Does the Adequacy of Clinicians' Diagnostic Practice in Routine Clinical Settings Matter?

Mark Zimmerman, MD

The report by Nakash and colleagues[1] raises concerns about the adequacy of the unstructured, unstandardized, clinical evaluation that is conducted in routine clinical practice. The results suggest that clinicians are not sufficiently thorough in determining whether a disorder's *DSM-IV* criteria are fully met. Are these findings a cause for concern?

More than 25 years ago, when I was still a medical student, my colleagues and I conducted a survey of psychiatrists' attitudes about *DSM-III* and *DSM-III-R*.[1] This was done in anticipation of the publication of *DSM-IV*. We asked psychiatrists why they used the *DSM*, and only a minority rated the *DSM*s as being very important for planning treatment, determining prognosis, managing patients, and understanding patients' problems.

Fast forward 25 years. I have been an attending psychiatrist for 20 years, as well as the principal investigator of the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project, a study in which we have integrated the assessment tools and procedures of researchers into a hospital-affiliated outpatient practice. One focus of our research has been on a comparison of diagnoses made by clinicians using unstandardized, unstructured interviews and researchers using standardized, structured interview schedules. An early report[3] found that diagnostic comorbidity was less frequently identified by a routine clinical evaluation compared to an evaluation including the Structured Clinical Interview for *DSM-IV* (SCID).[4] We examined the issue of comorbidity because comorbidity predicts poorer outcome for patients with depressive and anxiety disorders, and the presence of multiple psychiatric disorders is associated with greater levels of psychosocial impairment. We therefore assumed that it was important to recognize comorbid disorders and that failure to do so would result in poorer outcome. Subsequent to our publication, 3 independent research groups replicated our finding that more diagnoses are made when a semistructured interview is used compared to an unstructured clinical interview.[5–7]

Problems with diagnostic recognition are not limited to Axis I disorders. In another report from the MIDAS project, we found that diagnostic interviewers using the Structured Interview for *DSM-IV* Personality[8] were much more likely to diagnose borderline personality disorder than clinicians using

an unstructured interview.[9] Moreover, when the information from the semistructured interview was provided to the clinician, clinicians then diagnosed borderline personality disorder more frequently; thus, it was not simply a matter of clinicians' reluctance to diagnose borderline personality disorder during the initial diagnostic evaluation.

While several of the initial articles from the MIDAS project identified problems with the detection of disorders, with regard to the diagnosis of bipolar disorder, we observed an opposite phenomenon—clinician overdiagnosis. While several research reports have suggested that bipolar disorder is underrecognized, and that many patients, particularly those with major depressive disorder, in fact have bipolar disorder,[10–14] there is also evidence of overdiagnosis.[15,16] The largest study of overdiagnosis and underdiagnosis of bipolar disorder was done in the MIDAS project.[17] More than half of the patients who reported that they had been previously diagnosed with bipolar disorder were not diagnosed with bipolar disorder based on the SCID. Underdiagnosis of bipolar disorder also occurred, but 3 times as many patients were overdiagnosed with bipolar disorder than underdiagnosed. Family history analyses supported the validity of the diagnostic procedures. In a follow-up to our initial article on bipolar disorder overdiagnosis, we examined whether there was a particular diagnostic profile associated with bipolar disorder overdiagnoses.[18] The patients overdiagnosed with bipolar disorder were 4 times more likely to be diagnosed with borderline personality disorder compared to patients who were not diagnosed with bipolar disorder.

Because I am the director of the outpatient division in the Rhode Island Hospital Department of Psychiatry, you might ask whether we have instituted a policy in which all patients are evaluated with a semistructured diagnostic interview in order to enhance the comprehensiveness, reliability, and validity of our diagnostic procedure. In fact, I have not recommended such a change to our routine clinical practice because no research has yet examined the clinical significance of the gap between researchers' and clinicians' diagnostic practices. Specifically, I am not aware of any studies that have addressed the critical question of whether more accurate and comprehensive research diagnostic evaluations improve outcomes in a heterogeneous sample of patients. One study found that when the results of the SCID are presented to clinicians, treatment regimens often change,[19] but the study did not examine whether outcome was improved. No research group has done the important study in which patients are randomly assigned to receive a semistructured interview or an evaluation as usual and

then examined as to whether outcome is superior in the "intervention" group. The primary reason we have not sought funding for such a study is that we are skeptical that it would get funded in a climate that prioritizes biological research and that has increasingly deemphasized the importance of *DSM*-type diagnosis.[20]

One needs to consider that perhaps the clinicians surveyed 25 years ago were correct and the approach toward psychiatric diagnosis embodied by the recent iterations of the *DSM* is not particularly helpful in treating and managing the majority of psychiatric outpatients presenting for treatment. One could argue that patients' outcomes are *not* more likely to be worse, even in the face of missed diagnoses or misdiagnosis, because of the broad spectrum of activity of the new-generation medications. Medications such as selective serotonin reuptake inhibitors and serotonin-norepinephrine reuptake inhibitors have been found to be effective for depression, almost all anxiety disorders, eating disorders, impulse-control disorders, substance use disorders, attention deficit disorder, and some somatoform disorders. Atypical antipsychotics are helpful in nonbipolar as well as bipolar depression, and there is some evidence of benefit in anxiety disorders; of course, they also are effective in treating psychosis. In short, most of the disorders for which individuals seek outpatient care have been found to be responsive to at least 1 of the new generation of antidepressant or antipsychotic medications. Thus, it is possible that accurate and comprehensive diagnostic evaluations are not critical after gross diagnostic class distinctions (eg, psychotic disorder vs mood disorder vs substance use disorder) are made.

Nevertheless, on the grounds of common sense, it seems logical that greater diagnostic accuracy will improve outcome. If a clinician is unaware of a comorbid condition's presence, then isn't it less likely to be successfully treated? If borderline personality disorder is misdiagnosed as bipolar disorder, or bipolar disorder misdiagnosed as major depressive disorder, then isn't outcome likely to be poorer? More complete and accurate diagnostic evaluations might influence whether a medication is prescribed (eg, an antidepressant for major depressive disorder but not adjustment disorder), choice of medication (eg, a selective serotonin reuptake for a depressed patient with a comorbid obsessive-compulsive disorder), the class of medication prescribed (eg, a mood stabilizer rather than an antidepressant for bipolar depression), and the prescription of psychotherapy (eg, dialectical behavior therapy for borderline personality disorder; cognitive behavioral therapy rather than supportive therapy for a specific anxiety disorder).

One can also hypothesize that more complete and accurate knowledge of patients' psychiatric disorders might improve outcome independent of changes in treatment decisions. That is, better diagnostic practice may result in greater patient satisfaction with the diagnostic assessment, an improved alliance with the treating clinician, and subsequent greater compliance with treatment and better outcome.

If future research demonstrates that comprehensive research evaluations are better, either by improving outcome

or by improving the prediction of outcome, then what can or should be done to change the standard of care regarding diagnostic evaluations? One possibility is the incorporation of semistructured diagnostic evaluations into routine clinical practice. The advantage of using instruments such as the SCID or the Mini-International Neuropsychiatric Interview[21] is that they enhance diagnostic thoroughness, reliability, and validity. Even if greater reimbursement were provided for the additional time needed to conduct the evaluation, it might prove difficult to change how clinicians conduct their initial diagnostic interview in the absence of compelling data warranting such a change.

At the beginning of the 21st century, psychiatric diagnoses continue to be almost entirely determined by clinical history taking. Though this may change in the future, no biological tests, as yet, are sufficiently accurate to be used to make a psychiatric diagnosis. However, there are inexpensive low-tech paper-and-pencil tests, and higher tech computer-administered diagnostic assessments, that might improve clinicians' diagnostic practice. In light of the consistent evidence that there are problems with the routine clinical evaluation, clinicians should consider using such adjunctive assessment tools while waiting for the results of studies demonstrating their clinical significance.

So what have we done in our clinical practice to hopefully improve diagnostic performance? While we believe that diagnostic thoroughness and accuracy are important, at the outset of the MIDAS project we assumed that it was unlikely that semistructured interviews would routinely be incorporated into clinical practice. Therefore, one of the early goals of the MIDAS project was to develop a diagnosis-oriented self-report questionnaire that would help clinicians use their time more efficiently and maintain or improve their level of diagnostic accuracy. We developed the Psychiatric Diagnostic Screening Questionnaire (PDSQ), a self-report scale designed to screen for the most common Axis I disorders encountered in outpatient mental health settings.[22–24] The advantage of an empirically developed measure over a homegrown form is that the psychometric and diagnostic properties of scientifically studied instruments have been established, thereby guiding the interpretation of the results. Brief self-report questionnaires cannot substitute for clinical evaluations and render definitive diagnoses; therefore, the questionnaire we constructed is referred to as a screening, rather than as a diagnostic, instrument. It was our hope that the PDSQ would improve the efficiency of the diagnostic evaluation by guiding clinicians toward symptom areas that required more versus less assessment. The clinicians in our practice believe this to be the case and therefore ask patients to complete the measure before the intake evaluation. Whether measures such as the PDSQ can improve diagnostic accuracy, and consequently improve outcome, remains an empirical question.

In concluding, it is worth reflecting on the change in the discourse on diagnosis over the past 35 years. *DSM-III*, the first officially sanctioned diagnostic system to incorporate specified inclusion and exclusion criteria for psychiatric

diagnosis, was published a generation ago. The empirical justification for the radical change in how psychiatric disorders were defined was based on the studies documenting problems with diagnostic reliability when diagnoses were based on earlier systems[25] and other studies demonstrating that high levels of reliability could be achieved when diagnoses were derived from semistructured interviews and based on specific criteria.[26] Validity was not so much an issue, except for some studies that demonstrated that the more narrow definition of schizophrenia proposed for *DSM-III* was more valid than *DSM-II*'s broader definition.[27] The clinical utility of the new diagnostic system was assumed, though improvement in patient outcomes was not the focus of attention. Nearly 35 years after the publication of *DSM-III*, research suggests that significant problems remain with psychiatric diagnosis in routine clinical practice. The study by Nakash et al[1] is another to raise such questions. During the past 35 years, we have also witnessed a revolution in the treatment of psychiatric disorders. Pharmacotherapies and psychotherapies have been repeatedly demonstrated to be effective for a wide range of *DSM-III/DSM-III-R/DSM-IV/DSM-5*–defined disorders. This would suggest that it is important for accurate diagnoses to be made. However, the broad-based efficacy of various treatments suggests that diagnostic precision might not be so important. The next generation of research on diagnosis (eg, changes in diagnostic criteria or changes in diagnostic practice) will hopefully attend to the most salient aspect of psychiatric treatment—the outcome of care.

This commentary has focused on diagnostic precision and completeness and raised the question of whether improvement in diagnostic practice will result in improved outcome and/or outcome prediction. While a diagnostic determination is an important function of the intake evaluation, it is not the sole objective. Other integral functions of the intake evaluation include additional history taking (eg, past psychiatric history, prior treatment efforts, medical history, life events, social supports, coping style, family history, developmental history), education about the disorder and treatment options, establishment of a therapeutic alliance, and identification of obstacles of treatment. Striving for improved diagnostic practice should not come at a cost of sacrificing the important details of a patient's story in order to make a diagnosis or diagnoses.[28] It has been our experience during the MIDAS project that the nondiagnostic functions of the initial evaluation, such as the establishment of the therapeutic relationship, are enhanced, rather than undermined, by good diagnostic practice. Good diagnostic practice, therefore, is likely to be associated with greater patient satisfaction and increased retention in treatment. This too can, and should, be the subject of empirical study.

REFERENCES

1. Nakash O, Nagar M, Kanat-Maymon Y. Clinical use of the *DSM* categorical diagnostic system during the mental health intake session. *J Clin Psychiatry*. 2015;76(7):e862–e869.
2. Zimmerman M, Jampala VC, Sierles FS, et al. *DSM-III* and *DSM-III-R*: what are American psychiatrists using and why? *Compr Psychiatry*. 1993;34(6):365–374.
3. Zimmerman M, Mattia JI. Psychiatric diagnosis in clinical practice: is comorbidity being missed? *Compr Psychiatry*. 1999;40(3):182–191.
4. First MB, Spitzer RL, Williams JBW, et al. *Structured Clinical Interview for DSM-IV (SCID)*. Washington, DC: American Psychiatric Association; 1997.
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(10):1599–1605.
6. Miller PR, Dasher R, Collins R, et al. Inpatient diagnostic assessments, 1: accuracy of structured vs unstructured interviews. *Psychiatry Res*. 2001;105(3):255–264.
7. Shear MK, Greeno C, Kang J, et al. Diagnosis of nonpsychotic patients in community clinics. *Am J Psychiatry*. 2000;157(4):581–587.
8. Pfohl B, Blum N, Zimmerman M. *Structured Interview for DSM-IV Personality*. Washington, DC: American Psychiatric Press, Inc; 1997.
9. Zimmerman M, Mattia JI. Differences between clinical and research practices in diagnosing borderline personality disorder. *Am J Psychiatry*. 1999;156(10):1570–1574.
10. Manning JS, Haykal RF, Connor PD, et al. On the nature of depressive and anxious states in a family practice setting: the high prevalence of bipolar II and related disorders in a cohort followed longitudinally. *Compr Psychiatry*. 1997;38(2):102–108.
11. Ghaemi SN, Boiman EE, Goodwin FK. Diagnosing bipolar disorder and the effect of antidepressants: a naturalistic study. *J Clin Psychiatry*. 2000;61(10):804–808, quiz 809.
12. Ghaemi SN, Ko JY, Goodwin FK. "Cade's disease" and beyond: misdiagnosis, antidepressant use, and a proposed definition for bipolar spectrum disorder. *Can J Psychiatry*. 2002;47(2):125–134.
13. Hantouche EG, Akiskal HS, Lancrenon S, et al. Systematic clinical methodology for validating bipolar-II disorder: data in mid-stream from a French national multi-site study (EPIDEP). *J Affect Disord*. 1998;50(2–3):163–173.
14. Perugi G, Akiskal HS, Lattanzi L, et al. The high prevalence of "soft" bipolar (II) features in atypical depression. *Compr Psychiatry*. 1998;39(2):63–71.
15. Hirschfeld RM, Cass AR, Holt DC, et al. Screening for bipolar disorder in patients treated for depression in a family medicine clinic. *J Am Board Fam Pract*. 2005;18(4):233–239.
16. Goldberg JF, Garno JL, Callahan AM, et al. Overdiagnosis of bipolar disorder among substance use disorder inpatients with mood instability. *J Clin Psychiatry*. 2008;69(11):1751–1757.
17. Zimmerman M, Ruggero CJ, Chelminski I, et al. Is bipolar disorder overdiagnosed? *J Clin Psychiatry*. 2008;69(6):935–940.
18. Zimmerman M, Ruggero CJ, Chelminski I, et al. Psychiatric diagnoses in patients previously overdiagnosed with bipolar disorder. *J Clin Psychiatry*. 2010;71(1):26–31.
19. Kashner TM, Rush AJ, Surís A, et al. Impact of structured clinical interviews on physicians' practices in community mental health settings. *Psychiatr Serv*. 2003;54(5):712–718.
20. Insel T, Cuthbert B, Garvey M, et al. Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167(7):748–751.
21. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for *DSM-IV* and *ICD-10*. *J Clin Psychiatry*. 1998;59(suppl 20):22–33, quiz 34–57.
22. Zimmerman M, Mattia JI. The reliability and validity of a screening questionnaire for *13 DSM-IV* Axis I disorders (the Psychiatric Diagnostic Screening Questionnaire) in psychiatric outpatients. *J Clin Psychiatry*. 1999;60:677–683.
23. Zimmerman M, Mattia JI. The Psychiatric Diagnostic Screening Questionnaire: development, reliability and validity. *Compr Psychiatry*. 2001;42(3):175–189.
24. Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. *Arch Gen Psychiatry*. 2001;58(8):787–794.
25. Spitzer RL, Fleiss JL. A re-analysis of the reliability of psychiatric diagnosis. *Br J Psychiatry*. 1974;125(4):341–347.
26. Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria: rationale and reliability. *Arch Gen Psychiatry*. 1978;35(6):773–782.
27. Pope HG Jr, Lipinski JF Jr. Diagnosis in schizophrenia and manic-depressive illness: a reassessment of the specificity of "schizophrenic" symptoms in the light of current research. *Arch Gen Psychiatry*. 1978;35(7):811–828.
28. Tucker GJ. Putting *DSM-IV* in perspective. *Am J Psychiatry*. 1998;155(2):159–161.