy of acute-phase psychotherapy for depression: a meta-analysis of randomized trials. Depress Anxiety 2016;33:370-83.
20. Cuijpers P, van Straten A, Andersson G et al. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 2008;76:909-22.
21. Cuijpers P, Andersson G, Donker T et al. Psychological treatment of depression: results of a series of meta-analyses. Nord J Psychiatry 2011;65:354-64.
22. 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.
23. 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.
24. Cohen J. Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale: Erlbaum, 1988.

25. Hedges LV, Olkin I. Statistical methods for meta-analysis. Orlando: Academic Press, 1985.
26. Beck AT, Ward CH, Mendelson M et al. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561-71.
27. Beck AT, Steer RA, Brown GK. BDI-II, Beck depression Inventory: manual. San Antonio: Psychological Corporation, 1996.
28. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56-62.
29. Furukawa TA. From effect size into number needed to treat. Lancet 1999; 353:1680.
30. 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.
31. Higgins JPT, Thompson SG, Deeks JJ et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
32. Ioannidis JPA, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. BMJ 2007;335:914-6.
33. Orsini N, Higgins J, Bottai M et al. Heterogi: Stata module to quantify heterogeneity in a meta-analysis. Boston: Boston College Department of Economics, 2005.
34. 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.
35. Bolton P, Bass J, Neugebauer R et al. Group interpersonal psychotherapy for depression in rural Uganda: a randomized controlled trial. JAMA 2003; 289:3117-24.
36. Chan AS, Wong QY, Sze SL et al. A Chinese Chan-based mind-body intervention for patients with depression. J Affect Disord 2012;142:283-9.
37. Chan MF, Ng SE, Tien A et al. A randomised controlled study to explore the effect of life story review on depression in older Chinese in Singapore. Health Soc Care Community 2013;21:545-53.
38. Chen CH, Tseng YF, Chou FH et al. Effects of support group intervention in postnatally distressed women. A controlled study in Taiwan. J Psychosom Res 2000;49:395-9.
39. Chiang KJ, Chen TH, Hsieh HT et al. One-year follow-up of the effectiveness of cognitive behavioral group therapy for patients depression: a randomized, single-blinded, controlled study. Sci World J 2015;373149.
40. Cho HJ, Kwon JH, Lee JJ. Antenatal cognitive-behavioral therapy for prevention of postpartum depression: a pilot study. Yonsei Med J 2008;49:553-62.
41. 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.
42. 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.
43. Furukawa TA, Horikoshi M, Kawakami N et al. Telephone cognitive-behavioral therapy for subthreshold depression and presenteeism in workplace: a randomized controlled trial. PLoS One 2012;7:e35330.
44. García-Peña C, Vázquez-Estupiñan F, Avalos-Pérez F et al. Clinical effectiveness of group cognitive-behavioural therapy for depressed older people in primary care: a randomised controlled trial. Salud Ment 2015;38:33-9.
45. Hamdan-Mansour AM, Puskar K, Bandak AG. Effectiveness of cognitive-behavioral therapy on depressive symptomatology, stress and coping strategies among Jordanian university students. Issues Ment Health Nurs 2009;30:188-96.
46. Hou Y, Hu P, Zhang Y et al. Cognitive behavioral therapy in combination with systemic family therapy improves mild to moderate postpartum depression. Rev Bras Psiquiatr 2014;36:47-52.
47. Huang CY, Lai HL, Chen CI et al. Effects of motivational enhancement therapy plus cognitive behaviour therapy on depressive symptoms and health-related quality of life in adults with type II diabetes mellitus: a randomised controlled trial. Qual Life Res 2015;25:1275-83.
48. Jiang L, Wang ZZ, Qiu LR et al. Psychological intervention for postpartum depression. J Huazhong Univ Sci Technol Med Sci 2014;34:437-42.
49. Leung SS, Lee AM, Chiang VC et al. Culturally sensitive, preventive antenatal group cognitive-behavioural therapy for Chinese women with depression. Int J Nurs Pract 2013;19:28-37.
50. Liu ET-H, Chen W-L, Li Y-H et al. Exploring the efficacy of cognitive bibliotherapy and a potential mechanism of change in the treatment of depressive symptoms among the Chinese: a randomized controlled trial. Cogn Ther Res 2009;33:449-61.
51. Mukhtar F. Predictors of group cognitive behaviour therapy outcomes for the treatment of depression in Malaysia. Asian J Psychiatry 2011;4:125-8.
52. 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.
53. Nakimuli-Mpungu E, Wamala K, Okello J et al. Group support psychotherapy for depression treatment in people with HIV/AIDS in northern Uganda: a single-centre randomised controlled trial. Lancet HIV 2015;2:e190-9.
54. Ng SE, Tien A, Thayala JN et al. The effect of life story review on depression of older community-dwelling Chinese adults in Singapore: a preliminary result. Int J Geriatr Psychiatry 2013;28:328-30.
55. Ngai FW, Wong PWC, Leung KY et al. The effect of telephone-based cognitive-behavioral therapy on postnatal depression: a randomized controlled trial. Psychother Psychosom 2015;84:294-303.
56. 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.
57. Petersen I, Hanass Hancock J, Bhana A et al. A group-based counselling intervention for depression comorbid with HIV/AIDS using a task shifting approach in South Africa: a randomized controlled pilot study. J Affect Disord 2014;158:78-84.
58. 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.
59. 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.
60. Songprakun W, McCann TV. Evaluation of a cognitive behavioural self-help manual for reducing depression: a randomized controlled trial. J Psychiatr Ment Health Nurs 2012;19:647-53.
61. Sreevani R, Reddemma K, Chan CL et al. Effectiveness of integrated body-mind-spirit group intervention on the well-being of Indian patients with depression: a pilot study. J Nurs Res 2013;21:179-86.
62. Teichman Y, Bar-el Z, Shor H et al. A comparison of two modalities of cognitive therapy (individual and marital) in treating depression. Psychiatry 1995;58:136-48.
63. Vitriol VG, Ballesteros ST, Florenzano RU et al. Evaluation of an outpatient intervention for women with severe depression and a history of childhood trauma. Psychiatr Serv 2009;60:936-42.
64. 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.
65. 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.
66. 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;152-4:262-7.

DOI:10.1002/wps.20493

Reward-related cognitive vulnerability to bipolar spectrum disorders

Bipolar spectrum disorders (BSDs) are characterized by extreme swings of mood (euphoria or irritability versus sadness), cognition (grandiosity and racing thoughts versus worthlessness and concentration difficulties), and behavior (supercharged energy and excessive goal-striving versus anhedonia and lethargy) occurring within the same individual. They are prevalent, associated with significant disability, and occur on a continuum of severity, from milder cyclothymia to bipolar II to full-blown bipolar I disorder.

Individuals' cognitive styles (i.e., the general filters they use to process information and construe events in their lives) may provide vulnerability to BSD symptoms and episodes. Indeed, individuals with BSDs exhibit cognitive styles with certain unique reward-relevant features[1,2] and these cognitive styles have been shown to predict the onset and course of BSDs[3]. According to the reward hypersensitivity theory[3,4], individuals with or vulnerable to BSDs possess a hypersensitive reward system, linked to a dopaminergic fronto-striatal neural circuit subserving approach motivation and goal-directed behavior, that overreacts to goals or reward-relevant cues. This hypersensitivity leads to excessive approach-related affect and incentive motivation in response to life events involving goal-striving and attainment, which in turn leads to hypomanic/manic symptoms. It also can lead to excessive downregulation or decrease in approach-related affect and motivation in response to non-attainment of goals or rewards (e.g., irreconcilable losses or failures), which in turn leads to bipolar depressive symptoms.

Thus, a propensity toward excessive reward system activation and deactivation is the hypothesized vulnerability to BSDs in this model. The model also proposes that vulnerable individuals' reward hypersensitivity leads to behaviors that increase their exposure (via "stress generation" processes) to the very goal- and reward-relevant events that, in turn, precipitate excessive responses from their reward systems. To date, extensive self-report, behavioral, cognitive, life event, neurophysiological and neural evidence supports this reward hypersensitivity model of BSDs[3,4].

High reward sensitivity may be a temperament trait that contributes to the development of reward-relevant cognitive styles[1]. In line with this hypothesis, euthymic bipolar individuals have been found to exhibit a distinctive profile of cognitive styles characterized by perfectionism, self-criticism and autonomy rather than the dependency and approval-seeking styles observed among unipolar depressed individuals[1]. Additionally, controlling for current mood symptoms, individuals with BSDs exhibit higher achievement motivation, goal-attainment dysfunctional attitudes (e.g., "A person should do well at everything") and ambitious goal-striving styles than controls[1,2].

The strongest evidence confirming that reward-relevant cognitive styles provide vulnerability to BSDs comes from a prospective study[5], which found that, controlling for initial mood symptoms and family history of bipolar disorder, adolescents with no prior history of BSD who exhibited an ambitious goal-striving cognitive style at baseline had a greater likelihood and shorter time to first lifetime onset of BSD than those without that cognitive style. Additionally, a cognitive style characterized by ambitious goal-striving mediated the predictive association between high self-reported reward sensitivity and shorter time to first onset of BSD in this adolescent sample[5], further suggesting that ambitious goal-striving is a vulnerability trait to BSDs that may account for some of the risk associated with reward sensitivity.

Ambitious goal-striving cognitive styles, perfectionism, and a tendency to overgeneralize from success (rewards) have also been observed in individuals with no prior history of BSD but at behavioral risk for developing a bipolar disorder[6,7]. Further, controlling for baseline hypomanic symptoms, a cognitive style to overgeneralize from success interacted with self-reported reward hypersensitivity to predict increases in hypomanic symptoms among adolescents with no prior history of BSD[8].

Reward-relevant cognitive styles also affect the course of BSDs. In individuals with bipolar I disorder, ambitious goal-striving for financial success and popular fame predicted increases in manic symptoms over a three-month follow-up[2]. In addition, controlling for past history of mood episodes and baseline symptoms, late adolescents with bipolar II disorder or cyclothymia who possessed self-critical or autonomous reward-relevant cognitive styles at baseline had a greater likelihood of hypomanic and manic episodes over a three-year follow-up than adolescents who did not exhibit these styles[1]. Moreover, an autonomous cognitive style mediated the predictive association between self-reported reward hypersensitivity and prospective occurrence of hypomanic and manic episodes in this sample[1].

Finally, in the same sample, reward-relevant life events interacted with reward-related cognitive styles to predict bipolar mood symptoms[9]. Specifically, controlling for initial mood symptoms and total number of life events experienced, baseline perfectionistic and self-critical cognitive styles interacted with reward system-activating positive events to predict increases in hypomanic/manic symptoms, and with reward system-deactivating negative events (e.g., certain failures) to predict increases in depressive symptoms over follow-up[9].

Reward-relevant cognitive styles may not always be maladaptive. Indeed, the high achievement motivation and ambitious goal-striving may contribute to high levels of creativity and achievement also exhibited by many individuals with BSDs or at behavioral risk for developing a bipolar disorder[6].

The role of reward-relevant cognitive styles in the onset and course of BSDs has implications for psychosocial interventions for these disorders, particularly for cognitive-behavioral therapy (CBT), which has been shown to have efficacious prophylactic effects for BSDs[10]. There may be added value to CBT interventions that specifically target achievement, ambitious goal-striving, and reward-oriented cognitive schemas in man-

aging BSDs[10]. For example, the therapist might develop a plan in which surges of ambitious goal-setting and overconfidence are identified and challenged during prodromal periods to lessen the likelihood of a manic episode onset[10].

In summary, ambitious goal-striving cognitive styles appear to be involved in the vulnerability to onset and recurrences of mood episodes in individuals with BSDs. Thus, these styles may be an excellent target for preventive and therapeutic interventions for individuals with bipolar disorders.

Lauren B. Alloy[1], Robin Nusslock[2]
[1]Temple University, Philadelphia, PA, USA; [2]Northwestern University, Evanston, IL, USA

1. Alloy LB, Abramson LY, Walshaw PD et al. J Abnorm Psychol 2009;118:459-71.
2. Johnson SL, Carver CS, Gotlib IH. J Abnorm Psychol 2012;121:602-9.
3. Alloy LB, Nusslock R, Boland EM. Annu Rev Clin Psychol 2015;11:213-50.
4. Nusslock R, Alloy LB. J Affect Disord 2017;216:3-16.
5. Alloy LB, Bender RE, Whitehouse WG et al. J Abnorm Psychol 2012;121:399-51.
6. Murray G, Johnson SL. Clin Psychol Rev 2010;30:721-32.
7. Stange JP, Shapero BG, Jager-Hyman SG et al. Cogn Ther Res 2013;37:139-49.
8. Stange JP, Molz AR, Black CL et al. Behav Res Ther 2012;50:231-9.
9. Francis-Raniere EL, Alloy LB, Abramson LY. Bipolar Disord 2006;8:382-99.
10. Nusslock R, Abramson LY, Harmon-Jones E et al. Clin Psychol Sci Pract 2009;16:449-69.

DOI:10.1002/wps.20494

Prevention of child maltreatment: strategic targeting of a curvilinear relationship between adversity and psychiatric impairment

Child maltreatment – which includes physical, emotional and sexual abuse as well as neglect – is the single most influential known cause of lifetime mental health impairment that is preventable (the other high-impact causes being primarily genetic), with conservative estimates of prevalence of about 15% in high-income countries[1,2].

Its deleterious impact arguably accounts for 25% or more of the population-attributable risk for child psychopathology[1,3], and in severe cases can extend to the lack of the minimum requirements for normative human development (food, hygiene, human interaction), physical injury, sexual exploitation and mutilation, permanent brain injury and death[4], or be associated with perpetration of child abuse by victims when they reach adulthood[5].

Maltreatment most commonly first occurs in infancy, particularly when adult caregivers are too stressed or functionally incapacitated to attend to the needs of the children under their care. The long-term cost for each yearly cohort of children abused in the US alone has been conservatively estimated to exceed $124 billions[6].

Our ability to predict child maltreatment on the basis of risk indicators that can be feasibly ascertained on the first day of an infant's life (including indices of parental mental health or substance use impairment, concentrated poverty, and a range of socio-economic stress indicators) has considerably advanced[7], and specific risk profiles can be delineated identifying a subgroup of children who have an up to 70% likelihood of ultimately being detected in official governmental records for child abuse/neglect. In spite of this, hospitals and health agencies rarely systematically screen for child maltreatment risk.

Child maltreatment is preventable. Its prevention requires the coordinated application of interventions that address key lapses in "species-typical" mechanisms of protection of the young: caregiving knowledge and competence, resource acquisition, surrogacy (i.e., the family or adult "village" surrounding a child to assist when a parent needs help), and close surveillance of the child[3].

A prototypic, yet remarkably common risk scenario is that of a single parent with multiple young children, isolated by poverty, under-educated in the modeling of appropriate caregiving (or whose own experience in being parented was traumatic or deficient) and with either an untreated mental health impairment or substance use disorder.

An effective, evidence-informed approach to reduce the risk of child maltreatment imposed by this set of circumstances would include nurse (or paraprofessional) home visitation, parenting education, parental mental health care, a support resource for times of crisis, and reproductive health planning. This is analogous to the level of comprehensive intervention that is afforded to patients with complex medical disorders in most health systems, encompassing cost-efficient, evidence-based interventions that could be prioritized for families at risk and coordinated by efficient, targeted case management.

Yet, rarely does any family at risk receive a full complement of these necessary supports[3,8]. In the US, fragmentation across health agencies, state departments, and local bureaucracies, together with a lack of ownership of systematic risk surveillance by health systems, all but ensure that almost no family at risk ever receives this level of support. The end result is that child maltreatment is perpetrated at epidemic proportions: a conservative estimate of prevalence based on official records is that, in the US, one out of every six children is a victim[2].

Not all children succumb to the deleterious impact of maltreatment. Rather, the effects of trauma on brain and behavioral development are moderated by factors such as timing of occurrence over the course of childhood; severity, type and chronicity of maltreatment; and genotypic variation of victims. These factors render children more or less prone to becoming overwhelmingly biologically stressed by the adverse experience. It is the phenomenon of being stressed beyond capac-

ity to compensate or respond adaptively that is believed to be most salient in the neurobehavioral toxicity of trauma, so-called "toxic stress"[9].

Specific inherited profiles of temperament render some children susceptible to overwhelming anxiety in response to trauma, and others to patterns of impulsive aggression. Jonson-Reid et al[5] observed dose-response effects of the number of officially reported instances of abuse/neglect on an array of child and adult mental health outcomes in an unbiased, state-wide ascertainment.

An important paradox about the role of child maltreatment in the causation of psychiatric syndromes is the overarching conclusion from large twin and family studies that environmental variations – within the typical range observed in the general population – tend *not* to be as influential in the causation of serious mental health impairments as either genetic factors or severe environmental adversities outside of the typical range.

Construed graphically, if plotting severity of adversity (X axis) against degree of psychiatric impairment (Y axis), incremental increases in adversity *within the typical range* result in only minimal increases along the Y axis of psychiatric impairment. But at an inflection point – that varies for each child on the basis of genetic vulnerability or resilience – the Y axis "cost" steepens with further incremental increases in adversity, and levels off (forming a sigmoid curve) at a point beyond which impairment is so pronounced that further increases in stress have negligible additional effect.

Improved ability to operationalize "level of adversity" in increasingly precise ways, and thereby specify a curve for an individual child, stands to revolutionize targeted preventive intervention by ensuring that each child remains within his/her "safe zone" along these two critical axes.

In the meantime, efforts to prevent the occurrence of maltreatment in all children and families with appreciable elevated risk – either for maltreatment or its consequences – represents a feasible strategy for reducing the public health burden of psychopathology. Given demonstrated progress in the ability to predict and prevent child maltreatment, health and governmental systems around the world have a new opportunity (and an ethical mandate) to deploy evidence-based elements of intervention at developmentally sensitive times during the life course, targeting multiple risks, and building on existing delivery platforms for feasibility of scale-up[2,8,10].

Parents, caregivers and families need to be supported in providing nurturing care and protection in order for young children to achieve their developmental potential, and – as emphasized by Shonkoff[9] – it must be understood and put into practice that healthy brain development requires not only enrichment, but protection from toxic stress.

John N. Constantino
Washington University School of Medicine, St. Louis, MO, USA

This work was supported in part by a grant from the US Administration for Children and Families (90YR0054-04).

1. Gilbert R, Widom CS, Browne K et al. Lancet 2009;373:68-81.
2. Wildeman C, Emanuel N, Leventhal JM et al. JAMA Pediatr 2014;168:706-13.
3. Constantino JN. Child Adolesc Psychiatr Clin N Am 2016;25:157-65.
4. Shaley I, Heim CM, Noll JG. JAMA Psychiatry 2016;73:897-8.
5. Jonson-Reid M, Kohl PL, Drake B. Pediatrics 2012;129:839-45.
6. Ferrara P, Corsello G, Basile MC et al. J Pediatr 2015;167:1457-9.
7. Putnam-Hornstein E, Cederbaum JA, King B et al. Am J Epidemiol 2015; 181:496-503.
8. Britto PR, Lye SJ, Proulx K et al. Lancet 2017;389:91-102.
9. Shonkoff JP. JAMA Pediatr 2016;1;170:1003-7.
10. Constantino JN, Ben-David V, Navsaria N et al. Am J Psychiatry 2016;173: 566-73.

DOI:10.1002/wps.20495

Mental health of children living in war zones: a risk and protection perspective

Armed conflicts have a devastating impact on the mental health of affected populations. Post-traumatic stress disorder (PTSD) and depression are the most common mental disorders in the aftermath of war for both adults and children, occurring in up to one third of the people directly exposed to traumatic war experiences[1]. Exposure to traumatic events is the most important risk factor in this context. However, for children in particular, the detrimental effects of war trauma are not restricted to specific mental health diagnoses, but include a broad and multifaceted set of developmental outcomes that compromise family and peer relations as well as school performance and general life satisfaction.

To understand a child's development in a war or post-war environment, we have to apply a socio-ecological perspective[2], which takes into account not only the direct consequences of the war for the individual child, but also variables in the proximal and distal environments, including the family and the community[3]. Today's wars almost exclusively affect low-resource countries and are typically associated with a number of risk factors at various ecological levels, e.g. extreme poverty, a lack of resources for health provisioning, a breakdown of the school system, as well as increased rates of family and community violence. Children are particularly sensitive to such an accumulation of stressors: in fact, there is considerable evidence for a dose-response relation between the amount of stressors experienced by children and their impairments in different areas of adaptation, such as mental and physical health, academic achievement and social relationships[4].

Family functioning seems to play a key role in the interplay of risk and protection factors across ecological levels. War is associated with elevated levels of family violence against children[5] as well as increased rates of intimate partner violence

against women[6]. In addition, violence related to both the war and family conflicts contributes independently to children's psychopathology. This includes PTSD, depression symptoms as well as internalizing and externalizing behavior problems[4].

A key question refers to the mechanisms behind this "cycle of violence" in the aftermath of war. How are the exposure to violent conflict and increased rates of child maltreatment interlinked? So far, studies have focused mainly on intergenerational effects, i.e. parental trauma and psychopathology as potential mediators. Evidence suggests that exposure to organized violence and psychopathology associated with these experiences might act as a catalyst for domestic violence and child maltreatment. In particular, PTSD symptoms, such as irritability and outbursts of anger, as well as elevated rates of alcohol consumption in parents, may contribute to higher levels of child abuse. In line with this hypothesis, studies in post-war Sri Lanka and Uganda have shown that, next to parents' own experiences of child abuse, children's reports of maltreatment were associated with the parents' exposure to war and their PTSD symptom severity as well as with male guardian's alcohol consumption[7].

Research, so far, has neglected a further pathway by which war trauma could translate into increased levels of family violence. It might be the child's own war exposure and related psychopathology that increase the risk of experiencing violence at home. Children who grow up in the midst of war are at greater risk of developing challenging behavior problems associated with their traumatization, e.g. irritability, outbursts of anger, internalizing and externalizing symptoms. Their mental health problems are typically accompanied by functional impairments that compromise their ability to perform well at school, carry out household duties, and engage in social relationships. All of these difficulties could make war-traumatized children more challenging to manage for their parents, who, in turn, may apply more violent and coercive parenting strategies. Consistent with this hypothesis, a recent study with Tamil families in post-war Sri Lanka found that children's exposure to mass trauma and child psychopathology were the main predictors of children's self-reported victimization in their families, even after controlling for parental trauma and parental mental health[5].

The notion that stressors from different ecological contexts interact with each other is supported by earlier longitudinal data on maltreated children, which showed that children's externalizing behavior uniquely predicted later exposure to community violence[8]. These findings have important implications for future research with war-affected children and their families. Instead of focusing on mental health problems as a mere outcome of war trauma in children, they should be considered as a potential risk factor for the experience of further adversities at a different ecological level, i.e. the family.

Applying a risk and protection perspective to the study of child mental health in a post-war context requires considering potentially protective factors that, again, may be found at various ecological levels. The family in particular may not only act as a stressor, in the case of family violence, but also foster children's resilience through care and warmth. There is some evidence that this is also valid in war-torn populations. Sriskandarajah et al[9] showed that, in a context of multiple trauma caused by war and natural disaster, parental care moderates the relation between children's trauma severity and their internalizing behavior problems. Children who reported their parents to be highly caring did not show a significant increase in internalizing problems related to exposure to mass trauma. Likewise, data from families in post-war Uganda revealed that the effect of war trauma on children's psychopathology was partially mediated by lower child-perceived care from female guardians[10].

We can conclude that children and families living in or fleeing war regions have a high probability of suffering from mental health problems. This is because they are confronted with an accumulation of risk factors at different socio-ecological levels. Parenting practices seem to play a crucial role for children's psychological wellbeing in a war context, both as a risk and a protective factor. Consequently, adequate health care programs for war-traumatized communities require both individual and family level approaches. The latter would assess and address potential problems between parents as well as in parent-child relationships. This might halt a potential vicious circle of war trauma, psychopathology and dysfunctional family dynamics, including the maltreatment of women and children.

Claudia Catani
Department of Psychology, Bielefeld University, Bielefeld, Germany

1. Steel Z, Chey T, Silove D et al. JAMA 2009;302:537-49.
2. Bronfenbrenner U. The ecology of human development. Cambridge: Harvard University Press, 1979.
3. Reed RV, Fazel M, Jones L et al. Lancet 2017;379:250-65.
4. Catani C, Gewirtz AH, Wieling E et al. Child Dev 2010;81:1176-91.
5. Sriskandarajah V, Neuner F, Catani C. Soc Sci Med 2015;146:257-65.
6. Clark CJ, Everson-Rose SA, Suglia SF et al. Lancet 2017;375:310-6.
7. Saile R, Ertl V, Neuner F et al. Child Abus Negl 2014;38:135-46.
8. Lynch M, Cicchetti D. Dev Psychopathol 1998;10:235-57.
9. Sriskandarajah V, Neuner F, Catani C. BMC Psychiatry 2015;15:203.
10. Saile R, Ertl V, Neuner F et al. Dev Psychopathol 2016;28:607-20.

DOI:10.1002/wps.20496

Hikikomori: experience in Japan and international relevance

The appearance of people in Japan, especially young men, who stopped going to school or the workplace and spent most of the time withdrawn into their homes for months or years, came to be seen as an increasing social phenomenon called *Shakaiteki hikikomori* (social withdrawal) by the late 1990s[1].

A community-based survey published in 2010 reported that

the prevalence of hikikomori was approximately 1.2% of the Japanese population[2], and in 2016 a Japanese cabinet report estimated people with hikikomori to be about 541,000 within the age range of 15-39 years.

Early epidemiological studies were limited by not being based on strict diagnostic standards. In 2010, Japan's Ministry of Health, Labour and Welfare announced a guideline for hikikomori which included a definition ("a situation where a person without psychosis is withdrawn into his/her home for more than six months and does not participate in society such as attending school and/or work")[3]. More recently, in order not only to diagnose but also to assess the severity of the condition, we proposed even more precise diagnostic criteria based on the levels of physical isolation at home, avoidance of social interactions, and functional impairment or distress, as well as a sustained duration of six months or more[4].

The Japanese sociocultural background has been traditionally permeated by "*amae*" (accepting overdependent behaviors) and shame, which may underlie the culture-bound syndrome called *Taijin Kyofusho* (a severe form of social phobia) as well as hikikomori[5,6]. Parent-child relationships in Japan have long been considered less oedipal than in Western societies and marked by an absent father and an extremely prolonged and close bond to the mother, which may result in difficulty to become independent[7]. Especially in hikikomori, the development of basic interpersonal skills during the early stages of life seems to be insufficient, which can induce vulnerability to stress in later school/workplace environments and lead to escape from social situations[7].

On the other hand, hikikomori-like cases have recently been reported in other countries of varying sociocultural and economic backgrounds such as Hong Kong, Oman and Spain, and our studies based on structured interviews have revealed the existence of hikikomori in India, South Korea and the US[4]. Thus, hikikomori has now crossed the limits of a culture-bound phenomenon to become an increasingly prevalent international condition. A major contributing factor may be the evolution of communication from direct to increasingly indirect and physically isolating[8]. This is especially the case for social interactions which hitherto required face-to-face contacts in a mutual physical space but can now occur, at least partially, in a virtual world.

Through our recent study using the Structured Clinical Interview for DSM-IV Axis I Disorders, we have found that hikikomori may be comorbid with various psychiatric disorders, including avoidant personality, social anxiety disorder and major depression[9]. In addition, autistic spectrum disorders and latent or prodromal states of schizophrenia may have some overlapping symptomatology with hikikomori. Thus, hikikomori is now understood to have links to several mental illnesses, and we hypothesize that some common psychopathological mechanisms may exist in the act of "shutting-in" regardless of psychiatric diagnosis.

Currently, there are more than fifty government-funded community support centers for hikikomori located throughout the prefectures of Japan, providing services such as telephone consultations for family members, the creation of "meeting spaces" for affected people, and job placement support. In addition, various private institutions provide treatment for hikikomori sufferers. However, there is yet to be a unified evidence-based method for these public/private interventions. A 4-step intervention is recommended by the government guideline for hikikomori, including family support and first contact with the individual and his/her evaluation; individual support; training through an intermediate-transient group situation (such as group therapy); and social participation trial[3].

We have recently established a hikikomori clinical research unit in a university hospital to develop evidence-based therapeutic approaches in collaboration with public/private hikikomori support centers. As a first step, we are trying to establish an evidence-based educational program for parents of individuals with hikikomori, because in the majority of cases the first consultation is made by them. Due to prejudice and lack of knowledge, in many cases family members cannot respond directly to individuals with this problem, are unable to intervene at all, and tend to turn a blind eye for many years without seeking help. Thus, we believe that education of parents to deal with hikikomori sufferers is essential for early intervention.

Within decades, following further advances in Internet society, more and more people may come to live a hikikomori-like existence, which may or may not be seen as a pathological condition at that time. Hikikomori is still a hidden epidemic in many countries and, to grasp its worldwide relevance, diagnostic criteria should be included in ICD-11 and future DSM systems. In addition, evidence-based evaluation tools such as structured diagnostic interviews, screening instruments and online systems should be developed for international and population-level epidemiological surveys. Such tools will also help to evaluate risk factors and effectiveness of interventions.

Takahiro A. Kato[1], Shigenobu Kanba[1], Alan R. Teo[2]
[1]Department of Neuropsychiatry, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; [2]Department of Psychiatry, Oregon Health and Science University, Portland, OR, USA

1. Saito T. Shakaiteki hikikomori: owaranai shishunki. Tokyo: PHP Shinsho, 1998.
2. Koyama A, Miyake Y, Kawakami N et al. Psychiatry Res 2010;176:69-74.
3. Saito K. Hikikomori no hyouka: shien ni kansuru gaidorain. Tokyo: Japan's Ministry of Health, Labour and Welfare, 2010.
4. Teo AR, Fetters MD, Stufflebam K et al. Int J Soc Psychiatry 2015;61:64-72.
5. Kato TA, Tateno M, Shinfuku N et al. Soc Psychiatry Psychiatr Epidemiol 2012;47:1061-75.
6. Kato TA, Kanba S, Teo AR. Am J Psychiatry 2016;173:112-4.
7. Kato TA, Hashimoto R, Hayakawa K et al. Psychiatry Clin Neurosci 2016; 70:7-23.
8. Kato TA, Kanba S. Psychiatry Clin Neurosci 2016;70:1-2.
9. Teo AR, Stufflebam K, Saha S et al. Psychiatry Res 2015;228:182-3.

DOI:10.1002/wps.20497

Psychosis-risk criteria in the general population: frequent misinterpretations and current evidence

Research on early detection and intervention in psychosis is going into its third decade. Its results are increasingly transferred into the clinic[1,2], and a tentative syndrome modeled on the description of "attenuated psychotic symptoms" (APS) in the ultra-high risk criteria has been included in the Section III ("Conditions for further study") of the DSM-5.

A wealth of evidence suggests: a) that symptomatic psychosis-risk criteria – in particular the APS and "brief intermittent psychotic symptoms" (BIPS) criteria, as well as the basic symptom criterion "cognitive disturbances" (COGDIS) – are associated with a significantly increased risk for psychosis in clinical samples, even in comparison with patients without psychosis-risk criteria from the same services[1], and b) that specific psychological and pharmacological interventions reduce conversion rates to psychosis in adult patients with psychosis-risk criteria relative to control conditions and improve psychosocial functioning, although not to a significantly larger degree than control conditions[2].

Nevertheless, this indicated preventive approach continues to be criticized when considered through the lens of epidemiological findings[3]. Thereby, the focus is predominately on the APS criterion[4]. A main reason for the disparity between clinically and epidemiologically based viewpoints obviously originates from differences in assessments. Until recently, epidemiological studies have mainly used self-rating questionnaires or standardized lay-person interviews for the assessment of "psychotic-like experiences" (PLEs)[4,5], that are usually equaled to (attenuated) psychotic symptoms by critics[3]. Yet, questionnaire studies significantly overestimate the prevalence of PLEs already in comparison to lay-person interview studies[5], and even more in comparison to clinician-based evaluations of APS using psychosis-risk assessment instruments[6].

Thus, PLEs are not an adequate proxy for APS/BIPS examined in clinical studies, and the conclusion – based on a wrongly assumed equality of these phenomena – that psychotic experiences are "a transdiagnostic dimension of psychopathology" and "a marker for the severity of non-psychotic states"[3] must be regarded as unfounded when related to APS/BIPS assessed in clinical psychosis-risk research.

Moreover, epidemiological studies usually assess the presence of PLEs but not their course or frequency, thus ignoring crucial requirements of APS/BIPS criteria. Indeed, the first sufficiently representative general population study (N=2,683) of young adults (age 16-40) in whom psychosis-risk criteria[4] were assessed by trained clinicians – using established early detection instruments – documented that, although 11.96% (N=321) reported any lifetime and 7.53% (N=202) any current APS/BIPS, only 0.56% (N=15) fulfilled APS criteria (including onset or worsening within the past 12 months and at least weekly occurrence in the past month). One person meeting APS criteria also fulfilled BIPS criteria (including onset within the past three months and at least monthly occurrence for several minutes)[4].

Thus, the fact that (attenuated) psychotic phenomena may occasionally occur in persons of the general population with no past or present psychotic disorder does not rule out the significance of APS criteria as denoting a distinct and rather rare syndrome if additional course and frequency requirements are met. Such requirements are commonly part of the definition of mental disorders, when they are based on phenomena that might as well occur occasionally in everyday life, such as feeling low, sad and hopeless, or euphoric, or being afraid.

Another feature that psychosis-risk criteria share with mental disorders is their frequent co-occurrence with other disorders, in particular depressive and anxiety disorders, with a prevalence of 23-94% in clinical[1,2] and 45% in general population samples[4]. Arguing that APS are solely a marker of the severity of non-psychotic disorders[3] disregards the fact that depression and anxiety often arise from other mental or somatic disorders[7], including non-affective psychoses[8].

In fact, a review of psychiatric comorbidity across different stages of schizophrenia confirmed the frequent co-occurrence of anxiety and depressive disorders throughout the course of the illness, including the prodrome[8]. From this, it was concluded that depressive and certain anxiety symptoms are "intrinsic to the illness and import a poorer outcome"[8]. Thus, the co-occurrence of psychosis-risk criteria and other mental disorders might indeed be "summarized as baseline differences in the severity of multidimensional psychopathology"[3], yet with depression and/or anxiety rather than the far more infrequent APS/BIPS serving as transdiagnostic markers of severity.

Longitudinal community studies using valid psychopathological assessments that are well comparable with clinical ones are needed to shed more light on the interplay and sequence of psychosis-risk criteria and non-psychotic mental disorders over time.

Notably, in the critique of the psychosis-risk approach from an epidemiological viewpoint[3], the basic symptom approach – in particular COGDIS – that next to APS and BIPS was recently recommended for the assessment of psychosis risk in the context of a guidance paper of the European Psychiatric Association[1] – usually remains unmentioned. COGDIS and APS are equally common in clinical[1,9] and community samples[4], often co-occur and, together, convey a much increased risk for psychosis, but not for other mental disorders, in clinical samples[9]. Yet, for their distinct quality[1,4,10], cognitive basic symptoms cannot be referred to as "low-grade psychotic symptoms" or "psychotic experiences"[3] and, consequently, critiques based on findings on PLEs cannot be extended to them.

To conclude, while it is undisputed that more epidemiological as well as clinical research is needed to further improve pre-

diction and prevention of psychosis[1,2] and to disentangle the dynamic relationship between psychosis-risk criteria and mental disorders, much of the recent critique of the psychosis-risk approach[3] reflects prejudice and misperceptions of epidemiological but also of clinical findings and is not in agreement with current evidence.

Frauke Schultze-Lutter[1,2], Joachim Klosterkötter[3], Wolfgang Gaebel[2], Stefanie J. Schmidt[1,3]
[1]University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland; [2]Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany; [3]Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany

1. Schultze-Lutter F, Michel C, Schmidt SJ et al. Eur Psychiatry 2015;30:405-16.
2. Schmidt SJ, Schultze-Lutter F, Schimmelmann BG et al. Eur Psychiatry 2015;30:388-404.
3. van Os J, Guloksuz S. World Psychiatry 2017;16:200-6.
4. Schultze-Lutter F, Michel C, Ruhrmann S et al. Psychol Med (in press).
5. Linscott RJ, van Os J. Psychol Med 2013;43:1133-49.
6. Schultze-Lutter F, Renner F, Paruch J et al. Psychopathology 2014;47:194-201.
7. Melartin TK, Rytsälä HJ, Leskelä US et al. J Clin Psychiatry 2002;63:126-34.
8. Buckley PF, Miller BJ, Lehrer DS et al. Schizophr Bull 2009;35:383-402.
9. Michel C, Ruhrmann S, Klosterkötter J et al. Eur Arch Psychiatry Clin Neurosci (in press).
10. Schultze-Lutter F, Theodoridou A. World Psychiatry 2017;16:104-5.

DOI:10.1002/wps.20498

Bridging the dichotomy of actual versus aspirational digital health

The future of digital health is bright. Enthusiasm for smartphone apps, virtual reality, artificial intelligence, machine learning and more in health care is no longer a niche interest, but rather mainstream fascination. From patient groups creating technology solutions for long-term conditions[1], to large technology companies like Google entering the digital health space[2], digital health enables new voices and perspectives to partake in advancing health care. But with such broad enthusiasm and so many voices active in digital health, it can also be overwhelming to separate out actual from aspirational claims of progress.

Attempting to delineate digital health fact from fiction is somewhat futile, given that the border is constantly shifting with new discoveries. Rather, it is more productive to consider a spectrum of actual to aspirational digital health claims. Here we outline two simple factors to facilitate decisions about where claims lie on that spectrum: a) the appropriateness for the target population, and b) the incentives used to obtain any given results.

When evaluating any digital health claim, it is critical to consider if the population the technology claims to help matches the actual population that was studied. While on the surface this sounds trivial, in the age of Internet crowdsourcing research, it is a growing concern. One of the most difficult aspects of traditional clinical research is recruiting a sufficient number of patients to partake in a study. The Internet offers a potential solution, where an ad on Facebook or Craigslist can bring in hundreds of potential research subjects willing to fill out a survey or try a new health app[3]. But who are these online subjects, that are never actually even seen by the research team?

In mental health research, there has been a trend to offer simple screening tools, such as the Patient Health Questionnaire-9 (PHQ-9), never intended to diagnose, as diagnostic inclusion criteria[4]. Without verification or ruling out of other physical or mental conditions, it is increasingly easy to join an online mental health study even if someone does not actually have a mental health condition. A similar concern is digital health claims related to symptoms of an illness obtained from those likely without the illness. What does it mean to study individual symptoms of post-traumatic stress disorder from those who may not meet the diagnostic criteria of having the actual condition[5]?

While there is tremendous value in crowdsourced and online-based health research as well as in studying non-traditional metrics or classifications of illness, it is critical to be cognizant of the differences from traditional research. In some cases these online methods may actually be superior to traditional face-to-face research. However, determining if these results are actually applicable to established definitions of illness is important to consider when deciding if these novel approaches can genuinely improve patient outcomes in a clinical setting. Novel approaches creating new definitions of illness or identifying new populations at risk of illness are equally important, but by their nature require further validation and are less actionable today.

Another factor to consider when evaluating any digital health claim is the role of incentives in obtaining that result. Again, on the surface this sounds trivial, but the complexities of the digital landscape add new challenges. Digital health research offers participants incentives to partake that can range from new smartphones, money for using of the device, extra coaching sessions, and more. But what happens when the digital health platform enters the real world, when the incentives disappear and when there is no external attention or interest in one's use of the technology? A recent study of an asthma-monitoring app reported that, while over 49,000 people downloaded the study's app, only 175 (0.35%) were actively engaged with it at six months[5]. A successful study of an app for alcohol use disorder[6] reported less success when later deploying without broad incentives[7].

Thus, understanding the context and incentives used to generate a positive outcome from a novel digital health intervention is important in deciding if that intervention can realistically be implemented today, or if it is more aspirational, with new resources and efforts required to sustain engagement. That is not to say that outcomes from research projects which incentivize participants are invalid or redundant – rather, these findings provide valuable insights into what is necessary to make digital health work and how the health care system may have to evolve

to support it. However, these factors must be considered when deciding what is immediately implementable, versus that which requires a supportive framework which has yet to be created.

All digital health research and claims are informative. Some offer immediate solutions to health care that should be implemented today and others highlight the potential of what may be possible. However, blurring the line between actual and aspirational can be counterproductive. Claiming that aspirational digital health research is ready for immediate use can lead to immediate negative results and broad disappointment. It may even inadvertently contribute to digital health "hype" and foster undue skepticism for the field.

However, ignoring digital health technologies with good evidence for real-world implementation is a missed opportunity for improving patient outcomes. Appreciating how aspirational research can guide, inform, and inspire current efforts is also important. Likewise, appreciating the real world success of actualized efforts can help guide aspirational research to be more translatable into health care systems.

There is no superior designation, as both ends of the actual and aspirational spectrum have critical roles that cannot be separated. However, the value of both depends upon correct identification of where any given project lies on this spectrum – and further consideration of populations sampled and incentives used are critical to determining this.

John Torous[1], Joseph Firth[2]
[1]Department of Psychiatry and Division of Clinical Informatics, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; [2]NICM, School of Science and Health, Western Sydney University, Sydney, Australia

1. Lee JM, Hirschfeld E, Wedding J. JAMA 2016;315:1447-8.
2. Eyre HA, Singh AB, Reynolds C. World Psychiatry 2016;15:21-2.
3. Walch OJ, Cochran A, Forger DB. Sci Adv 2016;2:e1501705.
4. Kroenke K, Spitzer RL, Williams JB. J Gen Intern Med 2001;16:606-13.
5. Place S, Blanch-Hartigan D, Rubin C et al. J Med Internet Res 2017;19:e75.
6. Gustafson DH, McTavish FM, Chih MY et al. JAMA Psychiatry 2014;71:566-72.
7. Ford JH II, Alagoz E, Dinauer S et al. J Med Internet Res 2015;17:e201.

DOI:10.1002/wps.20464

Compulsive sexual behaviour disorder in the ICD-11

During the last decade, there has been heated debate regarding whether compulsive sexual behaviour should be classified as a mental/behavioural disorder. Compulsive sexual behaviour disorder has been proposed for inclusion as an impulse control disorder in the ICD-11[1]. It is characterized by a persistent pattern of failure to control intense, repetitive sexual impulses or urges, resulting in repetitive sexual behaviour over an extended period (e.g., six months or more) that causes marked distress or impairment in personal, family, social, educational, occupational or other important areas of functioning.

The pattern is manifested in one or more of the following: a) engaging in repetitive sexual activities has become a central focus of the person's life to the point of neglecting health and personal care or other interests, activities and responsibilities; b) the person has made numerous unsuccessful efforts to control or significantly reduce repetitive sexual behaviour; c) the person continues to engage in repetitive sexual behaviour despite adverse consequences (e.g., repeated relationship disruption, occupational consequences, negative impact on health); or d) the person continues to engage in repetitive sexual behaviour even when he/she derives little or no satisfaction from it.

Concerns about overpathologizing sexual behaviours are explicitly addressed in the diagnostic guidelines proposed for the disorder. Individuals with high levels of sexual interest and behaviour (e.g., due to a high sex drive) who do not exhibit impaired control over their sexual behaviour and significant distress or impairment in functioning should not be diagnosed with compulsive sexual behaviour disorder. The diagnosis should also not be assigned to describe high levels of sexual interest and behaviour (e.g., masturbation) that are common among adolescents, even when this is associated with distress.

The proposed diagnostic guidelines also emphasize that compulsive sexual behaviour disorder should not be diagnosed based on psychological distress related to moral judgments or disapproval about sexual impulses, urges or behaviours that would otherwise not be considered indicative of psychopathology. Sexual behaviours that are egodystonic can cause psychological distress; however, psychological distress due to sexual behaviour by itself does not warrant a diagnosis of compulsive sexual behaviour disorder.

Careful attention must be paid to the evaluation of individuals who self-identify as having the disorder (e.g., calling themselves "sex addicts" or "porn addicts"). Upon examination, such individuals may not actually exhibit the clinical characteristics of the disorder, although they might still be treated for other mental health problems (e.g., anxiety, depression). Additionally, individuals often experience feelings such as shame and guilt in relationship to their sexual behaviour[2], but these experiences are not reliably indicative of an underlying disorder.

The proposed diagnostic guidelines also assist the clinician in differentiating compulsive sexual behaviour disorder from other mental disorders and other health conditions. For example, although bipolar disorder has been found at elevated rates among individuals with compulsive sexual behaviour disorder[3], sexual behaviours must be persistent and occur independently of hypomanic or manic episodes to provide a basis for a possible diagnosis of the disorder. A diagnosis of compulsive sexual behaviour disorder should not be made when the behaviour can be explained by other medical conditions (e.g., dementia) or by the effects of certain medications prescribed to treat specific medical conditions (e.g., Parkinson's disease)[4] or is entirely attributable to the direct effects of illicit substances

on the central nervous system (e.g., cocaine, crystal methamphetamine).

Currently, there is an active scientific discussion about whether compulsive sexual behaviour disorder can constitute the manifestation of a behavioural addiction[5]. For ICD-11, a relatively conservative position has been recommended, recognizing that we do not yet have definitive information on whether the processes involved in the development and maintenance of the disorder are equivalent to those observed in substance use disorders, gambling and gaming[6]. For this reason, compulsive sexual behaviour disorder is not included in the ICD-11 grouping of disorders due to substance use and addictive behaviours, but rather in that of impulse control disorders. The understanding of compulsive sexual behaviour disorder will evolve as research elucidates the phenomenology and neurobiological underpinnings of the condition[7].

In the absence of consistent definitions and community-based epidemiological data, determining accurate prevalence rates of compulsive sexual behaviour disorder has been difficult. Epidemiological estimates have ranged up to 3-6% in adults[8], though recent studies have produced somewhat lower estimates of 1 to 3%[9]. The more restrictive diagnostic requirements proposed for ICD-11 would be expected to produce lower prevalence rates.

In general, men exhibit the disorder more frequently than women, although robust data examining gender differences are lacking. Additionally, higher rates of the disorder have been noted among individuals with substance use disorders. Among treatment seekers, the disorder negatively impacts occupational, relationship, physical health and mental health functioning. However, systematic data are lacking regarding the prevalence of the disorder across different populations and associated socio-cultural and socio-demographic factors, including among non-treatment seekers.

Growing evidence suggests that compulsive sexual behaviour disorder is an important clinical problem with potentially serious consequences if left untreated. We believe that including the disorder in the ICD-11 will improve the consistency with which health professionals approach the diagnosis and treatment of persons with this condition, including consistency regarding when a disorder should not be diagnosed. Legitimate concerns about overpathologizing sexual behaviours have been carefully addressed in the proposed diagnostic guidelines. We posit that inclusion of this category in the ICD-11 will provide a better tool for addressing the unmet clinical needs of treatment seeking patients as well as possibly reduce shame and guilt associated with help seeking among distressed individuals.

The proposed diagnostic guidelines will be tested in international multilingual Internet-based field studies using standardized case material, which will help to assess the generalizability of the construct across different regions and cultures, and clinicians' ability to distinguish it from normal variations in sexual behaviour and from other disorders. Additional field studies in clinical settings will provide further information about the clinical utility of the proposed diagnostic guidelines for the disorder among clinical populations.

Shane W. Kraus[1], Richard B. Krueger[2], Peer Briken[3], Michael B. First[2], Dan J. Stein[4], Meg S. Kaplan[2], Valerie Voon[5], Carmita H.N. Abdo[6], Jon E. Grant[7], Elham Atalla[8], Geoffrey M. Reed[9,10]

[1]Edith Nourse Rogers Memorial Veterans Hospital, Bedford, MA, USA; [2]Department of Psychiatry, Columbia University, College of Physicians and Surgeons and New York State Psychiatric Institute, New York, NY, USA; [3]Institute for Sex Research and Forensic Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [4]Department of Psychiatry, University of Cape Town and Groote Schuur Hospital, Cape Town, South Africa; [5]Department of Psychiatry, University of Cambridge, Cambridge, UK; [6]Department of Psychiatry, Faculty of Medicine, University of São Paulo, São Paulo, Brazil; [7]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA; [8]Primary Care and Public Health Directorate, Ministry of Health, Manama, Bahrain; [9]Department of Mental Health and Substance Abuse, World Health Organization, Geneva, Switzerland; [10]Global Mental Health Program, Columbia University Medical Center, New York, NY, USA

1. Grant JE, Atmaca M, Fineberg NA et al. World Psychiatry 2014;13:125-7.
2. Gilliland R, South M, Carpenter BN et al. Sex Addict Compulsivity 2011;18:12-29.
3. Raymond NC, Coleman E, Miner MH. Compr Psychiatry 2003:44:370-80.
4. Weintraub D, Koester J, Potenza MN et al. Arch Neurol 2010;67:589-95.
5. Griffiths MD. Addict Res Theory 2012:20:111-24.
6. Kraus SW, Voon V, Potenza MN. Addiction 2016;111:2097-106.
7. Kraus SW, Voon V, Potenza MN. Neuropsychopharmacology 2016;41:385-6.
8. Kuzma JM, Black DW. Psychiatr Clin N Am 2008;31:603-11.
9. Klein V, Rettenberger M, Briken P. J Sex Med 2014;11:1974-81.

DOI:10.1002/wps.20499

Decline in suicide mortality after psychiatric hospitalization for depression in Finland between 1991 and 2014

Depression is the most important mental disorder in terms of suicide mortality. Numerous studies over time have estimated the lifetime risk of suicide in depression, including a recent Danish national study[1]. Organization of services and treatment practices for depression have undergone major changes over the past decades, including remarkable growth in the use of antidepressants, emphasis on community-based services, and deinstitutionalization. Temporal trends in suicide mortality among psychiatric patients with depression can be expected, but have not been investigated.

We followed a Finnish population-based cohort of depressive patients (N=56,826), with a first lifetime hospitalization due to depression between 1991 and 2011, up to the end of the year 2014 (maximum follow-up: 24 years). Here we report both cumulative risk of suicide and temporal trends in suicide mortality.

This study stems from the MERTTU research project[2]. Complete data at individual level via the Finnish identity codes were linked from the Finnish Hospital Discharge Register, containing data on all inpatient treatments, and Statistics Finland's register on causes of death.

We identified from the Finnish Hospital Discharge Register all ≥18-year-old people with a psychiatric diagnosis admitted to a psychiatric hospital or a psychiatric ward of a general hospital between 1987 and 2011. We next obtained data on patients' hospitalizations with psychiatric diagnoses between 1980 and 2011. Baseline hospitalizations for depressive disorder (principal diagnosis) in 1991-2011 were identified through the Finnish ICD-9[3] codes 2961A-G and 2968A (used in 1987-1995) and ICD-10 codes F32-33. Since 1987, the national guidelines have stipulated applying operationalized criteria for clinical psychiatric diagnoses (the Finnish ICD-9[3] was based on the DSM-III-R criteria and the ICD-10 on the Diagnostic Criteria for Research). The Finnish Hospital Discharge Register displays complete coverage and good accuracy of mental health diagnoses[4].

We excluded patients with previous psychiatric hospitalizations since 1980, with a principal diagnosis of psychotic disorder at baseline, or who died by suicide during the baseline hospitalization.

We retrieved the dates and causes of death for all cases, and then identified suicides (ICD-9 codes E950A-K, E951A-E957A, E959A-C, E959X; ICD-10 codes X60-X76, X78, X80-X84, Z91.5, Y87.0). Finland has high medico-legal autopsy rates (performed by a forensic pathologist who identifies all suicide and unnatural deaths). Overall, the death investigation process leaves few undetermined deaths[5].

Patients were followed up from the day of discharge to death by suicide or other cause, or until December 31, 2014, whichever occurred first. Events other than suicides were treated as censored. Diagnostic conversions to a principal psychotic or bipolar disorder were ignored because of risk of inducing survival bias.

The survival function and cumulative risk of suicide were estimated with the Kaplan-Meier product limit estimator. For time-trend analyses, we formed consecutive 5-year cohorts by admission year. We estimated for each cohort (years 1991-1995 as reference) the age- and gender-adjusted proportional hazards for suicide over 3 years (equal length of individual follow-ups) and maximum 24-year follow-up (varying length of individual follow-ups). We used software packages R and Survo.

A national cohort of 56,826 patients (25,188 men and 31,638 women) with first lifetime hospitalization for depression was followed for 628,514 person-years (follow-up: mean 11.1 years, median 10.7 years, maximum 24 years). Of 15,063 patients who died during the follow-up, 2,567 (17.0%) died by suicide (1,598 men, 969 women). The cumulative risk of suicide was 6.13% overall (95% CI: 5.80-6.46%), 8.64% in men (95% CI: 8.00-9.27%), and 4.14% in women (95% CI: 3.83-4.45%). The suicide incidence rate in men was 23.05 per 1,000 person-years (95% CI: 21.20-25.02) for the first 12 months; 8.84 per 1,000 person-years (95% CI: 7.69-10.10) for 12-24 months; and 6.12 per 1,000 person-years (95% CI: 5.17-7.20) for 24-36 months. The corresponding rates in women were 9.73 per 1,000 person-years (95% CI: 8.68-10.87), 3.82 per 1,000 person-years (95% CI: 3.17-4.57), and 3.19 per 1,000 person-years (95% CI: 2.60-3.88).

Relative to baseline years 1991-1995, the age- and gender-adjusted hazard ratio for suicide within 3 years post-discharge was 0.69 (95% CI: 0.61-0.79, $p<0.0001$) in 1996-2000, 0.54 (95% CI: 0.47-0.63, $p<0.0001$) in 2001-2005, and 0.48 (95% CI: 0.42-0.56, $p<0.0001$) in 2006-2011. The corresponding hazard ratio for maximum 24-year follow-up was 0.70 (95% CI: 0.63-0.77, $p<0.0001$) in 1996-2000, 0.57 (95% CI: 0.51-0.64, $p<0.0001$) in 2001-2005, and 0.49 (95% CI: 0.43-0.55, $p<0.0001$) in 2006-2011.

These Finnish cumulative risks of 8.6% in men and 4.1% in women are slightly higher than the cumulative Danish incidences of 6.7% in male and 3.8% in female depressive in- and outpatients[1]. The difference likely reflects higher overall suicide mortality in Finland, and our inclusion of only inpatients. Overall, hospital samples may overestimate the suicide mortality in depression by about 30-50%[6].

Our findings show a considerable and consistent decline in long-term suicide mortality since 1991, in contrast with the results of a recent meta-analysis of discharged general psychiatric inpatients[7]. In Finland, the suicide rates peaked in 1990 and were approximately halved by 2014. Our data are thus consistent with this overall pattern.

Following worldwide trends, numerous changes have occurred in Finland since 1990. First, the national Suicide Prevention Project was implemented in the early 1990s. Second, the consumption of antidepressants has eight folded during the study period[8]. Third, per capita alcohol consumption rose in 1990-2005, showing a downtrend since 2007. Fourth, the deinstitutionalization process has resulted in a reduction of about 60% in the number of psychiatric beds (equalling in 2011 the average in countries which are part of the Organisation for Economic Co-operation and Development (OECD): 71 per 100,000 vs. 70 per 100,000)[8]. From 1994 to 2011, the number of inpatient days have halved, and the number of treated patients reduced by 10%[9]. Fifth, the availability of psychiatric outpatient care has improved and outpatient-oriented services are associated with lower suicide mortality[2].

We conclude that the cumulative suicide risk in depression depends on time period and context. The large downtrend in suicide mortality of psychiatric inpatients in Finland over current treatment era is encouraging for ongoing efforts to prevent suicides in depression.

Kari I. Aaltonen[1,2], Erkki Isometsä[1], Reijo Sund[3,4], Sami Pirkola[2,5]

[1]Department of Psychiatry, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; [2]Department of Public Health Solutions, Mental Health Unit, National Institute for Health and Welfare, Helsinki, Finland; [3]Centre for Research Methods, Faculty of Social Sciences, University of Helsinki, Helsinki, Finland; [4]Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland; [5]Faculty of Social Sciences, University of Tampere and Pirkanmaa Hospital District, Tampere, Finland

1. Nordentoft M, Mortensen PB, Pedersen CB. Arch Gen Psychiatry 2011;68: 1058-64.

2. Pirkola S, Sund R, Sailas E et al. Lancet 2009;373:147-53.
3. Kuoppasalmi K, Lönnqvist J, Pylkkänen K et al. Psychiatria Fennica 1989; 20:65-81.
4. Sund R. Scand J Publ Health 2012;40:505-15.
5. Ohberg A, Lonnqvist J. Acta Psychiatr Scand 1998;98:214-8.
6. Crump C, Ioannidis JP, Sundquist K et al. J Psychiatr Res 2013;47:1298-303.
7. Chung DT, Ryan CJ, Hadzi-Pavlovic D et al. JAMA Psychiatry 2017;74:694-702.
8. Organisation for Economic Co-operation and Development (OECD). http://stats.oecd.org.
9. Finnish National Institute for Health and Welfare. https://www.sotkanet.fi.

DOI:10.1002/wps.20501

Complex PTSD and its correlates amongst female Yazidi victims of sexual slavery living in post-ISIS camps

The atrocities committed by the Islamic State of Iraq and Syria (ISIS) are having vast psychological effects around the world[1,2]. The Yazidis, a Kurdish religious minority, have suffered the most at the hand of ISIS[3]. Many men have been executed, while many women have been captured and subjected to sexual slavery, experiencing repeated abuse and rape. Detrimental effects of torture and sexual abuse have been repeatedly documented in the literature[4,5], but the Yazidi genocide includes both elements, and has not been hitherto addressed. This preliminary study assessed post-traumatic stress disorder (PTSD) and complex post-traumatic stress disorder (CPTSD) among female Yazidi former captives residing in post-ISIS camps.

Following traumatic exposure, both PTSD and CPTSD may ensue. PTSD typically follows a single traumatic event, while CPTSD is associated with prolonged trauma where one's destiny is under another's control and escape is unfeasible[6]. According to the ICD-11 draft[6], PTSD comprises three symptoms: re-experiencing, avoidance and arousal. CPTSD includes three more symptoms pertaining to disturbances in self-organization, i.e., affective dysregulation, negative self-concept, and disturbed relationships. Previous data from refugees experiencing torture showed that while 19% had PTSD, 32% fulfilled CPTSD criteria[4]. It is important to estimate both PTSD and CPTSD, as these conditions may correlate with different variables and require distinct interventions[7,8].

Resettled female Yazidi captives (N=108, mean age 24.4 ± 5.7 years; mean education 2.8 ± 4.0 years; 45.4% married; mean duration of captivity 7.7 ± 2.7 months; mean times sold 4.3 ± 5.7; mean number of fellow captives 32.3 ± 80.0) were sampled from four post-ISIS camps in Northern Iraq/Kurdistan region during February-March 2017.

Dichotomous (yes/no) exposure items (witnessing mass killings, people being killed; experiencing injury, torture, shelling, shooting, sexual abuse, rape, physical abuse; family members injured or killed) were aggregated to produce an exposure score. We administered the ICD-11 PTSD questionnaire, including six items which addressed the three proposed ICD-11 criteria[6] (alpha=0.71), and the ICD-11 CPTSD questionnaire, including six additional items, addressing the three proposed ICD-11 criteria[6] (alpha=0.71). Factor structure for two related yet distinctive constructs (PTSD vs. CPTSD) was slightly better than for a single construct. We also assessed stress in the post-ISIS camp, including experiencing violence, physical abuse, sexual abuse and hunger. These four items were responded on a 5-point Likert scale (from 1=not at all to 5=very much so, alpha=0.79).

Items that were not already available in Arabic were translated and back translated into English, reviewed, analyzed and corrected. Two pilot studies (N=20) were conducted, and two items (referring to feeling worthless and guilt) were reworded to ensure comprehension. Maintenance of the original meaning was evaluated by five assessors. Questionnaires in Arabic were read by female interviewers (trained by research team).

Fifty-five (50.9%) women had probable CPTSD, while 23 (20.0%) had probable PTSD. Dividing the sample into those with no PTSD, only PTSD and CPTSD revealed no significant group differences in age or marital status, but a marginally significant difference in mean years of education: no PTSD=1.58, only PTSD=2.08, CPTSD=3.92; $F(2,92)=2.98$, $p=0.055$. The groups did not differ significantly in captivity duration, number of fellow captives, number of times sold, or exposure score.

The groups differed significantly in stress endured in post-ISIS camps as evaluated on the Likert scale: no PTSD=2.45, PTSD=2.77, CPTSD=3.78; $F(2,93)=53.37$, $p<0.0001$. Post-hoc Bonferroni tests revealed that, while the no PTSD and PTSD groups were statistically comparable, the CPTSD group reported significantly higher post-ISIS stress than the other two groups.

The CPTSD prevalence we found was higher than CPTSD estimates in samples experiencing captivity/torture alone[4] or sexual abuse alone[9], which reflects the unique type of endured trauma combining captivity with sexual slavery. Given the high CPTSD prevalence, Kurdish training/intervention centers in formation should focus on preparing suitable CPTSD interventions. For example, CPTSD requires a phase-based treatment[8] where safety is a central initial goal; such victims benefit less from traditional PTSD treatments typically focusing on fear reduction.

Indeed, safety seems most relevant to our population, as the very same camp conditions may be less safe for CPTSD women who feel socially cut-off, worthless and guilty. Another possibility aligns with the "straw that broke the camel's back" model, whereby the emergence of CPTSD may be triggered by post-ISIS camp stress, which is less severe than the focal trauma. The above possibilities may be relevant to different

women, as CPTSD may both be a catalyst for increasing risk of experiencing future stress, as well as increasing one's vulnerability to such exposure. These various options can be assessed in a future longitudinal study addressing PTSD/CPTSD immediately after captivity release and at different time points in the post-ISIS camps. In any case, fortifying such traumatized women with a safe environment along with psychoeducation targeting their increased sensitivity may be very helpful until suitable interventions are available.

Limitations of the current study include a cross-sectional cohort and a relatively small sample. Although alpha values exceeded the reliability benchmark, they were lower than in previous studies[1,2], perhaps due to cultural/educational factors, which markedly differed in our sample from usual ones. Yet the findings illuminate the psychological aftermath of perhaps the most extreme atrocity occurring in recent years. Results also indicate the need for greater awareness of post-captivity conditions.

Future large-scale studies are required to continue the assessment of Yazidi captives. This should be informative with regard to theoretical issues concerning CPTSD, its distinction from PTSD, as well as aiding the development of feasible, culturally relevant and effective interventions to help these survivors.

Yaakov S.G. Hoffman[1], Ephraim S. Grossman[2], Amit Shrira[1], Mordechai Kedar[3,4], Menachem Ben-Ezra[5], Mirza Dinnayi[6], Lee Koren[7], Rassul Bayan[8,9], Yuval Palgi[10], Ari Z. Zivotofsky[11]
[1]Interdisciplinary Department of Social Sciences, Bar-Ilan University, Ramat-Gan, Israel; [2]School of Communication, Bar-Ilan University, Ramat-Gan, Israel; [3]Department of Arabic, Bar-Ilan University, Ramat-Gan, Israel; [4]Begin-Sadat Center for Strategic Studies, Bar-Ilan University, Ramat-Gan, Israel; [5]School of Social Work, Ariel University, Ariel, Israel; [6]Luftbrücke Irak, Osnabrück, Germany; [7]Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel; [8]Erbil Psychiatric Hospital, Erbil, Kurdistan; [9]Emma Organization for Human Development, Erbil, Kurdistan; [10]Department of Gerontology, University of Haifa, Haifa, Israel; [11]Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, Israel

This research was funded by a grant awarded to Y.S.G. Hoffman and A.Z. Zivotofsky by the Research Center for the Middle East and Islam (under formation), Bar-Ilan University, Ramat-Gan, Israel. The authors thank J. Neurink and E. Wisman for their assistance.

1. Hoffman Y. Stress and Health (in press).
2. Ben-Ezra M, Leshem E, Goodwin R. Am J Psychiatry 2015;72:795-6.
3. Abdel-Razek O, Puttick M. Contemp Arab Aff 2016;9:565-76.
4. Nickerson A, Cloitre M, Bryant RA et al. Eur J Psychotraumatol 2016;7:33253.
5. Kessler RC, Sonnega A, Bromet E et al. Arch Gen Psychiatry 1995;52:1048-60.
6. Maercker A, Brewin CR, Bryant RA et al. World Psychiatry 2013;12:198-206.
7. Silove D, Ventevogel P, Rees S. World Psychiatry 2017;16:130-9.
8. Cloitre M, Stovall-McClough KC, Nooner K et al. Am J Psychiatry 2010;167:915-24.
9. Cloitre M, Garvert DW, Weiss B, et al. Eur J Psychotraumatol 2014;5:25097.

DOI:10.1002/wps.20475

Mental health policies in Commonwealth countries

The large global burden of mental health conditions[1] has led to an increased emphasis on improving access to mental health services in countries across the world. There is clear recognition that improving mental health governance is key to improving access to and quality of mental health services, and the existence of a mental health policy is an important component of improving mental health governance[2].

The World Health Organization (WHO)'s Mental Health Atlas 2014 found that 68% of countries had a mental health policy[3], while the WHO Mental Health Action Plan 2013-2020 set a target that 80% of countries should have developed or updated their mental health policies/plans in line with international and regional human rights instruments by 2020[2].

The Commonwealth is a voluntary association of 52 independent and sovereign states part of the erstwhile British Empire which, in spite of their geographical variations, generally have similar political, legal and governance systems. We reviewed the mental health policies of Commonwealth countries in order to compare them to standards developed by the WHO and to assess their compliance with international recommendations.

We identified and downloaded mental health policies of Commonwealth countries from WHO MINDbank[4]. If a mental health policy was not found there, we extended our search to official government websites of countries. In the case of federal countries such as Canada and Australia, we relied on a federal mental health policy and, if this was absent, we used the most recent mental health policy from any of the provinces or, if there were two policies from the same year, the mental health policy from the province with the larger population. We used WHO's Mental Health Policy checklist[5] to assess compliance of country mental health policies with international recommendations.

Eleven countries (21.1%) did not have a mental health policy. We were unable to find a mental health policy in 16 (30.8%) additional countries, although we found references in various documents to such a policy. We found a mental health policy in 25 countries (48.1%), of which Naaru and Zambia had a "final draft" policy and Uganda and St. Lucia had a draft policy.

Of the countries with a mental health policy, only seven (28.0%) had adopted it after 2011. In only two (8.0%), the policies contained an explicit reference to country data and research informing policy development.

While 15 policies (60.0%) indicated how funding would be used for financing mental health services, only four (16.0%) had a clear statement on providing equitable funding between mental and physical health, and only five (20.0%) explicitly stated that mental health should be included in health insurance. Seventeen policies (68.0%) promoted human rights, while only 14 (56.0%) specifically mentioned developing human rights oriented mental health legislation.

Only one policy (4.0%) had detailed intersectoral collaboration outlined, while four (16.0%) had it to some extent. Fourteen policies (56.0%) included a process to measure and improve the quality of mental health services. Only three policies (12.0%) either significantly or to some extent made a commitment to putting in place suitable working conditions for mental health providers. Twenty-one policies (84.0%) did, however, recognize that training in core competencies and skills was important for human resource development.

Fourteen policies (56.0%) promoted integration of mental health services into general health services, but only 11 policies (44.0%) promoted deinstitutionalization. Nearly half of the policies had provisions for promotion of mental health, prevention of mental disorders and rehabilitation of persons with mental illness. Only half of the policies emphasized the need for research and evaluation of services and assessment of the policy and strategic plan.

Thus, while the WHO reported that 68% of countries globally had a mental health policy, we were only able to find such a policy in 48% of Commonwealth countries. Most (72%) of the Commonwealth mental health policies were adopted prior to 2011. Thus, these countries were behind the global average of 47% of countries having adopted a mental health policy after 2010[3].

There are many possible reasons for these findings, including failure of mental health sector to advocate effectively, lack of technical skills in mental health policy development, and limited political interest due to stigma and discrimination.

Financing of mental health remains a concern. Although 60% of policies recognized the role of funding in equitable mental health services, only a tiny minority explicitly addressed equitable funding between services for mental and physical health.

We know anecdotally that in many countries health insurance from private (and at times also public) insurance providers excludes treatment of mental illness. It is a matter of serious concern that policies in only a fifth of the countries have a commitment to including mental health in general health insurance.

As we said, the WHO Mental Health Action Plan sets a global target for 80% of countries to develop or update their mental health legislation in line with international human rights instruments by 2020[2]. Although half of the Commonwealth country policies identify the development of new mental health legislation as a key policy action, the reality lies in actual delivery and whether countries which have new legislation as a key policy action do succeed in delivering equity.

It is now well accepted that addressing mental health problems requires intervention across many different sectors including health, social welfare, education, justice, employment to name just a few. It is therefore surprising to note the near total absence of any emphasis on intersectoral collaboration in these policies in Commonwealth countries.

We were also concerned to note that only half of the policies promoted integration of mental health with physical health services, deinstitutionalization, promotion of mental health, prevention of mental disorders and rehabilitation. The lack of emphasis on improving quality of mental health services in the policies of half of the countries is equally disconcerting, because it allows poor quality services to continue without change and perpetuates poor outcomes.

Only half of the policies mentioned the need for research and evaluation of policies and services, indicating a general lack of interest in an important component of future policy development. The absence of systematic evaluation of existing policies may result in ineffective services continuing to be offered and consequent failure to achieve policy goals.

In conclusion, our findings indicate that there is still some way to go in bringing about equity between physical and mental health care but also equity between Commonwealth countries. The informal links between these countries indicate the strength of the relationship and perhaps high-income countries need to consider mentoring and supporting low-income countries to ensure that mental health policies are developed appropriately in order to deliver best care which our patients need, deserve and will utilize.

Dinesh Bhugra[1], Soumitra Pathare[2], Rajlaxmi Joshi[2], Antonio Ventriglio[3]
[1]Institute of Psychiatry, King's College London, London, UK; [2]Centre for Mental Health Law and Policy, Indian Law Society, Pune, Maharashtra, India; [3]University of Foggia, Foggia, Italy

The study was funded by the WPA while D. Bhugra was the President.

1. Whiteford HA, Ferrari AJ, Degenhardt L et al. PLoS One 2015;10:e0116820.
2. World Health Organization. Mental health action plan 2013-2020. Geneva: World Health Organization, 2013.
3. World Health Organization. Mental health atlas 2014. Geneva: World Health Organization, 2015.
4. World Health Organization. WHO MiNDbank. http://www.who.int/mental_health/mindbank/en/.
5. World Health Organization. The WHO mental health policy and service guidance package. http://www.who.int/mental_health/policy/essentialpackage1/en/.

DOI:10.1002/wps.20502

Chance of response to an antidepressant: what should we say to the patient?

Talking with patients appropriately is considered as one of the key competences of psychiatrists. "I would like to prescribe you the antidepressant X against your depression. According to available evidence, it will be helpful for one in five patients of

your type". This seems to be the truth according to meta-analyses and Cochrane reviews[1,2]. Numbers needed to treat (NNTs) are considered as an illustrative and rather exactly calculable result of randomized controlled trials (RCTs) and meta-analyses which can easily be understood and communicated.

Still, a communication as mentioned above would sound disastrous to the patient. But, would it be ethical to withhold this information, although the doctor is well aware of it? The World Health Organization considers "health-related education and information" as an "underlying determinant of health"[3]. Most developed countries provide detailed legal frameworks on the duty of physicians to inform patients comprehensively on suggested treatments, and failure to do so might entail legal consequences.

Within the last decades, patient's rights have been increasingly strengthened in many countries, and jurisdiction has outlined explicitly in how much detail patients must be informed. Ethically, the respect for patient's autonomy requires informed consent for all medical procedures. On the other hand, the ethical principles of "beneficence" and "non-maleficence" demand that physicians should do what is helpful for their patients and avoid what might have negative consequences without further benefit[4].

Giving hope is one of the most powerful instruments of doctors, and that applies particularly to depression, a condition in which hopelessness is a key symptom[5]. The role of unspecific factors in the treatment of depression is indicated by the high placebo response in drug trials. On average, 36% of patients receiving placebo reached the defined "response" criterion in antidepressant trials after the year 2000, whilst the corresponding rate under antidepressant treatment was 46%[6]. Analyses show that the placebo effect has even increased within past decades, possibly due to heightened expectations of clinicians and patients[6]. Vice versa, a higher chance and subsequent expectation to receive a placebo in antidepressant trials decreases the placebo effect[7]. Hence, communications as the above mentioned that diminish or even intentionally erode expectations are likely to decrease the chance of a good outcome and thus mean harm for the patient. This poses a difficult ethical dilemma between the duty to inform the patient before obtaining informed consent and the necessity to provide the best available therapy which, in this case, includes giving hope.

Help for the clinician in reconciling this dilemma comes from an unexpected source: from statistics. The basic question is whether the NNT does indeed indicate how many patients respond to a certain treatment and how many do not. This is suggested by the term and rarely questioned. However, what NNTs are and what they are not has been recently clarified[7].

Valid response rates are mostly unknown and difficult to determine. Actually, repeated individual cross-over testing would be required to eliminate several sources of variance and determine the real response rate. Recently, it could be demonstrated that even in mice the response rate under treatment with selective serotonin reuptake inhibitors was significantly dependent on the circumstances of the environment[8]. In everyday practice, key predictors of the individual response rate are in fact the "health care system" and the variable "doctor"[7]. And this brings us back to our initial discussion: whether an antidepressant works in the next episode is not only dependent on the results of available RCTs, but also on the circumstances under which the treatment takes place, and, least but not last, what the doctor himself says or does.

RCTs are conducted under defined circumstances with the control group receiving considerable psychological attention and support through study workers. This is a common explanation for the small observed effect sizes in those studies. NNTs are an abstract measure which is based on arbitrary dichotomous definitions of "response", "recovery" or whatsoever, calculated from the differences between drug response and placebo response in these trials under artificial conditions. Even if we assume that all patients benefit from a certain therapy to an identical degree, some of them would achieve the defined "response" or "recovery" threshold and others would not, depending on the efficacy of the therapeutic intervention and the individual baseline level. This is substantially different from interventions with dichotomous outcomes, e.g. death vs. recovery. In conditions with a dimensional character, which is the nature of most mental disorders, remission rates do not indicate individual benefit.

Therefore, NNTs are basically a statistical construct which allows us to determine easily in illustrative figures how effective an intervention is and whether one is more effective than another. But they do not at all indicate an individual likelihood of a positive outcome. Thus, the clinician's dilemma can be solved successfully: never use NNTs or response rates to explain chances and risks to patients.

Tilman Steinert
Department of Psychiatry and Psychotherapy, Ulm University, Ulm; Centers for Psychiatry Suedwuerttemberg, Ravensburg, Baden-Wuerttemberg, Germany

1. Gregory A, Mallikarjun P, Upthegrove R. Br J Psychiatry 2017;211:192-204.
2. Wilkinson P, Izmeth Z. Cochrane Database Syst Rev 2016;9:CD006727.
3. World Health Organization. The right to health. Geneva: World Health Organization, 2008.
4. Beauchamp TL, Childress JF. Principles of biomedical ethics. Oxford: Oxford University Press, 2008.
5. Cuijpers P, de Beurs DP, van Spijker BA et al. J Affect Disord 2013;144:183-90.
6. Khan A, Mar KF, Faucett J et al. World Psychiatry 2017;16:181-92.
7. Senn S. Stat Med 2016;35:966-77.
8. Alboni S, van Dijk RM, Poggini S et al. Mol Psychiatry 2017;22:552-61.

DOI:10.1002/wps.20511

WPA-WHO Africa Mental Health Forum – recommendations and position statement

The WPA Action Plan 2017-2020[1] has been released by H. Herrman, new WPA President, during the recent World Congress of Psychiatry in Berlin.

An agreement that could assist in achieving the objectives of that Action Plan in an African context has been produced during the WPA-World Health Organization (WHO) Africa Mental Health Forum meeting that took place in November 2016 in Cape Town, South Africa. In particular, the agreement can be helpful with regard to the Action Plan's enabling activities, aimed at supporting psychiatrists to promote mental health and improve care capacity, and its partnership and collaboration activities, aimed to expand the reach and effectiveness of partnerships with service providers, service beneficiaries and policy makers.

The Forum, co-chaired by D. Bhugra (WPA Immediate Past-President) and S. Saxena (Director of WHO Department of Mental Health and Substance Abuse), was opened by M. Moeti, WHO Director for the African Region. The following recommendations were made by the four panels (1. Leadership and governance; 2. Health and social services; 3. Prevention and promotion; 4. Information, evidence and research):

- To involve all stakeholders in all (planning) meetings at all levels, including consumers, while enabling and supporting consumers to participate meaningfully.
- To achieve a systematized approach in mental health leadership and governance, so that not all effort and support depends on one individual in a particular Ministry – the approach should include different departmental officials, from the chief medical officer to administrative staff, but also reach beyond and across departments and governments.
- To obtain comprehensive data on all aspects in order to have information and provide evidence for the financing required for different mental health programs.
- To retain the "bigger picture" with regard to the United Nations Convention on the Rights of Persons with Disabilities, namely to achieve humane mental health care, and not to be side-tracked in the debate while considering applicable options for mental health in a step-by-step way.
- To mobilize resources for training in public mental health from national to district level, in order to have understanding that resources must be identified and systems created beyond hospital care, e.g. not only to advocate for hospitals, but for systems of care.
- To utilize "Mental Health Innovations – Africa" as a platform to continue discussion and communication between role players in Africa.
- To reorganize and reform the whole mental health care system by integrating available resources (e.g., psychiatrists in private practice with other role players), while clearly identifying the roles of mental health care workers involved.
- To achieve integration and role identity through training of current and future practitioners and students – all need to know more about each other; an integrated model of practice must be promoted (e.g., psychiatry and other disciplines, mental and physical health care).
- To broaden the treatment pyramid base through self-care and getting people to be able to care for themselves – at least with regard to minor problems, while people with severe psychiatric problems should still be further treated in specialized centers.
- To clarify the roles of the different role players in the field in a specific catchment area, while people in a certain catchment area must also be aware of what the referral route is for emergencies, or the correct way to address problems.
- To address communication and logistical aspects will require leadership, in order to achieve a reorganization and reformation of the mental health care system.
- To incorporate the interests of patients, which must be at the heart of all mental health care, including promotion and prevention – their voice must be recognized in order to bring the richness and strength of their experience to the table; particular areas of concern include that a holistic approach is adopted when addressing comorbid physical illnesses of patients, in view of the known increased risk of morbidity and mortality associated with being a mental health care user; and involvement in the evaluation of service provision in order to achieve services that care and support, rather than stigmatize.
- To achieve different competencies, such as cultural, (health) educational, service delivery and policy competency.
- To involve the media to address stigma, e.g. through advertisement, while also addressing cultural aspects of stigma and constantly recognizing the voice of patients.
- To revise training curricula of under- and post-graduate programs to ensure inclusion of the minimum required content on mental health, including promotion and prevention.
- To acknowledge the critical importance of collaboration and networks.
- To share information and experiences.
- To address stigma, including stigma in mental health workers and the systems in which they work.
- To incorporate the use of technology in screening and intervention delivery.
- To consider cultural idioms of distress and appropriate interventions.
- To accommodate the qualification of new cadres of mental health workers through creating posts and career paths.
- To teach research methods and dispel myths about research, while refocusing

the emphasis on scientific curiosity to answer questions.
- To embrace a range of research methods in mental health, from quantitative, systems, mixed to qualitative; from basic neuroscience to implementation research; also, to develop "clinician researchers".
- To conduct further epidemiological research, as there are relatively few data, for example, on the prevalence and associations of mental disorders in primary care settings in the African context.
- To conduct research on the effectiveness and cost-efficiency of integrated care and collaborative care in the African context, as well for further work on moderating and mediating factors.

From the vision stated by M. Moeti in her presentation, the proceedings and the concluding remarks by the co-chairs of the Forum, the following position statement was drafted on a continental alliance for integrated mental health care in Africa:

In order to achieve the communicated vision, objectives and targets for achieving the potential of mental health for all and integrated mental health care in Africa, we will need to work together with collective strength and active collaboration. Such an alliance for integrated mental health care in Africa, with emphasis on public mental health, includes: individual and collective psychiatrists, as well as all members of the multidisciplinary mental health team (psychologists, nurses, social workers, occupational therapists); other health professionals in primary and specialist health care; community mental health workers and self-help resources; our patients and their families; the public at large through the media; training institutions; as well as governments' Ministries of Health and private service providers of mental health care services. While different countries and groups may have different entry points, strengthening of this alliance must be sought within countries nationally, provincially and locally, but also on subcontinental and continental levels.

The full report on the Forum can be obtained from the WPA website.

Bernard Janse van Rensburg[1], Dinesh Bhugra[2], Shekhar Saxena[3]
[1]President, South African Society of Psychiatrists; [2]WPA Immediate Past-President; [3]Director, WHO Department of Mental Health and Substance Abuse

1. Herrman H. World Psychiatry 2017;16:329-30.

DOI:10.1002/wps.20510

WPA Scientific Sections activities in the triennium 2014-2017

The triennium 2014-2017 has seen a growing interest in the work of WPA's Scientific Sections. Whereas Sections continued with their contributions in the entire scientific field of mental health, they also emerged as important and integral components of WPA[1]. Their activities during this triennium supported WPA's remit and objectives for promotion and dissemination of scientific knowledge around the globe[2]. With a record number of 72 Sections in total, further interest for having more Sections to cover some other scientific disciplines continued.

Although Sections do differ in the nature and extent of their activities, they have generally participated and organized sessions in WPA sponsored and co-sponsored meetings. The scientific programme of the WPA International Conference in South Africa (November 2016) included 28 symposia/workshops/round table sessions from WPA Sections. Similarly, there has been a noticeable increase in the number of Sections organizing individual meetings/sessions at WPA sponsored and co-sponsored meetings. It is reassuring to note that Sections have organized more than 275 individual sessions, 115 intersectional programmes and 110 sessions at Member Societies conferences during the triennium. This is indeed a record number. Likewise, over the last three years, the numbers of joint section activities have increased significantly. Organization of Intersection Forums has equally emerged as an innovative practice for promoting inter-sectional collaboration, and this initiative has strengthened and reinforced the need for exploring common interests among various Sections.

Sections have continued with the publication of scientific reports, guidelines and their individual journals and bulletins[3-5]. At present, 14 Sections are publishing and supporting publication of journals to enhance the scientific knowledge in their respective fields[6].

Meetings of Section chairs organized at various WPA conferences during the triennium (mainly at WPA international meetings) have generated a great extent of interest among the Section officers for supporting a visible role in WPA work.

Keeping in view the WPA Action Plan for 2011-2014, Sections initiated various academic and educational activities[7]. A number of Sections are at present involved in updating educational programmes and revising WPA website information with recent advances in their fields[8,9].

Similarly, many Section members continued with their participation in the ICD-11 work and also submitted their publications to *World Psychiatry*[10-12]. We thank Prof. M. Maj for his leadership and guidance to the Section members in this regard.

As per suggestions from Section chairs and having approval from the Executive Committee, special teaching sessions have also been planned at WPA co-sponsored meetings during the triennium. We requested different Section chairs/officers to give an updated account of preferred scientific topics from the expertise of the Section in these teaching programmes. This experience has proved a real success especially among the young psychiatrists.

A new publication based on the Section's work, *Advances in Psychiatry – Volume 4*, is at the final stage of production.

Thirty-six Sections have contributed to this book and we are expecting its release in the next few months.

Although Sections' work has always been acknowledged as having an important influence on the scientific performance of WPA, there have been some concerns about the low level of activity or limited reporting of activities by some of the Sections. A wide variation has been observed in the contribution of the Sections based on different measures of their functioning.

Whereas Sections enjoy a great degree of independence within the framework of the statutes and by-laws of WPA, the WPA Operational Committee has recommended some revisions in the by-laws, reflecting the changing directions for Sections' work. There have also been some initial discussions for clustering of Sections on the basis of common interests and activities. This will hopefully promote further collaboration and links among different Sections.

There is a strong need for encouragement of Sections in their role of guiding and supporting WPA in their areas of expertise. Member Societies and Zone Representatives are expected to support and assist Sections in enrolment of new members and planning projects in the respective countries and regions.

Afzal Javed
WPA Secretary for Sections 2014-2017

1. Javed A. World Psychiatry 2017;16:222.
2. Bhugra D. World Psychiatry 2017;16:221-2.
3. Bhaumik S, Kiani R, Dasari MM et al. Int J Cult Ment Health 2016;9:417-29.
4. Moreira-Almeida A, Sharma A, Janse van Rensburg B et al. World Psychiatry 2016;15:87-8.
5. Fiorillo A, Pinto da Costa M, Takashi N et al. Middle East Curr Psychiatry 2016;23:3.
6. Riba M. World Psychiatry 2017;16:114-5.
7. Bhugra D. World Psychiatry 2014;13:328.
8. Stewart DE, Chandra PS. World Psychiatry 2017;16:223-4.
9. Villasenor-Bayardo S, Rojas-Malpica C, Romero A. Int Rev Psychiatry 2016;28:130-2.
10. Gureje O, Reed GM. World Psychiatry 2016;15: 291-2.
11. Jablensky A. World Psychiatry 2016;15:26-31.
12. Bucci P. World Psychiatry 2017;16:115-6.

DOI:10.1002/wps.20505

WPA scientific publications in the triennium 2014-2017

It has been a very busy and productive triennium in terms of WPA scientific publications.

The WPA Operational Committee on Scientific Publications was appointed in 2014 by WPA President D. Bhugra and has served to promote research and publishing capacity in WPA member countries with an emphasis on capacity building in the developing world[1]. Members of this hard working committee included: M.B. Riba, Chair (USA); D. Lecic Tosevski, Co-Chair (Serbia); R. Heun (UK), P. Tyrer (UK), P. Chandra (India), C. Szabo (South Africa), A. Cia (Argentina) and J.M. Castaldelli Maia (Brazil).

We applaud the success and importance of *World Psychiatry*, the WPA official journal, which, as of July 2017, is ranked no. 1 in terms of impact factor among not only all psychiatric journals, but also among all the journals included in the Social Science Citation Index (SSCI) published by Thomson Reuters. Under the leadership of Editor M. Maj, the journal reaches more than 50,000 psychiatrists worldwide, and is published in English, Spanish, Chinese, Russian and French and in partial translation in Romanian. In addition to a variety of scholarly papers, the journal regularly publishes news about the WPA initiatives[2-6].

The Operational Committee has organized a number of symposia at various WPA Congresses – Taipei, Taiwan, 2015; Tbilisi, Georgia, 2016; Cape Town, South Africa, 2016; and Berlin, Germany, 2017 – to assist early career psychiatrists develop as writers and scholars. At the Japanese Society for Psychiatry and Neurology Congress, held in Nagoya, Japan in June 2017, under the leadership of S. Kanba, T. Akiyama, R. Freedman, P. Wang and M.B. Riba, young psychiatrists were mentored regarding their research and publication opportunities.

The Operational Committee has tried to improve and provide easier access to journals by linking the WPA website with open-access online journals from all over the world[7]. The links were proposed by national psychiatric associations. We appreciate the help of Secretary General R. Kallivayalil in this project[8,9].

Under the direction and coordination of A. Cia, the Biblioteca Iberoamericana de Psiquiatria de la WPA has been an innovative project, approved and supported by the WPA Executive Committee. This library gathers psychiatry and mental health journals, and is published digitally in Spanish by participating countries and organizations, with free access for all interdisciplinary health team members that sign up through a simple procedure, covering 20 countries and expected to reach over 100,000 Spanish-speaking mental health professionals visiting the website.

We have also enhanced the image and respect of the WPA and the field of psychiatry and mental health worldwide by publishing a large number of books, including a highly regarded WPA series on *Integrating Psychiatry and Primary Care* (Editors D. Bhugra and M.B. Riba), with books on *Physician Mental Health and Well-being* (Brower and Riba, 2017)[10], *Motherhood in the Face of Trauma* (Muzik and Rosenblum, 2017)[11], and *Physical Exercise Interventions for Mental Health* (Lam and Riba, 2016)[12].

We have appreciated the opportunity to serve under WPA President D. Bhugra and President Elect H. Herrman and to have had the support and collaboration of the WPA Executive Committee, WPA Member Societies, WPA Board, Past Presidents, Section and Committee chairs and all WPA members.

Michelle B. Riba
WPA Secretary for Scientific Publications 2014-2017

1. Riba MB. World Psychiatry 2017;16:114-5.
2. Bhugra D. World Psychiatry 2017;16:221-2.
3. Herrman H. World Psychiatry 2017;16:329-30.

4. Stewart DE, Chandra PS. World Psychiatry 2017; 16:223-4.
5. Shields G, Ng R, Ventriglio A et al. World Psychiatry 2017;16:113-4.
6. Moreira-Almeida A, Sharma A, Janse van Rensburg B et al. World Psychiatry 2016;15:87-8.
7. Riba MB. World Psychiatry 2016;15:88.
8. Kallivayalil RA. World Psychiatry 2017;16:114.
9. Kallivayalil RA. World Psychiatry 2017;16:330-1.
10. Brower KJ, Riba MB (eds). Physician mental health and well-being. Berlin: Springer, 2017.
11. Muzik M, Rosenblum KL (eds). Motherhood in the face of trauma. Berlin: Springer, 2017.
12. Lam LCW, Riba M (eds). Physical exercise interventions for mental health. Cambridge: Cambridge University Press, 2016.

DOI:10.1002/wps.20506

ICD-11 sessions in the 17th World Congress of Psychiatry

Within the 17th World Congress of Psychiatry, held in Berlin from 8 to 12 October 2017, eight symposia, three workshops, one state-of-the-art lecture and several individual presentations focused on various aspects of the chapter on mental and behavioural disorders of the 11th edition of the International Classification of Diseases and Related Health Problems (ICD-11), which is expected to be approved by the World Health Assembly in May 2018[1].

As emphasized by many presenters, improving clinical utility of psychiatric diagnosis in ordinary practice is the main objective of the new diagnostic system. The clinical descriptions and diagnostic guidelines provided for the various mental disorders will guide clinicians in their diagnostic practice, but clinical judgment will have to be finally exercised in order to decide whether the features of an individual case approximates sufficiently one of the prototypes proposed in the manual in order to justify the corresponding diagnosis. Precise (or pseudo-precise) thresholds concerning the number or duration of symptoms will not be included in the system, unless they are convincingly validated by available research.

An effort has been made to harmonize the two main diagnostic systems existing in the psychiatric field – the ICD and the DSM – and indeed the organizational framework ("metastructure") will be the same in the ICD-11 as in the DSM-5. Nevertheless, several intentional differences between the two systems will remain.

In particular, some diagnostic categories will appear in the ICD-11 that are not included in the DSM-5. Examples are given by complex post-traumatic stress disorder (PTSD) and prolonged grief disorder. The category of complex PTSD is characterized by the three core elements of PTSD (i.e., re-experiencing the traumatic event in the present, deliberate avoidance of reminders likely to produce this re-experience, and persistent perceptions of heightened current threat) plus severe and pervasive problems in affect regulation; persistent beliefs about oneself as diminished, defeated or worthless; and persistent difficulties in sustaining relationships and in feeling close to others. The category of prolonged grief disorder is characterized by a pervasive grief response, persisting for an abnormally long period of time following the loss, clearly exceeding expected social or religious norms for the individual's culture and context, and causing significant social impairment.

On the other hand, some diagnostic categories that are included in the DSM-5 will not appear in the ICD-11. An example is given by disruptive mood dysregulation disorder, which will be replaced in the ICD-11 by the subtype "with chronic irritability-anger" of oppositional defiant disorder. This subtype is marked by prevailing, persistent angry or irritable mood, including often being "touchy" or easily annoyed, that is characteristic of the individual's functioning nearly every day and is observable across multiple settings or domains of functioning (e.g., home, school, social relationships) and is not restricted to the individual's relationship with his/her parents or guardians. The negative mood is often accompanied by regularly occurring severe temper outbursts that are grossly out of proportion in intensity or duration to the provocation.

A reflection of the ongoing debate about these and other controversial diagnostic topics can be found in recent issues of this journal[2-15].

Conditions related to sexual health and sleep-wake disorders will appear in chapters of the classification different from the one on mental disorders. This has been decided in order to address the criticism to the ICD-10 concerning the problematic distinction between "organic" and "non-organic" sexual dysfunctions (covered in the ICD-10, respectively, in the chapters on diseases of the genitourinary system and on mental and behavioural disorders), and between "organic" and "non-organic" sleep disorders (covered in that system, respectively, in the chapters on diseases of the nervous system and on mental and behavioural disorders).

The new diagnostic system has been tested through several field studies. There were first two large international surveys of views of psychiatrists and psychologists about the classification of mental disorders and the features that would increase its clinical utility. These were followed by so-called formative field studies, aimed to guide decisions about the basic structure and content of the classification by exploring clinicians' conceptualizations of the interrelationships among categories of mental disorders.

Internet-based field studies were then implemented through the Global Clinical Practice Network, including more than 13,000 psychiatrists and other health professionals from more than 150 countries, which used vignette methodologies to examine clinical decision-making in relationship to the proposed ICD-11 diagnostic categories and guidelines. Finally, clinic-based (or ecological implementation) field studies were conducted to assess the reliability and clinical utility of the diagnostic guidelines with real patients. The results of several of these field studies were presented at the World Congress (see also https://gcp.network/en/).

Corrado De Rosa
WHO Collaborating Centre for Research and Training in Mental Health, University of Naples SUN, Naples, Italy

1. Reed GM, First MB, Medina-Mora ME et al. World Psychiatry 2016;15:112-3.
2. Kendler KS. World Psychiatry 2016;15:5-12.
3. Jablensky A. World Psychiatry 2016;15:26-31.
4. Frances A. World Psychiatry 2016;15:32-3.
5. Phillips MR. World Psychiatry 2016;15:38-9.
6. Maj M. World Psychiatry 2016;15:193-4.
7. Reed GM, Drescher J, Krueger RB et al. World Psychiatry 2016;15:205-21.
8. Maciejewski PK, Maercker A, Boelen PA et al. World Psychiatry 2016;15:266-75.
9. Gureje O, Reed GM. World Psychiatry 2016;15:291-2.
10. Mann K, Fauth-Bühler M, Higuchi S et al. World Psychiatry 2016;15:297-8.
11. Leibenluft E. World Psychiatry 2017;16:100-1.
12. Bucci P. World Psychiatry 2017;16:115-6.
13. Lahey BB, Krueger RF, Rathouz PJ et al. World Psychiatry 2017;16:142-3.
14. Kessing LV, Bukh JD. World Psychiatry 2017;16:318-9.
15. Steardo L Jr. World Psychiatry 2017;16:331-2.

DOI:10.1002/wps.20507

Correction

It has been brought to our attention that in Table 2 of the paper "Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls", by Correll et al, published in the June 2017 issue of *World Psychiatry*, the numbers of cases and controls in the article by Gasse et al (2014) were reported incorrectly. They should be, respectively, 126,106 and 4,419,221.

Acknowledgement
This publication has been partially supported by an unrestricted
educational grant from Angelini Acraf S.p.A.,
which is hereby gratefully acknowledged.