Classification in psychiatry: ICD-10 v. DSM-IV

A response

MICHAEL B. FIRST and HAROLD ALAN PINCUS

The editorial by Andrews *et al* (1999) usefully calls attention to issues of compatibility between diagnostic classification systems but we believe that the editorial greatly overstates the compatibility problem as well as its implications. The article begins with the suggestion that the DSM-IV authors' position is to downplay the differences between DSM-IV and ICD-10. After stating that the American Psychiatric Association "felt sufficiently confident to publish a DSM-IV International Version in which the DSM-IV criteria are listed against the ICD-10 codes", the authors go on to report concordances between the classifications for the main mental disorders as ranging from a low of 33% (for substance harmful use or abuse) to 87% (for dysthymia), with an overall concordance of only 68%. The authors conclude that if this "unnecessary dissonance between the classification systems continues, patients, researchers and clinicians will be all the poorer". Although we acknowledge that there are a number of differences between the two systems, the authors fail to assess fully the sources, significance and solutions for this compatibility problem.

LACK OF CLARITY REGARDING THE VERSION OF ICD-10 BEING COMPARED TO DSM-IV

Throughout the editorial, the authors employ the term 'classification' when comparing and contrasting the two systems (e.g. "on the surface, the two classifications appear very similar"). This is misleading to the reader because although there is only one published version of the DSM-IV, the ICD-10 is actually a 'family of classifications', with multiple versions developed for different purposes. The DSM-IV criteria sets have been designed so that the same definitions can be used in all settings – clinical, research, forensic, adminis-

trative and educational. This has the advantage of facilitating communication across different categories of DSM-IV users. In contrast, several different parallel versions of the ICD-10 definitions of mental disorder have been developed, each one tailored for a specific use.

From an international compatibility perspective, there is only one 'official version' of the ICD-10: the version approved by the 43rd World Health Assembly in October 1989 (World Health Organization (WHO), 1992*a*), which provides a glossary definition for each mental disorder and is intended to "serve as a reference point for compatibility for other classifications" (WHO, 1992*b*). In the USA, the National Center for Health Statistics uses this official version as the starting point for the development of its 'clinical modification', the version used for diagnostic and classification purposes in health care facilities in the USA. During the preparation of DSM-IV, a great deal of effort was made to ensure that the DSM-IV criteria were fully compatible with these glossary definition. The appearance of ICD-10 codes alongside the DSM-IV disorders in the International Version of DSM-IV reflects the fundamental compatibility between the two systems at the 'official' glossary level.

The version of ICD-10 that is intended for "general clinical, educational, and service use" (WHO, 1992*b*) is the *Clinical Descriptions and Diagnostic Guidelines* (CDDG) version. It includes descriptions of the disorders and diagnostic guidelines "which indicate the number and balance of symptoms that are required before a confident diagnosis can be made" (WHO, 1992*b*). Finally, there are the ICD-10 research diagnostic criteria (RDC), which were produced for research purposes and were designed to be "deliberately restrictive; their use allows the selection of groups of individuals whose symptoms and other characteristics resemble each other in clearly stated ways" (WHO, 1993). It

should also be noted that, in addition to these three, which were not formally approved by the World Health Assembly, there are several other members of the ICD-10 family, including a primary care version and a multi-axial version.

Although the three versions of the ICD-10 definitions are "compatible" (WHO, 1993), they are not identical and do not identify the same group of individuals as having the disorder. Consider, for example, the various ICD-10 definitions of post-traumatic stress disorder (PTSD; see Appendix). The description derived from the 'official' version is the most broadly defined and requires only that the symptoms arise as "response to a stressful event or situation . . . of an especially catastrophic nature". Accompanying symptoms (e.g. reliving of the trauma) are described as 'typical'. The version from CDDG is a bit more restrictive, requiring "a repetitive, intrusive, recollection or re-enactment of the event". Other symptoms are listed as 'often present'. The RDC definition is the most restrictive of all, requiring remembering or reliving the event, avoidance of circumstances resembling the event and either partial amnesia or persistent symptoms of increased psychological sensitivity and arousal.

All of the authors' analyses regarding the concordance between DSM-IV and ICD-10, in fact, compare the DSM-IV criteria with the ICD-10 RDC. Because the ICD-10 RDC are deliberately restrictive and tend to employ complex diagnostic algorithms (e.g. see criterion D for PTSD in the Appendix), diagnostic differences are likely to be much greater than if the DSM-IV criteria had been compared to the ICD-10 CDDG version. It is likely that DSM-IV is more compatible overall with the ICD-10 CDDG than with the ICD-10 RDC. Consider, for example, the difference between the various ICD-10 definitions of PTSD and the DSM-IV criteria (see Appendix). As the authors discuss in their editorial, many of the cases diagnosed as PTSD in ICD-10 but not in DSM-IV resulted from the lack of a numbing of responsiveness criterion in the ICD-10 RDC, a symptom included in the CDDG description of the disorder. Thus, a number of non-cases as defined by the ICD-10 RDC would have been cases if the comparison was with the ICD-10 CDDG definition. Because the ICD-10 RDC are intended for use almost exclusively by the research community, the authors' findings

have limited relevance to clinical and other applications of the two classification systems.

METHODOLOGICAL PROBLEMS IN ASSESSMENT OF DIAGNOSTIC CONCORDANCE

The authors determined the diagnostic concordance between DSM–IV and ICD–10 by administering the Composite International Diagnostic Interview (CIDI; WHO, 1997) to a large group of individuals drawn both from the community and from clinical populations, and examining differences in the overall diagnostic prevalence as well as the percentage of cases in which the interview assigned the same DSM–IV and ICD–10 diagnosis. Fully structured interviews such as the CIDI are designed to be administered by lay interviewers – no clinical judgement is needed for the diagnostic assessment. In fact, in their study the community subjects were interviewed by lay interviewers using the CIDI-Auto – a computerised version of the CIDI – and the clinical sample filled in a self-administered form. Although the use of CIDI may have made practical sense in view of the large number of subjects, the use of rigidly worded self-report questions has the potential to conflate artificially the discordance between the two systems.

Take, for example, the reported rates of diagnostic discordance in the diagnosis of Generalised Anxiety Disorder. Andrews *et al* (1998) reported that of the 81 people who were positive for ICD–10 Generalised Anxiety Disorder but negative for DSM–IV, 63 were negative because they did not meet DSM–IV criterion A, which requires 'excessive anxiety and worry'. The CIDI assesses this DSM–IV criterion by having the patient make this judgement for him- or herself by directly asking the patient: "do you think your worrying was excessive, that is, much stronger than in other people?" The DSM–IV requirement for 'excessive' worry is a clinical judgement call on the part of the clinician, who must take into account the extent of the person's worry and the person's life context – it does not make sense to have it rest solely on the person's insight into whether the anxiety is excessive. Similarly, the majority of the DSM–IV criteria sets have a criterion that requires the disturbance to cause "clinically significant impairment or distress in one or

more areas of functioning", a requirement that is lacking in the ICD–10 (Pincus *et al*, 1998; Regier *et al*, 1998). According to the authors, the presence of this criterion in DSM–IV is often responsible for a large portion of the diagnostic discordance. For example, almost half of the discrepancy in the diagnosis of PTSD can be accounted for by this DSM–IV requirement. Many individuals, particularly those who tend to be stoic, are prone to minimise the effect that psychiatric symptoms may exert on their lives. The assessment of this criterion requires the clinician to evaluate objectively the impact of symptomatology on the person's functioning. Relying solely on a self-report instrument to determine this criterion is likely to lead to false-negatives, increasing the apparent disagreement between the two systems.

Although the authors acknowledge that the CIDI itself may be responsible for some of the disagreement, they claim to have eliminated "disagreement not necessarily related to the classifications . . . [by removing] . . . two sources of error that could have contributed to discrepancy, first in the construction of the CIDI questions, and second in the interpretation of the diagnostic criteria in the scoring algorithm". It is our belief that the inherent requirement in the CIDI for the elimination of clinical judgement makes it impossible to eliminate the CIDI as a significant source of diagnostic disagreement. Future studies aiming to compare diagnostic concordance between ICD–10 and DSM–IV should employ assessment methods that more closely mimic clinical practice by incorporating clinical judgement into the application of the diagnostic criteria. Using semi-structured interviews such as the Schedule for the Assessment of Clinical Neuropsychiatry (SCAN; WHO, 1994) would accomplish this goal and provide results that are both more accurate and clinically meaningful.

THE DIAGNOSTIC CRITERIA IN DSM – IV AND ICD – 10 ARE NOT ON AN EQUAL SCIENTIFIC FOOTING

We acknowledge that even if the CDDG were used as the comparison set and even with a methodology that allows for the application of clinical judgement in the application of the diagnostic criteria, some differences between the two systems would

be likely to persist. The crux of the matter is, given these differences, which criteria set is preferable for clinical and research use? In the absence of an available 'gold standard' by which we can measure comparative validity, another potential yardstick is the relative degree of empirical data behind the criteria sets.

The DSM–IV criteria sets represent the third iteration in an ongoing process to refine the DSM based on a comprehensive three-stage empirical review process that extended over five years. The first stage consisted of a comprehensive and systematic review of the published literature for each disorder (Widiger *et al*, 1990). The second stage involved a re-analysis of already collected data sets to answer some of the questions raised during the first stage and to help generate and pilot new proposals for criteria sets for DSM–IV (Widiger *et al*, 1991). The third stage involved issue-focused field trials to assess the extent to which proposed revisions would actually improve the reliability and/or validity of criteria sets and to address the issues identified by the literature reviews. Twelve focused field trials were conducted, involving a total of over 7000 patients in 81 sites in the USA and around the world. Finally, the entire DSM–IV revision process, from the establishment of criteria for making changes to the results of the literature reviews, data re-analyses and field trials, was painstakingly documented in the four-volume *DSM–IV Sourcebook* (Widiger *et al*, 1994, 1996, 1997, 1998).

In contrast, the ICD–10 criteria sets were created *de novo* using the 'expert consensus process' that informed the development of the DSM–III criteria sets and that was rejected by the DSM–IV Task Force as no longer adequate to the task. The clinical descriptions, diagnostic guidelines and diagnostic criteria for research were "prepared after extensive collaboration with panels of experts, national and international psychiatric societies, and individual consultants." (WHO, 1992b). Like the process that preceded the development of the DSM–III criteria sets, there is virtually no documentation of the rationale for any of the ICD–10 diagnostic guidelines or criteria sets. A researcher or clinician who is interested in tracking down the justification for a particular criterion has little choice but to depend on the fast-fading memories of the participants of the ICD–10 process.

Drafts of both the CDDG and the RDC were field-tested at over 150 clinical and

research centres in over 40 countries (Sartorius et al, 1993, 1995). Because of the wide variety of centres and in an effort to keep the cost low, relatively minimal design requirements were imposed on the study participants – essentially pairs (or larger groups) of clinicians or researchers were asked to make diagnostic assessments on unselected patients, a methodology that was similar to that of the general field trials preceding the publication of DSM–III (Williams & Spitzer, 1980). Like the DSM–III field trials (and unlike those of DSM–IV), the purpose was to establish the overall feasibility of the system, as opposed to improving the reliability or validity of any of the categories – in fact, the field trial results were not analysed until the clinical guidelines and diagnostic criteria for research were pretty much in their final form.

Let us compare again the DSM–IV and ICD–10 definitions of PTSD. Post-traumatic stress disorder was first introduced into the psychiatric nomenclature by DSM–III in 1980; there was no corresponding category in ICD–9. The initial definition of PTSD had three components: re-experiencing the trauma, numbing of responsiveness and an additional list of hyperarousal symptoms, of which two must have developed after the trauma (including avoidance of activities reminiscent of the trauma). With the publication of DSM–III–R in 1987, the second component was changed to include avoidance of stimuli associated with the trauma, and additional symptoms were added to the hyperarousal list. Hundreds of studies have used the DSM–III and DSM–III–R definitions, and a number of assessment tools have been developed as well (e.g. the Clinician-administered PTSD Scale (Blake et al, 1990) and the Mississippi Scale (Keane et al, 1988)). As part of the DSM–IV revision process, a focused field trial was conducted to explore issues regarding the best way to depict the stressor criterion as well as the content of the three symptom lists that made up the PTSD definition. The field trial demonstrated that the 17 PTSD symptoms had high internal consistency and acceptable sensitivity, specificity and predictive power.

Diagnostic validity was demonstrated further by the finding that the PTSD symptom picture as defined by DSM–III–R/DSM–IV rarely occurred unless the person experienced the type of stressor typically thought to be capable of producing PTSD (Kilpatrick et al, 1998). In contrast, the origins of the ICD–10 criteria set for PTSD are both obscure and undocumented. As was demonstrated above, the discordance between DSM–IV and ICD–10 with regard to PTSD is primarily at the diagnostic criteria level. A comparison between the DSM–IV criteria and the ICD–10 Chapter V glossary definition and CDDG reveals the differences to be virtually non-existent. Why the developers of the ICD–10 RDC choose to create their own criteria set and not simply adopt the widely used DSM–III–R criteria set is anybody's guess.

CONCLUSION

The editorial begins with the authors seemingly lamenting the diagnostic confusion stemming from the coexistence of the ICD–10 and DSM–IV (i.e. "in the English-speaking world, psychiatric classification used to be governed by one system, now there are two"). We would argue that, certainly in the world of research, the DSM system of specified diagnostic criteria is the de facto standard (Maser et al, 1991) and that the introduction of the ICD–10 criteria is the main source of confusion among researchers. Of course, some of the differences in the criteria sets stem from different conceptual frameworks underlying European v. American diagnostic concepts (e.g. in schizoaffective disorder, the American definition relies on the temporal relationship between mood and psychotic symptoms whereas the European definition stresses the relative significance of the mood and psychiatric symptoms). However, many (if not most) of the differences seem to exist for no good reason.

We propose for the next revision of DSM–IV and ICD–10 that all conceivable efforts be made to eliminate unnecessary differences. We have completed a methodical comparison of the differences between the DSM–IV and ICD–10 (details available from the authors upon request), dividing the disorders into four categories: disorders included in one classification but not the other (e.g. narcissistic personality disorder in DSM–IV, neurasthenia in ICD–10); disorders included in both classifications that are essentially identical (e.g. undifferentiated somatoform disorder); disorders included in both classifications with unfounded or arbitrary differences (e.g. PTSD); and disorders that are intentionally different based on conceptual or traditional differences (e.g. schizoaffective disorder, schizophrenia). Diagnostic differences that represent idiosyncratic operationalisation of diagnostic constructs should be eliminated based on a comprehensive review of the empirical data, the method pioneered by the DSM–IV revision process. We agree with Kendell (1991), as well as the authors, that for some disorders the maintenance of alternative criteria sets based on well-documented conceptual differences has the potential to spur on new research. Thus, the method for resolving these differences should be through gathering research data and subsequent publication of these data in the open scientific literature, rather than promulgation of newly developed criteria sets without empirical support.

APPENDIX

Definitions of post-traumatic stress disorder in ICD–10 and DSM–IV

ICD–10 Chapter V (F)

Post-traumatic stress disorder arises as a delayed or protracted response to a stressful event or situation (of either brief or long duration) of an exceptionally threatening or catastrophic nature, which is likely to cause pervasive distress in almost anyone. Predisposing factors, such as personality traits (e.g. compulsive, asthenic) or previous history of neurotic illness, may lower the threshold for the development of the syndrome or aggravate its course, but they are neither necessary nor sufficient to explain its occurrence. Typical features include episodes of repeated reliving of the trauma in intrusive memories ('flashbacks'), dreams or nightmares, occurring against the persisting background of a sense of 'numbness' and emotional blunting, detachment from other people, unresponsiveness to surroundings, anhedonia and avoidance of activities and situations reminiscent of the trauma. There is usually a state of autonomic hyperarousal with hypervigilance, an enhanced startle reaction and insomnia. Anxiety and depression are commonly associated with the above symptoms and signs, and suicidal ideation is not infrequent. The onset follows the trauma with a latency period that may range from a few weeks to months. The course is fluctuating but recovery can be expected in the majority of cases. In a small proportion of cases the condition may follow a chronic course over many years, with eventual transition to an enduring personality change.

ICD–10 Clinical Description and Diagnostic Guidelines (CDDG)

Post-traumatic stress disorder should not generally be diagnosed unless there is evidence that it arose within six months of a traumatic event of exceptional severity. A 'probable' diagnosis might still be possible if the delay between the event and the onset was

longer than six months, provided that the clinical manifestations are typical and no alternative identification of the disorder (e.g. as an anxiety or obsessive–compulsive disorder or depressive episode) is plausible. In addition to evidence of trauma, there must be a repetitive, intrusive recollection or re-enactment of the event in memories, daytime imagery or dreams. Conspicuous emotional detachment, numbing of feeling and avoidance of stimuli that might arouse recollection of the trauma are often present but are not essential for the diagnosis. The autonomic disturbances, mood disorder and behavioural abnormalities all contribute to the diagnosis but are not of prime importance.

ICD–I0 Diagnostic Criteria for Research (DCR)

A. The patient must have been exposed to a stressful event or situation (either short- or long-lasting) of exceptionally threatening or catastrophic nature, which would be likely to cause pervasive illness in almost anyone.

B. There must be persistent remembering or 'reliving' of the stressor in intrusive "flashbacks", vivid memories or recurring dreams, or in experiencing distress when exposed to circumstances resembling or associated with the stressor.

C. The patient must exhibit an actual or preferred avoidance of circumstances resembling or associated with the stressor, which was not present before exposure to the stressor.

D. Either of the following must be present:

(1) inability to recall, either partially or completely, some important aspects of the period of exposure to the stressor;

(2) persistent symptoms of increased psychological sensitivity and arousal (not present before exposure to the stressor), shown by any two of the following:

(a) difficulty falling or staying asleep;

(b) irritability or outbursts of anger;

(c) difficulty in concentrating;

(d) hypervigilance;

(e) exaggerated startle response.

E. Criteria B, C and D must all be met within six months of the stressful event or of the end of a period of stress. (For some purposes, onset delayed more than six months may be included, but this should be clearly specified.)

DSM–IV

A. The person has been exposed to a traumatic event in which both of the following were present:

(1) the person experienced, witnessed or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others;

(2) the person's response involved intense fear, helplessness or horror.

B. The traumatic event is persistently re-experienced in one (or more) of the following ways:

(1) recurrent and intrusive distressing recollections of the event, including images, thoughts or perceptions;

(2) recurrent distressing dreams of the event;

(3) acting or feeling as if the traumatic event were recurring (includes a sense of reliving the experience, illusions, hallucinations and dissociative flashback episodes, including those that occur on awakening or when intoxicated);

(4) intense psychological distress at exposure to internal or external cues that symbolise or resemble an aspect of the traumatic event;

(5) physiological reactivity on exposure to internal or external cues that symbolise or resemble an aspect of the traumatic event.

C. Persistent avoidance of stimuli associated with the trauma and numbing of general responsiveness (not present before the trauma), as indicated by three (or more) of the following:

(1) efforts to avoid thoughts, feelings or conversations associated with the trauma;

(2) efforts to avoid activities, places or people that arouse recollections of the trauma;

(3) inability to recall an important aspect of the trauma;

(4) markedly diminished interest or participation in significant activities;

(5) feeling of detachment or estrangement from others;

(6) restricted range of affect (e.g. unable to have loving feelings);

(7) sense of a foreshortened future (e.g. does not expect to have a career, marriage, children or a normal life span).

D. Persistent symptoms of increased arousal (not present before the trauma) as indicated by two (or more) of the following:

(1) difficulty falling or staying asleep;

(2) irritability or outbursts of anger;

(3) difficulty in concentrating;

(4) hypervigilance;

(5) exaggerated startle response.

E. Duration of the disturbance (symptoms in criteria B, C and D) is more than one month.

F. The disturbance causes clinically significant distress or impairment in social, occupational or other important areas of functioning.

MICHAEL B. FIRST, MD, Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute; HAROLD ALAN PINCUS, MD, American Psychiatric Association Office of Research, American Psychiatric Association, Washington, DC, USA

Correspondence: Dr Michael B. First, Biometrics Research Department, New York State Psychiatric Institute, 1051 Riverside Drive, Unit 60, New York, NY 10032, USA. Tel: 212-543-5531; Fax: 212-543-5525; e-mail: mbf2@columbia.edu

REFERENCES

Andrews, G., Slade, T., Peters, L., et al (1998) Generalized anxiety disorder, obsessive–compulsive disorder and social phobia: sources of dissonance between ICD–10 and DSM–IV. *International Journal of Methods in Psychiatric Research*, **7**, 110–115.

— , — & — (1999) Classification in psychiatry: ICD–10 versus DSM–IV. *British Journal of Psychiatry*, **174**, 3–5.

Blake, D. D., Weathers, F. W., Nagy, L. N., et al (1990) A clinician rating scale for assessing current and lifetime PTSD: the CAPS–I. *Behavior Therapist*, **18**, 187–188.

Keane, T. M., Caddell, J. M. & Taylor, K. L. (1988) Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: three studies in reliability and validity. *Journal of Consulting and Clinical Psychology*, **56**, 85–90.

Kendell, R. E. (1991) Relationship between the DSM–IV and the ICD–10. *Journal of Abnormal Psychology*, **100**, 297–301.

Kilpatrick, D. G., Resnick, H. S., Freedy, J. R., et al (1998) Posttraumatic stress disorder field trial: evaluation of the PTSD construct-criteria A through E. In *DSM–IV Sourcebook*, Volume 4 (eds T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross, M. B. First, W. Davis & M. Kline), pp. 803–844. Washington, DC: American Psychiatric Association.

Maser, J. D., Kaelber, C. & Weise, R. E. (1991) International use and attitudes toward DSM–III and DSM–III–R: growing consensus in psychiatric classification. *Journal of Abnormal Psychology*, **100**, 271–279.

Pincus, H. A., Zarin, D. A. & First, M. B. (1998) 'Clinical significance' and the DSM–IV [Letter]. *Archives of General Psychiatry*, **55**, 1145.

Regier, D. A., Kaelber, C. T., Rae, D. S., et al (1998) Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. *Archives of General Psychiatry*, **55**, 109–115.

Sartorius, N., Kaelber, C. T., Cooper, J. E., et al (1993) 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. *Archives of General Psychiatry*, **50**, 115–214.

— , Ustun, T. B., Korten, A., et al (1995) Progress toward achieving a common language in psychiatry, II: results from the international field trials of the ICD–10

diagnostic criteria for research for mental and behavioral disorders. *American Journal of Psychiatry*, **152**, 1427–1437.

Widiger, T. A., Frances, A. J., Pincus, H. A., et al (1990) DSM – IV literature reviews: rationale, process, and limitations. *Journal of Psychopathology and Behavioral Assessment*, **2**, 189–202.

—, —, —, et al (1991) Toward an empirical classification for the DSM – IV. *Journal of Abnormal Psychology*, **100**, 280–288.

—, —, —, et al (1994) *DSM – IV Sourcebook, Volume 1.* Washington, DC: American Psychiatric Association.

—, —, —, et al (1996) *DSM – IV Sourcebook, Volume 2.* Washington, DC: American Psychiatric Association.

—, —, —, et al (1997) *DSM – IV Sourcebook, Volume 3.* Washington, DC: American Psychiatric Association.

—, —, —, et al (1998) *DSM – IV Sourcebook, Volume 4.* Washington, DC: American Psychiatric Association.

Williams, J. B. W. & Spitzer, R. L. (1980) DSM – III field trials: interrater reliability and list of project staff and participants. In *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn) (DSM – III), pp. 467–481. Washington, DC: American Psychiatric Association.

World Health Organization (1992a) *International Statistical Classification of Diseases and Related Health Problems* (10th Revision) (Volume 1). Geneva: World Health Organization.

— (1992b) *ICD – 10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines.* Geneva: World Health Organization.

— (1993) *ICD – 10 Classification of Mental and Behavioral Disorders: Diagnostic Criteria for Research.* Geneva: World Health Organization.

— (1994) *Schedules for Clinical Assessment in Neuropsychiatry* (Version 2.0 Manual). Geneva: World Health Organization.

— (1997) *Composite International Diagnostic Interview* (Version 2, 12-month). Geneva: World Health Organization.