World Psychiatry

OFFICIAL JOURNAL OF THE WORLD PSYCHIATRIC ASSOCIATION (WPA)

Volume 11, Number 1 *February 2012*

ELSEVIER

NEW IMPACT FACTOR: 5.562

ISSN 1723-8617

The World Psychiatric Association (WPA)

The WPA is an association of national psychiatric societies aimed to increase knowledge and skills necessary for work in the field of mental health and the care for the mentally ill. Its member societies are presently 135, spanning 117 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 66 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Research Reports containing unpublished data are welcome for submission to the journal.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. Bathe KJ, Wilson EL. Solution methods for eigenvalue problems in structural mechanics. Int J Num Math Engng 1973;6:213-26.
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.

Managing Director and Legal Responsibility - Emile Blomme (Italy)
Published by Elsevier S.r.l., Via P. Paleocapa 7, 20121 Milan, Italy.

***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).

Bereavement-related depression in the DSM-5 and ICD-11

Mario Maj

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

The approach of the DSM-5 and ICD-11 to the issue of bereavement-related depression is going to attract a considerable attention from the mental health community and the general public. This issue, in fact, is closely linked to the more general question of what is a mental disorder, or what is the boundary between mental pathology and homeostatic reactions to major life events. It is not by chance that the DSM-IV, in its introduction (p. xxi), identifies as one of the components of the definition of mental disorder the fact that "the syndrome or pattern must not be merely an expectable and culturally sanctioned response to a particular event, for example, the death of a loved one".

Bereavement appears in the DSM-IV in the section "Other conditions that may be a focus of clinical attention", where it is stated that "as part of their reaction to the loss, some grieving individuals present with symptoms characteristic of a major depressive episode" and that "the bereaved individual typically regards the depressed mood as 'normal'".

The DSM-IV does not totally exclude the diagnosis of major depressive episode in the presence of bereavement. It just moves the threshold upward for that diagnosis, by requiring a longer duration, a more substantial functional impairment, or the presence of specific symptoms (morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation). The aim is clearly to reduce the chance of false positives (as well as to avoid a trivialization of the concept of mental disorder).

This approach of the DSM-IV is evidence-based. First, a major depressive syndrome is indeed an "expectable response" to the death of a loved one: in the US, its prevalence among bereaved people ranges from 29 to 58% one year after the loss, and about 50% of all widows and widowers meet criteria for the syndrome at some time during the first year of bereavement (1). Second, the syndrome is indeed a "culturally sanctioned response" to the event: bereaved people and their environment accept depressive symptoms as "normal", whereas patients with primary affective disorder experience their condition as "a change", "not usual self" (2). Third, psychomotor retardation, feelings of worthlessness and suicidal ideation are less likely to be experienced by bereaved people when they have a major depressive syndrome (1).

It has been claimed that the ICD-10 is silent concerning the issue of bereavement, and that the elimination of the bereavement exclusion in the DSM-5 would contribute to the harmonization between the two systems. This is not correct. The ICD-10 Clinical Descriptions and Diagnostic Guidelines (3, p. 150) state that "normal bereavement reactions, appropriate to the culture of the individual concerned and not usually exceeding 6 months in duration" should not be coded in the chapter on mental disorders, but in chapter XXI ("Factors influencing health status and contacts with health services"). That chapter corresponds to the section where bereavement is placed in the DSM-IV. It is true that no mention of bereavement is made in the definition of depressive episode (which, as almost all ICD-10 definitions, does not contain exclusion criteria), but this mention is likely to be made in the ICD-11 (so that a change in the DSM-5 might actually create a discrepancy between the two systems).

Given this background, and considering the criteria established for DSM-5 changes (4), the removal of the bereavement exclusion from the diagnosis of major depressive episode can only be justified by a strong and unequivocal new research evidence (5). Wakefield and First's review published in this issue of the journal (6) suggests that such a solid and consistent new evidence is not available.

Bereavement-excluded major depression has been associated with a significantly lower risk of subsequent depressive episodes in two recent independent studies (7,8), which is the kind of longitudinal data previously regarded as necessary to support the current diagnostic framework (9). Furthermore, even studies usually quoted as supporting the removal of the bereavement exclusion did find that bereavement-related depression is significantly less likely than other loss-related depression to be associated with treatment seeking (10) and substantial functional impairment (11), and is marked by significantly lower levels of neuroticism and guilt (10). These are data in line with the DSM-IV approach.

Further reflection seems therefore warranted before proceeding with the deletion of the bereavement exclusion, a move that may be criticized by the mental health community as not fulfilling the criteria for DSM-5 changes ("major changes should generally require consistency of support across validators") and is likely to be perceived by the general public as a further step in psychiatry's attempt to pathologize normal human processes. A refinement of the formulation of the bereavement exclusion may, however, be needed in order to increase its predictive validity (6,12).

Acknowledgements

The author is the Chairperson of the World Health Organization Working Group on the Classification of Mood and Anxiety Disorders, reporting to the International Advisory

Group for the Revision of ICD-10 Mental and Behavioural Disorders, of which he is also a member. Unless specifically stated, this paper represents the views of the author and not the official policies and positions of the World Health Organization.

References

1. Zisook S, Paulus M, Shuchter SR et al. The many faces of depression following spousal bereavement. J Affect Disord 1997;45:85-95.
2. Clayton PJ, Herjanic M, Murphy GE et al. Mourning and depression: their similarities and differences. Can Psychiatr Assoc J 1974; 19:309-12.
3. World Health Organization. The ICD-10 classification of mental and behavioural disorders. Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1992.
4. Kendler KS, Kupfer D, Narrow W et al. Guidelines for making changes to DSM-5. www.dsm5.org.
5. Maj M. Depression, bereavement, and "understandable" intense sadness: should the DSM-IV approach be revised? Am J Psychiatry 2008;165:1373-5.
6. Wakefield JC, First MB. Validity of the bereavement exclusion to major depression: does the empirical evidence support the proposal to eliminate the exclusion in DSM-5? World Psychiatry 2012;11: 3-10.
7. Mojtabai R. Bereavement-related depressive episodes: characteristics, 3-year course, and implications for the DSM-5. Arch Gen Psychiatry 2011;68:920-8.
8. Wakefield JC, Schmitz MF. Recurrence of bereavement-related depression: evidence for the validity of the DSM-IV bereavement exclusion from the Epidemiologic Catchment Area Study. J Nerv Ment Dis (in press).
9. Pies RW. Depression and the pitfalls of causality: implications for DSM-V. J Affect Disord 2009;116:1-3.
10. Kendler KS, Myers J, Zisook S. Does bereavement-related major depression differ from major depression associated with other stressful life events? Am J Psychiatry 2008;165:1449-55.
11. Wakefield JC, Schmitz MF, First MB et al. Extending the bereavement exclusion for major depression to other losses. Evidence from the National Comorbidity Survey. Arch Gen Psychiatry 2007;64:433-40.
12. Wakefield JC, Schmitz MF, Baer JC. Did narrowing the major depression bereavement exclusion from DSM-III-R to DSM-IV increase validity? Evidence from the National Comorbidity Survey. J Nerv Ment Dis 2011;199:66-73.

Validity of the bereavement exclusion to major depression: does the empirical evidence support the proposal to eliminate the exclusion in DSM-5?

JEROME C. WAKEFIELD[1,2,3,4], MICHAEL B. FIRST[4,5]

[1]Department of Psychiatry, School of Medicine, New York University, 550 First Avenue, New York, NY 10016; [2]Silver School of Social Work, New York, NY 10003; [3]InSPIRES (Institute for Social and Psychiatric Initiatives–Research, Education and Services), Bellevue Hospital/New York University, New York, NY 10016; [4]Department of Psychiatry, Division of Clinical Phenomenology, Columbia University College of Physicians and Surgeons, New York, NY 10032; [5]New York State Psychiatric Institute, 1051 Riverside Drive, New York, NY 10032, USA

The DSM-IV major depression "bereavement exclusion" (BE), which recognizes that depressive symptoms are sometimes normal in recently bereaved individuals, is proposed for elimination in DSM-5. Evidence cited for the BE's invalidity comes from two 2007 reviews purporting to show that bereavement-related depression is similar to other depression across various validators, and a 2010 review of subsequent research. We examined whether the 2007 and 2010 reviews and subsequent relevant literature support the BE's invalidity. Findings were: a) studies included in the 2007 reviews sampled bereavement-related depression groups most of whom were not BE-excluded, making them irrelevant for evaluating BE validity; b) three subsequent studies cited by the 2010 review as supporting BE elimination did examine BE-excluded cases but were in fact inconclusive; and c) two more recent articles comparing recurrence of BE-excluded and other major depressive disorder cases both support the BE's validity. We conclude that the claimed evidence for the BE's invalidity does not exist. The evidence in fact supports the BE's validity and its retention in DSM-5 to prevent false positive diagnoses. We suggest some improvements to increase validity and mitigate risk of false negatives.

Key words: Major depression, bereavement, grief, DSM-5, diagnosis, validity, harmful dysfunction

(World Psychiatry 2012;11:3-10)

The DSM-5 Mood Disorders Work Group has proposed eliminating in DSM-5 the major depression criterion E, "bereavement exclusion" (BE), which recognizes that depressive symptoms are sometimes normal in recently bereaved individuals (1,2). This proposal has become one of the more contentious issues regarding the DSM-IV revision (3-9).

Those favoring the BE's elimination argue that the empirical evidence demonstrates the BE's invalidity and supports its removal. For example, Zisook et al (10), reviewing studies "that bear on the validity of the 'bereavement' exclusion", conclude that "the preponderance of available data suggests that excluding recently bereaved individuals from the diagnosis of MDE… may no longer be justified"; and Lamb et al (11) assert that, since Zisook et al's review, "four other studies have been published that provide further evidence supporting the removal of the bereavement exclusion".

In this review, we examine whether these claims are justified. We evaluate the quality of the evidence put forward in the cited reviews, and also examine some more recent evidence bearing on the validity of the BE. Based on our results, we offer some recommendations for DSM-5.

THE BEREAVEMENT EXCLUSION

Prospective studies of bereavement (12-14) have demonstrated what physicians have long known (15,16), that normal grief frequently includes depressive symptoms such as sadness, difficulty sleeping, decreased appetite, fatigue, diminished interest or pleasure in usual activities, and difficulty concentrating on usual tasks. A considerable number of individuals reach the 5-symptoms-for-2-weeks level that satisfies diagnostic criteria for major depressive disorder (MDD), and many experience clinically significant distress or role impairment due to their grief. Yet, their bereavement-related depression may resolve over time without treatment and may not have the chronic and recurrent course seen in MDD. The overlap of symptoms between intense normal grief and MDD creates a potential false positive problem in which depressions that are part of normal bereavement may be misdiagnosed as MDD.

Excluding all bereavement-related depressions from MDD diagnosis is no solution. Severe emotional stressors such as bereavement can trigger genuine MDD (17). Consequently, the diagnostic challenge is to distinguish those bereavement-related depressions that are likely intense normal grief from those that have turned into pathological depressions. The BE, which has been in the DSM in varying forms since 1980, offers the clinician guidance in making this difficult discrimination. It excludes bereavement-related depressions from MDD diagnosis only when they are "uncomplicated", that is, they manifest certain duration and symptom features more consistent with normal grief than with mental disorder.

The BE first specifies that, to be included in MDD, a depression must "not be better accounted for by bereavement". That is, the clinician is asked to compare two rival hypotheses regarding the patient's depressive feelings, MDD versus depressive symptoms that are part of normal grief.

The BE goes on to operationalize what features would suggest the depression qualifies for the diagnosis of MDD.

If the depressive episode either lasts longer than 2 months or includes at least one of a series of features that are uncharacteristic of normal grief (i.e., marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation), then the episode should be diagnosed as MDD. Conversely, if the episode resolves within 2 months and does not include any of the uncharacteristic features, then it is consistent with normal grief and is excluded from the MDD diagnosis.

THE ZISOOK AND KENDLER (2007) REVIEW

The "rationale" section on the DSM-5 website's major depressive episode page explains that the reason for eliminating the BE is that "evidence does not support separation of loss of loved one from other stressors" (18). The website cites only one reference as the basis for this proposal, a review paper by Zisook and Kendler (19) that claims that bereavement-related depressions are generally similar to standard depression.

Zisook and Kendler ask: "Is bereavement-related depression the same or different from standard (non-bereavement-related) major depression?". To answer this question, they compared bereavement-related depression to "standard" major depression following other triggers or no triggers. They evaluated whether the two conditions are similar or different on a variety of variables divided into antecedent, concurrent, and predictive "validators", including demographic variables, family and past personal history of major depression, health and social support, associated clinical features, biological factors, persistence, and response to treatment, and claim they are similar on most validators. "Similarity" was not precisely defined, but seemed to be understood as having significant relationships to a variable in the same direction. Given that the bulk of standard major depression is clearly disordered, if bereavement-related depression and standard major depression share enough "validators", this was taken to imply they are likely the same pathological condition.

However, in terms of assessing the BE's validity, there is a fatal flaw to this review. Comparing all bereavement-related depressions to all standard major depressions has little to do with the evaluation of the BE. The point of the BE is to distinguish between excluded "uncomplicated" likely-normal bereavement-related depressions versus non-excluded likely-disordered ones. The BE implies at most only that excluded bereavement-related depression is different from standard major depression; the BE declares non-excluded bereavement-related depression to be pathological. Combining the two bereavement-related depression groups and finding similarity to standard major depression does not test the BE.

Zisook and Kendler acknowledge the problem. They note that an evaluation of the BE must distinguish between those "who are considered by the DSM-IV-TR to be experiencing normal bereavement" and those "whose symptoms are so severe or persistent that the DSM-IV-TR recommends considering the diagnosis of a true major depressive episode rather than just normal bereavement", and that their review largely fails to meet this requirement. In comparing all bereavement-related depressions, most of which the BE labels MDD and not normal grief, to standard major depressions, it is hardly surprising that Zisook and Kendler find similarity across a range of validators.

THE ZISOOK, SHEAR AND KENDLER (2007) REVIEW

A subsequent review attempted to overcome these difficulties and to specifically evaluate the validity of the BE. Zisook et al (10) acknowledge the weakness of the earlier review, focusing on its failure to observe the BE's duration requirement: "Since most of the studies reviewed did not describe or follow individuals with bereavement-related depression specifically within the first two months of bereavement (the period of time the DSM-IV-TR demarcates as excluding a diagnosis of major depressive episode), we were unable to draw definitive conclusions about the validity of the bereavement exclusion".

Zisook et al cite no new evidence, and conduct the same type of "similarity" analysis using the same variables as in the earlier review. However, they attempt to fix the problem with the earlier review by focusing on studies of depressive syndromes evaluated during the first two months of bereavement, referred to here as "early-phase bereavement-related depression", which they consider directly relevant to assessment of the BE. Finding many similar relationships to validators, they conclude that the BE "is not valid because, using validating criteria, bereavement-related depression within the first two months after the death of a loved one resembles non-bereavement-related depression".

Here, as in the earlier review, the concept of similarity remains fuzzy. Does "similar" mean that correlations must be of comparable size? Sometimes it seems so, with relationships declared to be "virtually identical". Or, does "similar" just mean that correlations must be in the same direction, even if very different? Few quantitative comparisons are made, so in effect the latter, weaker approach is taken.

Despite the authors' claims to the contrary, in fact the Zisook et al review (10) offers no more support for the BE's invalidity than did the earlier one (19). Throughout the paper, from the title ("Validity of the bereavement exclusion criterion for major depressive episode") to the conclusion, the paper is framed as though it is reviewing studies pertinent to the BE's validity. A careful examination however reveals that *not one of the cited studies actually examined cases that satisfy the BE*. The BE's duration limitation, that excludes episodes which end by ≤2 months, and its requirement that excluded episodes lack the three uncharacteristic symptoms are its "core features". Yet, not one study cited by

Zisook et al applies either the duration or symptom requirements to the studied group. Consequently, they, too, examined mixed groups of BE-excluded and (mostly) non-excluded bereavement-related depressions. Again, it is unsurprising that correlations with validators are in the same direction as standard major depression.

Instead of the BE's 2-month duration limitation for exclusion, the Zisook et al review substitutes the "early-phase" requirement that the bereavement-related depressions in a study must be assessed prior to 2 months post-loss, no matter what their ultimate duration. From a duration perspective, these cases satisfy the BE only provisionally. Some of these cases will resolve within two months and thus ultimately meet the BE's duration limit. Many others, however, will continue for far longer than 2 months and thus ultimately not meet BE criteria, and would be classified as true MDD cases. After all, every bereavement-related depression, excluded or non-excluded, has an early phase. Thus, including cases based on their being evaluated within the two-month window rather than resolving within two months creates a mixed group of cases, some of which (in retrospect) will meet the BE whereas others will not. The Zisook et al review thus repeats in altered form the central error of the Zisook and Kendler review of attempting to draw conclusions about the similarity of BE-excluded bereavement-related depressions to standard major depressions from studies of mixed groups of bereavement-related depressions that mostly consist of non-excluded cases classified by the BE as pathological. Zisook et al acknowledge that their samples mix together excluded and non-excluded bereavement-related depressions: "Early bereavement-related depression, as conceptualized in this paper, is likely a mixture of cases including: those with "bereavement" as defined by the DSM-IV; those that start out with DSM-IV "bereavement" and evolve into true major depressive episode". However, they fail to recognize that this undermines any claim to showing BE invalidity.

The other core BE requirement for exclusion is that the episode does not include the so-called "uncharacteristic symptoms" (i.e., suicidal ideation, a sense of worthlessness, or psychomotor retardation). The Zisook et al review completely ignores this component of the BE criterion, taking the duration requirement (in the mistaken form considered above) as an adequate approximation to the BE. Yet epidemiological evidence suggests the symptom criteria are important independently of duration as determinants of whether the BE is met. For example, in the National Comorbidity Survey, of all those who reported bereavement-related depressions that lasted a total of 2 months or less, only 50% qualified for the BE; the other 50% manifested one or more symptoms disqualifying them from exclusion (Wakefield and Schmitz, unpublished analysis).

In sum, although Zisook et al claim to establish the invalidity of the BE, not one of the studies they cite applied the BE's duration or symptom criteria. The reviewed articles are essentially irrelevant to claims about the BE's validity.

ADDITIONAL ISSUES RAISED BY THE ZISOOK, SHEAR AND KENDLER (2007) REVIEW

Are the validators indicators of disorder?

The Zisook et al review offers no support for the utility of the selected validators in distinguishing between normal distress and mental disorder. If excludable bereavement-related depressions and standard major depressions are similar in their correlations to a validator, that proves nothing about the disordered nature of the excluded bereavement-related depressions if the validator itself tends to correlate both with disorder and normal distress. For example, the fact that larger percentages of women than men have both standard major depressions and excluded bereavement-related depressions (assuming that would be shown in an examination of legitimate studies of excluded bereavement-related depressions) could just mean that women react with more emotional intensity both in normality and disorder. Similarly, biological variables such as immune, endocrinological, and sleep changes occur in a wide variety of disordered and normal stressful conditions, even for example before major examinations (20), and thus are not specific enough to standard major depression or to disorder to suggest any conclusion about whether bereavement-related depressions are disorders.

Confusingly, some of Zisook et al's seemingly more promising validators of pathology (10) are definitionally linked to the BE criteria in ways that make their use as validators incoherent. For example, their use of "clinical features" (i.e., "suicidal thoughts, feelings of worthlessness and psychomotor disturbances") and "persistence" of bereavement-related depressions as validators makes no sense because, by definition, BE-excluded cases cannot have suicidal thoughts, feelings of worthlessness and psychomotor disturbances, and cannot persist past 2 months.

Treatment response as a validator

Jan Fawcett, the Chair of the DSM-5 Mood Disorders Work Group, in reviewing proposed changes (2), credits treatment response as the sole reason for eliminating the BE, citing Zisook and Kendler (19), who in turn based their claim completely on a single 2001 study by Zisook et al (21). In this study, 22 bereaved individuals satisfying DSM-IV MDD criteria about 2 months post-loss were treated with bupropion-SR for 2 months; 13 subjects experienced a reduction of ≥50% on Hamilton Depression Rating Scale scores. Given the small sample size and the fact that Zisook et al's study contains no control group in a diagnostic area with notoriously high placebo response rates, the results are impossible to interpret. Furthermore, given that prospective studies reveal that without treatment bereavement-related depressions have precipitous drops in symptoms after 2 months post-loss, the "response rate" is consistent with the natural course of bereavement. Even if bereavement-related depressions should re-

spond to medication, it is unclear why treatment response would be a reason for considering a condition pathological, given that many normal conditions respond to medication.

The suicide risk argument

Some proponents of eliminating the BE raise the spectre of suicide in excluded bereavement-related depressions. For example, Zisook et al (10) cite a study showing an elevated rate of suicide in MDD among those without partners. Shear et al (22), in considering the BE, note that "bereavement may increase the risk of suicide" and emphasize the value of early treatment. This issue was also raised by Zisook in a National Public Radio interview (3), in which he is quoted as saying: "I'd rather make the mistake of calling someone depressed who may not be depressed, than missing the diagnosis of depression, not treating it, and having that person kill themselves" .

Some bereaved individuals do attempt suicide, whether depressed or not, and missed cases can occur in many contexts. However, cases excluded by the BE by definition lack suicidal ideation. There is no evidence for elevated suicide risk in excluded bereavement-related depressions, and evidence suggests the opposite. For example, among those individuals who had only DSM-IV-excludable bereavement-related depressions in the National Comorbidity Survey (N=31), not one reported a lifetime suicide attempt (Wakefield and Schmitz, unpublished data). The study Zisook et al cite to establish elevated suicide risk in those without partners (23) has as subjects many severely pathological inpatients with prior suicide attempts, a sample irrelevant to predicting behavior by individuals with typical excluded bereavement-related depressions.

LAMB, PIES AND ZISOOK (2010) REVIEW

In a review published in 2010, Lamb et al (11) claimed that several studies published since the earlier reviews support the BE's invalidity. Some of the studies they cite do examine actual BE-excluded cases. We discuss each of the cited studies.

Studies failing to apply the BE criteria

Kessing et al (24) used the Danish Psychiatric Central Research Register to compare first-onset MDD following bereavement (N=26) versus other stressors or no stressors. They reported that bereaved patients did not differ from the other two groups on several variables. However, they did not identify BE-excluded cases. Two-thirds of the sample were inpatients and subjects were required to have received antidepressant treatment for at least a week, making it exceedingly unlikely that many were BE-excluded cases. Furthermore, as might be expected in a largely inpatient sample, 73% of the sample (19 out of 26) displayed suicidal ideation, yet suicidal ideation disqualifies for BE exclusion. In sum, the study does not examine BE-excluded cases and does not address the BE's validity.

Corruble et al (25) claim to study BE-excluded cases diagnosed by French physicians, but the BE was inaccurately applied. In this and other studies (26,27), this group reports provocative findings supposedly showing that BE-excluded cases are as or more severe than standard MDD and non-excluded bereavement-related depressions across a variety of features, ranging from symptom severity and treatment response to cognitive impairment, concluding that the BE should be eliminated. These startling claims go against the logic of the BE, which is constructed to exclude severe cases and conflicts with findings from earlier empirical studies comparing excluded to non-excluded bereavement-related depressions (28).

In fact, close inspection of Corruble et al's results reveals that the so-called BE-excluded cases did not in fact meet the BE criterion. The study found, for example, that 70.5% of excluded bereavement-related depressions manifested psychomotor retardation, 66.8% worthlessness, and 36.0% suicidal ideation. Yet, such symptoms disqualify an episode from BE exclusion. Thus, it appears that the great majority of claimed BE-excluded individuals did not qualify for exclusion.

The likely explanation for this apparent contradiction is simple (29): Corruble et al asked physicians to judge whether patients were excluded by the BE without any special training or checklist, then took those judgments at face value without validating that they were accurate. Apparently, the vast majority were incorrect, most likely because they were confused by the BE's double-negative wording. Indeed, one of us (MBF) encounters the resulting confusion frequently when doing training sessions for the Structured Clinical Interview for DSM (SCID) (30), with novice SCID users often coding MDD criterion E oppositely to what they intend.

Consequently, the Corruble et al results are not based on a true BE-excluded sample, and are not generalizable to any sample to which the BE is correctly applied. The results thus have no implications for the evaluation of the BE's validity. At most, they indicate that the BE's current wording is confusing to novices and likely requires clarification.

Studies applying the BE criteria

Three studies cited by Lamb et al do examine samples of BE-excluded cases that satisfy both core BE criteria. Karam et al's (31) prospective community study of depression among Lebanese people exposed to civil war found no statistically significant difference in 2-year recurrence rates between the five cases of DSM-excluded bereavement-related depression (40% recurrence) and standard MDD (61% recurrence). However, given the exceedingly small sample size, one must agree with Karam et al's caution that "the number of DSM-IV

excluded episodes was too small to allow for generalization". Exposure to civil war may also have raised the rates of normal and disordered distress to a degree that obscured true recurrence rates, further limiting generalizability.

Wakefield et al (28) compared excluded to non-excluded bereavement-related depressions in the National Comorbidity Survey. They argued that the large differences found on the study's nine validators (number of symptoms, melancholic depression, suicide attempt, duration of symptoms, interference with life, recurrence, and three service use variables), supported the validity of the BE.

However, critics argued that some validators were too closely related to the defining features of complicated episodes to provide unbiased tests (e.g., the validators "interference with life" and "suicide attempt" are closely related to the BE components "marked impairment" and "suicidal ideation," respectively) (32). Thus, the critics argued, the demonstrated differences were due to these biases and in effect tautological. These criticisms have some merit, although they do not impact all the validators. Whether the claimed biases were actually responsible for the findings can be empirically evaluated, but no study has attempted such an analysis as of this writing, so the implications of the Wakefield et al study for BE validity remain uncertain.

Kendler et al (33) compared bereavement-related depressions and standard major depressions on a range of validators in a sample of Virginia twins evaluated for 1-year depression at 4 points over 10 years. Although they did identify BE-excluded episodes, they did not compare excludable bereavement-related depressions to non-excludable bereavement-related depressions or standard major depressions in general. Instead, they examined the relationship between excluded bereavement-related depressions and "excludable" standard major depressions (that is, standard major depressions satisfying the BE's duration and symptom criteria for exclusion). The rationale for this shift of focus was that the DSM currently classifies such "excludable" standard major depressions as disorders, so if excluded bereavement-related depressions are similar to excluded standard major depressions – which both their study and Wakefield et al (28) showed they are – they must be disorders too. Such similarity, they argued, shows that the DSM's exclusion of uncomplicated bereavement-related depressions but not uncomplicated standard major depressions is an inconsistency that must be resolved by removing the BE (7).

However, the dispute over whether the BE is valid must be distinguished from the separate question of whether similarly transient non-severe depressive reactions to other stressors – such as marital dissolution or job loss – are properly considered disorders or should be excluded from MDD as well. To address the latter question, the similarities and differences between excludable standard major depressions and other standard major depressions would have to be examined, a comparison Kendler et al do not pursue in their data. The introduction of the BE was based on an evaluation of the evidence that bereavement-related depressions are sometimes not MDD, whereas the inclusion of other BE-satisfying episodes within MDD occurred without specific evidential evaluation and is not asserted by the BE. In any event, the net effect of the Kendler et al interpretation was that they did not analyze their data in a way that might directly bear on BE validity.

In sum, three of the studies cited by Lamb et al do properly apply the BE criteria to a sample. However, for varying reasons, none of the three offer substantial evidence for or against the validity of the BE.

RECENT STUDIES OF MDD RECURRENCE AFTER EXCLUDED BEREAVEMENT-RELATED DEPRESSIONS

Perhaps the validator with the most face validity in evaluating the relationship between excluded bereavement-related depression and standard major depression is recurrence of depressive episodes. There is a well-established heightened risk of developing future depressive episodes in individuals suffering from standard major depression, whereas in normal emotional reactions one would plausibly expect less recurrence. Moreover, recurrence is not a BE criterion, so it can be used to compare excluded bereavement-related depression versus standard major depression without tautologically biasing the result.

Mojtabai's (34) recent prospective study, using the 2-wave National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) community sample, is the first to compare BE-excludable vs. standard major depression recurrence in a methodologically rigorous and adequately powered study. Mojtabai compared the risk of depression during a 3-year follow-up period in participants who at baseline had lifetime BE-excluded depressive episodes, those who had other kinds of depressive episodes, and those with no history of depression. He found that participants who at wave 1 had experienced a single lifetime DSM-excluded bereavement-related depression (N=162) were no more likely to have an MDD episode over a 3-year follow-up period than were those in the general population who had no lifetime history of MDD at baseline (4.3% vs. 7.5%, respectively). In comparison, participants who had experienced either single brief standard major depressions, single non-brief standard major depressions, or recurrent MDD, had significantly higher 3-year recurrence rates (14.7%, 20.1%, and 27.2%, respectively) than those without a depression history or those with BE-excluded episodes. Mojtabai concluded that "the findings support preserving the DSM-IV bereavement exclusion criterion for major depressive episodes in the DSM-V".

Wakefield and Schmitz (35) attempted to replicate Mojtabai's results in the 2-wave longitudinal Epidemiologic Catchment Area Study. They compared 1-year depression recurrence rates at wave 2 in four wave 1 lifetime-disorder baseline groups: excludable bereavement-related depression; brief standard major depression; non-brief standard major depres-

sion; and no history of depression. The BE-excluded 1-year recurrence rate (3.7%, N=25) was not significantly different from the rate in the no-depression-history group (1.7%), but significantly and substantially lower than rates for brief and non-brief standard major depression (14.4% and 16.2%, respectively).

These findings confirm Mojtabai's (34) results using a different data set and follow-up period, supporting generalizability and substantially strengthening the case for the BE's validity. The Mojtabai and the Wakefield and Schmitz studies contradict the central argument for BE elimination, that there is no evidence that BE-excludable bereavement-related depression differs relevantly in course from standard major depression.

RECOMMENDATIONS FOR IMPROVING THE BEREAVEMENT EXCLUSION IN DSM-5

Although the literature does not support the invalidity of the BE or its elimination from DSM-5, there are some changes that could improve its validity and limit its misuse.

Use of a "provisional" qualifier

In epidemiological surveys or when evaluating a patient's history, the full duration of bereavement-related depressions may be known retrospectively. However, in clinical practice, bereaved patients experiencing depressive symptoms for less than 2 months must be diagnosed before knowing how long the episode will endure. The BE's duration and symptom criteria create uncertainty for the diagnostician: will the depressive symptoms persist beyond 2 months, or one or more uncharacteristic symptoms develop, necessitating a revised diagnosis of MDD?

There are several examples in DSM-IV of disorders whose diagnostic criteria depend on the condition resolving before some upper durational limit, where the diagnosis changes if the condition continues beyond the specified point. If a diagnosis must be made before that limit has been reached, the diagnosis must be provisional, due to lack of certainty whether the condition will resolve within the allotted time.

For example, according to the DSM-IV-TR (36), schizophreniform disorder requires that "an episode of the disorder… lasts at least 1 month but less than 6 months". If the identical symptoms persist longer than 6 months, the diagnosis is schizophrenia. For patients who present with ongoing symptoms of more than 1 and less than 6 month duration, the clinician is instructed to qualify the diagnosis as "provisional", because it is not yet known whether the symptoms will resolve within the required 6-month window. If not, then the diagnosis would be revised from schizophreniform disorder to schizophrenia.

The DSM-IV-TR's (36) "Use of the Manual" section notes that this diagnostic principle applies to any situation "in which differential diagnosis depends exclusively on the duration of illness". Thus, for example, because transient intense fears are common in childhood, the DSM specifies that a child's fear can be diagnosed as a specific phobia only if it lasts for at least 6 months. Consequently, a child with intense fears of large animals of 2 month duration must be diagnosed provisionally as normal, with watchful waiting used to establish whether the fear endures past 6 months and thus qualifies as a phobia.

The diagnosis of excluded bereavement-related depressions evaluated shortly after loss of a loved one fits this schema. Exclusion requires that the duration be 8 weeks or less, but the clinician must often make the diagnosis before it is known whether the symptoms will resolve by 8 weeks. Thus, following DSM-IV principles, it would be useful to use the "provisional" modifier for cases of excluded bereavement in which depressive symptoms are ongoing. The addition of "provisional" will serve to alert the clinician that a definitive diagnosis depends on the collection of more information, in this case a determination of whether the depressive symptoms have resolved by the 8-week point without development of uncharacteristic symptoms. This change could prevent some false negatives that might occur due to premature assumptions about the final diagnosis.

"Past history of MDD" as a criterion disqualifying exclusion

In guiding the provisional judgment whether bereavement-related depression symptoms are better explained as MDD or normal grief, an improvement in BE criteria that would protect against missing genuine cases would be to incorporate into the criteria the requirement that past history of MDD disqualifies a bereavement-related depression for exclusion. Individuals with a past MDD history have a vulnerability to developing MDD that might easily be activated under the severe stress of experiencing the loss of a loved one. The literature suggests that, among those experiencing an early-phase bereavement-related depression, past MDD history strongly predicts persistence, severe symptoms, and non-excludability. This variable is impactful enough that research reports often separate outcomes according to past history (e.g., 37).

For example, in Zisook and Schuchter's (14,38) classic study of the course of bereavement, 89 individuals satisfied DSM MDD criteria by 2 months post-loss, and of those, 20 (22%) were depressed at 13 months and considered disordered. However, 14 individuals satisfying MDD criteria at 2 months had a history of prior MDD, and 14 individuals still satisfying MDD criteria at 13 months also had such a history. Presuming those are the same individuals, then if individuals with prior MDD histories had been removed from the group to which the BE might be applied provisionally at 2 months, the false negatives rate based on duration alone would have fallen from 22% to 8%. We thus propose that a personal history of MDD should mitigate against provisional BE exclusion.

Improved wording of the BE

As discussed above, the studies by Corruble et al (25), which purported to support the elimination of the BE, in fact indicate how prone the BE is to misinterpretation and misapplication by clinicians not specifically trained in its application. Much of the problem likely results from the potentially confusing double-negative wording.

The wording could easily be improved to reduce the chance of such confusion. As a beginning point for discussion, we offer the following rewording of criterion E, incorporating suggestions made above:

If the episode occurs in the context of bereavement, it presents at least one of the following features suggestive of major depression rather than normal grief: duration greater than 2 months; suicidal ideation; morbid preoccupation with worthlessness; marked psychomotor retardation; prolonged and marked global functional impairment; psychotic symptoms; or a history of major depressive disorder in circumstances other than bereavement.

Bereavement-related depressive episodes that have none of these features should be given a diagnosis of "normal bereavement-related depression, provisional".

Beyond these changes, there are many questions that might be raised about how to achieve the optimal validity of the BE. For example, should the current 2-month duration threshold for non-exclusion be lengthened, based on recent evidence that optimal validity may be achieved at greater durations (39, 40)? Are the current uncharacteristic symptoms optimally valid? Is impairment a useful addition? And finally, should similar reactions to other life stressors be placed within an expanded BE to create a "stressor exclusion"? These questions await further evidence for their resolution.

However, the question of whether there is empirical evidence that the BE is invalid can be resolved. The claim that there is such evidence is based on faulty interpretations of the literature and has no basis in scientific fact. Consequently, there is no scientific basis for removing the bereavement exclusion from the DSM-5.

References

1. American Psychiatric Association. Criteria for major depressive episode. www.dsm5.org.
2. Fawcett J. An overview of mood disorders in the DSM-5. Curr Psychiatry Rep 2010;12:531-8.
3. National Public Radio, Morning Edition. "Today in your health: bereavement". August 2, 2010.
4. Pies R. Depression and the pitfalls of causality: implications for DSM-V. J Affect Disord 2009;116:1-3.
5. Wakefield JC. Should uncomplicated bereavement-related depression be reclassified as a disorder in DSM-5?: Response to Kenneth S. Kendler's statement defending the proposal to eliminate the bereavement exclusion. J Nerv Ment Dis 2011;199:203-8.
6. Maj M. Depression, bereavement, and "understandable" intense sadness: should the DSM-IV approach be revised? Am J Psychiatry 2008;165:1373-5.
7. Kendler KS. A statement from Kenneth S. Kendler, M.D., on the proposal to eliminate the grief exclusion criterion from major depression. www.dsm5.org.
8. First MB. DSM-5 proposals for mood disorders: a cost–benefit analysis. Curr Opin Psychiatry 2011;24:1-9.
9. Frances A. Good grief. New York Times, August 15, 2010.
10. Zisook S, Shear K, Kendler KS. Validity of the bereavement exclusion criterion for the diagnosis of major depressive episode. World Psychiatry 2007;6:102-7.
11. Lamb K, Pies R, Zisook S. The bereavement exclusion for the diagnosis of major depression: to be or not to be. Psychiatry 2010;7:19-25.
12. Clayton P, Desmarais L, Winokur G. A study of normal bereavement. Am J Psychiatry 1968;125:168-78.
13. Clayton P, Halikas JA, Maurice WL. The depression of widowhood. Br J Psychiatry 1972;120:71-7.
14. Zisook S, Shuchter SR. Depression through the first year after the death of a spouse. Am J Psychiatry 1991;148:1346-52.
15. Horwitz AV, Wakefield JC. The loss of sadness: how psychiatry transformed normal sadness into depressive disorder. New York: Oxford University Press, 2007.
16. Jackson SW. Melancholia and depression: from Hippocratic times to modern times. New Haven: Yale University Press, 1986.
17. Parkes C. Recent bereavement as a cause of mental illness. Br J Psychiatry 1964;110:198-204.
18. American Psychiatric Association. Major depressive episode. Rationale. www.dsm5.org.
19. Zisook S, Kendler KS. Is bereavement-related depression different than nonbereavement-related depression? Psychol Med 2007;37:779-94.
20. Webster Marketon JI, Glaser R. Stress hormones and immune function. Cell Immun 2008;252:16-26.
21. Zisook S, Shuchter SR, Pedrelli P et al. Bupropion sustained release for bereavement: results of an open trial. J Clin Psychiatry 2001;62:227-30.
22. Shear MK, Simon N, Wall M et al. Complicated grief and related bereavement issues for DSM-5. Depress Anxiety 2011;28: 103-17.
23. Sokero TP, Melartin TK, Rytsälä HJ et al. Prospective study of risk factors for attempted suicide among patients with DSM-IV major depressive disorder. Br J Psychiatry 2005;186:314-8.
24. Kessing LV, Bukh JD, Bock C et al. Does bereavement-related first episode depression differ from other kinds of first depressions? Soc Psychiatry Psychiatr Epidemiol 2010;45:801-8.
25. Corruble E, Chouinard VA, Letierce A et al. Is DSM-IV bereavement exclusion for major depressive episode relevant to severity and pattern of symptoms? A case control, cross-sectional study. J Clin Psychiatry 2009;70:1091-7.
26. Corruble E, Falissard B, Gorwood P. DSM bereavement exclusion for major depression and objective cognitive impairment. J Affect Disord 2011;130:113-7.
27. Corruble E, Falissard B, Gorwood P. Is DSM-IV bereavement exclusion for major depression relevant to treatment response? A case-control, prospective study. J Clin Psychiatry 2010;72:898-902.
28. Wakefield JC, Schmitz MF, First MB et al. Should the bereavement exclusion for major depression be extended to other losses?: Evidence from the National Comorbidity Survey. Arch Gen Psychiatry 2007;64:433-40.
29. Wakefield JC, First MB. Treatment outcome for bereavement-excluded depression: results of the study by Corruble et al are not what they seem. J Clin Psychiatry 2011;72:1155.
30. First MB, Spitzer RL, Gibbon M et al. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research Department, 2011.
31. Karam EG, Tabet CC, Alam D et al. Bereavement related and non-bereavement related depressions: a comparative field study. J Affect

Disord 2009;112:102-10.
32. Kendler KS, Zisook S. Drs. Kendler and Zisook reply. Am J Psychiatry 2009;166:492-3.
33. Kendler KS, Myers J, Zisook S. Does bereavement-related major depression differ from major depression associated with other stressful life events? Am J Psychiatry 2008;165:1449-55.
34. Mojtabai R. Bereavement-related depressive episodes: characteristics, 3-year course, and implications for the DSM-5. Arch Gen Psychiatry 2011;68:920-8.
35. Wakefield JC, Schmitz MF. Recurrence of bereavement-related depression: evidence for the validity of the DSM-IV bereavement exclusion from the Epidemiologic Catchment Area Study. J Nerv Ment Dis (in press).
36. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. 4th ed. Washington: American Psychiatric Association, 2000.
37. Brent DA, Perper JA, Moritz G et al. Major depression or uncomplicated bereavement? A follow-up of youth exposed to suicide. J Am Acad Child Adolesc Psychiatry 1994;33:231-9.
38. Zisook S, Shuchter SR. Uncomplicated bereavement. J Clin Psychiatry 1993;54:365-72.
39. Wakefield JC, Schmitz MF, Baer JC. Did narrowing the major depression bereavement exclusion from DSM-III-R to DSM-IV increase validity? Evidence from the National Comorbidity Survey. J Nerv Ment Dis 2011;199:66-73.
40. Wakefield JC, Schmitz MF, Baer JC. Relation between duration and severity in bereavement-related depression. Acta Psychiatr Scand (in press).

An attachment perspective on psychopathology

Mario Mikulincer[1], Phillip R. Shaver[2]

[1]School of Psychology, Interdisciplinary Center Herzliya, P.O. Box 167, Herzliya 46150, Israel
[2]Department of Psychology, University of California at Davis, One Shields Avenue, Davis, CA 95616-8686, USA

In recent years, attachment theory, which was originally formulated to describe and explain infant-parent emotional bonding, has been applied to the study of adolescent and adult romantic relationships and then to the study of psychological processes, such as interpersonal functioning, emotion regulation, coping with stress, and mental health. In this paper, we offer a brief overview of the attachment perspective on psychopathology. Following a brief account of attachment theory, we go on to explain how the study of individual differences in adult attachment intersects with the study of psychopathology. Specifically, we review research findings showing that attachment insecurity is a major contributor to mental disorders, and that the enhancement of attachment security can facilitate amelioration of psychopathology.

Key words: Attachment, psychopathology, emotion regulation, security, interpersonal relations, self, mental health

(World Psychiatry 2012;11:11-15)

Attachment theory (1-3) has proven to be a very fruitful framework for studying emotion regulation and mental health. In particular, research on adult attachment processes and individual differences in attachment orientations has provided strong evidence for the anxiety-buffering function of what Bowlby (2) called the attachment behavioral system and for the relevance of attachment-related individual differences to coping with stress, managing distress, and retaining psychological resilience (4).

In this paper, we offer a brief overview of the attachment perspective on psychopathology. Following a brief account of attachment theory's basic concepts, we review research findings showing that attachment insecurities – called attachment anxiety and avoidance in the theory – are associated with mental disorders, and that increases in attachment security are an important part of successfully treating these disorders.

ATTACHMENT THEORY: BASIC CONCEPTS

Bowlby (2) claimed that human beings are born with an innate psychobiological system (the *attachment behavioral system*) that motivates them to seek proximity to significant others (*attachment figures*) in times of need. Bowlby (1) also outlined major individual differences in the functioning of the attachment system. Interactions with attachment figures who are available in times of need, and who are sensitive and responsive to bids for proximity and support, promote a stable sense of attachment security and build positive mental representations of self and others. But when a person's attachment figures are not reliably available and supportive, proximity seeking fails to relieve distress, felt security is undermined, negative models of self and others are formed, and the likelihood of later emotional problems and maladjustment increases.

When testing this theory in studies of adults, most researchers have focused on the systematic pattern of relational expectations, emotions, and behavior that results from one's attachment history – what Hazan and Shaver (5) called *attachment style*. Research clearly indicates that attachment styles can be measured in terms of two independent dimensions, attachment-related anxiety and avoidance (6). A person's position on the anxiety dimension indicates the degree to which he or she worries that a partner will not be available and responsive in times of need. A person's position on the avoidance dimension indicates the extent to which he or she distrusts relationship partners' good will and strives to maintain behavioral independence, self-reliance, and emotional distance. The two dimensions can be measured with reliable and valid self-report scales (e.g., 6), and they are associated in theoretically predictable ways with relationship quality and adjustment (4).

Mikulincer and Shaver (4) proposed that a person's location in the two-dimensional conceptual space defined by attachment anxiety and avoidance reflects both the person's sense of attachment security and the ways in which he or she deals with threats and distress. People who score low on these dimensions are generally secure and tend to employ constructive and effective affect-regulation strategies. Those who score high on either the attachment anxiety or the avoidance dimension (or both) suffer from insecurity and tend to rely on what Cassidy and Kobak (7) called secondary attachment strategies, either deactivating or hyperactivating their attachment system in an effort to cope with threats.

According to Mikulincer and Shaver (4), people scoring high on avoidant attachment tend to rely on deactivating strategies – trying not to seek proximity, denying attachment needs, and avoiding closeness and interdependence in relationships. These strategies develop in relationships with attachment figures who disapprove of and punish closeness and expressions of need or vulnerability (8). In contrast, people scoring high on attachment anxiety tend to rely on hyperactivating strategies – energetic attempts to achieve proximity, support, and love combined with lack of confidence that these resources will be provided and with resentment and anger when they are not provided (7). These reac-

tions occur in relationships in which an attachment figure is sometimes responsive but unreliably so, placing the needy person on a partial reinforcement schedule that rewards persistence in proximity-seeking attempts, because they sometimes succeed.

Individual differences in attachment styles begin in interactions with parents during infancy and childhood (e.g., 9). However, Bowlby (3) claimed that meaningful relational interactions during adolescence and adulthood can move a person from one region to another of the two-dimensional conceptual space defined by attachment anxiety and avoidance. Moreover, a growing body of research shows that attachment style can change, subtly or dramatically, depending on current context, recent experiences, and recent relationships (e.g., 10,11).

ATTACHMENT, MENTAL HEALTH, AND PSYCHOPATHOLOGY

According to attachment theory, interactions with inconsistent, unreliable, or insensitive attachment figures interfere with the development of a secure, stable mental foundation; reduce resilience in coping with stressful life events; and predispose a person to break down psychologically in times of crisis (3). Attachment insecurity can therefore be viewed as a *general* vulnerability to mental disorders, with the particular symptomatology depending on genetic, developmental, and environmental factors.

Mikulincer and Shaver (4) reviewed hundreds of cross-sectional, longitudinal, and prospective studies of both clinical and non-clinical samples and found that attachment insecurity was common among people with a wide variety of mental disorders, ranging from mild distress to severe personality disorders and even schizophrenia. Consistently compatible results have also been reported in recent studies. For example, attachment insecurities (of both the anxious and avoidant varieties) are associated with depression (e.g., 12), clinically significant anxiety (e.g., 13), obsessive-compulsive disorder (e.g., 14), post-traumatic stress disorder (PTSD) (e.g., 15), suicidal tendencies (e.g., 16), and eating disorders (e.g., 17).

Attachment insecurity is also a key feature of many personality disorders (e.g., 18,19). However, the specific kind of attachment insecurity differs across disorders. Anxious attachment is associated with dependent, histrionic, and borderline disorders, whereas avoidant attachment is associated with schizoid and avoidant disorders. Crawford et al (18) found that attachment anxiety is associated with what Livesley (20) called the "emotional dysregulation" component of personality disorders, which includes identity confusion, anxiety, emotional lability, cognitive distortions, submissiveness, oppositionality, self-harm, narcissism, and suspiciousness. Crawford et al (19) also found that avoidant attachment is associated with what Livesley (20) called the "inhibitedness" component of personality problems, including restricted expression of emotions, problems with intimacy, and social avoidance.

Another related issue concerning the associations between attachment insecurities and psychopathology is the extent to which attachment insecurities are a sufficient cause of mental disorders. In our view, beyond disorders such as separation anxiety and pathological grief, in which attachment injuries are the main causes and themes, attachment insecurities per se are unlikely to be sufficient causes of mental disorders. Other factors (e.g., genetically determined temperament; intelligence; life history, including abuse) are likely to converge with or amplify the effects of attachment experiences on the way to psychopathology.

Consider, for example, the relation between attachment-related avoidance and psychological distress. Many studies of large community samples have found no association between avoidant attachment and self-report measures of global distress (4). However, studies that focus on highly stressful events, such as exposure to missile attacks, living in a dangerous neighborhood, or giving birth to a handicapped infant, have indicated that avoidance is related to greater distress and poorer long-term adjustment (4).

Life history factors are also important. For example, the association between attachment insecurity and depression is higher among adults with a childhood history of physical, psychological, or sexual abuse (e.g., 21). Stressful life events, poverty, physical health problems, and involvement in turbulent romantic relationships during adolescence also strengthen the link between attachment insecurity and psychopathology (e.g., 22).

The causal links between attachment and psychopathology are also complicated by research findings showing that psychological problems can increase attachment insecurity. Davila et al (23), for example, found that late adolescent women who became less securely attached over periods of 6 to 24 months were more likely than their peers to have a history of psychopathology. Cozzarelli et al (24) found that women who moved in the direction of insecure attachment over a 2-year period following abortion were more likely than other women who had an abortion to have a prior history of depression or abuse. Solomon et al (25) assessed attachment insecurities and PTSD symptoms among Israeli ex-prisoners of war (along with a matched control group of veterans) 18 and 30 years after their release from captivity. Attachment anxiety and avoidance increased over time among the ex-prisoners, and the increases were predicted by the severity of PTSD symptoms at the first wave of measurement.

Overall, attachment insecurities seem to contribute non-specifically to many kinds of psychopathology. However, particular forms of attachment insecurity seem to predispose a person to particular configurations of mental disorders. The attachment-psychopathology link is moderated by a large array of biological, psychological, and socio-cultural factors, and mental disorders per se can erode a person's sense of attachment security.

THE HEALING EFFECTS OF ATTACHMENT SECURITY

If attachment insecurities are risk factors for psychopathology, then the creation, maintenance, or restoration of a sense of attachment security should increase resilience and improve mental health. According to attachment theory, interactions with available and supportive attachment figures impart a sense of safety, trigger positive emotions (e.g., relief, satisfaction, gratitude, love), and provide psychological resources for dealing with problems and adversities. Secure individuals remain relatively unperturbed during times of stress, recover faster from episodes of distress, and experience longer periods of positive affectivity, which contributes to their overall emotional well-being and mental health.

In some of our studies, we have examined the effects of increased security on various indicators of mental health by experimentally activating mental representations of supportive attachment figures (e.g., 26,27). These research techniques, which we (11) refer to as "security priming", include subliminal pictures suggesting attachment-figure availability, subliminal names of people designated by participants as security-enhancing attachment figures, guided imagery highlighting the availability and supportiveness of an attachment figure, and visualization of the faces of security-enhancing attachment figures.

Security priming improves participants' moods even in threatening contexts and eliminates the detrimental effects of threats on positive moods (e.g., 26). Mikulincer et al (28) found that subliminal priming with security-related words mitigated cognitive symptoms of PTSD (heightened accessibility of trauma-related words in a Stroop-color naming task) in a non-clinical sample. Admoni (29) found that priming the names of each participant's security providers mitigated two cognitive symptoms of eating disorders (distorted body perception and heightened accessibility of food-related words in a Stroop task) in a sample of women hospitalized for eating disorders.

There is also preliminary evidence that a sense of security provided by a psychotherapist improves a client's mental health. In a study based on data from the multi-site National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program, Zuroff and Blatt (30) found that a client's positive appraisals of his or her therapist's sensitivity and supportiveness predicted relief from depression and maintenance of therapeutic benefits over an 18-month period. The results were not attributable to patient characteristics or severity of depression. In a one-year prospective study of the effectiveness of residential treatment of high-risk adolescents, Gur (31) found that staff members' provision of a sense of attachment security in the adolescents resulted in lower rates of anger, depression, and behavioral problems. Although these preliminary findings are encouraging, there is still a great need for additional well-controlled research examining the long-term effects of security-enhancing therapeutic figures on clients' mental health.

MEDIATING PROCESSES

According to attachment theory (3), the linkage between attachment insecurities (whether in the form of anxiety, avoidance, or both) and psychopathology is mediated by several pathways. In this section, we will review the most important of these pathways.

Self-representations

According to attachment theory and research, lack of parental sensitivity and responsiveness contributes to disorders of the self, characterized by lack of self-cohesion, doubts about one's internal coherence and continuity over time, unstable self-esteem, and over-dependence on other people's approval (e.g., 32,33). Insecure people are likely to be overly self-critical, plagued by self-doubts, or prone to using defenses, such as destructive perfectionism, to counter feelings of worthlessness and hopelessness (e.g., 34). These dysfunctional beliefs about oneself increase insecure people's risk for developing mental disorders.

Attachment research has also shown that attachment insecurities are associated with pathological narcissism (e.g., 35). Whereas avoidant attachment is associated with *overt* narcissism or grandiosity, which includes both self-praise and denial of weaknesses (36), attachment anxiety is associated with *covert* narcissism, characterized by self-focused attention, hypersensitivity to other people's evaluations, and an exaggerated sense of entitlement (36).

Emotion regulation

According to attachment theory, interactions with available attachment figures and the resulting sense of attachment security provide actual and symbolic supports for learning constructive emotion-regulation strategies. For example, interactions with emotionally accessible and responsive others provide a context in which a child can learn that acknowledgment and display of emotions is an important step toward restoring emotional balance, and that it is useful and socially acceptable to express, explore, and try to understand one's feelings (37).

Unlike relatively secure people, avoidant individuals often prefer to cordon off emotions from their thoughts and actions. As a result, they tend to present a façade of security and composure, but leave suppressed distress unresolved in ways that impair their ability to deal with life's inevitable adversities. This impairment is particularly likely during prolonged, demanding stressful experiences that require active coping with a problem and mobilization of external sources of support (e.g., 38).

People who score high on attachment anxiety, in contrast, often find negative emotions to be congruent with their attachment-system hyperactivation. For them, "emotion regu-

lation" can mean emotion amplification and exaggeration of worries, depressive reactions to actual or potential losses and failures, and PTSD intrusion symptoms following traumas. Attachment anxiety is also associated with socially destructive outbursts of anger and impulsive, demanding behavior toward relationship partners, sometimes including violence (4).

Problems in interpersonal relations

According to attachment theory, recurrent failure to obtain support from attachment figures and to sustain a sense of security, and the resulting reliance on secondary attachment strategies (hyperactivation and deactivation), interfere with the acquisition of social skills and create serious problems in interpersonal relations. Bartholomew and Horowitz (32), using as an assessment device the Inventory of Interpersonal Problems (39), found that attachment anxiety was associated with more interpersonal problems in general. Secure individuals did not show notable elevations in any particular sections of the problems circle, but avoidant people generally had problems with nurturance (being cold, introverted, or competitive), and anxious people had problems with emotionality (e.g., being overly expressive). These problems seem to underlie insecure individuals' self-reported loneliness and social isolation (e.g., 40) and their relatively low relationship satisfaction, more frequent relationship breakups, and more frequent conflicts and violence (4).

CONCLUSIONS

Attachment insecurities are associated with a wide variety of mental disorders, ranging from mild negative affectivity to severe, disorganizing, and paralyzing personality disorders. The evidence suggests that insecure attachment orientations (whether anxious or avoidant) are fairly general pathogenic states. Although many of the research findings supporting these ideas are correlational, several studies show a prospective connection between attachment insecurities and vulnerability to disorders. From a therapeutic standpoint, we have reviewed preliminary evidence that situationally heightening people's sense of attachment security reduces the likelihood and intensity of psychiatric symptoms (e.g., PTSD, eating disorders). This evidence underscores the soothing, healing, therapeutic effects of actual support offered by relationship partners, including therapists, and the comfort and safety offered by mental representations of supportive experiences and loving and caring attachment figures. The research evidence causes us to be optimistic about the utility of clinical interventions that increase clients' sense of attachment security.

In the long run, research on attachment security and insecurity, and on the connections between insecurity and psychopathology, should contribute to a strongly social conception of the human mind and its vulnerability to pathologies. In a pioneering chapter on the social neuroscience of attachment processes, Coan (41) proposed what he calls social baseline theory. According to this theory, the human brain evolved in a highly social environment, and many of its basic functions rely on social co-regulation of emotions and physiological states. This means that, rather than conceptualizing human beings as separate entities whose interactions with each other need to be understood, it makes more sense to consider social relatedness and its mental correlates as the normal "baseline" condition. Using this as a starting point helps us to see why experiences of separation, isolation, rejection, abuse, and neglect are so psychologically painful, and why dysfunctional relationships are often the causes or amplifiers of mental disorders.

References

1. Bowlby J. Attachment and loss, Vol. 2. Separation: anxiety and anger. New York: Basic Books, 1973.
2. Bowlby J. Attachment and loss, Vol. 1. Attachment (2nd ed.). New York: Basic Books, 1982.
3. Bowlby J. A secure base: clinical applications of attachment theory. London: Routledge, 1988.
4. Mikulincer M, Shaver PR. Attachment in adulthood: structure, dynamics, and change. New York: Guilford, 2007.
5. Hazan C, Shaver PR. Romantic love conceptualized as an attachment process. J Pers Soc Psychol 1987;52:511-24.
6. Brennan KA, Clark CL, Shaver PR. Self-report measurement of adult romantic attachment: an integrative overview. In: Simpson JA, Rholes WS (eds). Attachment theory and close relationships. New York: Guilford, 1998:46-76.
7. Cassidy J, Kobak RR. Avoidance and its relationship with other defensive processes. In: Belsky J, Nezworski T (eds). Clinical implications of attachment. Hillsdale: Erlbaum, 1988:300-23.
8. Ainsworth MDS, Blehar MC, Waters E et al. Patterns of attachment: assessed in the Strange Situation and at home. Hillsdale: Erlbaum, 1978.
9. Waters E, Merrick S, Treboux D et al. Attachment security in infancy and early adulthood: a twenty-year longitudinal study. Child Dev 2000;71:684-9.
10. Baldwin MW, Keelan JPR, Fehr B et al. Social-cognitive conceptualization of attachment working models: availability and accessibility effects. J Pers Soc Psychol 1996;71:94-109.
11. Mikulincer M, Shaver PR. Boosting attachment security to promote mental health, prosocial values, and inter-group tolerance. Psychol Inq 2007;18:139-56.
12. Catanzaro A, Wei M. Adult attachment, dependence, self-criticism, and depressive symptoms: a test of a mediational model. J Pers 2010;78:1135-62.
13. Bosmans G, Braet C, Van Vlierberghe L. Attachment and symptoms of psychopathology: early maladaptive schemas as a cognitive link? Clin Psychol Psychother 2010;17:374-85.
14. Doron G, Moulding R, Kyrios M et al. Adult attachment insecurities are related to obsessive compulsive phenomena. J Soc Clin Psychol 2009;28:1022-49.
15. Ein-Dor T, Doron G, Solomon Z et al. Together in pain: attachment-related dyadic processes and posttraumatic stress disorder. J Couns Psychol 2010;57:317-27.
16. Gormley B, McNiel DE. Adult attachment orientations, depressive symptoms, anger, and self-directed aggression by psychiatric patients. Cogn Ther Res 2010;34:272-81.

17. Illing V, Tasca GA, Balfour L et al. Attachment insecurity predicts eating disorder symptoms and treatment outcomes in a clinical sample of women. J Nerv Ment Dis 2010;198:653-9.
18. Crawford TN, Livesley WJ, Jang KL et al. Insecure attachment and personality disorder: a twin study of adults. Eur J Personality 2007;21:191-208.
19. Meyer B, Pilkonis PA. An attachment model of personality disorders. In: Lenzenweger MF, Clarkin JF (eds). Major theories of personality disorder. New York: Guilford, 2005:231-81.
20. Livesley WJ. Classifying personality disorders: ideal types, prototypes, or dimensions? J Pers Disord 1991;5:52-9.
21. Whiffen VE, Judd ME, Aube JA. Intimate relationships moderate the association between childhood sexual abuse and depression. J Interpers Violence 1999;14:940-54.
22. Davila J, Steinberg SJ, Kachadourian L et al. Romantic involvement and depressive symptoms in early and late adolescence: the role of a preoccupied relational style. Pers Relationship 2004;11:161-78.
23. Davila J, Burge D, Hammen C. Why does attachment style change? J Pers Soc Psychol 1997;73:826-38.
24. Cozzarelli C, Karafa JA, Collins NL et al. Stability and change in adult attachment styles: associations with personal vulnerabilities, life events, and global construals of self and others. J Soc Clin Psychol 2003;22:315-46.
25. Solomon Z, Dekel R, Mikulincer M. Complex trauma of war captivity: a prospective study of attachment and post-traumatic stress disorder. Psychol Med 2008;38:1427-34.
26. Mikulincer M, Hirschberger G, Nachmias O et al. The affective component of the secure base schema: affective priming with representations of attachment security. J Pers Soc Psychol 2001;81: 305-21.
27. Mikulincer M, Shaver PR. Attachment theory and intergroup bias: evidence that priming the secure base schema attenuates negative reactions to out-groups. J Pers Soc Psychol 2001;81:97-115.
28. Mikulincer M, Shaver PR, Horesh N. Attachment bases of emotion regulation and posttraumatic adjustment. In: Snyder DK, Simpson JA, Hughes JN (eds). Emotion regulation in families: pathways to dysfunction and health. Washington: American Psychological Association, 2006:77-99.
29. Admoni S. Attachment security and eating disorders. Doctoral dissertation, Bar-Ilan University, Ramat Gan, Israel, 2006.
30. Zuroff DC, Blatt SJ. The therapeutic relationship in the brief treatment of depression: contributions to clinical improvement and enhanced adaptive capacities. J Consult Clin Psychol 2006;74:199-206.
31. Gur O. Changes in adjustment and attachment-related representations among high-risk adolescents during residential treatment: the transformational impact of the functioning of caregiving figures as a secure base. Doctoral dissertation, Bar-Ilan University, Ramat Gan, Israel, 2006.
32. Bartholomew K, Horowitz LM. Attachment styles among young adults: a test of a four-category model. J Pers Soc Psychol 1991;61: 226-44.
33. Park LE, Crocker J, Mickelson KD. Attachment styles and contingencies of self-worth. Pers Soc Psychol B 2004;30:1243-54.
34. Wei M, Heppner PP, Russell DW et al. Maladaptive perfectionism and ineffective coping as mediators between attachment and future depression: a prospective analysis. J Couns Psychol 2006;53:67-79.
35. Dickinson KA, Pincus AL. Interpersonal analysis of grandiose and vulnerable narcissism. J Pers Disord 2003;17:188-207.
36. Wink P. Two faces of narcissism. J Pers Soc Psychol 1991;61:590-7.
37. Cassidy J. Emotion regulation: influences of attachment relationships. Monogr Soc Res Child Dev 1994;59:228-83.
38. Berant E, Mikulincer M, Shaver PR. Mothers' attachment style, their mental health, and their children's emotional vulnerabilities: a seven-year study of children with congenital heart disease. J Pers 2008;76:31-66.
39. Horowitz LM, Rosenberg SE, Baer BA et al. Inventory of Interpersonal Problems: psychometric properties and clinical applications. J Consult Clin Psychol 1988;56:885-92.
40. Larose S, Bernier A. Social support processes: mediators of attachment state of mind and adjustment in late adolescence. Attach Hum Dev 2001;3:96-120.
41. Coan JA. Toward a neuroscience of attachment. In: Cassidy J, Shaver PR (eds). Handbook of attachment: theory, research, and clinical applications (2nd ed.). New York: Guilford, 2008:241-68.

Prototype diagnosis of psychiatric syndromes

Drew Westen

Department of Psychology and Psychiatry, Emory University, 36 Eagle Row, Atlanta, GA 30322, USA

The method of diagnosing patients used since the early 1980s in psychiatry, which involves evaluating each of several hundred symptoms for their presence or absence and then applying idiosyncratic rules for combining them for each of several hundred disorders, has led to great advances in research over the last 30 years. However, its problems have become increasingly apparent, particularly for clinical practice. An alternative approach, designed to maximize clinical utility, is prototype matching. Instead of counting symptoms of a disorder and determining whether they cross an arbitrary cutoff, the task of the diagnostician is to gauge the extent to which a patient's clinical presentation matches a paragraph-length description of the disorder using a simple 5-point scale, from 1 ("little or no match") to 5 ("very good match"). The result is both a dimensional diagnosis that captures the extent to which the patient "has" the disorder and a categorical diagnosis, with ratings of 4 and 5 corresponding to presence of the disorder and a rating of 3 indicating "subthreshold" or "clinically significant features". The disorders and criteria woven into the prototypes can be identified empirically, so that the prototypes are both scientifically grounded and clinically useful. Prototype diagnosis has a number of advantages: it better captures the way humans naturally classify novel and complex stimuli; is clinically helpful, reliable, and easy to use in everyday practice; facilitates both dimensional and categorical diagnosis and dramatically reduces the number of categories required for classification; allows for clinically richer, empirically derived, and culturally relevant classification; reduces the gap between research criteria and clinical knowledge, by allowing clinicians in training to learn a small set of standardized prototypes and to develop richer mental representations of the disorders over time through clinical experience; and can help resolve the thorny issue of the relation between psychiatric diagnosis and functional impairment.

Key words: Prototype, diagnosis, classification, ICD-11, DSM-5, categorical diagnosis, dimensional diagnosis, comorbidity

(World Psychiatry 2012;11:16-21)

Diagnosis includes two components: the way disorders are classified, and the way patients are diagnosed using that classification system. The DSM-III represented a pivotal moment in the evolution of both. First, it shifted from a classification system that had little grounding in empirical research to one that had at least modest grounding and, more importantly, created the conditions for an explosion of research on psychiatric disorders. Second, it shifted from a way of diagnosing patients with little reliability between any two clinicians or researchers to an approach that had high reliability for research purposes (using structured interviews) but continued to have considerable problems in clinical settings (see 1).

In the intervening decades, thousands of studies have focused on classification – e.g., whether adding, subtracting, or revising this or that diagnostic criterion might make some kind of difference in reliability or validity – yet little research has focused on how to make the diagnostic process more clinically useful, valid, and reliable. The assumption of the framers of subsequent editions of the DSM has been that clinicians need to change their ways and start diagnosing patients the way researchers do.

The problems with that assumption are multifold. DSM-IV-TR (2) is an 886-page manual. The idea that clinicians in everyday practice could, would, or should ask questions about each of hundreds of largely irrelevant criteria for hundreds of largely irrelevant disorders when a relatively high-functioning patient presents with, for example, anxiety symptoms and marital problems, is questionable at best. Further, many of the questions required to make a research diagnosis are unrelated to the tasks of clinical diagnosis and treatment. Whether a patient with bulimic symptoms has binged and purged twice a week every week for an arbitrarily specified period of time is far less useful to know clinically than that the patient is binging and purging frequently (e.g., daily, weekly, or multiple times a day) and that binge episodes seem to be preceded by feelings of rejection or abandonment.

The arbitrary nature of criteria for severity, duration, and number of symptoms met is not just a problem for clinical work but for research as well. In meta-analyzing the results of empirically supported therapies for some of the most prevalent disorders (e.g., mood and anxiety disorders), colleagues and I found that the average study excluded the majority of patients even considered for clinical trials because they did not meet rigid inclusion criteria or they had "comorbidities" that are in fact the norm, not the exception, in both research and clinical work (3). Further, clinical trials require categorical diagnoses as a prerequisite for entry into the study, yet virtually none uses them as a primary outcome measure, because a patient can lose just one or two symptoms of the disorder over the course of several weeks and thus appear to have "remitted" when he or she may in fact remain highly symptomatic. Instead, researchers use dimensional measures of constructs such as depression or anxiety as outcome criteria because they recognize that patients vary on *the extent* to which they are symptomatic, not just on whether they are symptomatic.

I could offer a long list of such concerns about the count-and-cutoff approach to diagnosis used in psychiatric diagnosis since 1980, such as the dif-

ficulty both clinicians and researchers have in remembering the criteria and complex diagnostic algorithms for even the most common disorders, and the fact that the modal patient receives a low-information "not otherwise specified"("NOS") diagnosis in nearly every domain of the diagnostic manual, but will not enumerate such a list here (see 4,5). Suffice it to say that it is perhaps no surprise that a method of diagnosing patients designed for research purposes that was never tested empirically in any way against any alternative other than the failed DSM-I/DSM-II approach would itself run into problems over time, particularly as conceptions of psychopathology have changed (e.g., understanding most disorders as spectrum disorders or as present in varying degrees). The framers of ICD-10 attempted to coordinate with their DSM counterparts, but where they wisely parted company was in creating a distinct manual for clinical diagnosis that built in considerably more flexibility and a much more user-friendly format. The problem with diagnostic flexibility, of course, is that different clinicians can exercise that flexibility differently, leading to problems in reliability of diagnosis in clinical practice.

We have developed an alternative, prototype-matching approach to diagnosis, in which diagnosticians compare a patient's overall clinical presentation to a set of diagnostic prototypes – for clinical use, paragraph-length descriptions of empirically identified disorders – and rate the "goodness of fit" or extent of match of the patient's clinical presentation to the prototype. Rather than inquiring about each of several hundred symptoms, assessing whether the patient "has" each symptom, and then adding or otherwise combining symptoms (e.g., 3 from column A, 5 from column B) to determine whether the patient crosses a diagnostic threshold for "caseness", the clinician uses all available data – including clinical observation, patients' answers to questions, chart data, data from informants or past treatments, and the narratives the patient offers about his or her problems and relationships – to determine *the extent to which* the patient matches diagnostic descriptions that weave together diagnostic criteria into a memorable *gestalt* designed to facilitate pattern recognition.

In our prototype-matching procedure for clinical diagnosis (4,7), the diagnostician rates the patient on a 5-point scale for degree of match to the prototype (Figure 1). The scale ranges from 1 (*little or no match*) to 5 (*very good match – patient exemplifies this disorder; prototypical case*), with ratings of 4 and 5 corresponding to categorical diagnosis and a rating of 3 indicating *subthreshold* or *clinically significant features* of the disorder (much as physicians measure blood pressure treated as a continuous variable but by convention refer to values in certain ranges as "borderline" or "high"). Thus, a single rating yields both a dimensional and a categorical diagnosis without relying on symptom counting. The default value for each diagnosis is 1 (*little or no match*), so that clinicians only expend time rating prototypes of disorders relevant to the patient, allowing rapid diagnosis. The easy translation of dimensional into categorical diagnosis (e.g., a 3 translating to *clinically significant features*) is of particular use for communication among professionals, who are unlikely to find it useful to describe a patient as "3 on major depression, 2

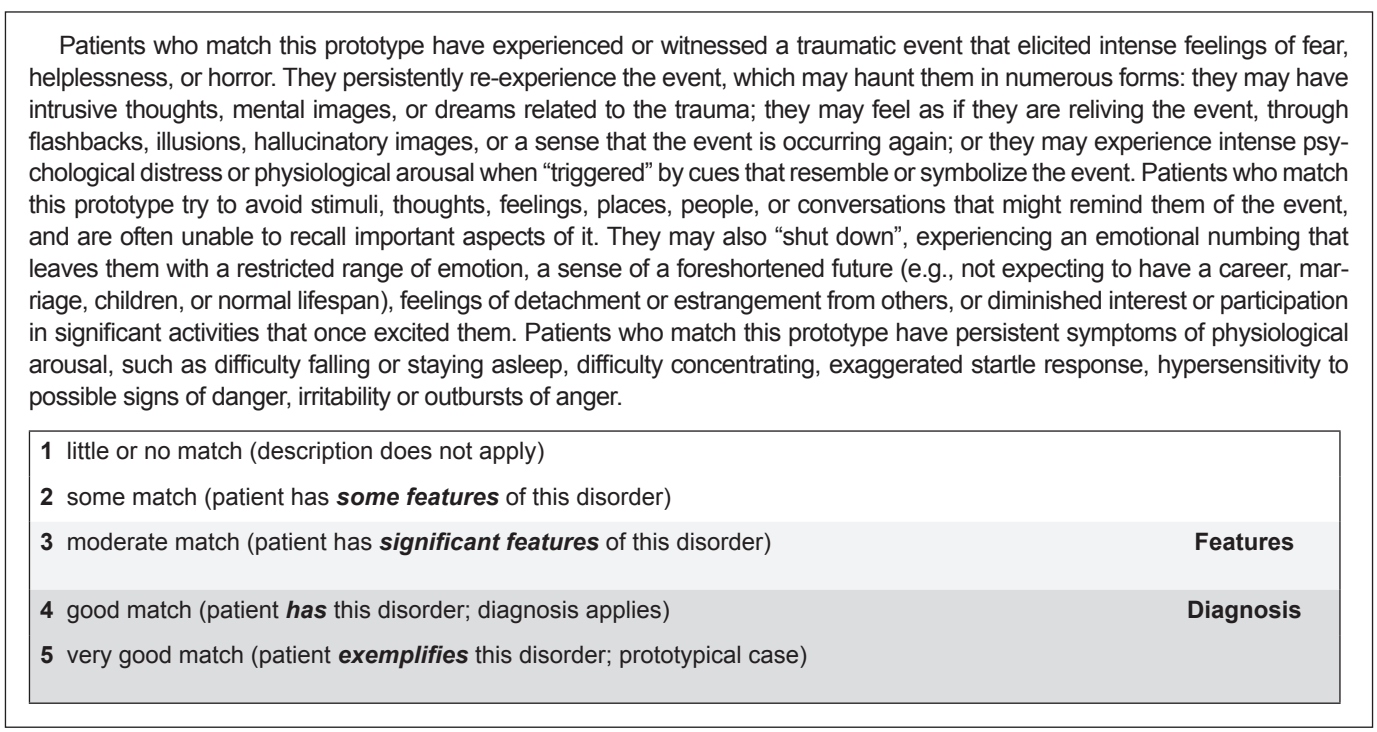

Figure 1 Post-traumatic stress disorder prototype

on panic" (one of the major limitations of potential dimensional approaches to psychiatric diagnosis). An emerging body of research suggests that clinicians can make prototype judgments of this sort for a wide range of syndromes, from mood and anxiety disorders to personality disorders, with high degrees of reliability (7-10). The complexity of this diagnostic method can be expanded as much as desirable, for example, by adding secondary ratings of severity or duration or empirically identified aspects of the disorder (e.g., severity of depressive phenomenology, vegetative symptoms, or melancholic symptoms for major depressive disorder; age of onset and severity of attentional deficits, hyperactivity, and impulsivity for attention-deficit/hyperactivity disorder).

Before briefly highlighting some of the advantages of this approach, three points are worth noting. First, both the polythetic diagnostic criteria built into DSM-IV (which require that a patient meet a certain number but not all of the criteria for a disorder) and particularly the clinical manual of the ICD-10 are essentially efforts to operationalize prototype matching. In the case of DSM-III-R and DSM-IV, that goal was explicit, based on early research on prototypes in the emerging field of cognitive science (11). The clinician version of the ICD-10 is already close to a prototype-matching procedure, in which clinicians are presented with what are usually paragraph-length descriptions of a disorder, often with an additional set of considerations, and are instructed to diagnose the patient based on their knowledge of the patient at the time (e.g., after a single session or months of treatment) with whatever degree of certainty they feel comfortable. What the current manual lacks is a way of operationalizing clinical judgment to maximize reliability, so that clinicians can feel confident that when a patient comes to them with a particular diagnosis – or when they diagnose a patient themselves – that diagnosis is as accurate and as clinically useful as practical.

Second, as Reed et al (12) have just shown in a WPA-World Health Organization (WHO) survey of nearly 5,000 psychiatrists across over 40 countries, most practicing psychiatrists, including users of both the ICD and the DSM, prefer a diagnostic method that has many of the features associated with prototype diagnosis, for reasons to be described shortly. Psychiatrists by large percentages preferred approaches that offer flexible rather than strict criteria-based diagnosis; keep the number of psychiatric diagnoses down to a manageable number (somewhere under 30-100); are clinician-friendly and clinically useful (e.g., in allowing clinicians to communicate their diagnostic judgments and to make useful treatment decisions based on them); and allow clinicians to represent dimensional aspects of the patient's presentation in ways that accurately capture clinical reality (e.g., diagnosing pathology that does not meet criteria for a categorical diagnosis but is nonetheless clinically significant).

Third, as the International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders for the WHO has recently noted (13), a diagnostic manual has many uses, including clinical, research, teaching and training, statistical, and public health. No approach is likely to be equally helpful or optimal for all of these uses, but a method of diagnosing patients should be reasonably useful for all of them, and in particular should have clinical utility, including utility in guiding treatment and public health. As will be seen shortly, prototype diagnosis has advantages in each of these domains.

ADVANTAGES OF PROTOTYPE DIAGNOSIS

Prototype diagnosis has a number of advantages. First, it better fits the ways humans naturally think and classify. People (in this case, diagnosing clinicians) tend to categorize complex, novel stimuli (in this case, patient presentations) through a probabilistic assessment of degree of match to a mental model they have formed (a prototype) or prominent exemplars of potentially relevant categories (14-16). Research in cognitive science suggests that, in everyday judgment and decision-making, people tend to *satisfice* (a cross between *satisfy* and *suffice*), that is, to make a "good-enough" assessment for their purposes, and to make more precise determinations based on explicit decision rules if the need arises (17-19). For example, rather than getting out the ICD-10 or DSM-IV to decide whether a patient with moderate panic symptoms once or twice a week meets formal criteria, most clinicians would diagnose the patient as suffering from moderate, clinically significant panic symptoms, whether or not the patient met formal diagnostic criteria. In light of the dearth of research on the treatment implications of clinical versus subthreshold symptoms and of data suggesting that subthreshold variants often produce similar levels of functional impairment (20,21), *satisficing* is not an irrational diagnostic strategy in clinical practice.

Compare this approach to the current diagnostic procedures, which were derived from the Research Diagnostic Criteria of the 1970s (22), and require clinicians to remember hundreds of lists of symptoms and, equally problematic, hundreds of distinct decision rules that differ for each disorder. Even putting aside the question of the validity of those lists and cutoffs, humans have trouble remembering long lists, let alone forming a coherent representation based on them. Expecting clinicians to remember how to combine the items from those lists to make a diagnosis is impractical, particularly when the precise number of symptoms from one subcategory or another may be relevant to making a "correct" diagnosis for research purposes but not for clinical practice. The DSM-IV diagnosis of post-traumatic stress disorder (PTSD), for example, requires at least one symptom of re-experiencing, three of avoidance/numbing, and two of hyperarousal. What matters to the clinician, in contrast, is the "gist" – that the patient experienced a traumatic event and is having some combination of these symptoms to varying degrees – in ways that should influence clinical decision making and treatment.

This leads to the second advantage of prototype diagnosis: clinical utility.

In multiple studies by multiple research teams, clinicians have rated prototype diagnosis as substantially more clinically useful than the more familiar DSM-IV system and alternative dimensional systems for a range of disorders on a range of measures, from utility in communicating with other clinicians to ease of use (5,7,8,23,24). Perhaps not surprisingly, a growing body of research finds that prototype diagnosis, unlike diagnosis using strict operational criteria, is highly reliable in clinical practice, with correlations typically ranging from .50 to .70 between two clinicians (10,25).

One of the reasons clinicians prefer prototype diagnosis is its third advantage, namely that it allows them to represent what they observe with their patients and to communicate it to other mental health providers both dimensionally and categorically, and to do so with relative ease. Whereas dimensional diagnosis is probably most precise in most cases, and categorical diagnosis is most familiar and feels most "natural", prototype diagnosis captures the advantages of both. Consider the DSM-IV category of eating disorders, which includes two diagnoses with two subtypes each – anorexia nervosa with restricting and binge-purging subtypes and bulimia nervosa with purging and non-purging subtypes – as well as an NOS diagnosis, for a total of five categories (and others reportedly on the way). This might be a relatively small number of disorders for an eating disorders specialist, but for the general practitioner – let alone the primary care practitioner – five disorders with multiple criteria and differing cutoffs for each is difficult to remember, particularly when a given eating disorder is just one of the hundreds of disorders in the manual. Even for specialists and researchers, the diagnostic challenge is substantial, as research suggests that this approach relegates roughly half of patients with clinically significant eating pathology to a nondescript NOS category; that over 60 percent of patients diagnosed with some variant of anorexia "switch" to a bulimia nervosa diagnosis at some point and vice versa; and that those with concurrent symptoms of both disorders are arbitrarily classified as a subtype of anorexia (3).

Perhaps not surprisingly, in a study of North American psychiatrists and psychologists treating at least one eating-disordered patient, clinicians overwhelmingly preferred prototype diagnosis to the more familiar DSM-IV system (see 8). In part this likely reflects ease of use, because prototype diagnosis vastly decreases the number of disorders that have to be included in the diagnostic manual. To cover the range of eating pathology, we presented clinicians with only two prototypes: anorexia nervosa (a syndrome characterized by self-starvation) and bulimia nervosa (a syndrome characterized by binging and purging), with both prototypes taken directly from DSM-IV criteria except without the arbitrary severity and duration criteria (clinicians also made secondary ratings such as severity of binging, severity of purging, and severity of weight loss or gain). Rather than counting symptoms and deciding whether the patient met arbitrary requirements and cutoffs (e.g., bingeing and purging at least twice a week for a minimum of 3 months), the clinician's task was simply to rate the extent to which the patient's condition matched each prototype. A score of 4 or 5 on the bulimia prototype meant that the patient's symptom picture strongly enough matched the diagnostic prototype to warrant a categorical diagnosis. A score of 3 on the anorexia prototype meant that the patient's symptom picture resembled the prototype but not enough to warrant a categorical diagnosis. A patient with both ratings would thus receive a categorical diagnosis of "bulimia nervosa with anorexic features".

A fourth advantage of prototype diagnosis is that it allows greater flexibility and validity not only in the diagnostic process but also in the definitions of disorders and the criteria that can be included in the prototypes. In the study described above, we simply combined the diagnostic criteria of each of the two major eating disorder syndromes thematically to create the prototypes. In other research, however, we have derived the disorders and criteria that create the prototypes empirically from large samples, using clinicians' detailed ratings of actual patients in their practice, and relied on statistical procedures such as factor analysis to identify both the disorders to be included and the criteria for those disorders. This allows not only for the development of empirically valid disorders without the need for committee wrangling over which disorders or criteria to include – including culturally relevant or culturally specific disorders that might emerge from a factor analysis in one culture but not in another – but also for clinically richer diagnostic descriptions. For example, Westen and Shedler (26) derived a set of personality prototypes in this way that in many respects resembled the DSM-IV and ICD-10 personality disorders but included a number of further subtle psychological diagnostic criteria that are important clinically and emerged empirically. These have been absent from official diagnostic criteria because they could not be readily self-reported by patients in structured interviews (e.g., for obsessive-compulsive personality disorder, "Is invested in seeing and portraying himself or herself as emotionally strong, untroubled, and emotionally in control, despite clear evidence of underlying insecurity, anxiety, or distress").

Developing prototypes this way substantially reduces artifactual "comorbidity", by identifying groupings of patients or criteria that are distinct from others. Even using prototypes derived from current overlapping diagnostic categories and criteria, prototype diagnosis inherently reduces artifactual comorbidity, because clinicians are making *configural* judgments, not judgments about isolated symptoms. Consider PTSD, which is often found to be comorbid with mood disorders. Some of that comorbidity is undoubtedly accurate. In other cases, however, that comorbidity is an artifact of current diagnostic methods. For example, dysphoria associated with anhedonia and general hopelessness is clearly part of a depressive picture and hence would contribute to a diagnosis of major depression or dysthymia using a prototype system. In contrast, dysphoria associated with persistent thoughts of a traumatic event or survivor guilt would likely be represented as part of

the patient's PTSD.

A fifth advantage of prototype diagnosis is its utility in integrating teaching, training, and subsequent clinical experience. The goal of prototype diagnosis is to help clinicians develop mental representations of different kinds of disorder and, equally important, to standardize those representations across diagnosticians. Instead of trying to memorize symptom lists, the goal is to form mental representations of coherent syndromes, in which signs and symptoms are functionally related (4,8,23). This approach parallels the way the brain functions, working with rather than against naturally occurring cognitive processes (16, 27-29).

The goal of teaching and training thus becomes to help trainees master a relatively small number of disorders grouped into a relatively small number of categories (e.g., psychotic disorders, mood disorders, substance use disorders, eating disorders, personality disorders, developmental disorders), with typically two to ten prototypes within each category (for a total of 30-100 disorders, depending on the number of non-overlapping disorders that emerges empirically). Formal training in diagnosis would entail learning the prototypes, perhaps with two or three exemplars for each disorder included in the diagnostic manual, and supervision on diagnosis would focus not only on diagnostic interviewing skills but also on learning to recognize the configurations, to rate them, and to make differential diagnoses. Rather than years of experience leading clinicians increasingly to ignore the symptom lists in the diagnostic manual as they learn the complexities of the disorders in real clinical practice, years of experiences would help clinicians "flesh out" and enrich their earlier mental prototypes, essentially "hanging" new experiences onto prototypes they first learned about while in training.

Finally, prototype diagnosis separates two closely related but non-identical questions addressed by the WHO's Advisory Board (13), namely the extent to which a patient has a given disorder and the extent of functional impairment. Prototype ratings are ratings of degree of match to a disorder. Although all disorders create some degree of dysfunction, they vary tremendously in how much and what kind of dysfunction they produce, which also vary by individual, depending on his or her social, psychological, and other resources. Thus, some patients can perform remarkably well occupationally despite being dysthymic, whereas for others dysthymia is debilitating. Capturing the degree of disability is not the same as capturing the degree to which a patient matches a diagnostic prototype, although we have found that prototype ratings can be extremely useful in measuring the extent to which patients match a prototype of psychological health (see 7).

CONCLUSIONS

By describing the advantages of prototype diagnosis, I do not mean to suggest that it is without limitations (although its detractors will no doubt lay those out with great clarity in the pages that follow). Perhaps the most important disadvantage of prototype diagnosis is that it could foster confirmatory biases and other heuristics that can lead clinicians, like all humans, to see what they expect to see, or to stick with hypotheses about a patient despite disconfirming information. Whether it does so more than the approaches in DSM-IV or ICD-10 is an empirical question, although it is certainly possible that encouraging clinicians to match patients to prototypes could make them more likely to gloss over disconfirming data or to adhere doggedly to early diagnostic hypotheses.

The best antidotes to these kinds of diagnostic biases are threefold, all of which we should be teaching young clinicians and practicing throughout our years of clinical experience, regardless of systems of diagnosis. The first is to understand the cognitive and emotional biases to which our minds are naturally prone, most of which are unconscious, and to exercise rigorous and continuous self-examination to try to minimize those biases in working with our patients. The second is to teach our trainees routinely to ask themselves (and to ask ourselves) to generate alternative hypotheses to those we are currently entertaining to understand what we are seeing clinically. The third is always to supplement descriptive diagnosis with *functional* diagnosis, where the question is not "What does this patient *have*", but "Under what conditions does this patient think, feel, and behave in this particular way (e.g., why is this patient depressed to this degree *now,* and under what circumstances does he function differently)?". By asking this last question, we not only challenge any diagnostic complacency we may have, but we also focus on what matters most in clinical diagnosis, namely on understanding how this particular patient's mind and brain work and under what circumstances they may work differently.

Acknowledgement

This research was supported by National Institute of Mental Health grants R01-MH62377 and R01-MH78100.

References

1. Maj M. Psychiatric diagnosis: pros and cons of prototypes vs. operational criteria. World Psychiatry 2011;10:81-2.
2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed., text revision. Washington: American Psychiatric Association, 2000.
3. Westen D, Novotny CM, Thompson-Brenner H. The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. Psychol Bull 2004;130:631-63.
4. Westen D, Bradley R. Prototype diagnosis of personality. In: Strack S (ed). Handbook of personology and psychopathology. New York: Wiley, 2005:238-56.
5. Westen D, Heim A, Morrison K et al. Simplifying diagnosis using a prototype-matching approach: implications for the next edition of the DSM. In: Beutler LE, Malik ML (eds). Rethinking the DSM: a psychological perspective. Washington: American Psychological Association, 2002:221-50.
6. Westen D, Shedler J. A prototype matching approach to diagnosing personality disorders: toward DSM-V. J Pers Disord 2000;14:109-26.

7. Westen D, Shedler J, Bradley R. A prototype approach to personality disorder diagnosis. Am J Psychiatry 2006;163:846-56.
8. Ortigo KM, Bradley B, Westen D. An empirically based prototype diagnostic system for DSM-V and ICD-11. In: Millon T, Krueger R, Simonsen E (eds). Contemporary directions in psychopathology: scientific foundations of the DSM-V and ICD-11. New York: Guilford, 2010:374-90.
9. DeFife JA, Peart J, Bradley B et al. The validity of prototype diagnosis in everyday practice. Submitted for publication.
10. DeFife JA, Westen D, Bradley B et al. Prototype personality diagnosis in clinical practice: a viable alternative for DSM-5 and ICD-11. Prof Psychol Res Pr 2010;41:482-7.
11. Widiger T, Frances A. The DSM-III personality disorders: perspectives from psychology. Arch Gen Psychiatry 1985;42:615-23.
12. Reed GM, Correia JM, Esparza P et al. The WPA-WHO global survey of psychiatrists' attitudes towards mental disorders classification. World Psychiatry 2011;10:118-31.
13. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. World Psychiatry 2011;10:86-92.
14. Folstein JR, Van Petten C. Multidimensional rule, unidimensional rule, and similarity strategies in categorization: event-related brain potential correlates. J Exp Psychol Learn Mem Cogn 2004;30:1026-44.
15. Medin DL. Concepts and conceptual structure. Am Psychol 1989;44:1469-81.
16. Rosch E, Mervis CB. Family resemblances: studies in the internal structure of categories. Cogn Psychol 1975;7:573-605.
17. Gigerenzer G, Goldstein DG. Reasoning the fast and frugal way: models of bounded rationality. Psychol Rev1996;103:650-69.
18. Gigerenzer G. Gut feelings: the intelligence of the unconscious. New York: Penguin Group, 2008.
19. Simon HA. Information-processing theory of human problem solving. In: Estes WK (ed). Handbook of learning and cognitive processes. Hillsdale: Erlbaum, 1978.
20. Fava GA, Mangelli L. Assessment of subclinical symptoms and psychological well-being in depression. Eur Arch Psychiatry Clin Neurosci 2001;251:47-52.
21. Marshall RD, Olfson M, Hellman F et al. Comorbidity, impairment, and suicidality in subthreshold PTSD. Am J Psychiatry 2001;158:1467-73.
22. Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria: rationale and reliability. Arch Gen Psychiatry 1978;35:773-82.
23. Rottman B, Ahn W, Sanislow C et al. Can clinicians recognize DSM-IV personality disorders from five-factor model descriptions of patient cases? Am J Psychiatry 2009;166:427-33.
24. Spitzer R, First M, Shedler J et al. Clinical utility of five dimensional systems for personality diagnosis: a "consumer preference" study. J Nerv Ment Dis 2008;196:356-74.
25. Westen D, Muderrisoglu S. Reliability and validity of personality disorder assessment using a systematic clinical interview: evaluating an alternative to structured interviews. J Pers Disord 2003;17:350-68.
26. Westen D, Shedler J. Revising and assessing Axis II, Part 2: toward an empirically based and clinically useful classification of personality disorders. Am J Psychiatry 1999;156:273-85.
27. Kim NS, Ahn W. Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. J Exper Psychol 2002;131:451-76.
28. Cantor N, Genero N. Psychiatric diagnosis and natural categorization: a close analogy. In: Millon T, Klerman GL (eds). Contemporary directions in psychopathology: toward the DSM-IV. New York: Guilford, 1986:233-56.
29. Ahn W. Effect of causal structure on category construction. Memory & Cognition 1999;27:1008-23.

Prototypes, syndromes and dimensions of psychopathology: an open agenda for research

Assen Jablensky
Centre for Clinical Research in Neuropsychiatry, School of Psychiatry and Clinical Neurosciences, University of Western Australia, Perth, Australia

The current process of revision of the DSM and ICD has generated requests to alter the criteria defining many individual disorders, to eliminate some and to add new "disorders". However, all definitional changes have serious disadvantages: they are confusing to clinicians; they create a situation in which the relevance of all previous clinical and epidemiological research into disorders hitherto defined is uncertain; and they involve tedious and often costly changes in the content and wording of diagnostic interviews, as well as in the algorithms to generate diagnoses from clinical ratings.

Most of psychiatry's disease concepts are merely working hypotheses and their diagnostic criteria are provisional. Psychiatric disorders are complex psychobiological entities and both extremes – a totally unstructured approach to diagnosis and rigid operationalization – should be avoided. Defining a middle range of operational specificity, which would be optimal for stimulating critical thinking in both clinical practice and research, but also rigorous enough to enable meaningful communication and comparisons between results of different studies in different contexts, is a better solution. This is where, with certain caveats, the prototype-matching approach could fit the bill (1). Drew Westen's proposals are to be welcomed for "biting the bullet" by bringing to attention an important alternative to current classificatory models, but the article leaves several key questions open for discussion.

What are the caveats? First, the concept of a prototype, intuitively attractive to clinicians, remains ambiguous and poorly operationalized. Prototypes represent the central ("perceptually salient") tendencies of categories (2). But could a prototype be synonymous, or partly overlapping, with a well-crafted narrative description of a "core" syndrome? In the complex psychiatric disorders, where aetiology is multifactorial, both research and everyday clinical practice could be considerably facilitated by a sharper delineation of the syndromal status of many current diagnostic categories. This provides a strong rationale for reinstating the syndrome (or its prototype template) as the basic unit of future versions of psychiatric classifications.

Secondly, should the concise, "one paragraph" formulation of a prototype contain, wherever relevant, pointers to likely aetiology, pathophysiology or associated features, e.g. cognition, which are not part of the presenting clinical picture? According to Hempel (3), membership in a prototype is defined by correlated features, not the necessary presence of all defining features.

Thirdly, how can a broad agreement be achieved on a universally "valid" prototype description of any particular disorder? Schizophrenia provides a relevant example. The description of the syndrome has undergone several metamorphoses since Bleuler's (4) original distinction between basic symptoms ("loosening of associations", uncoupling of affect from cognition, volitional ambivalence, autistic closure to reality) and accessory symptoms (delusions, hallucinations, catatonic phenomena). While the ICD-10 (5) clinical diagnostic guidelines have retained a remote echo of Bleuler's conceptualization ("a fundamental disturbance of personality…, involving its most basic functions which give the normal person a feeling of individuality, uniqueness and self-direction"), they attribute particular prominence to Schneider's (6) first-rank symptoms. In contrast, the diagnostic criteria of DSM-IV (7) and of the draft DSM-5 require various combinations of "positive" and "negative" symptoms but do not attempt to provide any prototype or *gestalt* of the characteristic imprint of the disorder.

Lastly, Westen's claim that a single rating on a 5-point scale for assessing the extent to which an individual matches the prototype could generate, in addition to a categorical statement, a meaningful dimensional score raises doubts. Whether psychiatric disorders can be better described dimensionally or categorically remains an open question for research. However, a major problem for dimensional models of psychopathology is the absence of an established, empirically grounded metrics. Most existing scales of symptom severity are of a psychometrically low level of measurement: they are either nominal or ordinal, where one is simply able to state that $a>b>c\ldots>n$ on some property, arbitrarily assigning numbers which indicate rank order and nothing more (8). It seems unlikely that equal-interval scaling or ratio scales will ever be developed for complex configurations of psychopathology.

Medical classifications are created with the primary purpose of meeting pragmatic needs related to diagnosis and treating people experiencing illness. Their secondary purposes is to assist in the generation of new knowledge relevant to those needs, though progress in medical research usually precedes, rather than follows, improvements in classification. The prototype matching approach has the potential of serving well both these purposes, provided that it is underpinned by sound decision rules and supported by evidence from multisite field trials.

References

1. Maj M. Psychiatric diagnosis: pros and cons of prototypes vs. operational criteria. World Psychiatry 2011;10:81-2.
2. Rosch E. Natural categories. Cogn Psychol 1973;4:328-50.
3. Hempel CG. Aspects of scientific explanation. New York: Free Press, 1965.
4. Bleuler E. Dementia praecox oder die Gruppe der Schizophrenien. Leipzig: Deuticke, 1911.
5. World Health Organization: The ICD-10 classification of mental and behavioural

disorders. Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1992.
6. Schneider K. Klinische Psychopathologie, 8th ed. Stuttgart: Thieme, 1950.
7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington: American Psychiatric Association, 1994.
8. Kerlinger FN. Foundations of behavioural research, 2nd ed. London: Holt, Rinehart & Winston, 1973.

Toward a clinically useful and empirically based dimensional model of psychopathology

Robert F. Krueger[1], Kristian E. Markon[2]
[1]University of Minnesota, Minneapolis, MN, USA;
[2]University of Iowa, Iowa City, IA, USA

In his target essay, Westen argues convincingly for moving from a criterion-based system for psychiatric diagnosis, to a system based on prototypes. There is much to admire in Westen's essay, and many points with which we can readily agree. The polythetic-categorical approach to diagnosis that frames modern DSMs was critical in the *zeitgeist* in which it was developed, but, as Weston eloquently describes, its limitations and conceptual conundrums are now well documented. We are therefore in complete agreement with Westen's overarching point: new approaches to conceptualizing psychiatric disorders are needed. In the remainder of this brief commentary, we note some areas where our approach might be somewhat different from Westen's, with the idea in mind of furthering discussion of these issues, and working collaboratively toward a novel and empirically-based psychiatric nosology.

Westen relies heavily on clinician report as the primary source of data to delineate empirically-based psychopathology constructs. Westen acknowledges this reliance on clinician reporters as a potential limitation of his prototype approach, but we believe there is an alternate way of handling the reporter issue that might better distinguish the prototype concept from the separate issue of the source of data on psychopathological signs and symptoms.

Fundamental psychopathological constructs must be delineated initially from clinician's experiences – there is no other place to begin to assemble a compendium of basic level diagnostic elements (e.g., a tendency to be manipulative, or to have culturally unusual beliefs and experiences). However, we believe a next critical step is to instantiate these constructs in instruments suitable for diverse reporters (e.g., patients, collateral informants, and treating clinicians). With these kinds of data in hand, one can then initiate an inductive-hypothetico-deductive process (1), in which data are collected and quantitative models are applied to these data through multiple rounds of data gathering and refinement, to arrive at an empirically based quantitative nosology built from data, from the ground level up (2).

One concern with focusing primarily or exclusively on clinicians is well-documented biases in clinical judgment. Clinicians, for example, have been taught systems that we know to be inaccurate (e.g., DSM-IV), and a sensible goal is to bootstrap a system that describes patient's actual experiences, as opposed to "pre-structuring" those experiences by DSM rubrics (whether consciously or unconsciously). Relatedly, prototypes can incorporate stereotypes, which may contribute to biases (e.g., racial or gender) in assessment (3). Obviously, data from other informants are also subject to limitations (e.g., less than perfect insight in self-report), so the idea is to not make a specific reporter exclusive. Rather, data from multiple reporters can and should always be taken into account in developing an empirically based nosology, to overcome the limitations of any given source (4).

Along these same lines, it is also important to distinguish a reporter's perspective from the "objective veridicality" of the report, which we can never really know in a definitive sense. For example, a person perceived by others as self-aggrandizing might not endorse "I am grandiose" but might describe his/her experience as "having to deal frequently with other people who are incapable of understanding my importance and talents". That is, regardless of any "objectively veridical" situation, the structure of psychopathology can be uncovered in data from different reporters, allowing comparison of how these structures are similar or different, as well as ways of combining information from different reporters in case formulation. Interestingly, some broad aspects of psychopathology structure seem consistent in data from various reporters: broad internalizing (anxiety and mood disturbance) and externalizing (substance use and antisocial behavior) spectrums are seen in clinician report (via the Shedler-Westen Assessment Procedure, SWAP (5)), collateral report (e.g., parents (6)) and self-report via structured interview (7).

Another deep issue Westen's commentary raises pertains to the ontological status of psychiatric diagnoses. Accompanying the prototype conception is the idea that psychiatric diagnoses exist as distinguishable, person-centered entities in nature. This conception works well only if there are discrete psychiatric diagnoses in nature, with discrete and separable accompanying etiologies and pathophysiologies, or at the very least, zones of rarity separating disorders in a descriptive space.

To date, these kinds of specific etiologies, pathophysiologies, and zones of rarity have proven highly elusive for psychopathology. Hence, although the prototype concept is extremely helpful clinically, prototypes need to be understood as salient combinations of constituent dimensions, as opposed to constructs that demarcate discrete groups of

persons (as such discreteness appears not to exist).

Combinations of dimensions that are independent do, nevertheless, combine in specific persons in a way that is clinically salient. For example, psychopathic personality entails at least three elements that do not tend to co-occur in people in general, but when they do co-occur *in a specific person*, the result is an unusual collision of dispositions that can be quite striking and pernicious (i.e., boldness or a lack or neuroticism combines with meanness or a tendency to be disagreeable, and with disinhibition or a lack of conscientiousness, to form a nasty and impulsive person who has no anxiety about their misdeeds (8)). In this way, prototypes can be understood as combinations of constituent dimensions that are clearly meaningful, but arbitrary in the sense that a nearly infinite set of combinations of constituent psychopathology dimensions exists in nature (9,10).

References

1. Cattell RB. Handbook of multivariate experimental psychology. Chicago: Rand McNally, 1966.
2. Krueger RF, Eaton NR. Personality traits and the classification of mental disorders: toward a more complete integration in DSM-5 and an empirical model of psychopathology. Personality Disorders: Theory, Research, and Treatment 2010;1:97-118.
3. Whaley AL, Geller PA. Toward a cognitive process model of ethnic/racial biases in clinical judgment. Rev Gen Psychol 2007; 11:75-96.
4. Oltmanns TF, Turkheimer E. Person perception and personality pathology. Current Directions in Psychological Science 2009;18:32-6.
5. Westen D, Shedler J. Personality diagnosis with the Shedler-Westen Assessment Procedure (SWAP): integrating clinical and statistical measurement and prediction. J Abnorm Psychol 2007;116:810-22.
6. Achenbach TM. Empirically based assessment and taxonomy: applications to clinical research. Psychol Assessment 1995; 7:261-74.
7. Krueger RF, Markon KE. Reinterpreting comorbidity: a model-based approach to understanding and classifying psychopathology. Ann Rev Clin Psychol 2006;2:111-33.
8. Patrick C, Fowles D, Krueger R. Triarchic conceptualization of psychopathy: developmental origins of disinhibition, boldness, and meanness. Dev Psychopathol 2009;21:913-38.
9. Turkheimer E, Ford DC, Oltmanns TF. Regional analysis of self-reported personality disorder criteria. J Pers 2008;76:1587-622.
10. Eaton NR, Krueger RF, South SC et al. Contrasting prototypes and dimensions in the classification of personality pathology: evidence that dimensions, but not prototypes, are robust. Psychol Med (in press).

A practical prototypic system for psychiatric diagnosis: the ICD-11 Clinical Descriptions and Diagnostic Guidelines

Michael B. First

Department of Psychiatry, Columbia University and Department of Clinical Phenomenology, New York State Psychiatric Institute, 1051 Riverside Drive, New York, Unit 60, NY 10032, USA

In the past 50 years, the DSM and ICD revision processes have focused almost exclusively on refining the diagnostic category definitions. The process by which patients are assigned a categorical diagnosis in routine clinical practice has received much less attention. Starting with the publication of DSM-I in 1952 and continuing with DSM-II and the mental disorders sections of ICD-8 and ICD-9, standardized glossary definitions were provided as a basis for assigning a diagnosis, essentially establishing a prototype-like approach as the standard method of psychiatric diagnosis. The problematic diagnostic reliability of glossary definitions (reviewed in 1) prompted researchers in the 1970s to develop explicit inclusion and exclusion criteria such as the Research Diagnostic Criteria (RDC) (2). The demonstration that the inter-rater reliability of these operationalized criteria was superior to the DSM-II glossary definitions (3,4) led to the provision of diagnostic criteria for every disorder in DSM-III, with the stated hope that they would "improve the reliability and validity of routine psychiatric diagnosis" (4).

However, despite the widespread utilization of the DSM diagnostic criteria by the research community, anecdotal and indirect evidence suggests that clinicians routinely fail to use them in everyday clinical practice. Although there have never been any studies examining how the DSM is actually used in clinical practice settings, several studies have demonstrated significant discrepancies between DSM diagnoses made by clinicians in practice and diagnoses made using structured diagnostic interviews that systematically evaluate the DSM criteria (5,6). Although the authors of these studies concluded that the problem was due to the misapplication of the diagnostic criteria and recommended clinicians should receive additional diagnostic training to improve their diagnostic accuracy, the more likely scenario is that experienced clinicians do not methodically evaluate every relevant DSM criterion but instead match the patient's symptomatology to a mental prototype. Studies demonstrating clinician preference for prototype matching over criterion counting (e.g., 7) further suggest that clinicians find prototype matching to be more concordant with the way they make psychiatric diagnoses. In recognition of the fact that diagnostic assessment using operationalized criteria is best suited to research settings, ICD-10 provides two versions of its classification of mental disorders: the Clinical Descriptions and Diagnostic Guidelines (CDDG) for clinical use, and a parallel system of diagnostic criteria for research use.

Although Westen's proposal to shift from a criterion-based to a prototype matching system is a step in the right direction with regard to improving user acceptability of the diagnostic system, from a practical perspective there are signifi-

cant drawbacks with its proposed implementation. The most problematic aspect concerns the requirement that clinicians rate the degree of prototype matching on a 5-point scale ranging from 1 (little or no match) to 5 (very good match). Recognizing the necessity of extracting a categorical diagnosis for both clinical and coding purposes, Westen proposes that the top two levels (i.e., 4 and 5) indicate the presence of the diagnosis, thus placing the entire disorder/non-disorder differentiation on the ability of the clinician to distinguish between a rating of 3 (the highest non-disorder rating, which he defines as "patient has *significant features* of this disorder") and a rating of 4 (defined as "patient *has* this disorder," italics in original) without any guidance provided as to how much of the patient's clinical features would have to match the prototype in order to justify the clinician's judgment that the patient has the disorder.

Given that prototype matching is considered to be the method used in DSM-II (8), a possible objection to the adoption of a prototype approach might be concerns about its reliability, given the conventional wisdom that DSM-II diagnoses are significantly less reliable than those made using DSM-III criteria, a contention based on comparisons of pre-DSM-III reliability data from the 1960s with reliability data using DSM-III criteria. In actual fact there is virtually no evidence demonstrating that DSM-III, when used in clinical settings, represents a significant improvement over DSM-II in terms of diagnostic reliability. Because differences in experimental design such as method of reliability determination, training of clinician participants, and base rates of diagnoses can significantly impact measured reliability, comparing DSM-II and DSM-III reliability obtained using different methodologies is like comparing apples and oranges. The only study that compared DSM-II and DSM-III diagnoses head-to-head (9) failed to demonstrate any differences in reliability. Moreover, most of the deficiencies in the DSM-II definitions cited by Spitzer et al (4) as likely sources of unreliability to be addressed in the construction of criteria sets, such as not distinguishing those features that are invariably present from those that are commonly but not invariably present, can easily be incorporated into clinical descriptions of disorders without the need to use operationalized criteria. Indeed, the reliability field trials of ICD-10 CDDG (10), which, like DSM-II, used clinical descriptions rather than diagnostic criteria, demonstrated that satisfactory reliability can be achieved without using diagnostic criteria.

The ICD-11 CDDG approach is similar to Westen's prototype matching in that it eschews defining disorders in terms of pseudoprecise criteria with arbitrary thresholds and instead involves the clinician deciding whether the diagnostic features outlined in the clinical description match those of the patient. However, unlike Westen's prototype matching system, which offers only a textual paragraph describing the typical patient and nothing else, the ICD-11 approach conveys a considerable amount of clinically-relevant information about the disorder, including both essential and typical features, differential diagnosis, the boundary with normality, typical course, and developmental and culture-specific features.

Uniform guidelines have been developed for the ICD-11 working groups with the goal of improving consistency across categories and increasing clinical utility. As noted above, diagnostic reliability comparable with what clinicians can achieve using diagnostic criteria can be attained utilizing diagnostic definitions expressed in terms of clinical descriptions, something which is planned to be tested again in the forthcoming ICD-11 field trials.

Acknowledgements

The author is a consultant to the World Health Organization (WHO) in the development of the chapter on mental and behavioural disorders of the ICD-11. However, the views expressed in this commentary are entirely those of the author and do not reflect the official views, positions, or policies of the WHO or the International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders.

References

1. Spitzer R, Fleiss J. A re-analysis of the reliability of psychiatric diagnosis. Br J Psychiatry 1974;125:341-7.
2. Spitzer R, Endicott J, Robins E. Research Diagnostic Criteria: rationale and reliability. Arch Gen Psychiatry 1978;35:773-82.
3. Spitzer R, Endicott J, Robins E et al. Preliminary report on the reliabiltiy of Research Diagnostic Criteria applied to psychiatric case records. In: Sudilovsky A, Gershon S, Beer B (eds). Predictability in psychopharmacology: preclinical and clinical correlations. New York: Raven, 1975:1-47.
4. Spitzer R, Endicott J, Robins E. Clinical criteria for psychiatric diagnosis and DSM-III. Am J Psychiatry 1975;132:1187-92.
5. Ramirez Basco M, Bostic JQ, Davies D et al. Methods to improve diagnostic accuracy in a community mental health setting. Am J Psychiatry 2000;157:1599-605.
6. Shear M, Greeno C, Kang J et al. Diagnosis of nonpsychotic patients in community clinics. Am J Psychiatry 2000;157:581-7.
7. Spitzer RL, First MB, Shedler J et al. Clinical utility of five dimensional systems for personality diagnosis: a "consumer preference" study. J Nerv Ment Dis 2008;196: 356-74.
8. Zimmerman M. A critique of the proposed prototype rating system for personality disorders in DSM-5. J Pers Disord 2011;25: 206-21.
9. Mattison R, Cantwell DP, Russell AT et al. A comparison of DSM-II and DSM-III in the diagnosis of childhood psychiatric disorders: II. Interrater agreement. Arch Gen Psychiatry 1979;36:1217-22.
10. Sartorius N, Kaelber C, Cooper J et al. Progress toward achieving a common language in psychiatry. Results from the field trial of the clinical guidelines accompanying the WHO classification of mental and behavioral disorders in ICD-10. Arch Gen Psychiatry 1993;50:115-24.

Prototypal diagnosis: will this relic from the past become the wave of the future?

Allen Frances
Department of Psychiatry and Behavioral Science, Duke University, Durham, NC, USA

History's ironic sense of humor is perhaps best displayed in its playful recursiveness. Don't throw away that dated old tie or dress; with a little patience you can expect it eventually to make a retro comeback as a new/old fashion fad. And so it is with prototypal diagnosis in psychiatry: made seemingly irrelevant not long ago by the new technology of criteria based diagnosis, but now regaining in popularity and legitimacy.

It is our misfortune that psychiatric diagnosis is stuck in a purely descriptive mode – we still require subjective word descriptions as our only tool for making diagnoses because we lack any objective biological tests. This shortcoming should be corrected for the dementias in the next five to ten years, but laboratory tests for the other psychiatric disorders seem even further off now than they did when we were preparing DSM-IV twenty years ago. The subsequent explosive neuroscience revolution has taught us a great deal about how the normal brain works, but perhaps its most profound lesson is that the ineluctable complexity of brain functioning will offer us no easy answers to the elusive riddles of psychopathology.

Our descriptive categories, however defined, are wildly heterogeneous in their underlying causes. There are likely to be hundreds of different ways to arrive at what we now lump together as schizophrenia. All of our labels, whether criteria based or prototypal, are only very rough first approximations. But they are all we have, so we have to make the best of them.

The defects of prototypal diagnosis are well known. It is completely unreliable in everyday clinical life and works well only in the most hothouse of conditions: with experienced, well trained clinicians having sufficient time to make multiple dimensional ratings on the easiest diagnostic distinctions. In real life, clinicians will not bother reading or rating the prototypes just as they often don't read the criteria sets.

DSM-I and DSM-II were prototypal systems with such low reliability that clinical psychiatry was becoming the laughing stock of medicine and research in psychiatry was virtually impossible. The radical solution was the provision of criteria-based definitions, introduced by Robert Spitzer into DSM-III. He had the vision to take what had been suggested as no more than a research tool to increase diagnostic reliability and to make it the new standard of everyday clinical practice.

The criteria set method had several great advantages. Under the right conditions, reliability was good to excellent. Clinicians and researchers were now reading off the same page, allowing an easier translation from research findings to clinical practice. Researchers around the world now spoke a common diagnostic language. It is no exaggeration to say the criteria-based diagnosis saved not just psychiatric diagnosis, but also psychiatry, from being regarded as a quaint clinical art not amenable to the newly emerging methods of evidence based medical science.

Of course, the reliability of DSM-III criteria-based diagnosis was oversold. Excellent reliability can be achieved in research studies with highly trained interviewers using a structured instrument on highly selected patients. Reliability evaporates to greater or lesser degrees depending on how much the testing conditions depart from the ideal. Reliability is no better now than it ever was for those clinicians who don't know or care about DSM diagnosis, who have only minutes to do the evaluation, and whose practice has unselected patients who are difficult to diagnose. And a criteria-based system with many narrowly defined categories is bound to have artificial "comorbidity", with each label providing no more than one building block of a complete diagnosis.

So where do we stand now? Drew Westen nicely lays out the advantages of prototypal diagnosis and his proposed method for accomplishing it. As he suggests, the best application of his system would clearly be in the diagnosis of personality disorders. We actually thought of including something similar as an option in DSM-IV, but decided not to because it seemed unlikely that busy clinicians would have the time or interest to carefully study the prototypes of each personality and provide the necessary multiple ratings.

That is certainly still the case today, but Westen's approach is far superior to the impossibly complicated and cumbersome dimensional systems that have been concocted by the DSM-5 personality disorders work group. And an even more simplified prototypal approach (without ratings) has been chosen for ICD-11, a return to the DSM-II fashion.

Complicated criteria-based and simple prototypal diagnosis can be complementary, not necessarily competitive. It is too bad that DSM-5 and ICD-11 were not developed in closer coordination. I would favor a nested, compatible system, with ICD-11 providing the simple, quick prototypal descriptions that are a shorthand for the complicated, time consuming DSM-5 criteria sets. The prototypes would be used in clinical settings where time and ease of use are essential; the criteria sets would be for more specialized clinicians and for all research. This would give us the best of both worlds while we await the likely slow progress toward diagnosis by biological tests.

Are you as smart as a 4th grader? Why the prototype-similarity approach to diagnosis is a step backward for a scientific psychiatry

Jerome C. Wakefield
Department of Psychiatry, School of Medicine, New York University, 550 First Avenue, New York, NY 10016, USA

Drew Westen proposes a prototype-similarity matching approach to diagnosis, in which clinicians classify a patient as having a disorder when the patient's symptoms are judged sufficiently similar to a described prototypical case of the disorder that contains all the standard symptoms. He claims that the prototype approach to diagnosis "better fits the ways humans naturally think and classify". I will argue that Westen's prototype approach is a step backward in light of research showing that people tend to classify by causal history, not similarity, consistent with psychiatry's goal of etiology-based diagnosis. I will also argue that the prototype approach undermines diagnostic validity by eliminating restraints on false positive diagnoses.

Inspired by Wittgenstein's "family resemblance" account of concepts, prototype-similarity theory challenges traditional notions that concepts consist of necessary-and-sufficient criteria for category membership. The theory holds instead that an entity is a member of a category if it is overall more similar to that category's "prototypical" member (e.g., robins for "bird") than to prototypes of competing categories (1).

Prototype-similarity theory was initially hailed as a breakthrough in the psychology of classification. Studies showed that people judge some members that are more similar to typical members as better examples of a category than others ("prototypicality effects"). For example, woodpeckers are better "bird" examples than ostriches due to more closely resembling robins. But ostriches are still birds because they resemble robins more than non-bird prototypes. Prototypicality effects supposedly indicated that category membership itself is dimensional and based on degree of similarity to a prototype.

Westen thus equates degree of diagnostic-category membership with degree of similarity to a clinical prototype. However, the prototype-similarity theory's fundamental assumption that prototypicality effects (i.e., degree of being a "good example" of a category and degree of similarity to a prototype) equate with degree of category membership has turned out to be incorrect. For example, the category "even number" is defined by necessary-and-sufficient criteria ("divisible by 2") that equally apply to all members, but nonetheless this category displays prototypicality effects; subjects judge some even numbers as more similar to the prototype "2" than others and as better examples of even numbers than others. Yet the same subjects, when asked, do not consider there to be degrees of category membership in "even number" (2). Whatever prototypicality effects represent psychologically, they are not equivalent to the way people think about category membership.

Research suggests that when people categorize, they are not merely similarity judgers but causal theoreticians who infer shared underlying natures (or "essences") as determinative of classification. Classic experiments show how similarity judgments and category membership judgments diverge based on causal knowledge. For example, subjects judge a 3-inch-diameter round object to be *more similar* to a quarter than to a pizza, but *more likely to be* a pizza than a quarter, understandably given the causal laws governing quarters and pizzas (3). When various-aged subjects are told that a raccoon was altered to appear similar to a skunk by painting it black with a white stripe down its back and adding odor sacs, kindergartners, using an approach analogous to prototype-similarity theory, categorize the creature as a skunk, but 4th graders and older subjects classify it as a raccoon based on the "etiology" of its features (4). Surely clinicians' diagnoses should be as conceptually sophisticated as 4th graders' "diagnoses" of the raccoon!

A causal-modeling/essentialist understanding of concepts explains why category membership so often deviates from prototype similarity (5). We classify steam, ice, and liquid water together as the substance "water" despite dissimilarity and despite steam and ice not at all resembling the presumed prototype of liquid water, based on shared molecular structure. Whales resemble prototypical fish more than prototypical mammals, yet are mammals based on evolutionary history. There are small round red carrots that look more like cherries, and elongated orange cherries that look more like carrots, all classified based on causal inferences, not superficial similarity. Throughout science and physical medicine, underlying essential properties determine category membership. This is the way we think and classify deliberatively, and this is the approach optimal for scientific progress.

In principle, the DSM works this way, with symptom criteria used to infer shared internal dysfunctions responsible for symptoms. Little is yet known about etiology, so at present the DSM relies heavily on syndromal similarity. However, the DSM's descriptive criteria are designed to be transitional until research reveals etiologically distinct disorders among current syndromes.

Such causal inferences occur implicitly now, for example, when clinicians classify prototypical depressive symptoms that occur during bereavement as normal and atypical depression as major depression, or adolescent antisocial behavior in the course of avoiding sexual abuse as normal (6). If sheer prototype similarity determined diagnosis, illiteracy and delinquency would likely be forms of dyslexia and conduct disorder,

respectively, rather than normal variations. Identifying distinct underlying dysfunctions is what scientific study of mental disorder is mostly about, so understanding diagnosis as inference about shared dysfunction is both how people naturally think conceptually and the optimal basis for scientific development of psychiatry (7). Diagnostic criteria allow for flexible development towards an etiologically based system, whereas prototype-similarity diagnosis freezes diagnosis at the symptom-similarity stage.

Couldn't the prototype view simply be extended to encompass etiological considerations? The problem is that etiology often forms a necessary-and-sufficient criterion for standard diagnostic category membership judgments, eliminating the relevance of the degree-of-similarity judgments. The prototype approach would thus be transformed into a criterial approach (8).

There is a tension between the fuzziness of disorder/normality boundaries and the need to validly distinguish disorders from normal distress. The DSM addresses this tension via artificially rigidly bounded high-threshold diagnostic criteria that limit false positives, combined with "not otherwise specified" (NOS) categories for subsyndromal conditions that would not be judged disorders based on symptoms alone but in context are judged to be disorders.

Westen criticizes the precision of DSM symptom and duration thresholds, but fails to explain how the prototype approach prevents false positives. False positives usually arise because normal distress and mental disorder resemble each other symptomatically. Prototype diagnosis, combining symptom similarity and very vague boundaries, is particularly liable to false positives. The only restraint lies in how the line is drawn between the global judgments "*has* this disorder" and "has *significant features* of this disorder" (but not the disorder). How this distinction is validly made based on global similarity judgments remains obscure.

Westen's response to the false-positives problem is to dismiss it: "In light of the dearth of research on the treatment implications of clinical versus subthreshold symptoms and of data suggesting that subthreshold variants often produce similar levels of functional impairment, *satisficing* is not an irrational diagnostic strategy in clinical practice." By "satisficing", Westen means "to make a 'good-enough' assessment". According to Westen, because normal/disorder boundaries are uncertain, prototype diagnosis can ignore the syndromal/subsyndromal boundary and thus largely ignore the disorder/normal-distress boundary, and still yield "good-enough" psychodiagnosis. More precise diagnosis using all available decision rules only occurs if "the need arises". But when does the need arise if not at initial diagnosis and treatment selection? And, who would knowingly entrust their psychological or medical fate to "satisficing" clinicians who prefer more fallible but easier-to-use heuristics to full deliberative analysis, especially when, as Westen notes, the prototype heuristic inherently biases towards certain errors ("encouraging clinicians to match patients to prototypes could make them more likely to gloss over disconfirming data…")?

Couldn't the prototype view simply be modified to encompass sharp cutoffs to prevent false positives? The problem is that adding sharp diagnostic boundaries strikes at the heart of the prototype view's equation of category membership with dimensional similarity judgments.

Westen argues that the prototype approach has clinical utility. Whether this is true depends on what you mean by "clinical utility". All else being equal, clinician-friendliness of criteria makes sense. But clinician-friendliness is secondary to validity and patient utility. False-positive diagnoses are neither good for the patient nor ethically defensible. So all else is not equal, and clinical utility becomes irrelevant.

References

1. Rosch E, Mervis CB. Family resemblances: studies in the internal structure of categories. Cogn Psychol 1975;7:573-605.
2. Armstrong SL, Gleitman LR, Gleitman H. What some concepts might not be. Cognition 1983;13:263-308.
3. Rips LJ. Similarity, typicality, and categorization. In: Vosniadou S, Ortony A (eds). Similarity and analogical reasoning. New York: Cambridge University Press, 1989:21-59.
4. Keil FC. Concepts, kinds, and cognitive development. Cambridge: Bradford Books/MIT Press, 1989.
5. Medin DL, Ortony A. Psychological essentialism. In: Vosniadou S, Ortony A (eds). Similarity and analogical reasoning. New York: Cambridge University Press, 1989:179-95.
6. Wakefield JC, Pottick KJ, Kirk SA. Should the DSM-IV diagnostic criteria for conduct disorder consider social context? Am J Psychiatry 2002;159:380-6.
7. Wakefield JC. Disorder as harmful dysfunction: a conceptual critique of DSM-III-R's definition of mental disorder. Psychol Rev 1992;99:232-47.
8. Wakefield JC. Evolutionary versus prototype analyses of the concept of disorder. J Abnorm Psychol 1999;108:374-99.

Nosological changes in psychiatry: hubris and humility

Oye Gureje

Department of Psychiatry, University of Ibadan, Nigeria

While many would agree that the current classifications of psychiatric disorders are not ideal and would be happy to see them improved, it is unclear how best to achieve this. This lack of clarity reflects a deeper problem with our understanding of the pathogenesis of most mental disorders. In the absence of that understanding, the empirical foundation upon which to erect a revised new set of descriptions of those disorders is either not available or at best noticeably weak (1). A common consensus therefore is that, at least for the time being, a conservative approach to the revision process makes the most sense (2).

This humility is not without basis. Many are still quite aware of the damage that the free-wheeling approach to diagnosis of mental disorders did to psychiatry as a scientific branch of medicine. Indeed, the derision with which the field is often held by both fellow medical colleagues and, sometimes, by the lay public is partly to do with what is perceived as the amorphous nature of our diagnostic entities. It is important that we should not lose sight of that recent past as we make efforts to build on what we currently have.

It is precisely in response to that past that the designers of DSM-III pitched their tent strongly in the achievable domain of improving reliability (3). The fashioning of explicit inclusion and exclusion criteria, and the specification of the number of symptoms as well as the duration of their occurrence, in making psychiatric diagnosis was entirely to achieve the goal of reliability. Most would agree that, largely, that goal has been achieved with our current classificatory systems. But the achievement of this common language has not been without its occasional discontents. Thus, even today, some researchers still indulge in the idiosyncratic definition of what constitutes some mental disorders. While it is probably less fashionable now to talk, for example, of "masked depression", it is still the case that, just by slicing what is decidedly a dimensional construct at an arbitrary point, some would claim to be describing a unique disease entity. So, if you cut the depression dimension at any arbitrary point and claim you are describing, for example, a unique African depressive syndrome, some, including prestigious academic journals, could be impressed. It is therefore clear that, even with the current rigid, albeit also arbitrarily defined symptom clusters upon which our consensus has been built, there are still instances when clear assault is launched at that common language, simply by going fishing, nosologically speaking.

This, for me, is the important backdrop against which to consider the proposal for the use of prototypes for making diagnosis of mental and behavioural disorders. Given our current state of knowledge and the limitations of our classificatory systems today, the question is worth asking: what do we seek or should we expect to achieve in a revision exercise? What gains must we protect even as we struggle to improve on what is obviously an imperfect system? Everyone would agree that we should seek to achieve improved validity for the disorders in our classificatory systems. However, most would also accept that the goal of validity remains a distant aspiration (4). Indeed, by the nature of the phenomena we deal with in psychiatry, many would probably acknowledge that there will always be a limit to which the road to the validity paradise could take us. However, while we strive to reach the validity nirvana, we have to keep an eye on the possibility of losing reliability and utility. It is precisely because our achievement in having reliably defined entities has also improved the clinical utility of our current classificatory systems, that there is now an admission that, at least for the ICD system, the pursuit of improved utility is the beginning of the revision wisdom (5). So, would the use of prototypes help the attainment of improved reliability and utility?

While Drew Westen presents a compelling reason for us not to dismiss the value of prototypes, embedded in his description is the source of my concern. May it lead us to the danger of losing some of our current gains whilst we are in pursuit of yet another forlorn diagnostic destination? For example, inferences from the same clinical presentation, in regard to clinician interpretations of that observation, can be protean and subject to a lot of subjectivity. Also, we may need several prototypes to capture the range of presentations that a disorder may have. For example, without explicit criteria, patients with schizophrenia with diverse admixtures of negative, positive and disorganization symptoms may not be unambiguously captured with a few prototype descriptions. Answering the question "Does this patient have personality disorder, schizophrenia, or delusional disorder?" may be problematic without some specific symptoms and the duration of their occurrence.

In many respects, the editorial by Mario Maj, presaging this forum (6), has identified some of the possible problems associated with the use of prototypes. I only wish to add that it may also embolden those who are keen on nosological exoticism to create "new" entities, hiding under the claim of taking cultural relativity into account, and develop versions of the template prototypes that suit their fancy. That way, we would have lost our imperfect but useful common language and created a Tower of nosological Babel.

References

1. Hyman SE. Can neuroscience be integrated into the DSM-V? Nat Rev Neurosci 2007;8: 725-32.
2. Hyman SE. The diagnosis of mental disorders: the problem of reification. Annu Rev Clin Psychol 2010;6:155-79.
3. American Psychiatric Association, Diagnostic and statistical manual of mental disorders, 3rd ed. Washington: American Psychiatric Association, 1980.
4. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry 2003;160:4-12.
5. International Advisory Group for the Revision of ICD-10 Mental and Behavioral Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioral disorders. World Psychiatry 2011;10:86-92.
6. Maj M. Psychiatric diagnosis: pros and cons of prototypes vs. operational criteria. World Psychiatry 2011;10:81-2.

Prototype diagnosis of psychiatric syndromes and the ICD-11

José L. Ayuso-Mateos
Universidad Autonoma de Madrid, Instituto de Investigacion Sanitaria Princesa, CIBERSAM, Madrid, Spain

Drew Westen presents an alternative approach to the diagnosis of psychiatric syndromes, based on prototype matching. First, the article thoroughly reviews the problems posed by the current polythetic or count/cutoff methods of psychiatric diagnostic procedure, which were derived from the Research Diagnostic Criteria of the 1970s. It then details the advantages of using the new prototype-matching approach to diagnosis, in which diagnosticians compare a patient's overall clinical presentation to paragraph-length descriptions of empirically identified disorders, and rate the "goodness of fit", or how well the patient's clinical presentation matches these prototypes. This method has been developed and tested by Westen and colleagues, mostly in mental health settings.

My comments will focus on the possibility of incorporating Westen's prototype-matching approach into the revision of ICD-10 Chapter V (F): Mental and Behavioural Disorders, which is now underway. As Westen correctly points out in his article, the Clinical Descriptions and Diagnostic Guidelines for ICD-10 Mental and Behavioural Disorders (the clinician version) have important similarities to a prototype-matching procedure, because they present what are usually paragraph-length descriptions of the clinical features of each disorder.

In my opinion, Westen's proposed model may have some advantages: namely, that it allows for greater flexibility and is more directly comparable to how clinicians think about patients. Moreover, it could be useful in research, teaching and training. However, I do not believe that the model can be incorporated as such into the revised chapter on mental and behavioural disorders in the forthcoming ICD-11. In particular, its rating procedure is problematic for a classification system like the ICD, which aims to encompass every kind of medical and mental condition, and targets a wide variety of users around the world.

The ongoing revision of the ICD-10 Chapter V (F): Mental and Behavioural Disorders is occurring within the context of the revision of the entire ICD-10. The overall revision process has established rules for presenting information and for coding the presence or absence of the different disorders, as well as uniform requirements for the description of every disorder within the entire system. An attempt to use a different system of description and scoring as the basis for the chapter on mental and behavioural disorders would be against the general rules of the classification system as a whole, and undermine the parity of psychopathology with the rest of the medical system for clinical, administrative, and financial purposes in health care.

In addition, Westen's proposed system, as presented, loses any apparent advantage in clinical utility if we consider that mental health professionals are not the only ones involved in the diagnosis and classification of mental disorders. In fact, only a very small percentage of individuals with mental disorders will ever see a psychiatrist or any other type of mental health professional. Therefore, psychiatrists, clinical psychologists and psychiatric nurses cannot be envisioned as the primary users and the sole professional constituency for the ICD classification system – many other professional groups will also be using the classification. This includes primary care physicians as well as lay health care workers who deliver the majority of primary and mental health care in some developing countries. Asking these professionals to differentiate between a score of "4) good match: patient *has* this disorder, diagnosis applies"; and "5) very good match, patient *exemplifies* this disorder, prototypical case" would likely create confusion and uncertainty and reduce the clinical utility of the system.

Furthermore, asking them to consider scores such as "2) some match, patient has *some features* of this disorder"; and "3) moderate match, patient has *significant features* of this disorder" could unnecessarily prolong the diagnostic procedure and lead to inflation in the diagnosis of subthreshold conditions. Reed et al (1) recently highlighted that the World Health Organization is concerned about the proliferation of diagnoses of mental disorders. As the International Advisory Group to the Revision of ICD-10 Mental and Behavioural Disorders has pointed out (2), all decisions concerning changes in the current classification should consider whether the proposed changes provide an improved basis for efficiently identifying people with the greatest mental health needs when they come into contact with health care systems. Although subthreshold conditions are increasingly being recognized as an important topic for research, this does not automatically mean that they should be defined as a disease, or included in the diagnostic formulation to the extent proposed by Westen.

The new ICD version needs to be simpler, and also needs to pay special attention to the differentiation of what is a disorder from what is not. The main challenge in developing the new classification of mental and behavioural disorders is to identify the relevant threshold that clearly signals the presence of a condition deserving clinical attention, and to establish differentiations among conditions that have clinical utility (3,4). Attempting to integrate Westen's approach would likely sidetrack this objective.

Some elements of the rationale described for the content of the prototypes in Westen's article could be relevant to revising the descriptions presented in the ICD system. One of the key elements in prototype-matching is that it centers on the fact that what matters to the clinician is the "gist": a set of salient symptoms which, when present, are "good enough" for the clinician to establish a diagnosis,

without the need to check the presence of other symptoms that are less relevant to making a diagnosis. The clinical descriptions incorporated into the ICD could take into account this need to emphasize the conditions' salient features, and give less weight to symptoms that are less relevant to determining a given diagnosis.

References

1. Reed GM, Dua T, Saxena S. World Health Organization responds to Fiona Godlee and Ray Moynihan. BMJ 2011;20:342.
2. International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. World Psychiatry 2011;10:86-92.
3. Maj M. When does depression become a mental disorder? Br J Psychiatry 2011;199: 85-6.
4. Reed GM, Ayuso-Mateos JL. Hacia una clasificación internacional de los trastornos mentales de la OMS de mayor utilidad clínica. Rev Psiquiatr Salud Ment 2011;4:113-6.

Prototype matching together with operational criteria would make a better approach to psychiatric classification

Pichet Udomratn

Department of Psychiatry, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand

In 1980, the American Psychiatric Association introduced panic disorder in the DSM-III and provided diagnostic criteria for that new disorder (1). At that time, several Asian psychiatrists manifested a resistance to use the new diagnosis, due to the lack of a corresponding terminology in their native languages (2). This would have not happened if in DSM-III a prototype had been provided as an example or illustrative case.

Actually, one prototype may not be sufficient to make a diagnostician fully understand a particular disorder, especially if it has a variety of clinical presentations, as is the case for panic disorder (the so-called "many faces of panic disorder" (3)). Moreover, prototypes from one country might not fit well in another country where culture and beliefs are different. For example, southern Thai women with gastrointestinal symptoms, who would be labeled as having panic disorder by Western psychiatrists, may present themselves as having *"rook lom"* ("wind illness") (4). This comes from the traditional belief that humans are composed of four elements: earth, water, wind *(lom)*, and fire.

In this particular context, the panic disorder prototype for Thai clinicians could be adapted as follows: "Patients who match this prototype have many clinical presentations. In southern Thailand, they may have initial symptoms of feeling the *"lom"* moving upward in the abdomen and the patient may try to push down on the abdomen in an attempt to force the *"lom"* out. Then the *"lom"* may still continue to move upward and compress the heart, causing their hearts pounding fast and hard, which may be followed by rapid and deep breathing, and later developing dizziness or headache. Most of them have a fear of *"lom"* moving into the head, which would cause unconsciousness, or death…". So, when applying the prototype approach for psychiatric diagnosis, the content needs to be adapted to suit specific cultures and contexts.

Regarding the clinical utility of prototype matching, I agree with Westen's view that it is clinically helpful and easy to use in everyday practice. Clinicians in both low and middle income countries, who see 50-100 or more cases per day in their outpatient settings, may find the prototype approach to be more practical than extensive interviews to explore fulfillment of the diagnostic criteria of conventional classifications. However, research psychiatrists may not prefer this alternative approach to recruit their patients, due to the necessity to collect homogeneous samples for research purposes and reduce research risk (5).

To sum up, using both the operational and prototype matching approaches for diagnosing psychiatric patients should be recommended. Prototype matching together with operational approaches such as the DSM-like format (diagnostic criteria with arbitrary cutoff points) would aid trainees in clearly understanding mental disorders and would help clinicians to develop mental representations of different kinds of disorders.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental Disorders, 3rd ed. Washington: American Psychiatric Association,1980.
2. Udomratn P. Panic disorder in Thailand: a report on the secondary data analysis. J Med Assoc Thai 2000;83:1158-66.
3. Ley R. The many faces of Pan: psychological and physiological differences among three types of panic attacks. Behav Res Ther 1992;30:347-57.
4. Udomratn P, Hinton DE. Gendered panic in southern Thailand. In: Hinton DE, Good BJ (eds). Culture and panic disorder. Stanford: Stanford University Press, 2009:183-204.
5. Yanos PT, Stanley BS, Greene CS. Research risk for persons with psychiatric disorders: a decisional framework to meet the ethical challenge. Psychiatr Serv 2009;60:374-83.

Generalizability of the Individual Placement and Support (IPS) model of supported employment outside the US

Gary R. Bond, Robert E. Drake, Deborah R. Becker

Dartmouth Psychiatric Research Center, Rivermill Commercial Center, 85 Mechanic Street, Lebanon, NH 03766, USA

While reviews of controlled studies of the Individual Placement and Support (IPS) model of supported employment for clients with severe mental illness have documented its effectiveness in the US, its generalizability to other countries has not been systematically evaluated. This is the first review to compare US to non-US studies. We identified 15 randomized controlled trials of IPS programs, 9 in the US and 6 outside the US. We examined competitive employment outcomes, including employment rate, days to first job, weeks worked during follow-up, and hours worked. We also considered noncompetitive employment, program retention, and nonvocational outcomes. IPS programs had significantly better outcomes across a range of competitive employment indicators and higher retention in services than control groups. The overall competitive employment rate for IPS clients in US studies was significantly higher than in non-US studies (62% vs. 47%). The consistently positive competitive employment outcomes strongly favoring IPS over a range of comparison programs in a group of international studies suggest that IPS is an evidence-based practice that may transport well into new settings as long as programs achieve high fidelity to the IPS model, but further research is needed on international adaptations.

Key words: Severe mental illness, competitive employment outcomes, vocational rehabilitation, systematic review

(World Psychiatry 2012:11:32-39)

Individual Placement and Support (IPS) is a systematic approach to helping people with severe mental illness achieve competitive employment (1). It is based on eight principles: eligibility based on client choice, focus on competitive employment, integration of mental health and employment services, attention to client preferences, work incentives planning, rapid job search, systematic job development, and individualized job supports (2). Systematic reviews have concluded that IPS is an evidence-based practice (3-12).

With the development of a strong evidence base for IPS in the US, mental health leaders in other countries have interest in the transportability of IPS to their countries. Generalizability of other evidence-based practices developed in the US has been variable and in some cases adoption has been curtailed after failures to replicate US findings (13).

The current review has two goals. First, given the growing international attention to IPS, we examined its effectiveness in studies conducted outside the US compared to US studies. Second, we expanded the scope of prior IPS reviews by adding recent randomized controlled trials (RCTs) and enlarging the range of outcome measures in order to examine the hypothesis that IPS yields better competitive employment outcomes across a range of measures than alternative vocational programs.

METHODS

The study inclusion criteria were as follows: RCT; comparison of IPS to a control condition not providing IPS; target population of clients with severe mental illness; longitudinal competitive employment outcomes; intervention monitored with the IPS Fidelity Scale (14).

We used a combination of search procedures including formal electronic searches, bibliographic searches of prior reviews and conference proceedings, and inquiries to vocational researchers, especially those in other countries. We cross-checked our findings with an exhaustive search conducted by Cochrane reviewers (9).

Consistent with the goal of IPS, the review's main focus was competitive employment, defined as permanent jobs paying commensurate wages in integrated community settings (i.e., employing nondisabled workers) and available to anyone (not just individuals with disabilities). Consistent with the goals of social inclusion, this definition excludes noncompetitive jobs, such as transitional and sheltered employment (15).

IPS researchers have not adopted a standardized measurement framework, although some indicators are common across studies. Competitive employment indicators include measures of job acquisition (e.g., percentage of clients obtaining competitive employment and time from study entry to first job start), duration (e.g., cumulative number of weeks worked in all jobs), intensity (e.g., percentage working at least 20 hours a week), and productivity (e.g., total hours worked/wages) (16).

Some vocational models place clients in noncompetitive jobs (e.g., sheltered employment, agency-run business, etc.). When reported, we summarize these noncompetitive employment outcomes. We also examined dropout rates from IPS and control programs. Finally, many studies also examined a range of outcomes outside the employment domain; we summarize these findings.

Data were recorded directly from published reports or calculated from information presented in the published studies. For the measure of job duration, we converted total weeks

worked to an annualized rate, reporting the findings for both the full intent-to-treat sample and the worker subsample (those who obtained at least one competitive job during follow-up).

Given the small number of studies, our comparisons between US and non-US studies relied on visual inspection. The one exception was competitive employment rate, where we combined samples within US and within non-US studies and used a 2x2 χ^2 to compare overall rates.

For each study we calculated the effect size for the difference in competitive employment rate between IPS and controls using the arc sine approximation (17). An unweighted overall effect size was calculated as the simple mean of the individual effect sizes. For hours of employment, we first converted data for each study to an annualized rate to accommodate the different follow-up periods. We next calculated the *d* effect size for the difference in means between IPS and controls (17). Finally, we calculated the unweighted overall effect size. For all other outcome measures, means are reported without standard deviations, because this statistic was usually unavailable from the original studies. Overall means were calculated weighting individual means by sample sizes.

RESULTS

We excluded 9 RCTs that evaluated a form of supported employment that either preceded the development of IPS (18-22) or reflected a vocational approach that was not IPS (23-26).

We identified 15 studies, 9 from the US and 6 outside the US, as shown in Table 1. Altogether, these studies enrolled 1063 IPS participants (mean = 70.9 per study) and 1117 control participants (mean = 74.5 per study). The mean length of follow-up was 18.4 months. Except for one three-group design (31), all studies used a two-group design (IPS vs. control). Ten studies were conducted at a single site, while five studies (7,27,29,32,33) had multiple sites. Two studies used nonintegrated supported employment control groups (31,33). Otherwise, all the control groups consisted of either treatment as

Table 1 Randomized controlled trials of individual placement and support

Study	Location	Study population	Control condition	Follow-up (months)	N (IPS)	N (Ctl)
Drake et al (33)	Manchester & Concord, NH	CMHC clients	Skills training, nonintegrated	18	73	67
Drake et al (37)	Washington, DC	Case management program clients	Traditional vocational services including sheltered workshop	18	74	76
Lehman et al (34)	Baltimore, MD	CMHC clients, including those without vocational goals	PSR	24	113	106
Mueser et al (31)	Hartford, CT	CMHC clients	Brokered SE; PSR	24	68	136
Gold et al (39)	Rural SC	CMHC clients	Sheltered workshop	24	66	77
Latimer et al (38)	Montréal, Canada	Clients receiving MH services	Traditional vocational services	12	75	74
Bond et al (32)	Chicago, IL	New admissions to PSR agency	Diversified placement approach	24	92	95
Burns et al (7)	6 European countries	Clients receiving MH services	Traditional vocational services	18	156	156
Wong et al (40)	Hong Kong	Hospital and community referrals	VR referral	12	46	46
Killackey et al (65)	Melbourne, Australia	Young adults with early psychosis	Traditional vocational services	6	20	21
Twamley et al (42)	San Diego, CA	Middle aged and older adults (≥45)	VR referral	12	28	22
Davis et al (28)	Tuscaloosa, AL	Unemployed veterans with PTSD	Standard VA vocational rehabilitation	12	42	43
Nuechterlein (30)	Los Angeles, CA	Young adults with early psychosis	VR referral	18	46	23
Heslin et al (29)	London, UK	Clients receiving outpatient care	Usual care	24	93	95
Michon et al (27)	4 cities in the Netherlands	Clients receiving MH services	Traditional vocational services	30	71	80

IPS – Individual Placement and Support; Ctl – control group; CMHC – community mental health center; MH – mental health; SE – supported employment; PSR – psychosocial rehabilitation; VR – State-federal vocational rehabilitation system; VA – Veteran Affairs; PTSD – post-traumatic stress disorder

usual or well-established alternative vocational models. All studies reported using standard methods to assess and monitor IPS fidelity.

In most studies, participants were recruited from clients receiving services from community mental health centers. In all the studies, participants were unemployed at the time of study admission. In all but one study (34), the study inclusion criteria included an expressed desire to work. Another eligibility criterion common across most studies was the absence of significant medical conditions, such as end-stage cancer, that would preclude working during the follow-up period or participating in assessment interviews. The Los Angeles study (35) required a two-to-three-month stabilization period before study entry because participants were often in a psychotic state at referral. In most studies, participants were required to attend two or more research information meetings explaining the study purpose (36).

Competitive employment outcomes

The competitive employment rate was significantly higher for the IPS condition than for controls in every one of the studies, as shown in Figure 1. In total, 592 (55.7%) of IPS participants obtained employment, compared with 253 (22.6%) control participants. Averaging the rates across studies, the competitive employment rate was 58.9% (median = 63.6%) for IPS compared to 23.2% (median = 26.0%) for controls. The mean difference in percentage employed between IPS and controls was 35.7%, ranging from 11.0% to 55.5%. The individual study effect sizes ranged from .30 to 1.18. The overall unweighted effect size was .77.

We next compared the competitive employment rates between the 9 US and 6 non-US studies. Combining samples across studies, 374 (62.1%) of 602 IPS clients from the 9 US studies obtained competitive employment, compared with 218 (47.3%) of 461 IPS clients from the 6 non-US studies, χ^2 (1) = 23.29, p<0.001. The comparison for the combined control samples was not significant: 150 (23.5%) of 645 control clients from the US studies obtained competitive employment, compared with 103 (21.8%) of 472 control clients from the non-US studies, χ^2 (1) = 0.32. For the US studies, the unweighted mean competitive employment rate was 65% for IPS and 25% for controls, with an overall unweighted effect size of .84. For the non-US studies, the mean competitive employment rate was 50% for IPS and 20% for controls, with an overall unweighted effect size of .67.

Among US studies, the competitive employment rate of 27% for the Maryland IPS sample (34) was an outlier – less than half the rate for IPS in the other 8 US studies and equal to the mean control group rate. With this study removed, the unweighted mean competitive employment rate for US stud-

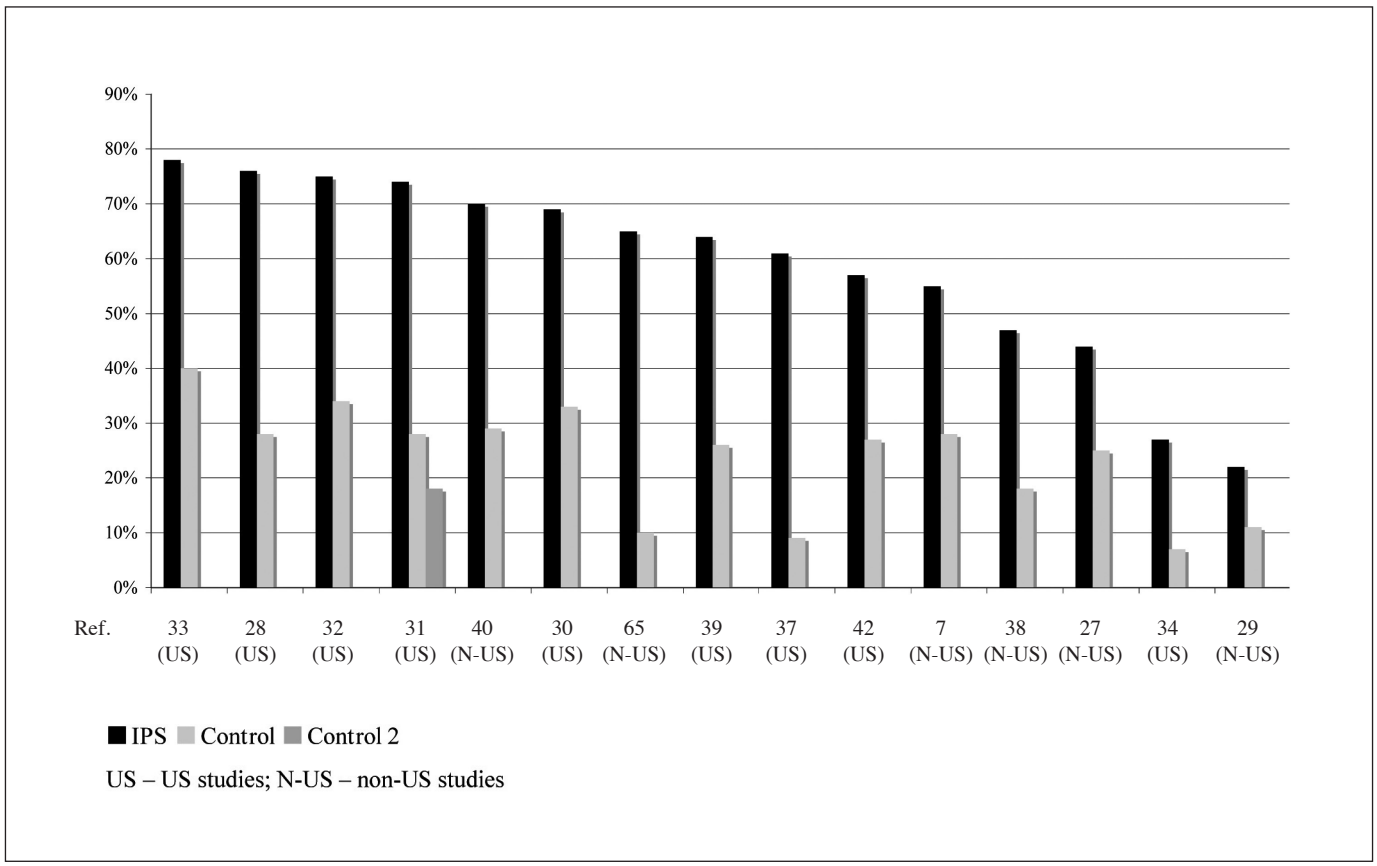

Figure 1 Competitive employment rates in 15 randomized controlled trials of Individual Placement and Support (IPS)

Table 2 Mean number of days to first competitive job in nine IPS studies

Study	IPS	Control
Wong et al (40)	72 (N = 32)	118 (N = 13)
Latimer et al (38)	84 (N = 51)	89 (N = 39)
Twamley et al (42)	93 (N = 16)	171 (N = 6)
Drake et al (37)	126 (N = 45)	293 (N = 7)
Gold et al (39)	133 (N = 42)	322 (N = 20)
Bond et al (32)	156 (N = 69)	193 (N = 32)
Lehman et al (34)	164 (N = 47)	287 (N = 12)
Mueser et al (31)	197 (N = 51)	277 (N = 31)
Heslin et al (29)	708 (N = 21)	698 (N = 11)
Total	**167.7 (N = 374)**	**236.3 (N = 171)**
Total without Heslin et al study	**135.6 (N = 353)**	**204.6 (N = 160)**

IPS – Individual Placement and Support

ies increased to 69% for IPS and 28% for controls. Similarly, among non-US studies, the competitive employment rate of 22% for the UK IPS sample (29) was an outlier – less than half the rate for IPS in the other 5 non-US studies and equal to the mean control group rate. With this study removed, the unweighted mean competitive employment rate for non-US studies was 56% for IPS and 22% for controls.

Four IPS studies (31-33,37) reported the proportion of participants who worked 20 hours or more per week. Aggregating across these studies, 134 (43.6%) of 307 IPS participants and 53 (14.2%) of 374 controls held such jobs, yielding an effect size of .67. One study reporting rates of full-time competitive employment found no difference (8.7% of IPS participants vs. 11.6% of controls) (32).

Number of days to first competitive job was reported in 9 IPS studies (6 US, 3 non-US), as shown in Table 2. The UK study was an extreme outlier (29), with mean of 680 days to first job. Excluding this outlier, the average time to first competitive job was 50% faster for IPS participants compared to controls (136 days versus 205 days). The other two non-US studies (from Hong Kong and Canada) had the shortest mean time to first job of the 9 studies.

The findings for mean hours worked per year in competitive employment for 5 US and 2 non-US studies are shown in Table 3. The variability across studies was substantial, from a mean of 656 hours for the Alabama (28) study to 126 hours for the Québec (38) study. Nonetheless, the overall unweighted effect size was large (d = .58), and the ratio of IPS to controls in hours worked was threefold overall. No obvious pattern was apparent for the comparison of US to non-US studies.

The findings for annualized weeks worked in competitive employment are reported for six US and two non-US studies in Table 3. Overall, the mean weeks worked per year in competitive employment for IPS was more than twice the mean weeks for controls. When the comparisons were limited to participants who obtained competitive employment during follow-up, the weeks worked were virtually the same for IPS and controls.

Other outcomes

Total paid employment outcomes, including noncompetitive jobs, were reported in seven studies (31-33,37-40). In six of these studies, the rate of noncompetitive employment for IPS was modest (11% or less of IPS participants), though in the Québec (38) study, 20% of IPS participants obtained a noncompetitive job. In three US studies (31,33,39) and one non-US study (40), inclusion of all paid employment did not materially affect the employment findings. Considering all paid employment outcomes, one US study (32) and one non-US study (38) showed no differences between IPS and controls in employment rates and on several other employment measures, while another US study reported no differences in overall earnings between IPS and controls (37).

Early program dropouts refer to clients who either discon-

Table 3 Mean hours worked per year in competitive jobs in seven IPS studies

	Follow-up (months)	IPS		Control		Ratio IPS/Ctl	Effect size
		Mean	SD	Mean	SD		
Davis et al (28)	12	656	661	236	494	2.78	0.72
Drake et al (33)	18	405	843	137	400	2.96	0.60
Bond et al (32)	24	298	836	143	723	2.09	0.40
Burns et al (7)	18	286	707	79	312	3.61	0.57
Drake et al (37)	18	215	569	19	125	11.5	0.72
Mueser et al (31)	24	187	516	36	231	5.22	0.86
Latimer et al (38)	12	126	267	73	252	1.73	0.20
Mean all studies		**284.3**		**86.1**		**3.30**	**0.58**

IPS – Individual Placement and Support

Table 4 Annualized weeks worked in competitive jobs in eight IPS studies

	All study participants		Working participants	
	IPS	**Control**	**IPS**	**Control**
Davis et al (28)	21.6 (N=42)	6.8 (N=43)	28.4 (N=32)	24.4 (N=12)
Latimer et al (38)	17.0 (N=75)	14.1 (N=74)	25.0 (N=51)	26.8 (N=39)
Bond et al (32)	16.2 (N=92)	8.2 (N=95)	21.6 (N=69)	24.3 (N=32)
Mueser et al (31)	14.9 (N=68)	2.3 (N=136)	19.8 (N=51)	9.8 (N=31)
Wong et al (40)	13.0 (N=46)	7.0 (N=46)	18.6 (N=32)	24.9 (N=13)
Drake et al (37)	10.1 (N=74)	0.8 (N=76)	16.6 (N=45)	8.7 (N=7)
Gold et al (39)	10.0 (N=66)	2.9 (N=77)	15.8 (N=42)	11.3 (N=20)
Lehman et al (34)	6.0 (N=113)	1.6 (N=106)	14.4 (N=47)	14.1 (N=12)
Total	**12.8 (N=576)**	**4.9 (N=653)**	**20.0 (N=369)**	**19.3 (N=166)**

IPS – Individual Placement and Support

tinue vocational services within an early time period or never make an initial contact. Studies reporting dropout (or attrition) rates did not have a standardized time period or common method of assessing discontinuation. For example, the Washington study (37) reported attrition rates after 2 months, the Illinois study (32) identified early program dropouts as clients who discontinue services within the first 6 months, and the Québec study (38) defined attrition as failure to have at least one contact with vocational staff in each of the first and the second three-month follow-up periods. Averaging across six studies (7,31-33,37,38), 9% of IPS participants were early program dropouts, compared to 42% of controls.

Nine of the studies included in the review also examined nonvocational outcomes, which most often included psychiatric symptoms, quality of life, and psychiatric hospitalizations (31-34,37-39,41,42). Some also included measures of self-esteem, social functioning, and social network. With rare exception, IPS participants did not differ from controls on any of these measures.

DISCUSSION

Rigorous evaluations of IPS suggest that 60% or more of IPS clients obtain competitive jobs, compared to about 25% of those who receive other types of vocational assistance. One way of interpreting this finding is that approximately 25% of clients who express an interest in competitive employment will succeed in obtaining a job in diverse and ineffective vocational programs or even without any vocational services, but IPS helps an additional 35% of the target group who otherwise would remain unemployed. The finding of a large and statistically significant beneficial impact of IPS is robust, upheld in all 15 studies. The effectiveness of IPS is also suggested by other measures of competitive employment outcome, including time to first job, job duration and total hours employed during the follow-up period. Most IPS clients work part-time, typically half-time; about two-thirds of those who obtain competitive employment work 20 hours or more per week. Few IPS clients work full-time, likely due to preferences, limited stamina, and/or fear of losing health insurance or other benefits. Consistent with the principle of rapid job search, the time to first competitive job for IPS participants is nearly 10 weeks sooner than for controls. The mean length of time to first job for IPS participants (19 weeks) is, however, still lengthy for a model that prescribes rapid job search.

This review advances over earlier reviews in several respects. First, it has the largest and most up-to-date collection of pertinent randomized controlled trials. Second, it expands the scope of outcomes examined. Third, it is limited to rigorous evaluations of IPS programs, giving the clearest picture of the potential for IPS. Fourth, it is the first review to systematically compare US to non-US studies.

Some comment is warranted about the inclusion of the two studies clearly identified as outliers (29,34). The Maryland (34) study clearly deviated from the other IPS studies in that it was the only study among those reviewed that did not require participants to have a goal of competitive employment. Many participants apparently joined the study to receive the research payments and not because of their interest in employment. This study's poor competitive employment outcomes are consistent with its lenient admission criteria. Regarding the UK study (29), we concur with two commentators (43,44) who noted this study's shortcomings in adhering to the IPS model, according to descriptions provided by the investigators (45).

Conversely, an outlier on the upper end was the IPS study of veterans with post-traumatic stress disorder (PTSD) (28). This study had outstanding outcomes on most employment indicators, suggesting that this target population may be especially amenable to IPS interventions, though replication is needed. While PTSD is not usually classified as a severe mental illness, some PTSD researchers have argued that it should be, given its long-term nature and the disability it often engenders (46). Systematic research is needed to determine which diagnoses and disabilities IPS is suited for.

An unresolved question is whether noncompetitive employment outcomes are equivalent to competitive jobs with respect to their utility for clients, program managers, funders of rehabilitation services, and society at large. The IPS model is based on the argument that competitive jobs are greatly preferred over noncompetitive ones by clients themselves (47). In addition, a sustained period of competitive employment has been associated with better nonvocational outcomes in some studies (41,48), whereas this has not been shown as clearly for noncompetitive jobs. We assume that the advantages of competitive jobs are best evaluated in long-term studies (49). Nonetheless, several studies in this review found that control interventions were equally effective as IPS in achieving a range of paid employment outcomes when noncompetitive jobs were included. Finally, the costs associated with developing and maintaining noncompetitive job programs should also be considered; anecdotal evidence suggests the costs are often enormous (15). Moreover, the societal burden of developing and maintaining noncompetitive jobs is unsustainable on a large scale, in that costs are usually entirely borne by governmental subsidies rather than by the private sector and clients typically do not pay taxes on noncompetitive jobs.

The low dropout rates reported in IPS studies are thought provoking. First, they are in contrast to an early review noting high dropout rates among supported employment clients (3). Consistent with the assertive outreach component of the model, IPS programs have exceptionally low dropout rates, less than 10% in most studies. Conversely, studies often report high dropout rates for control participants. The contrast in termination rates for IPS and control groups raises a different question, whether the superior employment outcomes for IPS can be attributed to attrition. In other words, would the intent-to-treat findings for IPS reported above hold up for treatment exposure analysis? That is, what if the analyses were repeated with program dropouts removed? One study that has conducted this analysis found that IPS exceeded controls in comparisons that excluded dropouts (32). However, treatment exposure analyses were not reported in the other studies. Of course, it could be argued that control participants who dropped out did so because they viewed the control intervention as ineffective. Clearly, this question warrants further study.

Enrollment in IPS per se does not improve nonvocational outcomes beyond services as usual. Improved nonvocational outcomes may only accrue for clients who work steadily over time in a competitive job (48). These relationships need further exploration within longitudinal studies.

Worldwide interest in the IPS model is suggested by the increased proportion of IPS studies conducted outside the US reported since 2007. One new finding to emerge from the current review was that competitive employment rates are stronger for the US studies than for non-US studies. In particular, the European and Canadian studies had poorer outcomes than the US studies, while the outcomes from the Hong Kong and Australian studies were comparable to those in the US. Understanding the reasons will be important for policy planners and service providers as IPS continues to be disseminated internationally (50). Diminished effectiveness for IPS, particularly in Europe, has been typically attributed to labor and disability policies that can impede returns to work, for example, what Burns et al (7) refer to as the "disability trap". A Swedish study of IPS currently in progress describes in detail the bureaucratic inertia and attitudinal barriers within the Swedish welfare system impeding the development of effective IPS services (51). A Dutch study has also described the challenges in implementing IPS (52). Qualitative studies suggest that these barriers are formidable and to some extent represent challenges not found in the US. IPS leaders in several other countries have pursued strategies to overcome these barriers (53,54). Further international studies are needed to examine the nature and strength of these policy factors and to determine what adaptations are needed. At present, too few international RCTs have been conducted to draw strong conclusions about the influence of policy and of economic, cultural, and societal factors.

An alternative explanation for the poorer employment outcomes in several non-US IPS studies is the lack of adequate technical assistance and training for staff, leading to substandard implementation. Without adequate fidelity, the effectiveness of a program is attenuated and the quality of the resulting evaluation is greatly compromised. We note that all of the US studies were either conducted by, or received consultation from, the developers of the IPS, whereas only one-third of the non-US studies (7,38) received direct input from the model developers. Geographic distance is likely a factor for this difference.

The quality of implementation of the non-US studies is generally difficult to evaluate because of the lack of process details contained in their published reports. Two non-US multisite studies reported substandard fidelity in a minority of sites (7,27).

How do we explain the high ratings for IPS fidelity reported in the UK (29) study? We notice that the ratings were not made by independent assessors familiar with IPS, and a wealth of research has shown that self-ratings by project staff are often inflated (55). Given the strong association between IPS fidelity and competitive employment outcome (14), we propose that future reviews be restricted to evaluations of high-fidelity IPS programs, as verified by independent fidelity reviewers trained in conducting these assessments.

The broader issue for the advancement of an evidence-based practice, both for practical reasons and for scientific rigor, is the criticality of adequately trained staff and access to appropriate technical assistance. While the field of implementation science is still in its infancy (56), some general findings are beginning to emerge. Widescale dissemination of IPS has been facilitated by expert technical assistance in the US (57). When IPS technical assistance has been absent, dissemination results have often been dismal (58-61). The critical need for training and quality assurance in implementation of a program model has led the developers of other evi-

dence-based models to insist that users agree to systematic training and technical assistance regimens to assure quality of implementation (62). As a guide for determining when to intervene, Becker et al (63) have suggested that programs with quarterly competitive employment rate under 33% should be considered still in startup phase or as failing programs in need of immediate technical assistance.

Of course, undue influence of model developers on evaluations of their own model has been criticized as introducing the bias of therapeutic allegiance (64). This suggests the continuing need for the training of a second generation of IPS experts to conduct studies independent of the model developers, work that has already begun and has been represented in the current set of studies.

To summarize, the question of IPS transportability outside the US remains unanswered. While the published studies suggest that the labor and disability laws in some European countries may make a direct replication of IPS difficult, there are also indications that IPS transports well to other countries, such as Australia and the Hong Kong region of China. Finally, before concluding that the IPS must undergo radical adaptations in another nation, IPS programs should receive sufficient training and guidance to implement the model with high fidelity.

References

1. Becker DR, Drake RE. A working life for people with severe mental illness. New York: Oxford University Press, 2003.
2. Drake RE, Bond GR, Becker DR. IPS supported employment: an evidence-based approach. New York: Oxford University Press (in press).
3. Bond GR, Drake RE, Mueser KT et al. An update on supported employment for people with severe mental illness. Psychiatr Serv 1997;48:335-46.
4. Bond GR, Becker DR, Drake RE et al. Implementing supported employment as an evidence-based practice. Psychiatr Serv 2001; 52:313-22.
5. Bond GR. Supported employment: evidence for an evidence-based practice. Psychiatr Rehabil J 2004;27:345-59.
6. Bond GR, Drake RE, Becker DR. An update on randomized controlled trials of evidence-based supported employment. Psychiatr Rehabil J 2008;31:280-90.
7. Burns T, Catty J, Becker T et al. The effectiveness of supported employment for people with severe mental illness: a randomised controlled trial. Lancet 2007;370:1146-52.
8. Crowther RE, Marshall M, Bond GR et al. Helping people with severe mental illness to obtain work: systematic review. BMJ 2001; 322:204-8.
9. Kinoshita Y, Furukawa TA, Omori IM et al. Supported employment for adults with severe mental illness (Protocol). Cochrane Database of Systematic Reviews 2010;1:CA008297.
10. Kreyenbuhl J, Buchanan RW, Dickerson FB et al. The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2009. Schizophr Bull 2010;36:94-103.
11. Twamley EW, Jeste DV, Lehman AF. Vocational rehabilitation in schizophrenia and other psychotic disorders: a literature review and meta-analysis of randomized controlled trials. J Nerv Ment Dis 2003;191:515-23.
12. Wang CC, Sung C, Hiatt E et al. Effectiveness of supported employment on competitive employment outcomes: a meta-analysis. Presented at the National Council of Rehabilitation Educators, Los Angeles, April 2011.
13. Burns T. Models of community treatments in schizophrenia: do they travel? Acta Psychiatr Scand 2000;102(Suppl. 407):11-4.
14. Bond GR, Becker DR, Drake RE. Measurement of fidelity of implementation of evidence-based practices: case example of the IPS Fidelity Scale. Clinical Psychology: Science and Practice 2011;18: 126-41.
15. Black BJ. Work and mental illness: transitions to employment. Baltimore: Johns Hopkins Press, 1988.
16. Bond GR, Campbell K, Drake RE. Domains of employment measurement in Individual Placement and Support research. Psychiatr Serv (in press).
17. Lipsey MW. Design sensitivity. Newbury Park: Sage, 1990.
18. Bond GR, Dietzen LL, McGrew JH et al. Accelerating entry into supported employment for persons with severe psychiatric disabilities. Rehabil Psychol 1995;40:91-111.
19. Chandler D, Meisel J, Hu T et al. A capitated model for a cross-section of severely mentally ill clients: employment outcomes. Commun Ment Health J 1997;33:501-16.
20. Gervey R, Bedell JR. Supported employment in vocational rehabilitation. In: Bedell JR (ed). Psychological assessment and treatment of persons with severe mental disorders. Washington: Taylor & Francis, 1994:151-75.
21. McFarlane WR, Dushay RA, Deakins SM et al. Employment outcomes in Family-aided Assertive Community Treatment. Am J Orthopsychiatry 2000;70:203-14.
22. Test MA, Allness DJ, Knoedler WH. Impact of seven years of assertive community treatment. Presented at the American Psychiatric Association Institute on Psychiatric Services, Boston, October 1995.
23. Macias C, Rodican CF, Hargreaves WA et al. Supported employment outcomes of a randomized controlled trial of ACT and clubhouse models. Psychiatr Serv 2006;57:1406-15.
24. Rogers ES, Anthony WA, Lyass A et al. A randomized clinical trial of vocational rehabilitation for people with psychiatric disabilities. Rehabil Counsel Bull 2006;49:143-56.
25. Shafer M. Employment Intervention Demonstration Project: Arizona site. Unpublished paper, 2005.
26. Tsang HW, Chan A, Wong A et al. Vocational outcomes of an integrated supported employment program for individuals with persistent and severe mental illness. J Behav Ther Exper Psychiatry 2009;40:292-305.
27. Michon HW, van Vugt M, van Busschbach J. Effectiveness of Individual Placement and Support; 18 & 30 months follow-up. Presented at the Enmesh Conference, Ulm, June 2011.
28. Davis LL, Leon AC, Toscano R et al. A randomized clinical trial evaluating the impact of supported employment among veterans with posttraumatic stress disorder. Psychiatr Serv (in press).
29. Heslin M, Howard L, Leese M et al. Randomized controlled trial of supported employment in England: 2 year follow-up of the Supported Work and Needs. World Psychiatry 2011;10:132-7.
30. Nuechterlein KH. Individual placement and support after an initial episode of schizophrenia: the UCLA randomized controlled trial. Presented at the International Conference on Early Psychosis, Amsterdam, November 2010.
31. Mueser KT, Clark RE, Haines M et al. The Hartford study of supported employment for persons with severe mental illness. J Consult Clin Psychol 2004;72:479-90.
32. Bond GR, Salyers MP, Dincin J et al. A randomized controlled trial comparing two vocational models for persons with severe mental illness. J Consult Clin Psychol 2007;75:968-82.
33. Drake RE, McHugo GJ, Becker DR et al. The New Hampshire study of supported employment for people with severe mental illness: vocational outcomes. J Consult Clin Psychol 1996;64:391-9.
34. Lehman AF, Goldberg RW, Dixon LB et al. Improving employment outcomes for persons with severe mental illness. Arch Gen

Psychiatry 2002;59:165-72.
35. Nuechterlein KH, Subotnik KL, Turner LR et al. Individual Placement and Support for individuals with recent-onset schizophrenia: integrating supported education and supported employment. Psychiatr Rehabil J 2008;31:340-9.
36. Drake RE, Becker DR, Anthony WA. The use of a research induction group in mental health services research. Hosp Commun Psychiatry 1994;45:487-9.
37. Drake RE, McHugo GJ, Bebout RR et al. A randomized clinical trial of supported employment for inner-city patients with severe mental illness. Arch Gen Psychiatry 1999;56:627-33.
38. Latimer E, Lecomte T, Becker D et al. Generalisability of the individual placement and support model of supported employment: results of a Canadian randomised controlled trial. Br J Psychiatry 2006;189:65-73.
39. Gold PB, Meisler N, Santos AB et al. Randomized trial of supported employment integrated with assertive community treatment for rural adults with severe mental illness. Schizophr Bull 2006;32:378-95.
40. Wong KK, Chiu R, Tang B et al. A randomized controlled trial of a supported employment program for persons with long-term mental illness in Hong Kong. Psychiatr Serv 2008;59:84-90.
41. Burns T, Catty J, White S et al. The impact of supported employment and working on clinical and social functioning: results of an international study of Individual Placement and Support. Schizophr Bull 2009;35:949-58.
42. Twamley EW, Narvaez JM, Becker DR et al. Supported employment for middle-aged and older people with schizophrenia. Am J Psychiatr Rehabil 2008;11:76-89.
43. Campbell K. Employment rates for people with severe mental illness in the UK not improved by 1 year's individual placement and support. Evidence-Based Mental Health 2010;doi:10.1136/ebmh1090.
44. Latimer E. An effective intervention delivered at sub-therapeutic dose becomes an ineffective intervention. Br J Psychiatry 2010;196: 341-2.
45. Howard LM, Heslin M, Leese M et al. Supported employment: randomised controlled trial. Br J Psychiatry 2010;196:404-11.
46. Friedman MJ, Rosenheck RA. PTSD as a persistent mental illness. In: Soreff S (ed). The seriously and persistently mentally ill: the state-of-the-art treatment handbook. Seattle: Hogrefe & Huber, 1996:369-89.
47. Bedell JR, Draving D, Parrish A et al. A description and comparison of experiences of people with mental disorders in supported employment and paid prevocational training. Psychiatr Rehabil J 1998;21:279-83.
48. Bond GR, Resnick SG, Drake RE et al. Does competitive employment improve nonvocational outcomes for people with severe mental illness? J Consult Clin Psychol 2001;69:489-501.
49. Penk W, Drebing CE, Rosenheck RA et al. Veterans Health Administration transitional work experience vs. job placement in veterans with co-morbid substance use and non-psychotic psychiatric disorders. Psychiatr Rehabil J 2010;33:297-307.
50. Rinaldi M, Killackey E, Smith J et al. First episode psychosis and employment: a review. Int Rev Psychiatry 2010;22:148-62.
51. Hasson H, Andersson M, Bejerholm U. Barriers in implementation of evidence-based practice: supported employment in Swedish context. Journal of Health Organization and Management 2011;25:332-45.
52. van Erp NH, Giesen FB, van Weeghel J et al. A multisite study of implementing supported employment in the Netherlands. Psychiatr Serv 2007;58:1421-6.
53. Boyce M, Secker J, Floyd M et al. Factors influencing the delivery of evidence-based supported employment in England. Psychiatr Rehabil J 2008;31:360-6.
54. Waghorn GR, Collister L, Killackey E et al. Challenges to implementing evidence-based supported employment in Australia. Journal of Vocational Rehabilitation 2007;27:29-37.
55. Lee N, Cameron J. Differences in self and independent ratings on an organisational dual diagnosis capacity measure. Drug and Alcohol Review 2009;28:682-4.
56. Fixsen DL, Naoom SF, Blase KA et al. Implementation research: a synthesis of the literature. Tampa: University of South Florida, 2005.
57. Becker DR, Drake RE, Bond GR et al. A mental health learning collaborative on supported employment. Psychiatr Serv 2011;62: 704-6.
58. Bond GR. Modest implementation efforts, modest fidelity, and modest outcomes. Psychiatr Serv 2007;58:334.
59. Liu HH. Policy and practice: an analysis of the implementation of supported employment in Nebraska. Dissertation, University of Nebraska, Lincoln, 2011.
60. McGrew JH. Final report to SECT Center: IPS fidelity survey of supported employment programs in Indiana. Indianapolis: Indiana University-Purdue University Indianapolis, 2008.
61. Menear M, Reinharz D, Corbière M et al. Organizational analysis of Canadian supported employment programs for people with psychiatric disabilities. Soc Sci Med 2011;72:1028-35.
62. Schoenwald SK, Heiblum N, Saldana L et al. The international implementation of multisystemic therapy. Evaluation and the Health Professions 2008;31:211-25.
63. Becker DR, Drake RE, Bond GR. Benchmark outcomes in supported employment. Am J Psychiatr Rehabil 2011;14:230-6.
64. Jacobson NS. The role of the allegiance effect in psychotherapy research: controlling and accounting for it. Clinical Psychology: Science and Practice 1999;6:116-9.
65. Killackey E, Jackson HJ, McGorry PD. Vocational intervention in first-episode psychosis: individual placement and support v. treatment as usual. Br J Psychiatry 2008;193:114-20.

Age at onset versus family history and clinical outcomes in 1,665 international bipolar-I disorder patients

Ross J. Baldessarini[1,2], Leonardo Tondo[1,2,3], Gustavo H. Vázquez[2,4], Juan Undurraga[1,5], Lorenza Bolzani[2,6], Aysegul Yildiz[2,7], Hari-Mandir K. Khalsa[2], Massimo Lai[8], Beatrice Lepri[3], Maria Lolich[4], Pier Mario Maffei[6], Paola Salvatore[1,2,9], Gianni L. Faedda[2,10], Eduard Vieta[2,5], Mauricio Tohen[2,11]

[1]Department of Psychiatry, Harvard Medical School, Boston, MA, USA; [2]International Consortium for Bipolar Disorder Research, McLean Hospital, Belmont, MA, USA; [3]Lucio Bini Mood Disorder Center, Cagliari, Italy; [4]Department of Neuroscience, Palermo University, Buenos Aires, Argentina; [5]Department of Psychiatry, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain; [6]Viarnetto Psychiatric Clinic, Lugano, Switzerland; [7]Department of Psychiatry, Dokuz Eylül University, Izmir, Turkey; [8]Clinique Lyon-Lumière, Lyon, France; [9]Section of Psychiatry, Department of Neuroscience, University of Parma, Italy; [10]Lucio Bini Mood Disorders Center, New York, NY, USA; [11]Department of Psychiatry, University of Texas Health Science Center, San Antonio, TX, USA

Early onset in bipolar disorder (BPD) has been associated with greater familial risk and unfavorable clinical outcomes. We pooled data from seven international centers to analyze the relationships of family history and symptomatic as well as functional measures of adult morbidity to onset age, or onset in childhood (age <12), adolescence (12-18), or adulthood (19-55 years). In 1,665 adult, DSM-IV BPD-I patients, onset was 5% in childhood, 28% in adolescence, and 53% at peak ages 15-25. Adolescent and adult onset did not differ by symptomatic morbidity (episodes/year, percentage of months ill, co-morbidity, hospitalization, suicide attempts) or family history. Indications of favorable adult functional outcomes (employment, living independently, marriage and children, and a composite measure including education) ranked, by onset: adult > adolescent > child. Onset in childhood versus adolescence had more episodes/year and more psychiatric co-morbidity. Family history was most prevalent with childhood onset, similar over onset ages 12-40 years, and fell sharply thereafter. Multivariate modeling sustained the impression that family history and poor functional, but not symptomatic, outcomes were associated with younger, especially childhood onset. Early onset was more related to poor functional outcomes than greater symptomatic morbidity, with least favorable outcomes and greater family history with childhood onset.

Key words: Bipolar disorder, adult functional status, morbidity, age at onset, outcome

(World Psychiatry 2012:11:40-46)

Age at onset of type-I bipolar disorder (BPD) typically averages 12-24 years, is older among patients with type-II BPD, and oldest in unipolar major depressive disorder (1-3). Reported onset ages probably vary by ascertainment methods, and possibly among different countries and cultures (1-6). Early onset of BPD in childhood or adolescence is of particular interest as it may help to define subgroups. Juvenile onset generally appears to be common, reportedly ranging from 39% to 65% of samples including patients evaluated as juveniles or adults (2,4-7). BPD patients with early onset may represent a phenotype of special interest for genetic and other biomedical research, as well as having potential clinical importance (8-11). Genetic interest arises in part from relatively high rates of familial mood disorders in BPD, generally, and particularly in association with young onset (5-11). Early onset cases also provide challenges of earlier recognition, prognosis, and treatment of young patients whose early illnesses may differ from typical adult types, and whose diagnosis and treatment are typically delayed for years (7,12-15).

Several investigators have applied the statistical method of admixture analysis of onset age distributions, which typically yield nodal prevalence in adolescent or early adult years with some skewing to younger ages (16-21). When large samples of type-I BPD probands have been evaluated with this method, onset ages typically have yielded three putatively independent, nearly normal Gaussian distributions, with ages averaging 17.1±1.7, 25.3±1.8, and 38.0±4.3 years (16-21). Findings in these studies were similar across various geographical regions (including Canada, France, Italy, the US, and Wales), suggesting some consistency despite likely ethnic and clinical heterogeneity. However, the contributions of the three computed onset age subgroups to the total varied widely (36% to 80%, 7% to 39%, and 13% to 25%, respectively). Usually, however, the youngest subgroup predominated and was associated with various adverse clinical outcome measures.

Most previous research has suggested that BPD following pre-adult onset may be particularly severe. Such severity reportedly is indicated by relatively high rates of rapid cycling or even chronic illness, prominent psychotic and anxiety features, substance abuse, and limited response to mood stabilizing treatment, compared to cases with adult onset (4,5,7,22-24). Poor outcomes may reflect: a) particularly virulent long-term illness following juvenile onset, b) effects of prolonged delay or refusal of treatment in childhood and adolescence, c) impressions arising from juvenile illnesses that may not be expressed in more clearly episodic, adult form, or d) possible destabilizing effects of antidepressants and stimulants commonly used to treat children and adolescents with behavioral symptoms (7,12,13). Another possibility is that juvenile, and particularly childhood onset disorders may have a severe impact on maturation, particularly to adult functional levels (25). However, not all reports support the hypothesis that early onset BPD follows a more severe course or outcome than adult onset BPD (26-28). Moreover, admixture analysis has yielded evidence of discrete subgroups that do not necessarily correspond to developmental periods of clinical inter-

est (childhood, adolescence, young adulthood, middle age, and late years).

We undertook the present study in view of the evident importance of early identification of potentially severe, disabling or even fatal BPD following early onset, and the relative rarity and inconsistency of comparisons of long-term symptomatic and functional outcomes in large numbers of female and male patients meeting standard diagnostic criteria for type-I BPD, followed into adulthood, and across multiple cultural settings. We endeavored to limit effects of potential geographic and cultural variance by pooling demographic and clinical data from 1,665 patients meeting DSM-IV diagnostic criteria as adults, from seven mood disorder centers in Argentina, Italy, Spain, Switzerland, Turkey, and the US. Study hypotheses were that: a) early onset would be followed by greater morbidity by most available measures, and b) family history of affective illness would be inversely and continuously related to onset age. We also planned specifically to consider possible differences between onset in childhood compared to adolescence and adulthood, and to compare measures of symptomatic versus functional outcomes.

METHODS

Subjects

We pooled data from a total of 1,665 patients diagnosed at adult ages with type-I BPD by DSM-IV criteria at seven sites affiliated with an International Consortium for Bipolar Research based at McLean Hospital/Harvard Medical School: Lucio Bini Mood Disorders Center, Cagliari, Italy (n=586); Argentine Network for Bipolar Disorders at Palermo University, Buenos Aires, Argentina (n=328); McLean Hospital, Boston, MA, USA (n=215); University Clinic, Barcelona, Spain (n=204); Viarnetto Psychiatric Clinic, Lugano, Switzerland (n=174); Department of Psychiatry, Dokuz Eylül University, Izmir, Turkey (n=134); and Lucio Bini Juvenile Mood Disorder Center, New York, NY, USA (n=24). The US sites sought to enhance representation of early onset cases by selecting cases with juvenile onset (age ≤18; New York) and first-episode patients (Boston), all followed prospectively into adult years, as at other sites. Patients were evaluated, treated clinically, and followed for at least 3-5 years, using methods detailed previously (29-35).

Subjects were assessed retrospectively for estimated onset age, based on first clinically appreciable syndromal illness, as indicated by patient history, recollections of family members, and medical records. Onset age was separated into onset groups of clinical interest: childhood (<12 years), adolescence (12-18 years), or adulthood (≥19 years), or onset age was considered as a continuous measure. For assessment of family history rates, all cases were included, with subgroups formed by decades of onset age, as well as during childhood, adolescence, early adult, and advanced ages.

Assessments

As continuous, symptomatic measures of morbidity, we considered the annual rate (episodes/year) of major DSM-IV BPD episodes (mania or hypomania, major depression, mixed states, or psychosis) from illness onset, as well as estimates of percent of months ill per year, both during exposure times limited to ≥2 years. We also considered several categorical clinical measures: presence of any DSM-IV Axis I psychiatric or substance use disorder co-morbidity; ever manifesting psychotic symptoms; ever being psychiatrically hospitalized; and ever having attempted suicide. The following functional or social outcomes were considered: having completed high school or higher education; ever having married; having children; being gainfully employed or a student or homemaker at last follow-up; and living independently, with a composite categorization based on these functional measures. The composite functional measure was based on the sum of weighted ratings of being employed (10 points), living independently (5 points), ever being married (2 points), having children (1 point), and having completed high school (1 point). "Poor functional outcome" was defined by a total score of zero. We also assessed family history of psychiatric illness (affective or substance use) in first-degree relatives.

Data analyses

We compared those with onset in childhood (<12 years), adolescence (12–18 years) or in juvenile years overall (≤18 years) versus adulthood (≥19 years) for the previously defined clinical measures. Family history rates were compared in the same subgroups as well as across decades of onset ages. Outcome assessments were limited to subjects followed-up to age ≥25 years to allow time to attain adult indices of functional accomplishment, as well as being at risk for at least 2 years to avoid exaggerating estimates with very short exposure times (35), and with onset age ≤55 years to avoid cases of secondary mania (36), yielding 1,368 cases (82.2% of the total).

Histographic analysis involved all onset ages observed. We also used contingency table-based (χ^2) comparisons of cases following onset in childhood (<12), adolescence (12-18), or young adult age (19-55 years) for comparison to categorical outcome measures. In addition, continuous measures, including onset ages (log-normalized), episodes/year and proportion of time ill, were compared between onset age categories by ANOVA methods (F). Degrees of freedom (df) are provided. We also used stepwise, multivariate logistic regression modeling for factors associated significantly and independently with the onset age subgroups, as well as multivariate linear regression modeling of representative factors associated with onset age as a continuous measure, in both cases limiting the sample to the 1,368 cases, as defined above. Averages are means ± standard deviation (SD) or medians with interquartile range (IQR), unless stated otherwise.

RESULTS

Overall subject characteristics

The 1,665 type-I BPD subjects included 52.3% women and onset age averaged 25.7±11.3 years. The median (IQR) onset age was 23.0 (13.0) overall. It was 22.0 (12.0) in men vs. 24.0 (14.4) in women (by log-normalized onset-age: F=9.21, p=0.002). Juvenile onset (age ≤18 years) involved 26.6% of subjects (n=477), of whom 83 (5.0% of all cases) were children. Current age averaged 40.8±14.4 years, and exposure time of illness from onset averaged 15.1±11.5 years. Proportions of subjects with juvenile onset (≤18 years) averaged 28.6% (it was 26.9% in Izmir, 24.7% in Buenos Aires, 22.0% in Cagliari, 21.6% in Barcelona, and 17.8% in Lugano based on hospitalized subjects). US patients had higher proportions of juvenile onset cases, owing to selection factors (63.7% in Boston, involving only first-episode patients followed-up prospectively, and 79.2% in New York, at a pediatric mood disorder center), with correspondingly young median (IQR) onset ages: 16.0 (7.8) and 13.5 (10.5), respectively.

Distribution of onset ages

Overall median (IQR) onset age was 24.0 (13.1) years, with moderate skewing toward ages 15-25 years, compared to a normal Gaussian distribution (Figure 1). Peak prevalence at ages 15-25 years accounted for a majority (53.0%) of all 1,665 cases, and prevalence was <5% at ages <15, and >45 years.

Characteristics by age subgroups

We compared various demographic and clinical characteristics in patients with onset in childhood (ages <12), adolescence (12-18 years), or adulthood (19-55 years) in a sample limited as defined above (onset age <55, ill ≥2 years, followed to ages ≥25). Statistical comparisons across onset age groups are provided for illustrative purposes and to guide subsequent multivariate modeling, and are unadjusted for multiple comparisons (Table 1). Notable relationships included the following. Family history was strongly associated with younger onset, and significantly greater among childhood than adolescent onset cases (χ^2 = 11.1, df=1, p=0.004). Episodes/year as a major measure of symptomatic morbidity was highest with childhood and lowest with adolescent onset (F=3.92, p=0.02), whereas the approximate proportion of months/year in a major episode of BPD tended to rise inversely with younger onset, but did not differ significantly between childhood and either adolescent or adult onset. There was also somewhat greater prevalence of psychiatric or substance use co-morbidity among childhood onset than adolescent onset patients (χ^2=2.67, p=0.07). Psychosis differed significantly between childhood and adolescent onset patients (χ^2=24.8, p<0.0001), but its prevalence *decreased*

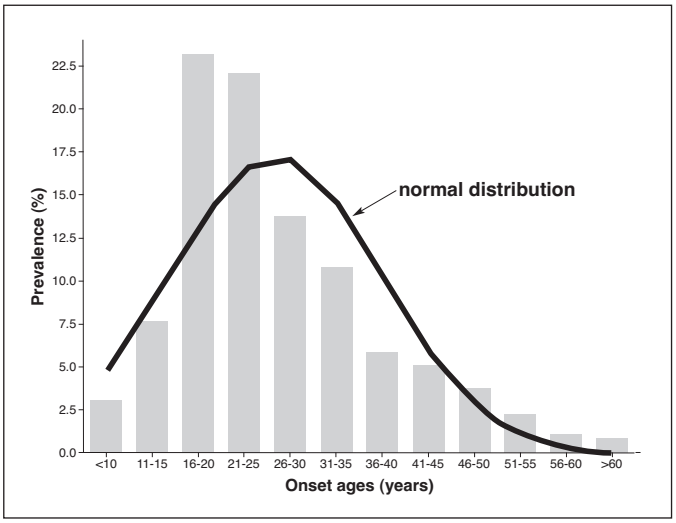

Figure 1 Histogram of onset ages in 1,665 bipolar-I disorder patients, with superimposed normal Gaussian distribution, indicating moderate skewing toward younger ages

with younger onset age. For the composite measure of adult functional status, there was a much greater risk of poor functional outcome with younger onset age, with significantly greater prevalence following onset in childhood than in adolescence (χ^2=31.5, p<0.0001; Table 1).

Since childhood onset (ages <12 years) represented a minority of cases (5.0%), we also compared adult outcomes among all subjects with juvenile onset (age ≤18 years; 33.0% of the total) and those with older onset ages in the restricted sample of 1,368 patients. Adult functional outcome measures again differed much more between juvenile and adult onset cases than did symptomatic measures. For example, the pooled index of successful functional outcome was 41% lower following younger onset (χ^2=17.0, p<0.0001), with similar differences for employment, independent living, and marital status. In contrast, among symptomatic measures, episodes/year, any co-morbidity, hospitalization, and psychotic features did not differ significantly between these onset age groups, although, with younger onset, percentage time ill/year was 31% greater (F=5.65, p=0.02) and suicide attempts 33% more frequent (χ^2=4.47, p=0.04). Moreover, the effect of onset age on adult functional outcomes was even greater among men than women (relative risk by juvenile/adult onset age = 1.68 versus 1.17).

The preceding findings were similar across geographic regions sampled. Notably, employment, a measure of functional outcome, was consistently lower among juvenile onset cases (by an average of 1.37±0.16 times) in both US and other centers.

Family history

The frequency of identified affective illness or substance abuse among first-degree relatives was strongly related to on-

Table 1 Comparisons of adult bipolar-I disorder patients with onset in childhood, adolescence or adulthood

	Onset age group				
	Child (<12)	Adolescent (12-18)	Adult (19-55)	χ^2 or F	p
Cases (n)	53	335	980	-	-
% women	34.0	53.1	57.4	11.6	0.003
Onset age (years)					
Mean ± SD	7.94±2.03	16.9±1.78	30.2±10.2	-	<0.0001
Median (IQR)	8.00 (4.00)	16.0 (2.00)	28.0 (13.1)	-	-
Current age (years, mean ± SD)	34.1±9.83	38.4±11.1	45.7±12.8	59.7	<0.0001
Years of illness (mean ± SD)	26.2±9.54	22.0±11.4	15.5±10.4	64.9	<0.0001
Family affective history (%)*	88.6	66.3	61.6	11.1	0.004
Episodes/year (mean ± SD)*	1.06±1.71	0.68±0.72	0.79±0.89	3.92	0.02
% of months/year ill (mean ± SD)	46.9±41.1	39.7±35.4	33.7±32.7	2.69	0.07
Ever hospitalized (%)	82.9	86.2	81.3	2.35	0.31
Ever psychotic (%)*	16.2	38.9	53.5	24.8	<0.0001
Ever attempted suicide (%)	27.3	27.8	20.4	4.45	0.11
Any psychiatric co-morbidity (%)*	90.9	54.1	57.4	5.71	0.06
Education ≥ high school (%)	53.8	61.6	63.1	1.05	0.59
Ever married (%)	37.0	44.7	58.4	24.2	<0.0001
With children (%)	10.0	29.2	38.1	6.52	0.04
Employed (%)	37.5	62.0	71.8	28.2	<0.0001
Living independently (%)	0.00	43.9	75.5	16.9	0.0002
Functionally impaired (%)*	60.9	42.7	30.2	31.5	<0.0001

*Significantly different in patients with onset in childhood versus adolescence

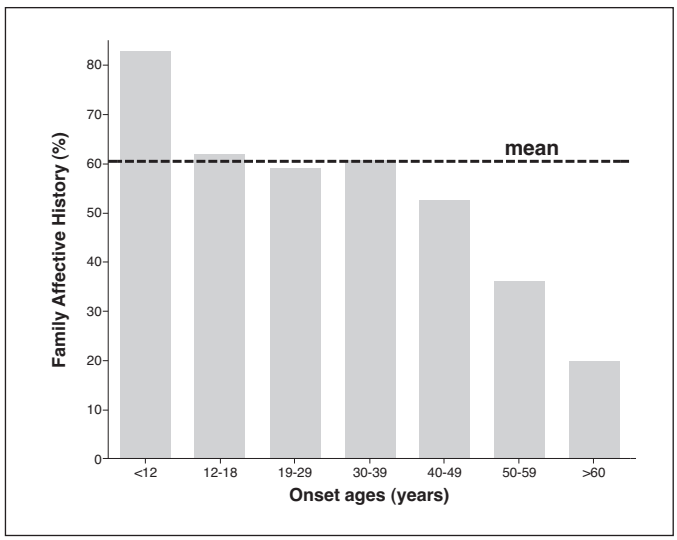

Figure 2 Prevalence (%) of family history of affective illness versus onset ages among 1,665 bipolar-I disorder patients

set age, overall (Figure 2). Prevalence of family history was highest with childhood onset (83.1%), similar from ages 12 through 39 (averaging 60.7±1.4%) years, and declined with higher onset ages, to 52.8% at 40-49, 39.4% at 50–59, and only 20.0% at ≥60 years.

Multivariate analyses of factors associated with onset age

We used both linear (actual onset age) and logistic (juvenile vs. adult, and childhood vs. adolescent onset) multivariate regression modeling to test for significant and independent associations of selected factors with onset age (Table 2). We were particularly interested in testing the hypothesis that functional outcome and family history, but not a measure of symptomatic morbidity (episodes/year), would be associated with earlier onset age. This prediction was sustained with both models. The impression that outcomes were generally even less favorable, and family history greater, with onset in childhood versus adolescence was also supported (Table 2).

DISCUSSION

The present findings indicate that several demographic, clinical, and functional outcome factors were associated variably with onset age in a large, international sample of DSM-IV, type-I BPD patients. Outcome analyses were limited to patients with onset age below 55 years and followed for at least 2 years into adult years (≥25) to limit potential confounding effects of immaturity, of very short exposure times (36), and risk of secondary mania (37).

Generally, measures of symptomatic morbidity were sur-

Table 2 Multivariate regression modeling: factors associated with onset age

Factors associated with juvenile onset: logistic regression			
Factors	**Odds ratio (95%CI)**	**χ^2**	**p**
Poor functional outcome	2.00 (1.34 to 2.95)	8.82	0.003
More family history	1.71 (1.20 to 2.43)	5.86	0.015
More episodes/year	1.08 (0.89 to 1.33)	0.63	0.43

Factors associated with onset in childhood vs. adolescence: logistic regression			
Factors	**Odds ratio (95%CI)**	**χ^2**	**p**
Poor functional outcome	2.70 (1.08 to 6.77)	4.51	0.03
More family history	4.65 (1.04 to 20.8)	4.04	0.04
More episodes/year	4.85 (1.05 to 22.4)	0.63	0.04

Factors associated with younger onset age: linear regression			
Factors	**β coefficient (95% CI)**	**t**	**p**
More family history	2.78 (1.05 to 4.41)	3.35	0.001
Poor functional outcome	2.61 (1.01to 4.21)	3.20	0.001
Episodes/year	0.62 (–1.90 to 1.43)	1.50	0.13

prisingly similar following onset in adolescent versus adult years, whereas functional outcomes were more favorable with older onset, sometimes differing significantly between childhood and adolescent onset. Family history was more prevalent with childhood onset, similar from onset ages of 12-40, and fell sharply at later onset ages. The general tendency for family history to be greater with younger onset is well known (4-11), but the striking differences among cases of childhood, adolescent to middle aged, and older onset ages appear to be a novel finding. Multivariate modeling sustained the impression that family history and poor functional, but not symptomatic, outcomes were associated with juvenile onset, with somewhat greater morbidity as well as familial risk with onset in childhood versus adolescence.

It is particularly noteworthy that the present findings indicate that cases of BPD with juvenile onset may have different outcomes, in that childhood onset appears to differ from both adolescent and adult onset and to be a particularly virulent form of the illness. Moreover, there was evidence that adult functional outcomes were even more impaired than symptomatic and other clinical measures following juvenile onset. If these hypothesis-generating findings are valid, they may be particularly important clinically in suggesting that developmental or maturational effects of early, and especially childhood onset BPD, may be especially great.

It is tempting to speculate that the observed effects on functional disability may be associated with the depressive-dysphoric components of BPD, which remain highly prevalent and a major therapeutic challenge (32,38). However, both depression (7) and mania/hypomania (39) have been reported to occur in excess with early onset BPD. Cognitive impairment in BPD also can occur and may contribute to the observed adverse functional outcomes associated with early onset (40). Whatever the bases of poor outcomes with very young onset may be, we propose that it is likely that morbidity-associated developmental delays are involved.

The observed lack of evidence of more severe symptomatic morbidity among juvenile onset cases of BPD followed-up for ≥2 years into adult years seems inconsistent with much (4,7,22-24), but not all (26-28), of the literature reviewed above. Inconsistencies may reflect differences in inclusion of cases with onset in childhood versus adolescence. Another possible factor is that most studies involving onset of putative BPD in childhood or adolescence considered diagnosis and morbidity in young ages with variable later verification of diagnosis in adult life, and few have included direct and systematic comparisons of cases involving juvenile versus adult onset, or childhood versus adolescent onset (4,7,23). Patients with BPD of juvenile, and especially childhood onset, continue to present severe diagnostic uncertainties, and tend to differ from adult onset forms of BPD, by lacking clear episodes or following a rapidly fluctuating, chaotic, or chronic course, and having high rates of psychotic features and co-morbidities that include anxiety, attentional, conduct, and substance-related disorders (2,3,12,19,23,24). These characteristics of juvenile onset BPD patients may contribute to the impression that such illnesses may be more severe than in adults. Whether that hypothesis is valid or not, it seems clear that earlier recognition, diagnosis, and improved treatments for early onset BPD are required (7,13-16). Our observations suggest that BPD starting in childhood may differ in its clinical implications not only from adult onset cases, but possibly also from adolescent onset patients – another hypothesis that requires further testing.

Limitations of this study include potential differences among study sites in methods of case and morbidity ascertainment, although onset age (except for samples selected for early onset in the US sites) and representative outcome measures were similar across geographic regions. Another limitation was that estimates of morbidity factors as well as onset ages were largely retrospective, including the need for recall by patients or families over many years; such recall might differentially impact cases with earlier onset. Also, some measures (suicidal acts, co-morbidity, psychosis) were present in relatively low prevalence, which may limit the reliability of their estimates. In addition, case selection biases may be involved, in that patients who were followed for long times may not be representative of others who were less accessible to, or cooperative with long-term follow-up and treatment.

Despite efforts to limit their impact, there may still be effects of years-at-risk, and years of adult life. Since duration of illness was longer with earlier onset, it may be that longer exposure times increased risk of some clinical outcomes (such as hospitalization, suicide attempts, psychosis, anxiety disorder, or other co-morbid psychiatric or substance use disorders). However, such an effect would tend to limit the observed lack of association of more of such outcomes with

earlier onset. In addition, longer exposure times with younger onset would likely have limited, rather than increased, measures of morbidity-per-time (such as episodes/year) (36). Treatment was clinical and uncontrolled, and may have modified long-term morbidity, though presumably randomly. A noteworthy aspect of this study was that DSM-IV diagnostic criteria for type-I BPD were met in adult life, to avoid the diagnostic complexity and uncertainties of pediatric diagnoses (41). Despite their potential limitations, reported findings were strikingly consistent across several methods of analysis and among geographic regions.

In conclusion, the findings presented raise the intriguing possibility that particular illness-related factors and outcomes may be associated differentially with onset of type-I BPD in childhood versus both adolescence and adulthood. We found especially strong relationships of juvenile onset with adverse social and functional outcome measures, including not being employed, not living independently, being unmarried and not having children, whereas most measures of symptomatic morbidity were much less related to onset age. We also found especially high prevalence of family history among childhood onset cases. If these observations are valid and replicable, they may suggest a particularly important impact of juvenile, and especially childhood, onset of BPD at the level of maturation and functional success in attaining major adult roles of later life, with lesser impact on the symptomatic expression of BPD. The particularly strong association of childhood-onset with high rates of reported family history further supports efforts to identify specific phenotypic subgroups of interest for genetic and other biomedical investigations. Early versus late onset has also been proposed as a course specifier for DSM-5 (42). Finally, the present findings further encourage earlier diagnosis and development of interventions aimed particularly at limiting maturational-functional impairments among young persons with BPD.

Acknowledgements

This study was supported in part by a grant from the Bruce J. Anderson Foundation and by the McLean Private Donors Research Fund for Bipolar Disorders, and by the Lucio Bini Private Donors Research Fund.

References

1. Baldessarini RJ, Bolzani L, Cruz N et al. Onset-age of bipolar disorders at six international sites. J Affect Dis 2010;121:143-6.
2. Larsson S, Lorentzen S, Mork E et al. Age at onset of bipolar disorder in a Norwegian catchment area sample. J Affect Disord 2010; 124:174-7.
3. Tondo L, Lepri B, Cruz N et al. Age at onset in 3014 Sardinian bipolar and major depressive disorder patients. Acta Psychiatr Scand 2010;121:446-52.
4. Perlis RH, Miyahara S, Marangell LB et al. Long-term implications of early onset in bipolar disorder: data from the first 1000 participants in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BP). Psychiatry 2004;55:875-81.
5. Goodwin RK, Jamison KR (eds). Manic-depressive illness, 2nd ed. New York: Oxford University Press, 2007.
6. Beesdo K, Höfler M, Leibenluft E et al. Mood episodes and mood disorders: patterns of incidence and conversion in the first three decades of life. Bipolar Disord 2009;11:637-49.
7. Post RM, Leverich GS, Kupka RW et al. Early-onset bipolar disorder and treatment delay are risk factors for poor outcome in adulthood. J Clin Psychiatry 2010;71:864-72.
8. Leboyer M, Henry C, Paillere-Martinot ML et al. Age at onset of bipolar affective disorders: a review. Bipolar Disord 2005;7:111-8.
9. Pavuluri MN. Effects of early intervention on the course of bipolar disorder: theories and realities. Curr Psychiatry Rep 2010;12:490-8.
10. Faraone SV, Glatt SJ, Su J et al. Three potential susceptibility loci shown by a genome-wide scan for regions influencing the age at onset of mania. Am J Psychiatry 2004;161:625-30.
11. Priebe L, Degenhardt FA, Herms S et al. Genome-wide survey implicates the influence of copy number variants (CNVs) in the development of early-onset bipolar disorder. Mol Psychiatry (in press).
12. Faedda GL, Baldessarini RJ, Glovinsky IP et al. Pediatric bipolar disorder: phenomenology and course of illness. Bipolar Disord 2004;6:305-13.
13. Birmaher B, Axelson D. Course and outcome of bipolar spectrum disorder in children and adolescents: review of the existing literature. Dev Psychopathol 2006;18:1023-35.
14. Leverich GS, Post RM, Keck PE Jr et al. The poor prognosis of childhood-onset bipolar disorder. J Pediatr 2007;150:485-90.
15. Suominen K, Mantere O, Valtonen H et al. Early age at onset of bipolar disorder is associated with more severe clinical features but delayed treatment-seeking. Bipolar Disord 2007;9:698-705.
16. Bellivier F, Golmard JL, Rietschel M et al. Age at onset in bipolar I affective disorder: further evidence for three subgroups. Am J Psychiatry 2003;160:999-1001.
17. Lin PI, McInnis MG, Potash JB et al. Clinical correlates and familial aggregation of age at onset in bipolar disorder. Am J Psychiatry 2006;163:240-6.
18. Manchia M, Lampus S, Chillotti C et al. Age at onset in Sardinian bipolar I patients: evidence for three subgroups. Bipolar Disord 2008;10:443-6.
19. Hamshere ML, Gordon-Smith K, Forty L et al. Age-at-onset in bipolar-I disorder: mixture analysis of 1369 cases identifies three distinct clinical sub-groups. J Affect Disord 2009;116:23-9.
20. Ortiz A, Bradler K, Slaney C et al. An admixture analysis of the age at index episodes in bipolar disorder. Psychiatry Res 2011;188:34-9.
21. Tozzi F, Manchia M, Galwey NW et al. Admixture analysis of age at onset in bipolar disorder. Psychiatry Res 2011;185:27-32.
22. Schürhoff F, Bellivier F, Jouvent R et al. Early and late onset bipolar disorders: two different forms of manic-depressive illness? J Affect Disord 2000;58:215-21.
23. Craney JL, Geller B. A prepubertal and early adolescent bipolar disorder-I phenotype: review of phenomenology and longitudinal course. Bipolar Disord 2003;5:243-56.
24. Geller B, Tillman R, Bolhofner K et al. Child bipolar I disorder: prospective continuity with adult bipolar I disorder; characteristics of second and third episodes; predictors of 8-year outcome. Arch Gen Psychiatry 2008;65:1125-33.
25. Goldstein TR, Birmaher B, Axelson D et al. Psychosocial functioning among bipolar youth. J Affect Disord 2009;114:174-83.
26. McGlashan TH. Adolescent versus adult onset of mania. Am J Psychiatry 1988;145:221-3.
27. Jairam R, Srinath S, Girimaji SC et al. Prospective 4–5-year follow-up of juvenile onset bipolar disorder. Bipolar Disord 2004;6:386-94.
28. Birmaher B, Axelson D, Strober M et al. Clinical course of children and adolescents with bipolar spectrum disorders. Arch Gen Psychiatry 2006;63:175-83.
29. Tohen M, Zarate CA Jr, Hennen J et al. The McLean-Harvard First-Episode Mania Study: prediction of recovery and first recurrence.

Am J Psychiatry 2003;160:2099-107.
30. Tondo L, Lepri B, Baldessarini RJ. Risks of suicidal ideation, attempts and suicides among 2826 men and women with types I and II bipolar, and recurrent major depressive disorders. Acta Psychiatr Scand 2007;116:419-28.
31. Yildiz A, Guleryuz S, Ankerst DP et al. Protein kinase C inhibition in the treatment of mania: a double-blind, placebo-controlled trial of tamoxifen. Arch Gen Psychiatry 2008;65:255-63.
32. Baldessarini RJ, Salvatore P, Khalsa HM et al. Morbidity in 303 first-episode bipolar I disorder patients. Bipolar Disord 2010;12: 264-70.
33. Tondo L, Lepri B, Baldessarini RJ. Reproduction in women and men with major affective disorders. Acta Psychiatr Scand 2011;123: 283-9.
34. Vázquez GH, Lolich M, Leiderman EA et al. Age-at-onset in 648 patients with major affective disorders: clinical and prognostic implications. Mind Brain J Psychiatry (in press).
35. Vieta E. Bipolar disorder units and programs: are they really needed? World Psychiatry 2011;10:152.
36. Baldessarini RJ, Tondo L, Baethge C et al. Effects of treatment latency on response to maintenance treatment in manic-depressive disorders. Bipolar Disord 2007;9:386-93.
37. Shulman KI, Tohen M, Kutcher S (eds). Mood disorders across the life span. New York: Wiley, 1996.
38. Baldessarini RJ, Vieta E, Calabrese JR et al. Bipolar depression: overview and commentary. Harv Rev Psychiatry 2010;18:143-57.
39. Bauer M, Glenn T, Rasgon N et al. Association between age of onset and mood in bipolar disorder: comparison of subgroups identified by cluster analysis and clinical observation. J Psychiatr Res 2010;44:1170-5.
40. Harvey PD, Wingo AP, Burdick K et al. Cognition and disability in bipolar disorder: lessons from schizophrenia research. Bipolar Disord 2010;12:364-75.
41. Faedda GL, Baldessarini RJ, Suppes T et al. Pediatric-onset bipolar disorder: a neglected clinical and public-health problem. Harv Rev Psychiatry 1995;3:171-95.
42. Colom F, Vieta E. The road to DSM-V: bipolar disorder episode and course specifiers. Psychopathology 2009;42:209-18.

Lessons learned in developing community mental health care in North America

Robert E. Drake[1], Eric Latimer[2]

[1]Dartmouth Psychiatric Research Center and Dartmouth Medical School, Rivermill Complex, 85 Mechanic St., Lebanon, NH 03766, USA
[2]Douglas Mental Health University Institute and McGill University, Montreal, Canada

This paper summarizes the findings for North America of the WPA Task Force on Steps, Obstacles and Mistakes to Avoid in the Implementation of Community Mental Health Care. Community mental health has evolved over five decades in the United States and Canada. The United States has led the world in innovation and spending, but provide variable quality of care; Canada has steadily developed a more uniform public health system for less cost. Lessons learned from North America include: team-based approaches and other evidence-based practices, when implemented with high fidelity, can improve outcomes in routine mental health care settings; recovery ideology and peer support enhance care, though they have not been studied rigorously; effective community-based care for people with serious mental disorders is expensive.

Key words: Community mental health care, United States, Canada, evidence-based practices, recovery, team-based care, psychiatric rehabilitation

(World Psychiatry 2012;11:47-51)

This paper is part of a series describing the development of community mental health care in regions around the world (see 1,2), produced by a Task Force appointed by the WPA as part of its Action Plan 2008-2011 (3,4). The WPA Guidance on Steps, Obstacles and Mistakes to Avoid in the Implementation of Community Mental Health Care, developed by this Task Force, has been previously published in the journal (5). In this article, we describe these issues in relation to North America.

The paper reviews the evolution of community mental health care in the United States and Canada and highlights several principles based on the experiences of these two close neighbors. In spite of their geographical proximity and common language (the province of Québec excepted), the United States and Canada have distinct cultures and significantly different health care systems. These differences extend to the organization of community mental health care. Accordingly, we describe community mental health care separately for each country. The discussion of lessons learned brings together observations drawn from the experiences of both countries.

THE UNITED STATES COMMUNITY MENTAL HEALTH SYSTEM

The United States has numerous mental health systems, independent agencies, and single providers. A mixture of primary care providers and private practice therapists delivers most of the mental health care to people with non-severe mental disorders. For those with severe disorders, such as schizophrenia, bipolar disorder, and chronic depression, each of the 50 states oversees a public mental health program. Larger states often devolve responsibility to county or city authorities, so that there are actually multiple mental health care authorities and systems within many states. The only administrative commonality across states is funding by the federal Medicaid and Medicare programs for those who are impoverished, disabled, or aging. Medicaid is, however, administered differently in each state according to a variety of rules, regulations, and waivers. In addition to state public mental health programs, the federal government runs separate health care systems for active members of the military, for retired and disabled members of the military, and for Native Americans. For people with substance use disorders, the system is somewhat simpler, because the private sector is relatively small and the federal government supports most of the public care, with states contributing different amounts.

The organization of these many systems, programs, and providers is typically based on funding and profits rather than on public health needs, the preferences of users of the mental health services, or research. Few of the systems, programs, or individual practitioners collect data on quality or outcomes. Because so many providers, programs, and intermediaries (e.g., insurance agencies and managed care organizations) participate in the context of a largely private, for-profit system, health care costs are very large. The United States health care system expenditures were approximately 2.5 trillion dollars in 2009 (6). Of this total, 5-12% has been devoted to behavioral health care in recent years (7). The United States also spends much more than any other country on medical research. One advantage of the tremendous variation of programs and expenditures across regions has been the opportunity for innovation and research.

History

The community mental health movement began in the United States in 1963, when President John Kennedy signed the Community Mental Health Act and community mental health centers arose in towns and cities throughout the country (8). Initially, these centers assumed too broad an agenda, including all mental health problems and prevention as well as treatment. By the 1970s, community mental health programs narrowed their goals to treatment of persons with long-term and disabling illnesses and facilitated deinstitutionalization of this population. Many long-term patients were actually transferred to group homes, nursing homes, and other institutions in the community, but the deinstitutionalization philosophy did result in significant downsizing of large state hospitals and of the total hospitalized population. The population in large public mental hospitals dropped from over 500,000 to less than 150,000 (8).

During the 1980s and 1990s, two movements strongly influenced community mental health care in the United States. The evidence-based practice movement arose from effectiveness research and evidence-based medicine, and, somewhat later, the recovery movement arose from the experiences of users of the mental health system.

Models of care and evidence-based practices

The initial plan for community care, developed by the National Institute of Mental Health and termed the Community Support Program, centrally featured professional case managers, who would coordinate and broker all of the services for people with severe and persistent mental disorders in the community (9). In the 1970s and 1980s, many challenges of caring for people in the community began to be apparent. Common concerns included integration and continuity of services for those with the most complex needs, appropriate housing, family burden, substance abuse and dependence, victimization, and violence (9). More recently, unemployment, criminalization, and early mortality of people with mental illnesses have emerged as major concerns. All of these problems were exacerbated by poverty, reductions in housing subsidies, and shunting of people with mental illnesses into inner-city areas plagued by unemployment, crime, and drugs.

Many models of care were developed to address the special problems of people with severe mental disorders living in the community (10). For integration and continuity of care, assertive community treatment, intensive case management, clinical case management, and other models appeared. To address the need for housing, foster care, Fairweather Lodge, residential continuum, and supportive and supported housing models emerged. Likewise, other concerns were addressed by a variety of family interventions, treatments for co-occurring disorders, and so on. Research has supported some of these models and not others. Research-based models of care became identified as evidence-based practices. Various government reviews (11,12) and systematic reviews (13-15) have identified specific interventions as evidence-based practices.

An additional concern has been the general failure to implement effective services in routine mental health treatment settings (16). In 1997, the Robert Wood Johnson Foundation, the Substance Abuse and Mental Services Administration, several State Departments of Mental Health, and additional private foundations initiated a national demonstration to implement six specific evidence-based practices that were deemed essential community mental health services: systematic medication management, assertive community treatment, supported employment, family psychoeducation, illness management and recovery, and integrated treatment for co-occurring disorders (17). Because of research showing that faithfulness to evidence-based practices was strongly related to outcomes, the project emphasized implementation and fidelity. Outcomes showed that, with training and supervision for one year, most programs were able to implement and sustain high-quality evidence-based practices (18,19). Nevertheless, the degree of implementation of these practices varies widely from state to state (20).

Recovery

Approximately half of all people with severe and persistent mental illnesses in the United States have received no mental health services in the past year, often because they have rejected the available services (21). Many others who have received mental health services have expressed dissatisfaction with the services. Users of the mental health system (variously called patients, clients, users, consumers, or survivors) have lodged strong objections to the existing mental health system. They have also argued that professionals' goal of stabilization does not correspond to their aspirations for "recovery" (22), a concept defined by each individual, but which typically encompasses opportunities for education, work, friendship, independent living, and community participation (11). They also argued for meaningful roles in making decisions and in delivering mental health care, and for the elimination of coercion in the contexts of hospitalization, prescribing of medications, and outpatient treatment.

The recovery movement has influenced numerous changes in community mental health care. Many states embrace recovery at the level of philosophy and mission, even if they have varying levels of success implementing its tenets. Rehabilitative services are more widely available, and many mental health programs have decreased the use of coercive measures, such as seclusion and restraint. The impact of the recovery movement has been much greater in some states than others (20).

Recent developments

In the 2000s, community mental health care in the United States was dominated by attempts to control costs. These included managed care, fee-for-service systems, and Medicaid audits, which resulted in the government demanding that millions of dollars be returned. The financial recession severely affected state budgets and led to numerous cycles of financial cuts. The 15% of citizens without insurance (higher for those with mental illnesses) had great difficulty accessing even minimal care (23). The net result has been a dramatic deterioration of community mental health care for people with the most severe disorders (11,20,24).

Very recently, parity legislation and health care reform legislation offer hope that people with mental illnesses in the United States will more easily acquire insurance and that mental health disorders will be treated in the same manner as physical health disorders. How these two pieces of legislation are enacted over the next decade remains to be seen.

THE CANADIAN COMMUNITY MENTAL HEALTH SYSTEM

Analogously to the situation in the United States, each of Canada's ten provinces and three (Northern) territories has its own mental health care system: health care for the great majority of the population falls under provincial and territorial jurisdiction. Nonetheless, several factors, including various institutional features common across all provinces, common proximity to the United States, which has had a major influence on service development in Canada (25,26), federally-managed equalization payments from richer to poorer provinces, and various mechanisms of exchange of information across provinces, have resulted in provincial mental health care systems bearing fairly close resemblance to each other. Partly due to equalization payments, perhaps partly also due to greater homogeneity in outlook concerning the resources needing to be allocated to the care of people with mental illness, per capita levels of funding for mental health are more similar to each other across Canadian provinces than they are across the United States. Health care spending per capita in Canada is about half what it is in real terms in the United States. A recent report estimated that in 2007/2008 total behavioral health spending in Canada was 7.2% of total health care spending, with per capita spending for inpatient mental health care, physician mental health services, and pharmaceuticals together varying from $156 in Saskatchewan to $240 in New Brunswick (27).

Common features of the Canadian mental health system include: a) a mix of institutionally-based services delivered by unionized professionals and less regulated, non-unionized voluntary sector providers; b) universal coverage of hospital and physician services, as well as those of voluntary sector providers – but no public coverage of psychologists practicing independently; c) public coverage of medications for seniors and social assistance recipients, with some provinces providing varying levels of coverage for other people not covered by supplementary employer-based private insurance; d) physicians, including hospital-based psychiatrists, who are paid directly by provincial governments, mostly on a fee-for-service basis; e) now with the notable exception of Alberta, regionalization of care delivery, resulting in further differentiation of community mental health services within individual provinces.

As in the United States, deinstitutionalization began in the late 1950s or early 1960s and was followed by the development of psychiatry departments within general hospitals. Voluntary sector providers quickly emerged to offer community-based care to the growing number of people with severe mental illness living in the community. Psychiatry departments and psychiatric hospitals gradually followed, adding community mental health services to their programming. This process has been slow, following the evidence base with a lag measured in decades, and even today a number of psychiatry departments have hardly embarked on it. Access to evidence-based practices remains, as in the United States, very limited.

As in the United States, there has been much discussion of recovery. Peer-support workers have become more commonly embedded into clinical services, and the notion that people with lived experience of mental illness should participate in administrative decisions and in research projects that have implications for them has gradually gained ground.

Some long-standing features of Canadian mental health care systems impede the development of high-quality community mental health care. Psychiatrists, who necessarily play a key role in community mental health care delivery, are paid directly by provincial governments independently of the quality of care they provide and have limited accountability. Union rules often seem designed to protect the privileges of members, especially those with more seniority, rather than serve the needs of clients. Funding for medications is open-ended while funding for psychosocial services is severely constrained.

In 2006, after extensive consultation, the Standing Senate Committee on Social Affairs, Science and Technology tabled an influential report that painted a grim picture of the state of mental health and addiction services in Canada (28). Services have not been well integrated and it is difficult, indeed often impossible, for people with mental illness and their caregivers to navigate successfully through them. Affordable, decent supportive or supported housing is often unavailable, as are integrated services for people with concurrent (mental illness and substance use) disorders and employment services. Much more often than should be the case, what care is accessed is of questionable quality. Stigma is a common, disabling experience.

Following a key recommendation of the report, in 2007 the federal government established the Mental Health Commission of Canada, whose mandate is "to help bring into being an integrated mental health system that places people living with mental illness at its centre". Whatever its ultimate

impact on provincial and territorial mental health policies and services, the Commission has stimulated an unprecedented sharing of ideas and perspectives across a broad range of stakeholders throughout the country.

LESSONS LEARNED IN NORTH AMERICAN COMMUNITY MENTAL HEALTH FROM 1960 TO 2010

During five decades of developing community mental health programs in North America, several robust and durable concepts have emerged. These concepts contrast with many transient ideas that were never implemented broadly in real-world systems of care. The concepts correspond only loosely with research because research-based interventions sometimes fail to interest clients, providers, or payers. Some interventions take root and blossom over decades without much of a research base.

Team-based care

For people who have the most severe illnesses, the greatest disabilities, and the fewest family or community supports, integration of treatment, rehabilitation and support services needs to be achieved at the clinical level (29). Team-based care is the most straightforward way to insure access, continuity, and integration of services. Teams make it possible to offer clients an individualized, coherent, and long-term program of medical, psychiatric, housing, financial, vocational, family, and social services to help them in achieving their own goals. This insight emerged more than 30 years ago during early development of the assertive community treatment model and remains valid today.

Recovery

The recovery movement has spawned several important implications for the service delivery system: a) it calls for widespread participation of peer support workers within all types of services, and indeed, in the administration of and research on those services; b) it emphasizes choice and self-determination, so that the focus of services becomes helping clients achieve their own goals, as far as possible in their own way; c) it demands the replacement of unnecessarily coercive practices in favor of more clinically-skilled, creative interactions within the context of mutually respectful, collaborative partnerships; and d) it also calls for funding consumer-run programs of various types that may have a weaker evidence base but are favored by many people with mental illnesses.

Psychiatric rehabilitation and evidence-based practices

With its attention to the individual's goals, values, and preferences, the recovery movement has been consonant with parallel developments in the mental health field called psychiatric rehabilitation. One seminal paper described recovery as "the lived experience of rehabilitation" (22). Within the rehabilitation movement, values (people first) and goals (successful adjustment in one's psychosocial areas of preference) have been consistent, methods and outcomes measurement have evolved, and evidence-based practices have emerged (10). Earlier stepwise approaches to rehabilitation based on lengthy psychotherapy and training programs have gradually been replaced by more modern approaches that involve helping clients to reach their goals rapidly by providing highly individualized supports. Exemplars of the current approaches are supported housing, supported employment, supported education, and strengths case management.

Peer support

Social and instrumental support among peers with similar conditions has a long and robust history in North America. Alcoholics Anonymous, for example, has grown steadily since its development in the 1940s, and millions of people with alcoholism now attend Alcoholics Anonymous meetings in cities, suburbs, and towns throughout the region. Peer support programs for people with severe and persistent mental disorders have proliferated for several decades (30). Like Alcoholics Anonymous, they have been supported by the strong endorsement of people with these disorders themselves rather than by randomized controlled trials. Some 10% to 20% of people with severe mental disorders find these services helpful as a substitute or an adjunct to community mental health services. The forms of peer services continue to evolve, but the concept has been enduring.

Economic consequences of neglect

In any mental health service system, a minority of individuals with mental illness consume a disproportionate share of expensive resources such as inpatient care, emergency room visits, incarcerations, and so on. Providing comprehensive, team-based care to these individuals is expensive, but it mitigates personal suffering, is no more costly, and avoids shifting the costs to families, communities, and the criminal justice system (31-33).

Implementation

As evidence-based practices have been defined with increasing precision and model fidelity, the inadequacy of brief training of front-line staff and team leaders has become clear. High fidelity implementation requires ongoing training and supervision for six months to a year, and maintaining high quality services also requires ongoing supervision (34-36).

CONCLUSIONS

The story of the development of community-based mental health services in the United States has been one of both hope and disappointment. It has been one of hope in that many of the most important innovations in the field have arisen there, from the most long-standing, extensive and influential program of mental health services research in the world, as well as from a plethora of consumer initiatives; and also in that a few states have been trailblazers in showing how these advances can be turned into effective programs that actually help people with mental illnesses to recover. It has also been one of disappointment in that the political will and the funding levels needed to ensure development of these programs across the country have more often than not been lacking. Health care reform initiated by President Obama may change these trends.

Canada has imported many ideas from its larger neighbor to the south. While funding levels are more uniform than in the United States, and may be about adequate for good quality care to be delivered throughout, various institutional arrangements, combined with lack of political will, have impeded the integration of state-of-the-art practices into routine care. Recent recognition of these failings by the federal government and the establishment of the Mental Health Commission of Canada may now lead to a period of accelerated progress.

References

1. Hanlon C, Wondimagegn D, Alem A. Lessons learned in developing community mental health care in Africa. World Psychiatry 2010;9: 185-9.
2. Semrau M, Barley E, Law A et al. Lessons learned in developing community mental health care in Europe. World Psychiatry 2011;10:217-25.
3. Maj M. The WPA Action Plan 2008-2011. World Psychiatry 2008;7: 129-30.
4. Maj M. Report on the implementation of the WPA Action Plan 2008-2011. World Psychiatry 2011;10:161-4.
5. Thornicroft G, Alem A, Dos Santos RA et al. WPA guidance on steps, obstacles and mistakes to avoid in the implementation of community mental health care. World Psychiatry 2010;9:67-77.
6. Centers for Medicare and Medicaid Services. National health care expenditures data. Baltimore: Centers for Medicare and Medicaid Services, 2010.
7. Frank R, Goldman H, McGuire T. Trends in mental health cost growth: an expanded role for management? Health Affairs 2009;28: 649-59.
8. Dixon L, Goldman H. Forty years of progress in community mental health: the role of evidence-based practices. Aust N Z J Psychiatry 2003:37:668-73.
9. Turner J, TenHoor W. The NIMH community support program: pilot approach to a needed social reform. Schizophr Bull 1978;4:319-48.
10. Corrigan P, Mueser K, Bond G et al. The principles and practice of psychiatric rehabilitation. New York: Guilford, 2008.
11. New Freedom Commission on Mental Health. Achieving the promise: transforming mental health care in America. Rockville: U.S. Department of Health and Human Services, 2003.
12. U.S. Department of Health and Human Services. Mental health: a report of the Surgeon General. Rockville: U.S. Department of Health and Human Services, 1999.
13. Chambless D, Ollendick T. Empirically supported psychological interventions: controversies and evidence. Annu Rev Psychol 2001; 52:685-716.
14. Cook J. Blazing new trails: using evidence-based practice and stakeholder consensus to enhance psychosocial rehabilitation services in Texas. Psychiatr Rehab J 2004;27:305-6.
15. Kreyenbuhl J, Buchanan R, Dickerson F et al. The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2009. Schizophr Bull 2010;36:94-103.
16. Torrey W, Drake R, Dixon L et al. Implementing evidence-based practices for persons with severe mental illnesses. Psychiatr Serv 2001;52:45-50.
17. Drake R, Goldman H, Leff H et al. Implementing evidence-based practices in routine mental health service settings. Psychiatr Serv 2001;52:179-82.
18. McHugo G, Drake R, Whitley R et al. Fidelity outcomes in the National Implementing Evidence-Based Practices Project. Psychiatr Serv 2007;58:1279-84.
19. Swain K, Whitley R, McHugo G et al. The sustainability of evidence-based practices in routine mental health agencies. Commun Ment Health J 2010;46:119-29.
20. National Alliance on Mental Illness. Grading the States: a report on America's health care system for serious mental illness. Arlington: National Alliance on Mental Illness, 2006.
21. Mojtabai R, Fochtmann L, Chang S et al. Unmet need for mental health care in schizophrenia: an overview of literature and new data from a first-admission study. Schizophr Bull 2009;35:679-95.
22. Deegan P. Recovery: the lived experience of rehabilitation. Psychosoc Rehab J 1988;11:11-9.
23. Cunningham P, McKenzie K, Taylor E. The struggle to provide community-based care to low-income people with serious mental illness. Health Aff 2006;25:694-705.
24. Glied R, Frank R. Better but not best: recent trends in the well-being of the mentally ill. Health Affairs 2009;28:637-48.
25. Goering P, Wasylenki D, Durbin J. Canada's mental health system. Int J Law Psychiatry 2000;23:345-59.
26. Rochefort D, Goering P. "More a link than division": how Canada has learned from U.S. mental health policy. Health Affairs 1998;17: 110-27.
27. Jacobs P, Dewa C, Lesage A et al. The cost of mental health and substance abuse in Canada. Alberta: Institute of Health Economics, 2010.
28. Standing Senate Committee on Social Affairs. Out of the shadows at last: transforming mental health, mental illness and addiction services in Canada. Ottawa: Government of Canada, 2006.
29. Lehman A, Kreyenbuhl J, Buchanan R et al. The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2003. Schizophr Bull 2004;30:193-217.
30. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehab J 2004;27: 392-401.
31. Culhane D. The costs of homelessness: a perspective from the United States. Eur J Homelessness 2008;2:97-114.
32. Larimer M, Malone D, Garner M et al. Health care and public service use and costs before and after provision of housing for chronically homeless persons with severe alcohol problems. JAMA 2009; 301:1349-57.
33. Latimer E. Economic impacts of assertive community treatment: a review of the literature. Can J Psychiatry 1999;44:443-54.
34. Drake R, Bond G, Essock S. Implementing evidence-based practices for people with schizophrenia. Schizophr Bull 2009;35:704-13.
35. Fixsen D, Naoom S, Blase K et al. Implementation research: a synthesis of the literature. Tampa: University of South Florida, 2005.
36. Rapp C, Etzel-Wise D, Marty D et al. Evidence-based practice implementation strategies: results of a qualitative study. Commun Ment Health J 2008;44:213-24.

Mental health services in the Arab world

Ahmed Okasha[1], Elie Karam[2], Tarek Okasha[1]

[1]Okasha Institute of Psychiatry, Faculty of Medicine, Ain Shams University, Cairo, Egypt
[2]St. George Hospital University Medical Center, Faculty of Medicine, Balamand University, Beirut, Lebanon

This paper summarizes the current situation of mental health services in the Arab world. Out of 20 countries for which information is available, six do not have a mental health legislation and two do not have a mental health policy. Three countries (Lebanon, Kuwait and Bahrain) had in 2007 more than 30 psychiatric beds per 100,000 population, while two (Sudan and Somalia) had less than 5 per 100,000. The highest number of psychiatrists is found in Qatar, Bahrain and Kuwait, while seven countries (Iraq, Libya, Morocco, Somalia, Sudan, Syria and Yemen) have less than 0.5 psychiatrists for 100,000 population. The budget allowed for mental health as a percentage from the total health budget, in the few countries where information is available, is far below the range to promote mental health services. Some improvement has occurred in the last decade, but the mental health human resources and the attention devoted to mental health issues are still insufficient.

Key words: Arab world, mental health services, resources, primary care

(World Psychiatry 2012;11:52-54)

The Arab world is taken to mean the 22 members of the Arab League, accounting for 280 million people. The region has the largest proportion of young people in the world: 38% of Arabs are under 14. Life expectancy has increased by 15 years over the past three decades, and infant mortality has dropped by two-thirds. Around 12 million people, or 15% of the labor force, are unemployed. The quality of education has recently deteriorated, and there is a severe mismatch between the labor market and the education system. Adult illiteracy rates have declined, but are still very high: 65 million adults are illiterate, almost two-thirds of them women. Some 10 million children still have no schooling at all.

The health expenditure estimated as a percentage of gross domestic product is highest in Palestine (13.5%), followed by Lebanon (8.8%), Jordan and Djibouti (8.5%) and Egypt (6.4%) (1). Health services in all Arab countries are provided by public (government) and private sector facilities and out of pocket (this last category representing 63.4% of the total in Sudan, 58.7% in Egypt, 58% in Yemen, 56.1% in Morocco and 54.9% in Syria). In some countries insurance systems contribute to the provision of the service. Non-governmental organizations (NGOs) have come to be recognized as an important actor in the provision of health services, especially in countries with internal instability (in particular, Lebanon in late 1980s and Palestine now).

The mental health expenditure as a percentage of total health expenditure is not available in most Arab countries and not reported by the officials. Only three Arab countries have provided an estimate: Qatar (1%), Egypt (less than 1%) and Palestine (2.5%).

There are no projections on the burden of mental disorders specific to the Arab world. Only two countries (Lebanon and Iraq) so far conducted national studies using comparable methodology, based on the World Health Organization (WHO) World Mental Health Surveys (2,3). Two other studies were conducted in Morocco (4) and Egypt (5) using different methodologies. The lifetime prevalence of any anxiety disorder among adults was 16.7% in the Lebanese study and 13.8% in the Iraqi survey; that of any mood disorder was, respectively, 12.6% and 7.5%. The study carried out in Morocco reported a point prevalence of 9.3% for generalized anxiety disorder and 26.5% for major depressive disorder, while the Egyptian study reported a point prevalence of 4.8% for anxiety disorders and 6.4% for mood disorders.

Table 1 shows the data concerning the availability of mental health policies in the various Arab countries, obtained through ministries of health, the Eastern Mediterranean Region (EMRO) office of the WHO, national psychiatric societies and national psychiatric leaders. Six out of 20 countries do not have a mental health legislation and two do not have a mental health policy. There is no information for Mauritania and Comoros.

As shown in Table 2, three countries (Lebanon, Kuwait and Bahrain) had in 2007 more than 30 psychiatric beds per 100,000 population, while two (Sudan and Somalia) had less than 5 per 100,000. A substantial reduction of psychiatric beds with respect to our 1998 survey (6) occurred in Iraq, Jordan, Kuwait, Libya, Oman, Qatar and Palestine.

The highest number of psychiatrists is found in Qatar, Bahrain and Kuwait, while seven countries (Iraq, Libya, Morocco, Somalia, Sudan, Syria and Yemen) have less than 0.5 psychiatrists per 100,000 population. Although there is a mental hospital in Djibouti, yet there are no psychiatrists, and general practitioners with special interest in mental health look after those patients (Table 2). The number of psychiatrists decreased with respect to the 1998 survey in Libya, Saudi Arabia and Sudan, while there was a substantial increase in several other countries.

Psychiatric nurses per 100,000 population range from 23 in Bahrain and 22.5 in Emirates to 0.09 in Yemen and 0.03 in Somalia. The number of nurses increased in almost all

Table 1 Mental health policies in Arab countries

Country	Mental health policy (year)	Substance abuse policy (year)	National mental health programme (year)	Mental health legislation (year)
Algeria	Yes (?)	Yes (1990)	Yes (2001)	Yes (1998)
Bahrain	Yes (1993)	Yes (1983)	Yes (1989)	Yes (1975)
Djibouti	No	No	No	An old French legislation
Egypt	Yes (1978)	Yes (1986)	Yes (1986)	Yes (2009)
Emirates	Yes (?)	Yes (?)	Yes (1991)	Yes (1981)
Iraq	Yes (1981)	Yes (1965)	Yes (1987)	Yes (1981)
Jordan	Yes (1986)	Yes (2000)	Yes (1994)	Yes (2003)
Kuwait	Yes (1957)	Yes (1983)	Yes (1997)	No
Lebanon	No	No	Yes (1987)	No
Libya	Yes (?)	No	Yes (1988)	Yes (1975)
Morocco	Yes (1972)	Yes (1972)	Yes (1973)	Yes (1998)
Oman	Yes (1992)	Yes (1999)	Yes (1990)	Yes (1999)
Palestine	Yes (2004)	Yes (2004)	Yes (2004)	Yes (2004)
Qatar	Yes (1980)	Yes (1986)	Yes (1990)	No
Saudi Arabia	Yes (1989)	Yes (2000)	Yes (1989)	No
Somalia	Yes (?)	Yes (?)	Yes (?)	No
Sudan	Yes (1998)	Yes (1995)	Yes (1998)	Yes (1998)
Syria	Yes (2001)	Yes (1993)	Yes (2001)	Yes (1965)
Tunisia	Yes (1986)	Yes (1969)	Yes (1990)	Yes (2003)
Yemen	Yes (1986)	No	Yes (1983)	No

Table 2 Mental health resources in Arab countries

Country	Psychiatric beds per 100,000		Psychiatrists per 100,000		Psychiatric nurses per 100,000		Psychologists per 100,000		Social workers per 100,000	
	1998	2007	1998	2007	1998	2007	1998	2007	1998	2007
Algeria	14	25	1.1	2.2	1.1	4.2	0.8	0.2	0	0.4
Bahrain	33.8	33	3.7	5	13.3	23	0.5	0.8	1	1.5
Djibouti	N.A.	7	0	0	0	0.2	0	0	0	0
Egypt	12.5	13	0.9	0.9	2	2	0.3	0.4	0.09	0.1
Emirates	N.A.	14	0.9	2	N.A.	11	0.9	1	0.6	1.2
Iraq	7	6.3	0.1	0.7	0.1	0.1	N.A.	0.05	0.05	0.2
Jordan	20	15.7	1.1	1	0	2	0.2	0.6	0.5	2
Kuwait	47	34	2.6	3.1	16.2	22.5	0.9	1.4	0.4	0.4
Lebanon	47	75	1.2	2	0.9	5.03	1.9	0.6	0.6	1.5
Libya	56	10	0.3	0.2	N.A.	0.05	0.3	5	0.2	1.5
Morocco	7.6	7.8	N.A.	0.4	N.A.	2.02	N.A.	0.03	N.A.	0.007
Oman	5.5	4.9	0.2	1.4	0.2	5	0	0.2	0.1	0.5
Palestine	14.2	8.8	0.8	0.9	3.2	3.4	1.7	1	0.7	1.1
Qatar	37.9	9.7	0.8	3.4	7.4	10	1.4	1.2	1.7	10
Saudi Arabia	6.5	11.8	2.4	1.1	6.3	6.4	0.5	1	0.9	2.4
Somalia	N.A.	4	0.5	0.06	0.03	0.03	0	0	0	0.2
Sudan	0.1	2	0.2	0.09	N.A.	0.2	0.01	0.2	0.01	0.1
Syria	7.8	8	N.A.	0.5	N.A.	0.5	N.A.	0	N.A.	0
Tunisia	9.6	11.3	0.8	1.6	3.3	0.2	0.1	0.6	0	N.A.
Yemen	N.A.	18.5	0.1	0.5	N.A.	0.09	3.2	1.2	0.01	0.04

N.A. – not available

countries compared to the 1998 survey. The same applies to psychologists and social workers, with the most substantial increase observed in Bahrain, Emirates, Jordan, Egypt, Kuwait, Libya, Saudi Arabia and Yemen (Table 2).

Recent years have seen significant changes in the field of mental health in the countries of the Arab Region. Psychiatric services, which were earlier totally confined to a few large mental hospitals, are now gradually being replaced by psychiatric units with both inpatient and outpatient facilities in general hospitals. In some countries, the process of decentralization has been taken still further, and psychiatric services are being provided at district hospitals and smaller peripheral units, along with other general health services. Training programmes in mental health for general practitioners, non-physicians and health personnel working at primary health care level have started in a large number of countries as a part of in-service skills enhancement programmes (7).

Although a majority of the countries of the region have agreed in principle to integrate mental health into the pri-

mary health care delivery system, implementation so far has been limited. Globally, the mental health infrastructure and services in most countries is grossly insufficient for the large and growing needs.

Currently, most of the Arab countries are exposed to conflicts, wars, terrorism and fundamentalism, which may be the seeds for many behavioral and mental disorders.

Cultural beliefs of possessions and the impact of sorcery or the evil eye affect interpretation of mental symptoms. In this context, the first resort for the families of mental patients is not even the general practitioner, but the traditional healers, who acquire a special importance because of their claim of dealing with the "mystical" and the "unknown". In the majority of Arab countries there is no interaction between the medical profession and the traditional healers. In Jordan, there is some kind of a relationship, which remains informal and unorganized. In Saudi Arabia, however, they constitute part of the staff, using religious text and recitation in management.

In conclusion, our data show that, in the Arab world, health and education budget assignment is below the recommended requirements far better quality of life. The budget allowed for mental health as a percentage from the total health budget, in the few countries where information is available, is far below the range to promote mental health services. The mental health human resources and the inefficient data collection by the official agencies are incompatible with the gross domestic product of Arab countries. An appeal for implementing mental health in primary care as stipulated as a policy in many Arab countries and to prioritize mental health in the agenda of politicians is urgently needed.

References

1. World Health Organization. The work of WHO in the Eastern Mediterranean Region. Annual report of the regional director. Cairo: WHO Regional Office for the Eastern Mediterranean, 2009.
2. Karam EG, Mneimneh ZN, Karam AN et al. 12 month prevalence and treatment of mental disorders in Lebanon. A national epidemiologic survey. Lancet 2006;367:1000-6.
3. Alhasnawi A, Sadik S, Rasheed M et al. The prevalence and correlates of DSM-IV disorders in the Iraq Mental Health Survey (IMHS). World Psychiatry 2009;8:97-109.
4. Kadri N, Agoub M, Assouab F et al. Moroccan national study on prevalence of mental disorders: a community-based epidemiological study. Acta Psychiatr Scand 2010;121:71-4.
5. Ghanem M, Gadallah M, Meky FA et al. National survey of prevalence of mental disorders in Egypt: preliminary survey. East Mediterr Health J 2009;15:65-75.
6. Okasha A, Karam E. Mental health services and research in the Arab world. Acta Psychiatr Scand 1998;98:406-13.
7. World Health Organization. Mental health in the Eastern Mediterranean Region. Reaching the unreached. Cairo: WHO Regional office of the Eastern Mediterranean, 2006.

The crisis of psychiatry – insights and prospects from evolutionary theory

Martin Brüne[1], Jay Belsky[2], Horacio Fabrega[3], Jay R. Feierman[4], Paul Gilbert[5], Kalman Glantz[6], Joseph Polimeni[7], John S. Price[8], Julio Sanjuan[9], Roger Sullivan[10], Alfonso Troisi[11], Daniel R. Wilson[12]

[1]Department of Psychiatry, Psychotherapy and Preventive Medicine, University of Bochum, LWL University Hospital, Alexandrinenstrasse 1, D-44791 Bochum, Germany; [2]University of California, Davis, CA, USA and Birkbeck University of London, UK; [3]University of Pittsburgh, PA, USA; [4]University of New Mexico, NM, USA; [5]University of Derby, UK; [6]Private Practice, Boston, MA, USA; [7]University of Manitoba, Winnipeg, Manitoba, Canada; [8]Oxford Regional Health Authority, UK; [9]University of Valencia, Spain; [10]California State University, CA, USA; [11]Tor Vergata University, Rome, Italy; [12]Creighton University, Ohama, NE, USA

Darwin's emphasis on natural selection has had a transformative influence on how biological and medical sciences are conceptualized and conducted. However, the relevance of his ideas for the understanding of psychiatric conditions is still under-appreciated. Modern understanding of disease has required appreciation of the dialectical give and take between environmental influences, life history theory imperatives, human behavioral ecology, and characteristics of adaptive processes at all levels of the individual. This has enabled a better comprehension of metabolic disturbances, cancers, auto-immune disease, inherited anemias, and vulnerability to infectious disease (1). Here we propose that a contemporary and scientifically satisfying understanding of psychiatric conditions requires adopting a similar logic of inquiry, by taking into consideration the influence of environmental contingencies and natural selection in sculpting not just brain based mechanisms and processes germane to clinical neurosciences, but also diverse characteristics of behavior.

One approach to understand psychiatric disorders in an evolutionary perspective builds upon Nobel laureate Nikolaas Tinbergen's ideas, suggesting that, for a full understanding of any given phenotypic trait, one needs to detect the development and nature of its mechanisms, construed as the "proximate causes", and, in addition, its evolutionary (or phylogenetic) history and adaptive value (2). Studying the proximate mechanisms is standard in psychiatry and the clinical neurosciences, but the questions pertaining to the phylogeny of traits have largely been ignored.

Admittedly, placing dysfunctional cognitive, emotional and behavioral processes in the context of possible adaptation is not straightforward at first sight. The clinical directive requires that "disorder" represent the appropriate focus. However, a "disorder" – by definition – is counter-intuitive in the context of adaptation. By *adaptation* we mean a genetically-mediated structural or behavioral trait, which when possessed, increased survival and reproductive success in the environment in which the trait evolved. Were psychiatry's focus be placed on "traits" (i.e., cognitive processes, emotions, and behaviors), problems which are clinically relevant could more satisfactorily be understood as distorted expression of mechanisms that in earlier environments provided answers to problems of adaptive significance, but which currently interfere in light of prevailing environmental contingencies (3).

Important to the understanding of a particular phenotype is the evolutionary concept of variation. Without variation, no evolution by natural selection could take place. Mainstream psychiatry has largely ignored the fact that variation is the rule, not the exception, and this creates conceptual tensions. Psychiatry conceptualizes "disorder" as a statistical deviation from a normative statistical mean, yet handles it as a category. In other words, both "normalcy" as well as "disorder" with regard to psychological or behavioral functioning are burdened with the connotation of low variation.

Phenotypic variation is the result of a complex interplay of genotype and environment, including epigenetic mechanisms that are decisively shaped by experience over the individual lifespan. These issues translate to providing a clinician with a rationale for explaining why, how, and when adaptive behavior is compromised and constrained; that is, when social, cultural, or ecological conditions and circumstances pose hindrances or risks which interfere with achievement of best solutions to socio-biological problems, and which may require a modification of a strategy of coping, selection of an alternative strategy, and/or the setting of more realistic biological goals. This integrative view of psychopathology, we believe, can have profound effects on how psychiatry conceptualizes disorders, which shall be illustrated briefly in three examples.

GENETICS

One presumption of how to explain the nature and causes of psychiatric conditions pertains to the idea that individuals carry variations of genes that make them vulnerable to develop a disorder, commonly referred to as the "diathesis-stress-model". Evolutionarily informed research into the genetics of psychiatric disorders now demonstrates that while such alleles can predispose to developing a psychiatric condition under adverse environmental conditions such as childhood maltreatment, they can also protect, and in fact can allow enhanced coping upon encountering favorable environmental conditions during early stages of development. For example, the "short" allele of the serotonin transporter coding gene is associated with greater risk for depression if linked with early childhood adversities, yet the same version of the

gene is associated with *reduced* risk for depression if carriers grow up in emotionally secure conditions (5). This suggests that selection favored plasticity or "open programs" (4) that render individuals more susceptible to environmental contingencies – for better *and* worse (6).

Similarly, psychiatrists guided by evolutionary theory have recognized that antagonistic pleiotropy may play a role in psychiatric disorders – genes that convey fitness advantages in one domain, while having potentially maladaptive value in another domain, a concept that was originally put forth with regard to senescence (7,8). Nowadays, examples for antagonistic pleiotropy can be pinned down to even single genes such as the catecholamine-O-methyltransferase coding gene, of which one particular allele is associated with poorer working memory performance but superior empathy (9).

Taken together, these insights offer an answer to the question of why natural selection designed bodies that are – under specific circumstances – vulnerable to disease (10). In addition, speaking of genetic "vulnerability" in one-sided ways that are common in psychiatry seems to be incomplete if not simplistic, and requires reformulation considering complex gene-environment interactions, and trade-offs between different functional aspects.

EXPRESSIONS OF EMOTIONS

Contemporary psychiatry has minimized the functional significance of non-verbally expressed emotions (11). This is an unfortunate development, because it makes psychiatry a "science" relying largely on subjective self-report and clinician-generated rating scales. What is overlooked is that the biology of social interaction is based on facial expressions, gesture and body language, complemented by verbal language. However, it has repeatedly been shown that not only can psychiatric patients reliably be distinguished from non-clinical individuals on the basis of their non-verbal behavior. In addition, the study of non-verbal behavior can be more informative in terms of response to treatment and relapse compared to standard psychopathological scores (12). Changing patterns of behavior, e.g., a reduction in frequency of defensive body positions, can be linked to clinical improvement, even before the patient (or clinician) becomes subjectively aware of it. Conversely, an increase of "displacement activities" related to motivational conflict can alert clinicians to monitor for clinical deterioration, because such patterns may be indicative of impending suicidal behavior. These examples of behavioral analyses based on ethological methodology explicitly assume that behaviors found in clinical conditions are not qualitatively distinct from behaviors in healthy individuals but different by degree, i.e. intensity, frequency or contextual inappropriateness (13).

PSYCHOTHERAPY

Environmental conditions include the behavioral ecology in which human cognition, emotions, and behavior developed, and the adaptive nature of psychological mechanisms that evolved to solve recurring biosocial problems such as eliciting from and providing care to others of a relevant group, forming cooperative alliances, finding a mate, and attaining an acceptable rank in the social hierarchy. An inability to achieve relevant biosocial goals is at the core of many psychiatric conditions. For example, depression-like behaviors have been likened to a de-escalating strategy to avoid ongoing conflict (14). In many if not all psychiatric disorders, alternative psychological mechanisms play a prominent role in shaping the actual manifestations or phenotype, which often include defenses against perceived threat, such as in social anxiety (disorder), obsessive-compulsive rituals, or paranoid ideation (15).

Accordingly, therapy ought to help patients understand the bio-ecological bases inherent and communicated through their symptoms and provide motivations for giving up unprofitable behavioral strategies or defenses. For example, a recently developed method termed "compassion focused therapy" (CFT) draws upon attachment theory (the first evolutionarily-grounded theory of psychopathology and therapy) and other sources (16,17). CFT aims to provide patients with healing environments which promote feelings of warmth, understanding, and kindness toward themselves and others in light of and despite burdens imposed by evolutionarily based motivations and emotions.

An evolutionarily informed psychiatry also proposes that psychotherapy needs to be individually tailored as regards sex, age, and environmental differences, which shape psychosocial goals, needs, and behavior (18). Moreover, insights from gene-environment interaction in phenotypic development open the promising perspective that behavioral plasticity can be used constructively in the therapeutic process to reduce and avoid suffering and emotional pain by encouraging patients to use their potential for change and enlightening them about the evolutionary significance of behaviors and symptoms.

CONCLUSIONS

The search for a coherent and comprehensive scientific understanding of psychiatric disorders has long been ignored by "mainstream" psychiatry. Even "biological" psychiatry has long failed to take into account those aspects of human experience and behavior that have been formed during the ancestral past of *Homo sapiens*. Instead, theory and practice of psychiatry has developed in response to human health problems tied to a comparatively recent segment of human history. Political, economic, ecological, scientific, and cultural contingencies prevailing in modern Anglo-European societies had the effect of directing inquiry to population health problems towards an emphasis on mental phenomena.

Here it is proposed that a Darwinian approach may advance the endeavor to formulate optimal ways of conceptualizing and explaining psychopathology. It necessitates rigorous analyses of how environments have and continue to shape and constrain adaptive behavior, producing different varieties of signs, symptoms, and responses. The latter represent the data on which contemporary clinical sciences have built their disciplines in conformance to a process that has been repeated throughout ancestral and recorded human history (19).

Building upon life-history theory, behavioral ecology, ethology (not to be confused with ethnology), developmental psychology, and evolutionary genetics, ideas germane to evolutionary theory enable formulation of testable predictions about the causation, unfolding ("natural history") and significance of psychiatric disorders. For example, it has recently been shown that maternal-neonate separation has tremendous effects on the autonomic activity and sleep quality of newborns compared to mother-neonate co-sleeping (20), which in turn may have profound impact on one's ability to cope with stress (21) and interpersonal orientation in terms of attachment (22). This is exactly the way gene-environment interaction should be studied in light of evolutionary constraints on human behavior.

Likewise, in view of current controversies about how to conceptualize and categorize psychiatric disorders (23), which is currently occurring as the DSM-5 and ICD-11 are taking shape, it is likely that psychiatric nosology will need a reshuffling of categories. We suggest that it is worth considering a reclassification of disorders according to the evolutionary significance of behavior that is expressed in malfunctioning ways, given conditions germane to modern environments compared to ancestral ones. Several conceptualizations have been published in the recent past, including the "harmful dysfunction analysis" (24) and an "evolutionary taxonomy of treatable conditions" (25), but none of them satisfactorily addresses the problem of reductionism (26). Accordingly, historical aspects of psychiatric nosology and findings from neuroscience have been proven difficult to reconcile, and similar obstacles will arise for any attempt to develop a psychopathological system based on insights from evolutionary theory (27). In any event, if such a prospect shall be successful at all, it would need to involve analyses by researchers with expertise in evolutionary social sciences.

It seems that the old claim by one of the founding fathers of the "new synthesis", Theodosius Dobzhansky, "nothing in biology makes sense, except in the light of evolution", is obviously true for psychiatric neuroscience, if not medicine and the life sciences in general. It is time not just to rethink but to implement such an integrative approach in research, clinical practice and medical education.

References

1. Gluckman P, Beedle A, Hanson M. Principles of evolutionary medicine. Oxford: Oxford University Press, 2009.
2. Tinbergen N. On aims and methods of ethology. Zeitschr Tierpsychol 1963;20:410-33.
3. Fabrega H Jr. Phylogenetic and cultural basis of mental illness. New Brunswick: Rutgers University Press, 2002.
4. Mayr E. What evolution is. New York: Basic Books, 2001.
5. Belsky J, Pluess M. Beyond diathesis stress: differential susceptibility to environmental influences. Psychol Bull 2009;135:885-908.
6. Belsky J, Jonassaint C, Pluess M et al. Vulnerability genes or plasticity genes? Mol Psychiatry 2009;14:746-54.
7. Williams GC. Pleiotropy, natural selection and the evolution of senescence. Evolution 1957;11:391-411.
8. Nesse RM. Life table tests of evolutionary theories of senescence. Exp Gerontol 1988;23:445-53.
9. Heinz A, Smolka MN. The effects of catechol-O-methyltransferase genotype on brain activation elicited by affective stimuli and cognitive tasks. Rev Neurosci 2006;17:359-67.
10. Nesse RM, Williams GC. Why we get sick. The new science of Darwinian medicine. New York: Times Books, 1994.
11. Feierman JR. The ethology of psychiatric populations II: Darwinian neuropsychiatry. Clin Neuropsychiatry 2006;3:87-106.
12. Troisi A. Ethological research in clinical psychiatry: the study of nonverbal behavior during interviews. Neurosci Biobehav Rev 1999;23:905-13.
13. Geerts E, Brüne M. Ethological approaches to psychiatric disorders: focus on depression and schizophrenia. Aust N Z J Psychiatry 2009;43:1007-15.
14. Price JS. Darwinian dynamics of depression. Aust N Z J Psychiatry 2009;43:1029-37.
15. Troisi A. The concept of alternative strategies and its relevance to psychiatry and clinical psychology. Neurosci Biobehav Rev 2005; 29:159-68.
16. Bowlby J. Attachment and loss. Vol. 1. Attachment. New York: Basic Books, 1969.
17. Gilbert P. Introducing compassion-focused therapy. Adv Psychiatr Treat 2009;15:199-208.
18. Glantz K, Pearce J. Exiles from Eden: psychotherapy from an evolutionary perspective. New York: Norton, 1989.
19. Fabrega H Jr. The need for an ethnomedical science. Science 1975; 189:969-75.
20. Morgan BE, Horn AR, Bergman NJ. Should neonates sleep alone? Biol Psychiatry (in press).
21. Porges SW. The polyvagal perspective. Biol Psychol 2007;74:116-43.
22. Rilling JK. A potential role for oxytocin in the intergenerational transmission of secure attachment. Neuropsychopharmacology 2009;34:2621-2.
23. Maj M. Psychiatric diagnosis: pros and cons of prototypes vs. operational criteria. World Psychiatry 2011;10:81-2.
24. Wakefield JC. Evolutionary versus prototype analyses of the concept of disorder. J Abnorm Psychol 1999;108:374-99.
25. Cosmides L, Tooby J. Toward an evolutionary taxonomy of treatable conditions. J Abnorm Psychol 1999;108:453-64.
26. Banzato CEM. Classification in psychiatry: the move towards ICD-11 and DSM-V. Curr Opin Psychiatry 2004;17:497-501.
27. Brüne M. Textbook of evolutionary psychiatry. The origins of psychopathology. Oxford: Oxford University Press, 2008.

Neurophysiology of a possible fundamental deficit in schizophrenia

Judith M. Ford[1,2], Veronica B. Perez[2], Daniel H. Mathalon[1,2]

[1]Psychiatry Service, San Francisco VA Medical Center; [1]Department of Psychiatry, University of California, San Francisco, CA 94121, USA

To survive, all animals must be able to distinguish between sensations resulting from their own actions and sensations coming from outside sources (1). To make this distinction, the sensory areas of the brain are sent an advance copy, or forward model, of the expected sensation resulting from action. This can be thought of as a carbon copy (Cc) from "self" to sensory cortex, saying "It's just me". And, so warned, sensory cortex dampens or even cancels the sensation. Importantly, it also notes the sensation as coming from "self". This forward model system has been labeled the "efference copy" and/or "corollary discharge" system. Although the terms are often used interchangeably, we define efference copy as a copy of a motor plan sent from motor to sensory cortical areas, and corollary discharge as the expected sensory consequences generated by the arrival of the efference copy.

In his seminal paper, Feinberg (2) proposed that auditory hallucinations are caused by a defect in the efference copy and corollary discharge systems of thoughts and ideas. He suggested that an abnormality in these systems could result in an inability to distinguish self-initiated neural activity from neural activity resulting from external stimulation. Thus, patients with schizophrenia might be unable to distinguish ideas or thoughts produced in their own minds from voices or influences coming from the environment. Feinberg noted that Hughlings Jackson, the 19th century British neurologist, considered thinking as the most complex of our motor acts, which "might conserve and utilize the computational and integrative mechanisms evolved for physical movement". Feinberg suggested that thinking might therefore retain successful control mechanisms present at lower levels of integration. In 1987, Frith (3) expanded this concept and prompted a series of behavioral experiments confirming corollary discharge dysfunction in schizophrenia. Evidence for dysfunction of the corollary discharge system in schizophrenia has been documented in auditory and somatosensory modalities (see 4 for a review).

About 10 years ago, we began to develop electrophysiological tools to address the relationship between auditory hallucinations and biological assays of the efference copy and corollary discharge mechanisms. Taking advantage of the writings of Hughlings Jackson, we decided to link the motor system of thoughts to the motor system of talking. Accordingly, we developed an electroencephalography (EEG)-based event related potential (ERP) method to study the putative action of these mechanisms by recording the response of auditory cortex to the spoken sound as it is being spoken (see 4 for details). Because the N1 component of the ERP, peaking at 100ms after the onset of a sound, is generated in primary and secondary auditory cortices, we have been using it to assess the responsiveness of auditory cortex to onset of the speech sound. Our paradigm and methods are similar to those used by others with human and non-human primates to study the action of the efference copy and corollary discharge mechanisms during vocalization (see 4).

In Figure 1 we illustrate the paradigm using a cartoon profile of a man talking (saying "ah") and listening (hearing "ah") to a recording of that speech played back. Above the cartoons, we show ERPs recorded from the vertex of the head, elicited by the onset of the speech sound (at time 0 ms, vertical dotted line) during talking (dashed lines) and listening (solid lines). As can be seen on the left of the Figure, where we show ERP data from 75 healthy control subjects, N1 to the speech sound is suppressed during talking relative to listening (5). ERPs are plotted with amplitude (microvolts, μV) on the y-axis and time (milliseconds, ms) on the x-axis. Negativity recorded at the vertex of the head (Cz) is plotted down. In the cartoon, the intention to say "ah" is indicated as a "thought bubble" over speech production areas in the frontal lobe. The curved arrow pointing to auditory cortex indicates the transmission of the efference copy of the motor plan, which produces a corollary discharge (burst) of the expected sensation in auditory cortex. When the expected sensation (corollary discharge) matches the actual sensation (sensory re-afference) in auditory cortex (gray burst), perception is suppressed. Strong responses in auditory cortex during listening are depicted as black, and weaker responses during talking are depicted as gray.

We have collected data from several samples of patients and controls with this paradigm and showed that auditory cortical responsiveness to sounds is dampened during talking in healthy controls but less so in patients with schizophrenia (5-9). Data from 75 patients are shown on the right of the Figure, where a relative failure to suppress N1 during talking is seen in the ERP (5).

Counter to our expectations, the amount of suppression of the N1 to the speech sounds during talking was not related to auditory hallucinations (6-8). However, neural synchrony in the EEG, 100ms before speech onset, was related to them. Because pre-speech neural synchrony was related to subsequent suppression of the N1 during talking in controls, we suggested EEG synchrony preceding speech may reflect the action of the efference copy of the motor com-

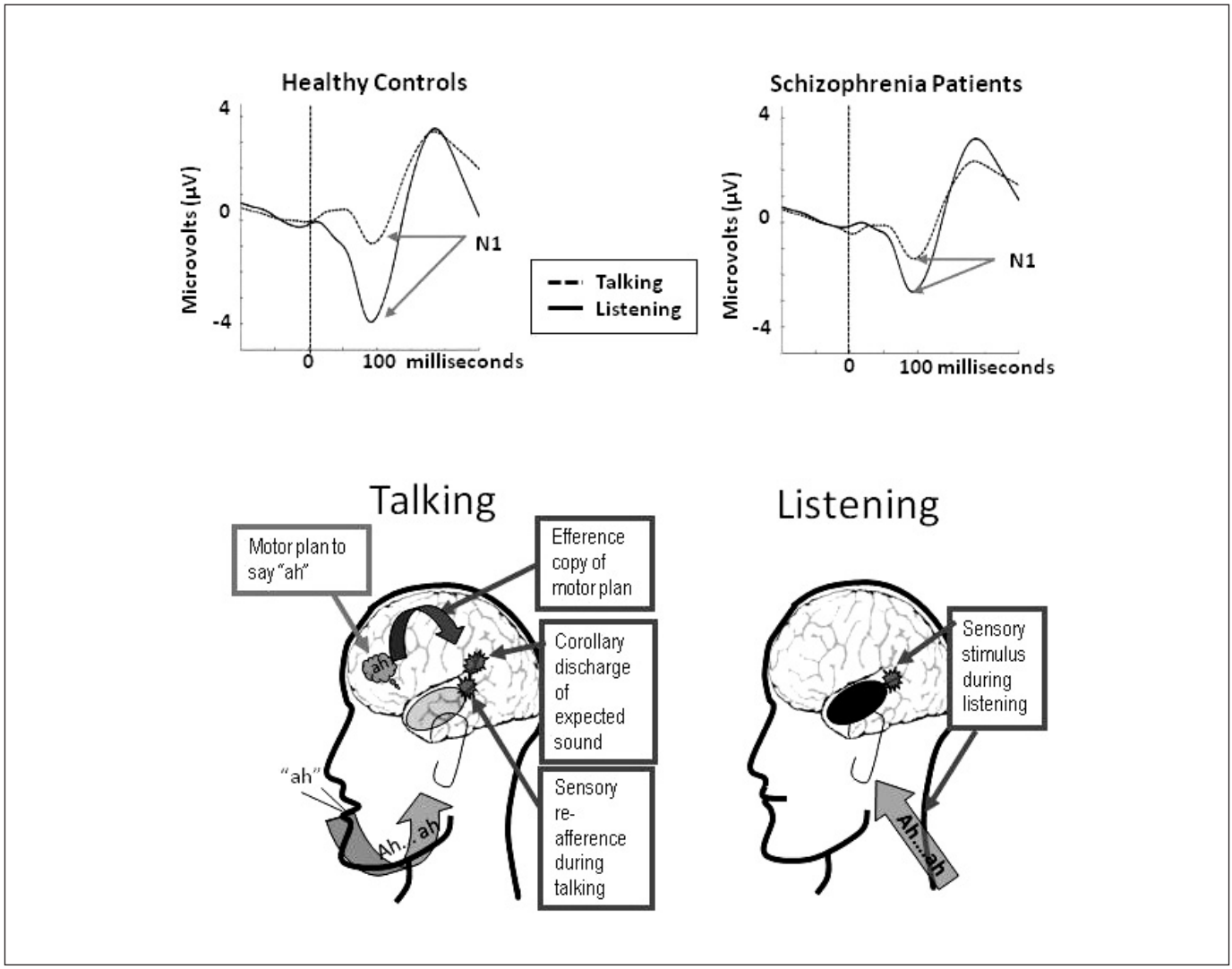

Figure 1 Schema of Talk/Listen Paradigm and resulting event related potentials from healthy controls and patients with schizophrenia

mand to speak (7). Perhaps, deficits in tagging the thought or idea as coming from self, and not deficits in suppressing the experience, map onto auditory hallucinations.

Initially, our studies were focused on whether dysfunction of the efference copy and corollary discharge mechanisms could explain auditory verbal hallucinations. However, in doing a "control experiment" with simple button pressing, we found that dysfunction of the efference copy of the motor command, preceding the button press, can also be observed in the somatosensory system in patients with schizophrenia, and its neurobiological manifestation was shown to selectively map onto symptoms in the motor domain, such as avolition and apathy (10). Schizophrenia is a pan-cerebral illness, affecting almost every modality, function, and brain region studied. While each symptom and function might have its own failed mechanism, parsimony encourages us to find an elemental mechanism that could be at the root of at least some of the dysfunctions observed. We suggest that dysfunction of the forward model system may reflect an elementary deficit in schizophrenia patients.

If our measures of these elementary mechanisms are reliable, valid, and not affected by antipsychotic medications, they might represent a major new domain of electrophysiological measurement sensitive to a fundamental and ubiquitous pathophysiological process in schizophrenia. These electrophysiological signals, reflecting abnormal feed-forward motor-sensory circuitry, could generate aberrant identification and experience of the sensory consequences of self-generated actions in schizophrenia. As such, they may underlie the aberrant experiences that may give rise to the variety of symptoms characteristic of patients with schizophrenia, from auditory verbal hallucinations to avolition and apathy. Accordingly, these electrophysiological measures may serve as novel neurophysiological outcome measures for developing and testing novel treatments in schizophrenia. They may also represent a novel endophenotype for studies of risk for schizophrenia, as we have shown that people at clinical high risk for the illness show N1 suppression during

talking that is intermediate between healthy controls and patients with a confirmed diagnosis (5). To the extent that they precede the onset of psychosis itself, they may enhance the prediction of psychosis onset in individuals with prodromal symptoms.

Most important, mechanistic studies such as these offer translation to bench neuroscience and translation to other species, and hence they can open the door to invasive manipulations that are not possible with in vivo human studies. For example, studies of the efference copy mechanism, like the ones reviewed above in humans, can be applied to animals that make social calls, such as song-birds and non-human primates. In such experiments, perturbations of the neurotransmitters implicated in schizophrenia might produce a pattern in the neural signature of the mechanism that resembles the pattern seen in schizophrenia patients who hallucinate.

Acknowledgements

This work was supported by VA Merit Review (J.M. Ford) and National Institute of Mental Health grants T32 MH089920 (V. Perez), K02 MH067967 and R01 MH58262 (J.M. Ford), and R01 MH076989 (D.H. Mathalon).

Permission to reproduce elements of Figure 1 was granted by Schizophrenia Bulletin.

References

1. Crapse TB, Sommer MA. Corollary discharge across the animal kingdom. Nat Rev Neurosci 2008;9:587-600.
2. Feinberg I. Efference copy and corollary discharge: implications for thinking and its disorders. Schizophr Bull 1978;4:636-40.
3. Frith CD. The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. Psychol Med 1987;17:631-48.
4. Ford JM, Roach BJ, Mathalon DH. How to assess the corollary discharge in humans using non-invasive neurophysiological methods. Nature Protocols 2010;5:1160-8.
5. Perez VB, Ford JM, Roach BJ et al. Auditory cortex responsiveness during talking and listening: early illness schizophrenia and patients at clinical high-risk for psychosis. Schizophr Bull (in press).
6. Ford JM, Mathalon DH, Heinks T et al. Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. Am J Psychiatry 2001;158:2069-71.
7. Ford JM, Roach BJ, Faustman WO et al. Synch before you speak: auditory hallucinations in schizophrenia. Am J Psychiatry 2007;164: 458-66.
8. Ford JM, Gray M, Faustman WO et al. Dissecting corollary discharge dysfunction in schizophrenia. Psychophysiology 2007;44: 522-9.
9. Heinks-Maldonado TH, Mathalon DH, Houde JF et al. Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. Arch Gen Psychiatry 2007;64:286-96.
10. Ford JM, Roach BJ, Faustman WO et al. Out-of-synch and out-of-sorts: dysfunction of motor-sensory communication in schizophrenia. Biol Psychiatry 2008;63:736-43.

Dear Editor,

Stigma blocks recovery from mental illness. Advocates have known this for more than a century and been eloquently arguing to audiences that would listen for much of that time. Psychiatrists have come later to the fray, most prominently in the past fifteen years which brought the authority of medicine and science to battles against the injustice. Unfortunately, research often showed psychiatrists of yore, among most mental health service providers, frequently added to the stigma of mental illness by endorsing notions of unrelentingly poor prognosis and need of institutionalization. In this light, a WPA task force recently posed stigma experienced by psychiatrists as another area of prejudice and discrimination (1). The report noted the "nefarious consequences" of psychiatrist stigma that undermine training needs of students and residents, lessen professional prestige, impact good salaries, and restrict possible resources. We feel that prioritizing psychiatrist stigma undermines psychiatry's moral authority in an argument where it had been noticeably missing. At best, posing psychiatrist stigma distracts from the core of the injustice, its harm on people with serious mental illnesses. At worst, it muddies psychiatry's role in promoting recovery.

In their saltatory definition, Link and Phelan (2) described power and status loss as primary to stigma. People labeled as mentally ill are shamed and have far fewer opportunities due to fellow citizens in their society. Many are unable to work, live independently, develop relationships, or enjoy health to their full potential. Perhaps psychiatrists suffer status loss too, especially in the light of superior acting medical colleagues. But the loss is miniscule compared to those about whom the stigma movement is built.

Consider two historical examples. For thousands of years people with leprosy were chased out of their homes into colonies of shame and deprivation. Good people often stepped up to care for those with leprosy, typically suffering the same kind of social opprobrium as those with the disease. Alternatively, the first 250 years of Europe's history in the New World were a time when Africans were brutalized as slaves. Caucasians of conscience spoke out about this moral plague, some working in the underground railroad, offering refuge and nourishment to slaves who were fleeing north. The freedom riders risked arrest, jail, and the murderous rage of neighbors. It is hard to focus on the unjust discrimination of carers in leper colonies or freedom fighters on the underground railroad, viewing both as a preposterous distraction. Perhaps our comparison is overstated. But then using the stigma of mental illness to describe the woes of psychiatry may be exaggerated. Psychiatrist stigma might gain legitimacy if it had the support of grass roots advocates with serious mental illnesses, those leading the charge against mental illness stigma. This is an interesting empirical question to be sure, one perhaps in need of a study. We hypothesize most from this group may not concur with the psychiatrist stigma agenda.

Patrick Corrigan[1], Matthias Angermeyer[2]
[1]Illinois Institute of Technology, Chicago, IL, USA
[2]Center for Public Mental Health, Gösing am Wagram, Austria

References

1. Sartorius N, Gaebel W, Cleveland H-R et al. WPA guidance on how to combat stigmatization of psychiatry and psychiatrists. World Psychiatry 2010;9:131-44.
2. Link B, Phelan J. Conceptualizing stigma. Annu Rev Sociol 2001; 27:363-85.

Dear Editor,

Thank you for letting us see the letter by P. Corrigan and M. Angermeier.

We think that the guidance on how to combat stigmatization of psychiatry and psychiatrists (1) should be seen within the context of WPA activities concerning stigma. The WPA has implemented the world's largest international collaborative project against the stigma of mental illness (2), a project which generated anti-stigma action in 18 countries of Africa, the Americas, Asia and Europe. Groups that have participated in this project continue to work together and see the reduction of stigma of mental illness and the stigmatization of people who suffer from them as their goal. The WPA, convinced of the importance of this work, also created a scientific section dealing with stigma and sponsored four major conferences (in Leipzig, Queenstown, Istanbul and London) that brought together researchers, members of user and carer organizations, health decision makers and practitioners of medicine and psychiatry. WPA's programs resulted in a large number of publications dealing with various aspects of stigma and its reduction. WPA's congresses over the past two decades regularly included symposia and other events dealing with stigma and ways of reducing it.

It was to complement these activities that the WPA included a project dealing with the stigma of psychiatry in its Action Plan 2008-2011 and consequently created a Task Force that was requested to produce guidelines that might help in reducing the stigma that marks psychiatry and psychiatrists. Reducing that emanation of the stigma of mental illness to psychiatry as a profession and a service would increase the interest of students of health professions in psychiatry and improve the prestige of the profession, as P. Corrigan and M. Angermeier say, but would – more importantly – also reduce the reluctance of people with mental health problems to seek help from psychiatric services and signifi-

cantly contribute to the potential of psychiatrists and other mental health specialists to advocate the needs of people with mental illness and to seek a higher priority for mental health programs in national health plans.

Both P. Corrigan and M. Angermeier have helped WPA in its fight to reduce the stigma of mental illness and its consequences in the past. We hope that WPA can count on their continued collaboration in all of its anti-stigma activities, including those that might improve the image of psychiatry and the psychiatrists.

Norman Sartorius[1], on behalf of the WPA Task Force convened to recommend measures that might reduce the stigmatization of psychiatry and psychiatrists
[1]Association for the Improvement of Mental Health Programmes, Geneva, Switzerland

References

1. Sartorius N, Gaebel W, Cleveland H-R et al. WPA guidance on how to combat stigmatization of psychiatry and psychiatrists. World Psychiatry 2010;9:131-44.
2. Sartorius N, Schulze H. Reducing the stigma of mental illness: a report from a Global Programme of the World Psychiatric Association. Cambridge: Cambridge University Press, 2005.

ERRATUM

It has been brought to our attention that in Figure 1 of the paper "Prediction and prevention of schizophrenia: what has been achieved and where to go next?", by Klosterkötter et al, published in the October 2011 issue of *World Psychiatry*, there was a factual error, already present in the submitted manuscript: the abbreviation for "Late At-Risk of Psychosis State" is not ERPS, but LRPS.

Papers and documents available on the WPA website

Many papers of interest to clinicians, researchers and educators working in the field of psychiatry and mental health can be found on the WPA website (www.wpanet.org).

The WPA has produced during the past triennium four guidance papers, available on the website in several languages. They deal with steps, obstacles and mistakes to avoid in the implementation of community mental health care (1), how to combat stigmatization of psychiatry and psychiatrists (2), mental health and mental health care in migrants (3), and protection and promotion of mental health in children of persons with severe mental disorders (4).

Three sets of slides dealing with the recognition, epidemiology, pathogenesis, cultural aspects, medical costs and management of the comorbidity of depression with diabetes, heart disease and cancer have been also produced. These slides are available on the WPA website in 18 different languages.

The WPA has developed an educational module on physical illness in patients with severe mental disorders (5,6), available on the website in several languages. Two sets of slides based on this educational module have been also posted on the website.

The WPA Committee on Ethics has produced a set of recommendations for relationships of psychiatrists, health care organizations working in the psychiatric field and psychiatric associations with the pharmaceutical industry (7). An international task force has produced a set of WPA recommendations on best practices in working with service users and family carers (8). Another task force has developed a WPA template for undergraduate and postgraduate education in psychiatry and mental health. Two surveys have been conducted in collaboration with WPA Member Societies, exploring their views about various issues concerning diagnosis and classification of mental disorders (9) and strategies to reduce the treatment gap for mental disorders (10). Several papers and documents have been produced by the WPA Early Career Psychiatrists Council (11-13). Reports have been provided on various activities implemented within the WPA Programme on Disasters (14-17). All these documents are available on the website.

References

1. Thornicroft G, Alem A, Dos Santos RA et al. WPA guidance on steps, obstacles and mistakes to avoid in the implementation of community mental health care. World Psychiatry 2010;9:67-77.
2. Sartorius N, Gaebel W, Cleveland H-R et al. WPA guidance on how to combat stigmatization of psychiatry and psychiatrists. World Psychiatry 2010;9:131-44.
3. Bhugra D, Gupta S, Bhui K et al. WPA guidance on mental health and mental health care in migrants. World Psychiatry 2011;10:2-10.
4. Brockington I, Chandra P, Dubowitz H et al. WPA guidance on the protection and promotion of mental health in children of persons with severe mental disorders. World Psychiatry 2011;10:93-102.
5. 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.
6. De Hert M, Cohen D, Bobes J et al. Physical illness in patients with severe mental disorders. II. Barriers to care, monitoring and treatment guidelines, plus recommendations at the system and individual level. World Psychiatry 2011;10:138-51.
7. Appelbaum P, Arboleda-Florez J, Javed A et al. WPA recommendations for relationships of psychiatrists, health care organizations working in the psychiatric field and psychiatric associations with the pharmaceutical industry. World Psychiatry 2011;10:155-8.
8. Wallcraft J, Amering M, Freidin J et al. Partnerships for better mental health worldwide: WPA recommendations on best practices in working with service users and family carers. World Psychiatry 2011;10:229-37.
9. Reed GM, Correia JM, Esparza P et al. The WPA-WHO Global Survey of Psychiatrists' Attitudes Towards Mental Disorders Classification. World Psychiatry 2011;10:118-31.
10. Patel V, Maj M, Flisher AJ et al. Reducing the treatment gap for mental disorders: a WPA survey. World Psychiatry 2010;9:169-76.
11. Fiorillo A, Lattova Z, Brahmbhatt P et al. The Action Plan 2010 of the WPA Early Career Psychiatrists Council. World Psychiatry 2010;9:62-3.
12. Fiorillo A, Brahmbhatt P, Elkholy H et al. Activities of the WPA Early Career Psychiatrists Council: the Action Plan is in progress. World Psychiatry 2011;10:159.
13. Fiorillo A, Luciano M, Giacco D et al. Training and practice of psychotherapy in Europe: results of a survey. World Psychiatry 2011;10:238.
14. van Ommeren M, Jones L, Mears J. Orienting psychiatrists to working in emergencies: a WPA-WHO workshop. World Psychiatry 2010;9:121-2.
15. Ravenscroft K. My experience in Haiti: a brief report. World Psychiatry 2010;9:191-2.
16. Bromet EJ. Lessons learned from radiation disasters. World Psychiatry 2011;10:83-4.
17. Akiyama T. Addressing the mental health consequences of the Japan triple catastrophe. World Psychiatry 2011;10:85.

The new WPA leadership

Resulting from the elections held in Buenos Aires on September 21 during the General Assembly, the composition of the new Executive Committee of the WPA is the following:

President: Pedro Ruiz (USA)
President-Elect: Dinesh Bhugra (UK)
Secretary General: Levent Kuey (Turkey)
Secretary for Finances: Tsuyoshi Akiyama (Japan)
Secretary for Meetings: Tarek Okasha (Egypt)
Secretary for Education: Edgard E. Belfort (Venezuela)
Secretary for Publications: Michelle B. Riba (USA)
Secretary for Sections: Afzal Javed (UK)

The composition of the new Board of the Association is the following:

Zone 1 (Canada): Donna E. Stewart (Canada)

Zone 2 (United States of America): John S. McIntyre (USA)
Zone 3 (Mexico, Central America and the Caribbean): Mauricio A. Sanchez (Nicaragua)
Zone 4 (Northern South America): Fabrizio Delgado Campodonico (Ecuador)
Zone 5 (Southern South America): José Geraldo Vernet Taborda (Brazil)
Zone 6 (Western Europe): Linda Gask (UK)
Zone 7 (Northern Europe): Henrik Wahlberg (Sweden)
Zone 8 (Southern Europe): Miquel Roca Bennasar (Spain)
Zone 9 (Central Europe): Jiri Raboch (Czech Republic)
Zone 10 (Eastern Europe): Petr Morozov (Russia)
Zone 11 (Northern Africa): Driss Moussaoui (Morocco)
Zone 12 (Middle East): Walid Sarhan (Jordan)
Zone 13 (Western and Central Africa): Joseph Adeyemi (Nigeria)
Zone 14 (Southern and Eastern Africa): Solomon Rataemane (South Africa)
Zone 15 (Western and Central Asia): S. Ahmad Jalili (Iran)
Zone 16 (Southern Asia): Manickam Thirunavukarasu (India)
Zone 17 (Eastern Asia): Min-Soo Lee (Republic of Korea)
Zone 18 (Australasia and South Pacific): Francis Agnew (New Zealand)

Acknowledgement
This publication has been supported by an unrestricted educational grant from AstraZeneca, which is hereby gratefully acknowledged.

€ 17,67 per issue

Printed in Italy by LEGO SpA,
via Galilei, 11 - 38015 Lavis, TN