DSM-III: Use of the Multiaxial Diagnostic System in Clinical Practice

Anne S. Bassett, M.D.¹ and Morton Beiser, M.D.²

¹Assistant Professor, Department of Psychiatry, University of Toronto, Toronto, Ontario

²Professor, Department of Psychiatry, University of British Columbia, Vancouver, British Columbia; Canada Health and Welfare National Health Research Scientist

Abstract

The authors report on the use of the DSM-III, several years after its introduction, in the clinical diagnosis of 154 subjects with first onset psychosis. Clinicians usually assigned Axis I diagnoses but used the remainder of the multiaxial system less than one time in three; if a standard recording form was in place, the multiaxial system was used more often. Trainees used the DSM-III most, followed by psychiatrists affiliated with a university and community based clinicians. Agreement between researchers and clinicians on diagnoses was fair to poor. The authors discuss the implications of the acceptance of the complex diagnostic system in routine clinical practice.

Since its publication in 1980, the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition* (DSM-III) (1) has had a major impact on psychiatric research (2). However, as Klerman noted, “in the long run, the acceptance of DSM-III by clinical practitioners will be the main determinant of its impact, independently of scholarly and research activities” (3). In recent surveys, American and Canadian psychiatrists have reported that they routinely use DSM-III terminology and two or more of the five axes (4–7). However, no publications to date have recorded the actual acceptance, as opposed to reported acceptance of the DSM-III multiaxial system in clinical practice. The accuracy with which the system is used in routine practice also remains an important area of inquiry. In this paper, we report patterns of use of DSM-III diagnoses in everyday clinical practice.

Method

The population chosen for this study consisted of 175 subjects from the University of British Columbia Markers and Predictors (MAP) study, all of whom experienced their first episode of psychosis during the two year period from 1982 to 1984. Details of the patient’s characteristics and the recruitment process used in the prospective cohort study are reported elsewhere (8). A research clinician conducted a semi-structured clinical interview, the Present State Examination (PSE) Ninth Edition (9), with each subject. In addition, research assistants obtained extensive information from each subject, from significant others and from referral sources. Data were evaluated by at least two clinicians using a “best estimate”

approach (10) to record the presence or absence of all the criteria relevant to each possible DSM-III diagnosis. Then, using a strict application of DSM-III rules, diagnoses were assigned to each subject. Research diagnoses were not communicated to the treating staff.

Typically, our research subjects encountered several treatment agencies during the first episode of their illness. For example, an initial assessment in a hospital emergency room was followed by transfer to an inpatient setting and finally by discharge to a community agency. In each of these settings, the patient probably received an admission and discharge assessment, thus generating a great many diagnoses during the episode, although these were not necessarily independent of one another. All hospital and community mental health agency records of the initial illness were obtained for 154 of the 175 MAP subjects. Charts were unavailable or could not be located for 21 subjects. Each admission and discharge diagnosis found on these charts was recorded. All diagnoses obtained in this manner were examined by the principal author to determine whether or not they could be classified into DSM-III categories and how systematically the multi-axial system was employed. Criteria used to arrive at these judgements were generous. For example, if it was unclear whether the principal diagnosis was made using the ICD-9 or the DSM-III, the diagnosis was assumed to reflect a DSM-III category. Any indication of use, including a notation of “no diagnosis” on an axis, was counted as an instance of use of the axis.

In addition to determining the frequency of use of the DSM-III, the clinical diagnoses recorded on 147 of the 154 subjects were used to assess agreement with the researchers' diagnoses. For seven of the subjects, no DSM-III diagnosis was recorded. For this evaluation, we assumed that the researchers' diagnoses represented the best possible application of DSM-III criteria, i.e., diagnoses arrived at by an extensive accumulation of data and strict adherence to DSM-III guidelines. Community based clinicians' diagnoses were compared to the researchers' assessments.

We also gathered information about diagnosticians and the settings in which the diagnoses were made. By checking diagnostician names against published professional listings, we established characteristics such as age, sex, professional background, level of training and whether or not the diagnostician was affiliated with a university. One of the hospitals introduced a standardized form to record DSM-III diagnoses mid-way through the intake period of the MAP study. This made possible a comparison of DSM-III use before and after the form was instituted in this particular setting.

Over half the diagnosticians were psychiatric residents and senior medical student interns in a clerkship rotation in psychiatry. All residents and medical students were based at university teaching hospitals: 43% were females and almost all were under age 40. In contrast, the majority of the practising psychiatrists who made diagnoses were over 40 years of age, only 15% were female and many worked in community or non university settings. Compared with psychiatrists affiliated with a university, those without university appointments were older, more likely to have been in practice longer and to have completed their psychiatric training outside North America. Almost all were working in clinical settings not affiliated with the university. A small number of diagnoses were made by nurses, social workers and psychologists.

Results

There were 587 opportunities (at admission, transfer and discharge) for the 154 patients in this study to receive a diagnosis. Clinicians actually recorded diagnoses on 523 occasions (89.1%). Table I examines the extent to which the DSM-III was used by the person diagnosing the subject — for example, all diagnoses made by psychiatrists in practice who had university appointments are arranged in column 1 and so forth. There were significant differences between groups of diagnosticians recording Axis I diagnoses. Residents, medical students and psychiatrists with a university affiliation almost always used DSM-III Axis I. In contrast, about one-third of the diagnoses made by psychiatrists without faculty appointments and by clinicians (usually nurses) in the “other” category either recorded no diagnosis or used another classification system, such as the ICD-9 or an idiosyncratic “impression”.

In the 523 cases where Axis I was apparently used, we examined the diagnosticians’ use of each of the axes independent of their use of the DSM-III as a whole. Thus, the denominator for the Axis II to V comparisons is the number of times the DSM-III was used, i.e., the number of Axis I diagnoses recorded.

Overall, diagnosticians used Axis II in less than one-third of the cases and Axis III just over one-quarter of the time. Axes IV and V were used less than one time in five. Medical students as a group most frequently used the multiaxial system, recording on Axes II through V over twice as often as psychiatrists with a university appointment. Psychiatrists not affiliated with a university and other clinicians rarely used any axis other than Axis I.

As noted, a standardized form for recording multiple axes of DSM-III was introduced in one of the settings. Use of Axes I and II or more increased from 27% ($n = 12$) to 55% ($n = 18$) after introduction of the form (chi-square = 6.31, $df = 1$, $p < 0.005$). At the same time, use of Axis I without other axes decreased from 73% ($n = 33$) to 45% ($n = 15$).

One estimate of the accuracy of DSM-III use is the frequency of appropriate recordings on each axis. Frequency of inappropriate use was 22.9% ($n = 120$) for Axis I, 5.4% ($n = 9$) for Axis II, 2.9% ($n = 4$) for Axis III, 29.0% ($n = 27$) for Axis IV and 14.6% ($n = 13$) for Axis V. Examples of inappropriate use of Axis I were terms such as “psychosis NYD” (psychosis not yet diagnosed) and “paranoid psychosis,” usually followed by a differential diagnosis using conventional DSM-III terminology. On Axes II through V, occasionally diagnoses belonging to other axes were recorded, for example, alcohol abuse on Axis III. Axes IV and V frequently contained lists of stressors and level of functioning respectively without any ratings. This was not considered inappropriate use in this study.

DSM-III Axis I is intended to be used for multiple recordings such as schizophrenic disorder or substance abuse (1). However, clinicians rarely made such additional diagnoses, even when a “substance-induced psychosis” was recorded in the differential diagnosis.

We compared the frequency with which major diagnostic categories were recorded by clinicians with the MAP researchers’ diagnoses. Results, presented in Table II, suggest that in routine clinical practice, the diagnosis of schizophrenia is assigned more often than in a

research setting with strict adherence to DSM-III criteria, whereas the researchers diagnosed affective disorder more often.

We also examined the level of agreement between community based clinicians' diagnoses and researchers' diagnoses. Because most patients had an inpatient hospital stay and were referred to the MAP project during that admission, DSM-III discharge diagnoses from those settings were selected as most comparable to the research diagnoses. Since all patients had had a psychotic episode, the four major categories of psychotic disorders were analyzed. Clinicians' diagnoses in each category were compared with researchers' diagnoses. Agreement was assessed using the kappa statistic (11). Results are presented in Table III. For comparison, we also present kappa coefficients obtained in field trials of the DSM-III and reported in the DSM-III handbook (1).

There was not much agreement between the clinicians and the researchers diagnoses and there was considerable variability among major diagnostic classes. The reliability of the clinicians' assessments of Axis I affective disorders was fair ($k = 0.53$). However, it was poor for schizophrenic disorders and very poor for psychotic disorders not elsewhere classified and paranoid disorders. The interrater reliability was also poor for Axes II and III. Kappa coefficients were not calculated for Axes IV and V because of their low numbers.

Discussion

Our findings suggest that clinicians use the multiaxial DSM-III system in their practices far less than they claim when surveyed. In clinical practice Axes II and III are used in less than one-third of the Axis I diagnoses, and Axes IV and V are rarely used. The frequency of Axes II to V use could be artificially lowered if many diagnoses assumed to be Axis I were actually ICD-9 diagnoses. However, maximum values from this study for use of multiple axes would still only be in the 55% to 60% range. A recent study of an Oregon state hospital reports a similar frequency for use of Axis III (12).

Clinicians in this study used non DSM-III generic labels such as "psychosis NYD" quite frequently. This is comparable to those in surgical specialties employing the term "abdominal pain NYD," a widely used but discouraged practice of identifying the primary complaint instead of the diagnosis. More distressing was the frequent omission of important additional diagnoses such as alcohol abuse (13). Treece (14) has suggested a separate assessment for each patient on substance related disorders. Instead of adding an axis, however, a standard DSM-III recording form might help overcome this problem by prompting the clinician to consider such additional diagnoses.

Mezzich (15) proposed a standard recording form for the DSM-III in 1982. Our study results suggest provision of a form does increase routine use of the DSM-III, as predicted by Williams (16) and implied in a recent article (12). The recording form serves as a prompt and encourages the use of the multiaxial system.

The estimates of reliability were better for affective disorders than for schizophrenic disorders and psychosis not elsewhere classified. This could be because patients with affective disorders present with more clear-cut or recognizable symptom profiles, or that

clinicians find the criteria for DSM-III affective disorders easier to use than for other psychotic conditions. Our results are consistent with other studies (6,17,18) which report that patients are often assigned diagnoses even if they do not satisfy all the necessary criteria. Traditionally, it has been difficult to obtain a consensus when defining schizophrenia (17,19). This study suggests that clinicians may be having difficulties using the DSM-III operational criteria intended for schizophrenic disorders. This is an even greater problem with atypical psychotic disorders having less established criteria. For example, schizophreniform disorder was a new classification in the DSM-III and the criteria probably needed refining for the DSM-III-R listing (20,21). Kappa coefficients for Axes II and III were also low.

Several factors limit the extent to which the findings of this study can be generalized. First, the estimate of the extent of use could be inflated because of the methods employed. When it was unclear whether a recorded diagnosis was made using the DSM-III or the ICD-9, it was assumed to have been the DSM-III. Diagnoses made subsequent to the initial recording may have been based on that first diagnosis. Inflated results could also be affected by the fact that these were all cases of first episode psychosis, a situation in which one might expect clinicians to exert the greatest care and attention when making a diagnosis. On the other hand, these results may underestimate current use, since they are from 1982 to 1984, when administrators in many community settings still expected psychiatrists to use the ICD-9. This study may reflect clinical preference of the multiaxial system, the use of which is reinforced in many settings.

Secondly, since the study is based on one Canadian centre, we may not be able to generalize results to other areas of Canada or the United States. While the Canadian Psychiatric Association does not officially endorse any one diagnostic classification system, survey results indicate DSM-III is the preferred system by Canadian psychiatrists (6,7). DSM-III manuals were widely distributed throughout Canada and the US from 1980 on. The University of British Columbia Department of Psychiatry embraced the DSM-III wholeheartedly in 1980, teaching only this system of classification to medical students and residents. The findings in this study reflect this, as the DSM-III has had its most noticeable impact on the diagnostic practices of medical students and residents, a result consistent with survey reports from the United States (5) and Canada (6).

The comparison of the clinicians' discharge diagnoses with the researchers diagnoses is limited by several considerations. Since research and clinical assessments did not take place at exactly the same time, clinicians and research teams may not have exactly the same data available upon which to base judgements. Differences in interviewer qualities are also known to be important factors in interrater reliability (22). Standard interviews were not used in clinical settings. Without standard interviews, Spitzer (23) has suggested that residents' diagnoses are probably not accurate enough for research purposes, and it is likely this is also true for practising psychiatrists (24). Some of the clinicians preferred diagnoses, assumed to be Axis I, could actually have been ICD-9 diagnoses. Each of these factors compromises interrater reliability.

Clinicians working in university affiliated settings use the DSM-III more often than clinicians in other settings. Perhaps over time there will be a fanning-out effect beginning with academics, progressing to clinicians affiliated with a university, and finally to practitioners in the community. Two to four years after the introduction of a new diagnostic manual may be too early for clinicians based in the community to embrace the system. Since this study was done, psychiatrists trained in university programs have taken key administrative positions in non university settings in greater Vancouver. These clinicians have encouraged the use of the DSM-III classification system for all patients. It is likely that an increased rate of use of the multiaxial system in non academic settings would be evident if this study were to be repeated now.

A majority of psychiatrists have indicated the multiaxial system is too complex or time consuming for day-to-day practice (6). Lack of knowledge about the clinical value of the information recorded on the multiaxial system may contribute to the low rate of use. Recent studies (15,21) suggest the Axis V rating is strongly correlated with decisions about hospital admissions and is a valuable predictor of future functioning in patients with a schizophreniform disorder. There is a widely held belief that a comprehensive approach to evaluation, such as the multiaxial system, is clinically useful (2,25,26). However, there is a need for more research to assess the actual value of each axis (1,21,22,27).

The primary goal of a diagnostic classification system in a clinical setting is to communicate information useful in making treatment and management decisions about patients. Despite drawbacks, the DSM-III is by common consensus the most comprehensive diagnostic system yet devised and the one most likely to improve the reliability of diagnosis and the clarity of communication. However, if busy clinicians are to adopt a complex system of diagnosis and use it consistently and appropriately, they must learn to use it, have confidence in its clinical value and have a standard recording form or compact visual reminder summarizing its intricacies. Our results indicate lower than expected rates of use and accuracy of the DSM-III, considering the high level of acceptance of the system in the academic setting. This suggests that, for the DSM-III to be clinically useful, the operational criteria and the multiaxial system need to be simplified.

We recommend replication of this study in other North American clinical settings. Future studies will be complicated by the recent introduction of the DSM-III-R (28). However, DSM-III-R retains the basic structure of the multiaxial diagnostic system. It remains important to determine how widely accepted the DSM-III format is for routine clinical practice and to estimate the time it takes for a new classification system to gain acceptance by clinicians in the community. Such data would undoubtedly be helpful in the planning of the DSM-IV, especially if major structural changes are proposed for the ever-evolving psychiatric nosological system.

Acknowledgments

The authors thank Janice Husted and Janet B.W. Williams for comments on the manuscript.

References

1. Diagnostic and statistical manual of mental disorders. 3. Washington DC: American Psychiatric Press, Inc; 1980.
2. Kendell, RE. DSM-III: a major advance in psychiatric nosology. In: Spitzer, RL. Williams, JBW., Skodol, AE., editors. International perspectives on DSM-III. Washington DC: American Psychiatric Press, Inc; 1983.
3. Klerman, GL. The significance of DSM-III in American psychiatry. In: Spitzer, RL. Williams, JBW., Skodol, AE., editors. International perspectives on DSM-III. Washington DC: American Psychiatric Press, Inc; 1983.
4. Jampala VC, Sierles FS, Taylor MA. Consumers' views of DSM-III: attitudes and practices of U.S. psychiatrists and 1984 graduating psychiatric residents. *Am J Psychiatry*. 1986; 143:148–153. [PubMed: 3946645]
5. Williams JBW, Spitzer RL, Skodol AE. DSM-III in residency training: results of a national survey. *Am J Psychiatry*. 1985; 142:755–758. [PubMed: 4003600]
6. Junek W, Leichner PP. Opinions of Canadian residents regarding the DSM-III. *Journal of Psychiatric Education*. 1986; 10:87–91.
7. Engels ML, Ghadirian AM, Dongier M. DSM-III: the view from the Canadian academe. *Psychiatr J Univ Ottawa*. 1985; 10:37–40.
8. Beiser M, Waxler-Morrison N, Iacono WG, et al. A measure of the “sick” label in psychiatric disorder and physical illness. *Soc Sci Med*. 1987; 25(3):251–261. [PubMed: 3629299]
9. Wing, JK., Cooper, JE., Sartorius, N. The measurement and classification of psychiatric symptoms. London: Cambridge University Press; 1974.
10. Leckman JF, Sholomskas D, Thompson WD, et al. Best estimates of lifetime psychiatric diagnosis: a methodologic study. *Arch Gen Psychiatry*. 1982; 39:879–883. [PubMed: 7103676]
11. Spitzer RL, Cohen J, Fleiss JL, et al. Quantification of agreement in psychiatric diagnosis: a new approach. *Arch Gen Psychiatry*. 1967; 17:83–87. [PubMed: 4952165]
12. Maricle R, Leung P, Bloom JD. The use of DSM-III Axis III in recording physical illness in psychiatric patients. *Am J Psychiatry*. 1987; 144:1484–1486. [PubMed: 3674233]
13. Skodol AE, Williams JBW, Spitzer RL, et al. Identifying common errors in the use of DSM-III through diagnostic supervision. *Hosp Community Psychiatry*. 1984; 35:251–255. [PubMed: 6706327]
14. Treece C. DSM-III as a research tool. *Am J Psychiatry*. 1987; 139:577–583.
15. Mezzich JE, Coffman GA, Goodpastor SM. A format for DSM- III diagnostic formulation: experience with 1,111 consecutive patients. *Am J Psychiatry*. 1982; 139:591–596. [PubMed: 7072843]
16. Williams JBW. The multi-axial system of DSM-III: where did it come from and where should it go? II. Empirical studies, innovations and recommendations. *Arch Gen Psychiatry*. 1985; 42:181–186. [PubMed: 3883940]
17. Lipkowitz MH, Idupuganti S. Diagnosing schizophrenia in 1980: a survey of U.S. psychiatrists. *Am J Psychiatry*. 1983; 140:52–55. [PubMed: 6847985]
18. Lipkowitz MH, Idupuganti S. Diagnosing schizophrenia in 1982: the effect of DSM-III. *Am J Psychiatry*. 1985; 142:634–637. [PubMed: 3985203]
19. Kendell RE. The choice of diagnostic criteria for biological research. *Arch Gen Psychiatry*. 1982; 39:1334–1339. [PubMed: 7138235]
20. Diagnostic and statistical manual of mental disorders. 3. Washington DC: American Psychiatric Press, Inc; 1987. revised
21. Beiser M, Fleming JAE, Iacono WG, et al. Refining the diagnosis of schizophreniform disorder. *Am J Psychiatry*. 1988; 145:695–700. [PubMed: 3369555]
22. Robins LN, Helzer JE. Diagnosis and clinical assessment: the current state of psychiatric diagnosis. *Ann Rev Psychol*. 1986; 37:409–432. [PubMed: 3963782]
23. Spitzer RL, Skodol AE, Williams JBW, et al. Supervising intake diagnosis: a psychiatric “Rashomon”. *Arch Gen Psychiatry*. 1982; 39:1299–1305. [PubMed: 7138231]

24. Helzer JE, Coryell W. More on DSM-III: how consistent are precise criteria? *Biol Psychiatry*. 1983; 18:1201–1202. [PubMed: 6652161]
25. Mezzich JE, Fabrega H, Mezzich AC, et al. International experience with DSM-III. *J Nerv Ment Dis*. 1985; 173:738–741. [PubMed: 4067596]
26. Williams JBW. The multiaxial system of DSM-III: where did it come from and where should it go? I. Its origins and critiques. *Arch Gen Psychiatry*. 1985; 42:175–180. [PubMed: 3883939]
27. Mezzich JE, Evanczuk KJ, Mathias RJ, et al. Admission decisions and multiaxial diagnosis. *Arch Gen Psychiatry*. 1984; 41:1001–1004. [PubMed: 6477051]
28. Rey JM. DSM-III-R: too much too soon? *Aust N Z J Psychiatry*. 1988; 22:173–182. [PubMed: 3041959]

Table 1

Overall Use of the DSM-III and its Multiaxial Systems by Type of Diagnostician

	Type of Diagnostician (%)				
	Practising Psychiatrists		Residents (n = 261) [†]	Medical Students (n = 78) [†]	Others (n = 63) [†]
	University Appointment (n=99) [†]	No University Appointment (n=86) [†]			
DSM-III (Axis I) used (n = 523; 89%)	90 (91)	59 (69)	256 (98)	77 (99)	41 (65)
DSM-III not used (n = 64; 11%)	9 (9)	27 (31)	5 (2)	1 (1)	22 (35)
Axis I (n= 523; 100%)	90* (100)	59* (100)	256* (100)	77* (100)	40* (100)
Axis II (n = 167; 32%)	26 (29)	4 (7)	82 (32)	48 (62)	7 (18)
Axis III (n = 140; 27%)	20 (22)	1 (2)	79 (31)	39 (51)	1 (3)
Axis IV (n = 93; 18%)	14 (16)	0 (0)	50 (20)	28 (36)	1 (3)
Axis V (n = 89; 17%)	12 (13)	0 (0)	50 (20)	27 (35)	0 (0)

[†] number of diagnoses assigned, chi square = 102, df = 4, p < 0.001

* number of Axis I diagnoses used as denominator

Table II

Frequency of Axis I Major Diagnoses

<u>Diagnosis</u>	<u>Clinical Settings</u>	<u>MAP Research</u>
Schizophrenia	48	39
Major affective disorder	47	66
Other psychosis	34	35
Paranoid psychosis	5	7
Other diagnosis	13	0

Table III

Comparison of Kappa Coefficients

<u>Diagnosis</u>	<u>Kappa Coefficients</u>	
	<u>Current Study</u>	<u>DSM-III Field Trials Phase One</u>
Axis I		
• Schizophrenic disorders	0.260	0.810
• Affective disorders	0.530	0.690
• Psychotic disorders not elsewhere classified	0.084	0.640
• Paranoid disorders	0.046	0.660
• Overall (above disorders)	0.210	
Axis II	0.240	0.560
Axis III	0.410	not done