

# Diagnosis and methodology

Wolfgang Hiller

Clinic Roseneck, Centre for Behavioural Medicine, Prien, Germany

The concepts, definitions and delineations of mental disorders have become much more standardized and uniform worldwide through the development of new classification systems: DSM-IV and ICD-10. The reliability of the diagnostic process has improved by the introduction of a number of well evaluated instruments. New data have been presented for some established scales designed to measure the severity of syndromes, and some new measures have also been developed for specific areas of psychopathology.

Current Opinion in Psychiatry 1995, 8:391-394

## Introduction

Efforts to develop empirically based concepts of mental disorders and scientifically adequate instruments have played a central role in psychiatry and clinical psychology during the past 2 decades. New landmarks were set with the development of operationalized diagnostic criteria. Today, the new classification systems of the DSM-IV and ICD-10 seem to represent an end-point of these innovations. Both systems will hopefully be valid for the next 10-15 years. A number of well evaluated instruments exist to guide the diagnostician through the complex task of assessing the correct diagnosis.

## The change of classification systems

The fourth edition of the *American Diagnostic and Statistical Manual* DSM-IV [1] and the tenth revision of the *International Classification of Diseases* (ICD-10) [2] have been adapted to the current clinical and scientific knowledge of mental disorders. DSM-IV is based on extensive literature reviews, data reanalyses and field trials, although only minor changes were made for the major diagnostic categories and no paradigmatic changes (compared with the former DSM-III and DSM-III-R systems) were introduced. ICD-10 can be regarded as an approach of the World Health Organization to facilitate the international language and understanding of mental disorders [3]. ICD-10 has adopted the primarily descriptive and criteria-related approach known from the DSM systems since DSM-III, and the general structure of both classifications is now very similar. However, complete congruence between ICD-10 and DSM-IV has not yet been reached despite close consultations between the developers of ICD-10 and DSM-IV. The remaining (though mostly minor) differences reflect incompatible national traditions and to some degree the unwillingness or inability of the

responsible committees to compromise on diagnostic formulations and definitions [4].

Whereas ICD-10 departs substantially from its predecessors ICD-8 and 9, the changes from DSM-III (1980) to DSM-III-R (1987) to DSM-IV (1994) were relatively small, and it was criticized repeatedly that three revisions within 14 years were unnecessary and premature. Zimmerman *et al.* [5] confirmed these concerns with data from a mail survey of US psychiatrists who were asked in 1989 about their training in, use of, and attitudes about DSM-III and DSM-III-R. Even 2 years after the publication of DSM-III-R, 30.6% of the practitioners and 22.8% of the researchers were still using the former DSM-III as their primary diagnostic reference. Less than 50% of the DSM-III-R users believed that the revision was needed, although the majority recognized the improvements of DSM-III-R over DSM-III. The authors [5] concluded from these results that a quick acceptance of DSM-IV should not be expected and that the mental health community may further split up into subgroups using different diagnostic criteria.

## Reliability and comparability of the systems

No reliability studies using the new DSM-IV diagnoses have been published yet, but it can be expected that the results of such studies will be generally comparable to the known reliabilities of DSM-III and DSM-III-R diagnoses. The inter-rater reliabilities of ICD-10 diagnoses found in a series of worldwide field trials were published by Sartorius *et al.* [3]. The values were satisfactory for most categories.

The empirical overlap between DSM-III-R and ICD-10 was investigated by Hiller *et al.* [6], who applied both sets of diagnostic criteria to a sample of 100 psychiatric in-patients. Overall reliability was 0.53 (kappa) for

## Abbreviations

HDRS—Hamilton Depression Rating Scale; QEWP—Questionnaire of Eating and Weight Patterns;  
SCID—Structured Clinical Interview for DSM-III-R.

DSM-III-R diagnoses and 0.59 for ICD-10. Agreement was best for the affective disorders, moderate for schizophrenia and unacceptable in both systems for schizoaffective disorder. ICD-10 yielded moderately higher reliabilities for all main disorders except for bipolar disorder. A subsequent analysis of the nature of intersystem discrepancies [7••] demonstrated a severe impact of the different duration criteria for schizophrenia (6 months in DSM-III-R and only 1 month in ICD-10). The base rate of schizophrenia was about 40% higher in ICD-10 than in DSM-III-R, and no more than 57.5% of schizophrenics according to ICD-10 received the same diagnosis according to DSM-III-R. It was further shown that the delineation between schizophrenia, on the one hand, and schizoaffective and bipolar disorder, on the other, was insufficient in both classification systems. A more detailed analysis [8•] of depressive disorders demonstrated a very good congruence between the internationally accepted concept of major depression and the new ICD-10 categories of 'depressive episode' and 'recurrent depressive disorder'.

### Instruments related to the new classifications

The use of (semi)structured interviews to assess symptoms and diagnoses more reliably has become a standard method, especially in clinical research. An evaluative review for one of the most popular diagnostic instruments, the Structured Clinical Interview for DSM-III-R (SCID), was worked out by Segal *et al.* [9•]. All SCID reliability studies available were summarized and the authors confirmed that the SCID yields highly reproducible diagnoses for most Axis I and Axis II disorders. It is, therefore, assumed that the new SCID version for DSM-IV will continue to enjoy widespread popularity. New data for the SCID-II were reported by Ekselius *et al.* [10•] to demonstrate that a high agreement of 0.78 (kappa) between the SCID screening questionnaire and the interview could be reached to diagnose personality disorders if the cutoff level for the questionnaire is adjusted. This approach could prove to be a methodological solution for the repeatedly demonstrated problem that self-reported personality traits tend to be overinclusive and thus inadequate to determine an exact diagnosis [11].

Other instruments have been developed within the crosscultural framework of the new ICD-10 classification. Janca *et al.* [12•] summarize the basic characteristics of the Composite International Diagnostic Interview [13], the Schedules for Clinical Assessment in Neuropsychiatry [14] and the International Personality Disorder Examination [15••]. These instruments have now been published for general use after many years of developing and field-testing in 30 centres in different parts of the world.

As an alternative to highly standardized interview techniques, a set of check lists based on the ICD-10 diagnostic criteria has been released by the World Health Organization. It consists of the ICD-10 Symptom

Checklist and Glossary developed by Janca *et al.* [16] and the International Diagnostic Checklists developed by Hiller *et al.* [17]. Checklists seem to be less time-consuming and cumbersome in clinical practice and they enable the clinician to be more flexible in adjusting the interview procedure to the level of the patients' intelligence, education and clinical picture.

### Other instruments to assess psychopathology

A number of new studies evaluated the psychometric and clinical characteristics of observer or selfrating scales designed to assess either specific or more general aspects of psychopathology. Such instruments are commonly used to quantify the severity of a syndrome, to measure treatment effects or to serve as a screening method to identify subjects with a presumed mental disorder.

Increasing and alarming evidence shows that the extensively used Hamilton Depression Rating Scale (HDRS) is only a weak index of depressive syndrome severity. Gibbons *et al.* [18] evaluated the internal construct validity of the HDRS using data from 370 patients with major depressive disorder. They found five distinct factors, but only the first factor seemed well defined and clinically interpretable as a primary measure of depression. Somatic and vegetative symptoms did not consistently load on this dimension and may thus represent something qualitatively different. The authors [18] strongly recommended from these and other results that one should be cautious when using total HDRS scores in treatment or clinical studies and advocated the use of other additional measures.

Discrepancies between self and observer-ratings of depression were studied by Domken *et al.* [19] using the Inventory for Depressive Symptomatology, a combined clinician and selfrating scale. The correlation between both measures of depression was only 0.65 in a sample of 48 patients with major depressive disorder. Domken *et al.* [19] found that neuroticism, dysfunctional attitudes and low selfesteem were significantly related to the differences between the self and clinician ratings.

Silverstone [20] evaluated the Hospital Anxiety and Depression Scale and found a positive predictive power of only 17–29% to detect medical and psychiatric patients diagnosed as having a DSM-III-R depressive disorder. Parker *et al.* [21] introduced a new, brief 11-item questionnaire assessing both mood state and social impairment. This measure is thought to be largely independent from diagnostic subtyping features such as melancholic or psychotic depression.

A study published by Thompson *et al.* [22] dealt with different scaling methods for the 18-item Brief Psychiatric Rating Scale. The authors compared the 1–7 scaling system (used in the USA) with the 0–6 scaling system (used in most other countries) in a sample of 396 patients with schizophrenia, and demonstrated that the 1–7 scaling system may result in a serious underestimation of treatment responders. The 0–6

scaling system is advocated for the Brief Psychiatric Rating Scale as well as for other rating scales such as the Positive and Negative Syndrome Scale for Schizophrenia.

Confirmatory factor analyses were reported by Axelrod *et al.* [23] for the Negative Symptom Assessment. This 26-item rating scale covers multiple facets of residual schizophrenic features and six factors were derived from data of 223 inpatients describing communication, emotion/affect, social involvement, motivation, gross cognition and retardation.

The ability to discriminate between depressive, negative and extrapyramidal symptoms in patients with schizophrenia was studied by Addington *et al.* [24\*] using the Calgary Depression Scale as well as other specific measures. Although a weak correlation of 0.33 was found between the Calgary Depression Scale and negative symptoms in their inpatient subsample, the authors concluded from a Lisrel analysis that the depression scores achieved a sufficient level of separation from negative and extrapyramidal symptoms in the entire sample of 150 schizophrenic inpatients and outpatients.

A new 15-item measure of manic symptomatology, the Clinician-Administered Rating Scale for Mania, was constructed by Altman *et al.* [25\*]. This instrument was primarily derived from the Schedule for Affective Disorders and Schizophrenia and ratings are based on a brief semistructured interview. Another new approach is the Mania-Depression Scale developed by Mazmanian *et al.* [26], a global one-item scale that provides a score ranging from -5 (depressive stupor), through 0 (euthymic) to +5 (manic delirium).

The factorial structure of the short 10-item Yale-Brown Obsessive-Compulsive Scale was analyzed by Kim *et al.* [27\*] using baseline and post-treatment data from 204 patients with obsessive-compulsive disorder. A stable three-factor solution discriminated well between obsessions and compulsions and an additional distinct dimension was found describing efforts to resist the symptoms. Kaplan [28\*] described the development of a Self-Rated Scale for Obsessive-Compulsive Disorder consisting of 35 items. A correlation of 0.62 with the Yale-Brown Obsessive-Compulsive Scale was found for this new scale.

The usefulness of the Brief Symptom Inventory as an outcome measure of psychiatric treatment was demonstrated by Piersma *et al.* [29] with data from 217 inpatients. The Brief Symptom Inventory is a 53-item self-report inventory of psychopathology and stressors that was constructed as a briefer form of the widely used Hopkins Symptom Check List. Piersma *et al.* [29] demonstrated how clinical significance indices can be calculated from the Brief Symptom Inventory with regard to changes made by individual patients.

The eight diagnostic items of the Questionnaire of Eating and Weight Patterns (QEWP) were used by Nangle *et al.* [30\*] to identify individuals meeting the new proposed DSM-IV criteria of binge eating disorder.

The authors were successful in discriminating clinical and nonclinical binge eaters and the diagnoses based on the QEWP were found to be moderately stable over a 3-week period. A fairly high level of agreement between SCID diagnoses of binge eating disorder and the corresponding QEWP classification was reported by de Zwaan *et al.* [31]. Loeb *et al.* [32] demonstrated that once the diagnosis of bulimia nervosa has been established using an interview, self-report data, such as the Eating Disorder Examination, provide a useful basis to evaluate the progress and outcome of treatment studies.

## Conclusion

The reliable assessment of diagnoses and psychopathological syndromes is an important prerequisite for almost any kind of clinical research. The classification systems of DSM-IV and ICD-10, as well as the accompanying instruments, reflect the increasing advances in the field of diagnosis and methodology that have been made in the past 2 decades. These developments seem to constitute a solid and common framework for future research.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*, edn 4. Washington: American Psychiatric Association; 1994.
2. World Health Organization: *The ICD-10 Classification of Mental and Behavioural Disorders*. Geneva: World Health Organization; 1993.
3. Sartorius N, Kaelber CT, Cooper JE, Roper MT, Rae DS, Gulbinat W, Üstün TB, Regier DA: **Progress toward achieving a common language in psychiatry. Results from the field trial of the clinical guideline accompanying the WHO classification of mental and behavioural disorders in ICD-10.** *Arch Gen Psychiatry* 1993, 50:115-124.
4. Nathan PE: **DSM-IV: empirical, accessible, not yet ideal.** *J Clin Psychol* 1994, 50:103-109.
5. Zimmerman M, Jampala VC, Sierles FS, Taylor MA: **DSM-III and DSM-III-R: what are American psychiatrists using and why?** *Compr Psychiatry* 1993, 34:365-374.
6. Hiller W, Dichtl G, Hecht H, Hundt W, von Zerssen D: **An empirical comparison of diagnoses and reliabilities in ICD-10 and DSM-III-R.** *Eur Arch Psychiatry Clin Neurosci* 1993, 242:209-217.
7. Hiller W, Dichtl G, Hecht H, Hundt W, von Zerssen D: **Testing**
  - **the comparability of psychiatric diagnoses between ICD-10 and DSM-III-R.** *Psychopathology* 1994, 27:19-28.

Examines the magnitude and nature of differences between diagnoses according to ICD-10 and DSM-III-R in 100 psychiatric in-patients with severe mental disorders.

8. Hiller W, Dichtl G, Hecht H, Hundt W, Mombour W, von Zerssen D: **Evaluating the new ICD-10 categories of depressive episode and recurrent depressive disorder.** *J Affective Disorder* 1994, 31:49-60.

An empirical evaluation of the new ICD-10 concept of depressive disorder based on the severity of depressive syndromes (as in DSM-IV) rather than on the former endogenous versus neurotic dichotomy.

9. Segal DL, Hersen M, van Hasselt VB: **Reliability of the Structured Clinical Interview for DSM-III-R: an evaluative review.** *Compr Psychiatry* 1994, 35:316-327.  
A comprehensive review of the existing research concerning the reliability and validity of one of the most popular clinical interviews.
10. Ekselius L, Lindström E, von Knorring L, Bodlund O, Kullgren G: **SCID-II interviews and the SCID screen questionnaire as diagnostic tools for personality disorders in DSM-III-R.** *Acta Psychiatr Scand* 1994, 90:120-123.  
This analysis examines the usefulness of the SCID screening questionnaire and proposes a new adjustment method in order to predict the existence of one or more personality disorders.
11. Carey KB: **Use of the Structured Clinical Interview for DSM-III-R personality questionnaire in the presence of severe axis I disorders: a cautionary note.** *J Nerv Ment Dis* 1994, 182:669-671.
12. Janca A, Üstün TB, Sartorius N: **New versions of World Health Organization instruments for the assessment of mental disorders.** *Acta Psychiatr Scand* 1994, 90:73-83.  
The existing versions of the new crosscultural instruments released by the World Health Organization are described, which include the Composite International Diagnostic Interview, Schedules for Clinical Assessment in Neuropsychiatry, and the International Personality Disorder Examination.
13. World Health Organization: *Composite International Diagnostic Interview (CIDI)*. Washington: American Psychiatric Press; 1994.
14. World Health Organization: *Schedules for Clinical Assessment in Neuropsychiatry (SCAN)*. Washington: American Psychiatric Press; 1994.
15. Loranger AW, Sartorius N, Andreoli A, Berger P, Buchheim P, Channabasavanna SM, Coid B, Dahl A, Diekstra RFW, Ferguson B *et al.*: **The International Personality Disorder Examination (IPDE): The WHO/ADAMHA international pilot study of personality disorders.** *Arch Gen Psychiatry* 1994, 51:215-224.  
The design and results of an evaluation study are described. The International Personality Disorder Examination was administered by 58 psychiatrists and clinical psychologists to 716 patients at 14 centres in 11 countries.
16. Janca A, Üstün TB, Early TS, Sartorius N: **The ICD-10 symptom checklist: a companion to the ICD-10 classification of mental and behavioural disorders.** *Soc Psychiatry Psychiatr Epidemiol* 1993, 28:239-242.
17. Hiller W, Zaudig M, Mombour W, Bronisch T: **Routine psychiatric examinations guided by ICD-10 diagnostic checklists (International Diagnostic Checklists).** *Eur Arch Psychiatry Clin Neurosci* 1993, 242:218-223.
18. Gibbons RD, Clark DC, Kupfer DJ: **Exactly what does the Hamilton Depression Rating Scale measure?** *J Psychiatr Res* 1993, 27:259-273.
19. Domken M, Scott J, Kelly P: **What factors predict discrepancies between self and observer ratings of depression?** *J Affect Disord* 1994, 31:253-259.
20. Silverstone PH: **Poor efficacy of the Hospital Anxiety and Depression Scale (HAD) in the diagnosis of major depressive disorder in both medical and psychiatric patients.** *J Psychosom Res* 1994, 38:441-450.
21. Parker G, Hadzi-Pavlovic D, Sengoz A, Boyce P, Mitchell P, Wilhelm K, Hickie I, Brodaty H: **A brief self-report depression measure assessing mood state and social impairment.** *J Affect Disord* 1994, 30:133-142.
22. Thompson PA, Buckley PF, Meltzer HY: **The Brief Psychiatric Rating Scale: effect of scaling system on clinical response assessment.** *J Clin Psychopharmacol* 1994, 14:344-346.
23. Axelrod BN, Goldman RS, Woodard JL, Alphas LD: **Factor structure of the Negative Symptom Assessment.** *Psychiatry Res* 1994, 52:173-179.
24. Addington D, Addington J, Maticka-Tyndale E: **Specificity of the Calgary Depression Scale for schizophrenics.** *Schizophr Res* 1994, 11:239-244.  
Deals with the difficult differentiation between depressive, negative and extrapyramidal symptoms in patients with schizophrenia.
25. Altman EG, Hedeker DR, Janicak PG, Peterson JL, Davis JM: **The Clinician-Administered Rating Scale for Mania (CARS-M): development, reliability, and validity.** *Biol Psychiatry* 1994, 36:124-134.  
A new scale to assess manic syndromes is described, and advantages and disadvantages of different existing scales for mania are discussed.
26. Mazmanian D, Sharma V, Persad E, Kueneman K, Burnham J, Franklin J, Hemmings M, Leiska G: **Development and validation of a scale for rating mood states of psychiatric inpatients.** *Hosp Community Psychiatry* 1994, 45:238-241.
27. Kim SW, Dysken MW, Pheley AM, Hoover KM: **The Yale-Brown Obsessive-Compulsive Scale: measures of internal consistency.** *Psychiatry Res* 1994, 51:203-211.  
A thorough investigation of the dimensions of obsessive-compulsive disorder as represented in the Yale-Brown Scale.
28. Kaplan SI: **A self-rated scale for obsessive-compulsive disorder.** *J Clin Psychol* 1994, 50:564-574.  
An alternative approach to assess symptoms of obsessive-compulsive disorder by means of a questionnaire.
29. Piersma HL, Reaume WM, Boes JL: **The Brief Symptom Inventory (BSI) as an outcome measure for adult psychiatric in-patients.** *J Clin Psychol* 1994, 50:555-563.
30. Nangle DW, Johnson WG, Carr-Nangle RE, Engler LB: **Binge eating disorder and the proposed DSM-IV criteria: psychometric analysis of the Questionnaire of Eating and Weight Patterns.** *Int J Eating Disord* 1994, 16:147-157.  
An interesting paper that evaluates the value of a questionnaire related to eating behaviour and weight for the new proposed category of binge eating disorder.
31. de Zwaan M, Mitchell JE, Specker SM, Pyle RL, Mussell MP, Seim HC: **Diagnosing binge eating disorder: level of agreement between self-report and expert-rating.** *Int J Eating Disord* 1993, 14:289-295.
32. Loeb KL, Pike KM, Walsh BT, Wilson GT: **Assessment of diagnostic features of bulimia nervosa: interview versus self-report format.** *Int J Eating Disord* 1994, 16:75-81.

Wolfgang Hiller, Clinic Roseneck, Centre for Behavioural Medicine, D-83209 Prien, Germany.