The Reliability and Validity of a Screening Questionnaire for 13 DSM-IV Axis I Disorders (the Psychiatric Diagnostic Screening Questionnaire) in Psychiatric Outpatients

Mark Zimmerman, M.D., and Jill I. Mattia, Ph.D.

Background: The purpose of this study was to examine the reliability and validity of a new multidimensional screening instrument for 13 DSM-IV Axis I disorders.

Method: The Psychiatric Diagnostic Screening Questionnaire (PDSQ) is a 90-item self-administered questionnaire that screens for 13 DSM-IV disorders in 5 areas (eating, mood, anxiety, substance use, and somatoform disorders). A consecutive series of 500 psychiatric outpatients completed the PDSQ immediately before their intake evaluation. Seventy-four patients completed the scale a second time less than a week after the initial administration, and 51 patients completed a booklet of questionnaires that included established measures of the same symptom domains assessed by the PDSQ.

Results: The PDSQ subscales achieved moderate-to-high levels of internal consistency (mean Cronbach's α coefficient = 0.82) and test-retest reliability (mean correlation coefficient = 0.84). Subscale scores were significantly associated with blind clinical diagnoses, and individual PDSQ items correlated much more highly with their own subscale than with other subscales. The PDSQ subscales were much more highly correlated with established measures of the same symptom domain (mean correlation coefficient = 0.72) than with measures of other types of psychopathology (mean correlation = 0.17).

Conclusion: The PDSQ is a reliable and valid measure of multiple DSM-IV disorders that is brief enough to be incorporated into routine clinical outpatient practice without disruption, yet lengthy enough to be a psychometrically sound instrument.

(J Clin Psychiatry 1999;60:677–683)

Received July 8, 1996; accepted Dec. 29, 1998. From the Department of Psychiatry and Human Behavior, Brown University School of Medicine, Rhode Island Hospital, Providence.

Supported in part by National Institute of Mental Health grant MH-48732 (Dr. Zimmerman).

The authors thank Ava Nepaul, Melissa Torres, and Sharon Younkin for assistance in collecting the data.

Reprint requests to: Mark Zimmerman, M.D., Bayside Medical Center, 235 Plain St., Providence, RI 02905 (e-mail: MZimmerman@lifespan.org).

A thorough psychiatric evaluation covering all the major domains of Axis I psychopathology as well as family, medical, developmental, and psychosocial history is time consuming. An adequate mental status examination must have a sufficiently broad basis to include questions about the symptoms of many specific mental disorders to determine which disorders are present.

Because completeness is time consuming, it is incompatible with cost containment, the impetus behind the recent changes in the delivery of health care services. Insurance company strictures are resulting in shorter diagnostic interviews in ambulatory mental health settings, and it is logical that the frequency of diagnostic errors will increase if insufficient time is permitted for a comprehensive mental status examination. Because of the pressures to spend less time evaluating patients, clinicians could benefit from a tool that would help them use their time more efficiently and maintain or improve their level of diagnostic accuracy.

An inexpensive method of collecting reliable and valid clinical information is the self-administered questionnaire. Questionnaires are commonly used by physicians in all branches of medicine to collect medical histories prior to a patient's initial evaluation. Questionnaires have long been used in the mental health field to evaluate personality, mood, psychosocial functioning, and so on. During the past 10 years, some questionnaires have been designed to screen for or "diagnose" single DSM Axis I disorders such as major depression or bulimia.[1,2] Several questionnaires have been developed to assess all of the DSM Axis II disorders.[3–8] However, no scales have been developed to screen psychiatric patients efficiently for a broad range of DSM-IV Axis I disorders.

A clinically useful multidimensional screening instrument should be brief enough to be completed by patients in a timely manner before their intake evaluation, yet comprehensive enough to cover the most common disorders for which patients seek treatment. It should also be long enough so that it is psychometrically sound. Finally, its scoring and organization should be simple enough so that a clinician or office worker can rapidly review and score the inventory and obtain clinically useful informa-

tion. The present article is the first report of the reliability and validity of such a scale, the Psychiatric Diagnostic Screening Questionnaire (PDSQ), in a psychiatric outpatient sample.

METHOD

Patients presenting for an intake evaluation at the Rhode Island Hospital Department of Psychiatry outpatient clinic (Providence, R.I.) were asked to complete the PDSQ as part of their initial paperwork. When scheduling their appointments, the patients were told to arrive early to complete some standard forms. The PDSQ took approximately 10 to 15 minutes to complete, and its administration did not disrupt routine clinical practice. Because we were planning to test the PDSQ's validity by examining the relationship between its subscale scores and the clinicians' diagnoses, the clinicians were kept blind to the patients' responses on the measure.

Almost all (96%, 480/500) intake evaluations were conducted by board certified or board eligible psychiatrists. The remaining evaluations were conducted by clinical nurse specialists or master's level social workers. Diagnoses were based on DSM-IV criteria. Patients' charts were reviewed by research assistants who recorded demographic information, Axis I diagnoses, and Global Assessment of Functioning (GAF) ratings.

Test-retest reliability of the PDSQ was examined in 93 patients evaluated by the senior author and one other psychiatrist in the department. These patients were given the scale at the conclusion of the intake evaluation and asked to mail it back in a preaddressed postage-paid envelope. They were told that the purpose of the second administration was to test the performance of the scale, not to question the truthfulness or accuracy of their responses. Twenty-nine patients completed the second PDSQ later on the same day of their intake evaluation, whereas the longest test-retest interval was 33 days by 1 patient. The completion date of the second test was the date marked on the questionnaire by the patient. The mean ± SD interval between completion of the 2 PDSQs was 3.5 ± 4.9 days. Because several studies have demonstrated that test-retest reliability of state measures of psychopathology decreases when the testing interval increases,[9–11] we examined reliability for the 74 patients who completed the second PDSQ less than a week after the first testing. Patients provided informed consent to complete the scale a second time.

To examine the concurrent validity of the PDSQ subscales, 51 patients completed a booklet of questionnaires at home that included measures of symptoms related to bulimia (Eating Disorder Inventory Bulimia Subscale[12]), depression (Diagnostic Inventory for Depression; available from the author on request), social phobia (Brief Fear of Negative Evaluation Scale,[13] Fear Questionnaire-social phobia subscale[14]), agoraphobic fears (Fear Questionnaire-agoraphobia subscale,[14] Social Phobia and Anxiety Inventory-agoraphobia subscale[15]), posttraumatic stress (Posttraumatic Stress Disorder Scale[16]), obsessive-compulsive behavior (Obsessive Compulsive Disorder Scale[17]), cognitions common in generalized anxiety (Penn State Worry Questionnaire[18]), anxiety symptoms common in panic attacks (Beck Anxiety Inventory[19]), alcohol use (Michigan Alcoholism Screening Test[20]), drug use (Drug Abuse Screening Test[21]), hypochondriasis (Whitely Index[22]), and somatization (Somatic Symptom Index[23,24]). Most of these scales have been commonly used in research, and their reliability and validity have been well established. For 3 validity indices (social phobia, agoraphobia, and somatization), we used the combined score of 2 brief measures of the construct. The depression and obsessive-compulsive disorder measures are recently developed scales currently being studied by their authors, and preliminary unpublished findings suggest that they are both reliable and valid. Patients provided informed consent to complete the booklet of questionnaires.

Table 1. Composition of the Psychiatric Diagnostic Screening Questionnaire (PDSQ)

PDSQ Subscale	No. of Items
Bulimia nervosa	3
Major depressive disorder	22
Obsessive-compulsive disorder	9
Posttraumatic stress disorder	6
Panic disorder	4
Agoraphobia	10[a]
Social phobia	10[a]
Alcohol abuse/dependence	6
Drug abuse/dependence	6
Generalized anxiety disorder	5
Somatization disorder	4
Hypochondriasis	3
Body dysmorphic disorder[b]	4

[a]Two items are common to these 2 subscales. See text for further description.
[b]The body dysmorphic disorder subscale was an experimental subscale that was appended to the PDSQ.

The PDSQ is a 90-item self-report questionnaire that screens for 13 DSM-IV disorders in 5 areas: eating disorders, mood disorders, anxiety disorders, substance use disorders, and somatoform disorders. Earlier versions of the scale were studied in 2 large samples of primary care medical patients.[25–27] Table 1 lists the number of items on each of the PDSQ subscales.

The depression subscale assesses each of the 9 DSM-IV symptom criteria for major depressive disorder. For compound depression criteria that have multiple components (e.g., the DSM-IV sleep disturbance criterion refers to both increased and decreased sleep), the PDSQ includes a separate question for each component. The agoraphobia and social phobia section on the PDSQ begins with a stem phrase inquiring about situations that are so anxiety provoking as to cause problems, followed by a list of 16 common agoraphobic and social phobic situations. This is fol-

lowed by 2 questions that assess overall fear and avoidance and contribute to the total score of both subscales. The 9-item obsessive-compulsive disorder scale covers the varied ways this disorder can present itself and asks specific questions regarding the more common cognitions and behaviors that can occur with the disorder. The remaining subscales assess the core features of their related DSM-IV disorders and are between 3 and 6 items in length. All questions are answered yes or no.

For 4 disorders (bulimia, depression, obsessive-compulsive, and panic), the PDSQ's questions refer to the past 2 weeks. For phobic, substance use, generalized anxiety, and somatoform disorders, the time frame of the questions is the past 6 months. Two of 6 screening questions for posttraumatic stress disorder (PTSD) refer to a lifetime history of experiencing or witnessing a traumatic event, and the remaining 4 questions inquire about PTSD symptoms within the previous 2 weeks.

For each of the PDSQ subscales, we examined 2 types of reliability. Because the PDSQ's questions refer to current (and recent) psychiatric state, we limited the analysis of test-retest reliability to the 74 patients who completed the scale the second time less than a week after the initial administration. Internal consistency, which is an estimate of scale homogeneity, was evaluated for each subscale with Cronbach's alpha.

The items of a multidimensional scale assessing distinct dimensions need to demonstrate discriminant and convergent validity at both the item and subscale level. At the item level, the correlation between an item and its parent subscale should be greater than the correlation between the item and the other subscales. For example, each item on the obsessive-compulsive disorder (OCD) subscale should be more highly correlated with the total OCD subscale score than with the total scores of other subscales, such as alcohol abuse or somatoform disorder. To determine item-convergent and item-discriminant validity, we calculated item-subscale correlation coefficients between each item and every PDSQ subscale. When computing the item-subscale coefficient between the item and its own parent subscale, the contribution of the item score to the subscale score was eliminated. Across all items on a particular subscale, we computed the mean of the item–parent subscale correlations and the mean of the item–other subscale correlations. Thus, for the OCD subscale we compared the mean of the 9 correlations between each OCD item and the OCD total score (item–parent subscale) with the mean of the 108 correlations between each OCD item and the other 12 subscale total scores (item–other subscale).

We also examined discriminant validity of subscales in the more traditional sense[28] by comparing the correlation between the PDSQ subscales and other measures of the same construct and different constructs. For example, the PDSQ bulimia subscale should correlate more highly with the Eating Disorder Inventory Bulimia Scale than with measures of alcoholism, depression, phobias, and so on.

To compare patients with and without a particular clinical diagnosis on the diagnosis-specific PDSQ subscale score (e.g., the patients with and without major depressive disorder were compared on the PDSQ depression scale), t tests were used. Subscale scores were counted as missing if more than 20% of the items were not answered by the patients.

Table 2. Demographic and Clinical Characteristics of 500 Psychiatric Outpatients

Characteristic	N	%
Gender		
Male	187	37.4
Female	313	62.6
Education[a]		
< 12 years	40	9.3
High school graduate or GED	116	27.1
Some college	134	31.3
College graduate	138	32.3
Marital status		
Married	247	49.4
Living with someone	21	4.2
Widowed	19	3.8
Separated	29	5.8
Divorced	66	13.2
Never married	118	23.6
Race		
White	472	94.4
Black	11	2.2
Hispanic	8	1.6
Asian	3	0.6
Other	6	1.2
Age (y)		
18–29	108	21.6
30–44	248	49.6
45–64	109	21.8
65 and over	35	7.0
DSM-IV diagnosis[b]		
Major depressive disorder	258	51.6
Bipolar depression	11	2.2
Dysthymic disorder	54	10.8
Generalized anxiety disorder	31	6.2
Panic disorder without agoraphobia	17	3.4
Panic disorder with agoraphobia	45	9.0
Social phobia	16	3.2
Specific phobia	4	0.8
Obsessive-compulsive disorder	12	2.4
Posttraumatic stress disorder	36	7.2
Adjustment disorder	48	9.6
Schizophrenia	1	0.2
Bulimia nervosa	5	1.0
Alcohol abuse/dependence	27	5.4
Drug abuse/dependence	15	3.0
Attention deficit/hyperactivity disorder	13	2.6

[a]Data missing on education for 72 patients.
[b]Individuals could be given more than one diagnosis.

RESULTS

During the 7 months of the study, 96.2% (537/558) of the patients who were asked to complete the PDSQ did so. Nine patients refused, 2 were not English speaking, 1 was

Table 3. Reliability of the PDSQ in Psychiatric Outpatients

PDSQ Subscale	Internal Consistency[a]	Test-Retest Correlation[b]
Bulimia nervosa	0.69	0.87
Major depressive disorder	0.89	0.90
Obsessive-compulsive disorder	0.79	0.86
Posttraumatic stress disorder	0.83	0.86
Panic disorder	0.82	0.73
Agoraphobia	0.81	0.91
Social phobia	0.82	0.93
Alcohol abuse/dependence	0.87	0.94
Drug abuse/dependence	0.90	0.80
Generalized anxiety disorder	0.84	0.72
Somatization disorder	0.74	0.81
Hypochondriasis	0.74	0.83
Body dysmorphic disorder	0.86	0.74

[a]Cronbach's alpha is calculated only when all of the items on a scale are answered. Because of missing items, the sample sizes for the Cronbach's alpha computation were as follows: bulimia nervosa, N = 477; major depressive disorder, N = 405; obsessive-compulsive disorder, N = 471; posttraumatic stress disorder, N = 472; panic disorder, N = 457; agoraphobia, N = 438; social phobia, N = 429; alcohol abuse/dependence, N = 488; drug abuse/dependence, N = 491; generalized anxiety disorder, N = 481; somatization disorder, N = 469; hypochondriasis, N = 481; body dysmorphic disorder, N = 490.
[b]Seventy-four patients completed the PDSQ within a week of the first administration. Because of missing items on some of the PDSQ subscales at one of the timepoints, the sample sizes varied between 64 and 74. All test-retest correlation coefficients are significant at $p < .001$.

Table 4. Discriminant and Convergent Validity of the PDSQ Items in 500 Psychiatric Outpatients

	Mean of Item-Scale Correlations	
PDSQ Subscale	Own Scale[a]	Other Scales[b]
Bulimia nervosa	0.53	0.14
Major depressive disorder	0.47	0.16
Obsessive-compulsive disorder	0.49	0.21
Posttraumatic stress disorder	0.61	0.17
Panic disorder	0.65	0.26
Agoraphobia	0.49	0.19
Social phobia	0.49	0.18
Alcohol abuse/dependence	0.68	0.04
Drug abuse/dependence	0.74	0.07
Generalized anxiety disorder	0.64	0.24
Somatization disorder	0.54	0.21
Hypochondriasis	0.58	0.19
Body dysmorphic disorder	0.72	0.24

[a]The value represents the mean of the item-scale correlations between each item and its own scale. When computing the total scale score, the contribution of the item to the total was eliminated.
[b]The value represents the mean of the item-scale correlations between each item on the subscale and all of the other subscales. For example, the bulimia subscale includes 3 items, and each item was correlated with the other 12 subscales. The mean of these 36 correlations was 0.14.

mentally retarded, 7 had visual or other physical limitations, and 2 were too confused or mentally ill to complete the scale. Another 37 questionnaires were excluded from the analysis because more than 10% of the items were not completed. In one third (13/37) of these cases, the patients failed to complete the back of the 2-sided questionnaire. This left a final sample of 500 usable forms. There were no demographic or clinical differences between the patients who did and did not complete the scale.

Table 2 shows the demographic and diagnostic characteristics of the sample. The majority of the patients were white, female, married or never married, and had some college education. The mean ± SD age of the sample was 39.6 ± 13.1 years, and the mean GAF score (rated in 375 patients) was 57.4 ± 11.1. The most frequent DSM-IV diagnoses assigned by clinicians were major depressive disorder (51.6%), dysthymic disorder (10.8%), adjustment disorder (9.6%), panic disorder (12.4%) and PTSD (7.2%). More than one third (38.4%) of patients received more than 1 Axis I diagnosis, and 7.6% received 3 or more diagnoses. Other diagnoses that were made but are not summarized in Table 2 include eating disorder not otherwise specified (NOS) (N = 3, 0.6%), cyclothymia (N = 1, 0.2%), depressive disorder NOS (N = 21, 4.2%), depressive disorder due to a general medical condition (N = 9, 1.8%), anxiety disorder due to a general medical condition (N = 1, 0.2%), anxiety disorder NOS (N = 7, 1.4%), schizoaffective disorder (N = 3, 0.6%), psychotic disorder NOS (N = 2, 0.4%), pedophilia (N = 1, 0.2%), and intermittent explosive disorder (N = 3, 0.6%).

Reliability

The PDSQ subscales had adequate-to-good levels of internal consistency (Table 3). Cronbach's alpha was greater than 0.70 for all subscales except the bulimia subscale, and 9 of the 13 subscales achieved alpha values above 0.80. The mean of the alpha coefficients was 0.82.

The test-retest correlation coefficients were greater than 0.70 for all subscales and greater than 0.80 for 10 of the 13 subscales (see Table 3). The mean of the test-retest correlation coefficients was 0.84.

Item-Discriminant and Item-Convergent Validity

To examine the discriminant and convergent validity of the diagnosis-specific subscales, each PDSQ item was correlated with its own parent subscale (removing the variance contributed by that individual item to the subscale total) and with all other subscales. Almost all items (96.7%) had a higher correlation with their parent subscale than with any other subscale. The mean item–parent subscale correlations were 2 to 10 times higher than the item–other subscale correlations (Table 4). Across all 13 scales, the mean correlation between the items and their own scale was 0.59, and the mean correlation with the other scales was 0.18.

Discriminant and Convergent Validity of the PDSQ Subscale Scores

Fifty-one patients completed a package of questionnaires at home a mean of 2.3 ± 3.2 days after the intake evaluation. All but 2 of these patients were in the test-retest reliability subsample; thus, they completed a second PDSQ as part of the questionnaire packet. For most

Table 5. Convergent and Discriminant Validity of PDSQ Subscales

PDSQ Subscale	Correlation With Validity Scale[a]	Mean Correlation With Other Scales[b]
Bulimia nervosa	0.77	0.12
Major depressive disorder	0.87	0.27
Obsessive-compulsive disorder	0.84	0.27
Posttraumatic stress disorder	0.64	0.07
Panic disorder	0.57	0.17
Agoraphobia	0.86	0.29
Social phobia	0.71	0.28
Alcohol abuse/dependence	0.68	0.02
Drug abuse/dependence	0.38	0.03
Generalized anxiety disorder	0.91	0.19
Somatization disorder	0.65	0.21
Hypochondriasis	0.79	0.14

[a]All correlation coefficients are significant at $p < .001$. The respective validity scales are listed in the Method section. Data for the body dysmorphic disorder subscale were not available.
[b]The value represents the mean of the 11 correlations between each PDSQ subscale and the scales used to validate the other 11 PDSQ subscales.

Table 6. Diagnosis-Specific PDSQ Subscale Scores in Patients With and Without the Diagnosis

Clinical Diagnosis	N With Diagnosis[a]	Scores for Patients With the Diagnosis		Scores for Patients Without the Diagnosis	
		Mean	SD	Mean	SD
Bulimia nervosa	5	2.8	0.4	0.3[d]	0.7
Depression[b]	266	12.4	4.7	8.0[d]	5.1
Obsessive-compulsive disorder	11	3.8	1.9	1.8[d]	2.1
Posttraumatic stress disorder	36	4.4	1.6	1.5[d]	1.9
Panic disorder[c]	57	2.9	1.4	1.6[d]	1.5
Agoraphobia	44	4.3	2.7	1.9[d]	2.3
Social phobia	16	5.4	2.7	2.1[d]	2.4
Alcohol abuse/dependence	26	4.0	2.0	0.4[d]	1.1
Drug abuse/dependence	15	2.7	3.0	0.2[d]	0.9
Generalized anxiety disorder	30	4.0	1.6	2.9[d]	1.9
Somatization disorder	1				
Hypochondriasis	0				
Body dysmorphic disorder	0				

[a]Only patients with a definite diagnosis were included. If the diagnosis was described as "probable" or "rule out," then it was counted as absent. The sample sizes are less than sample sizes reported in Table 2 because of missing data on the PDSQ.
[b]Includes major depressive disorder and bipolar depression.
[c]Includes panic disorder with or without agoraphobia.
[d]The patients with the diagnosis scored higher than the patients without the diagnosis at the $p < .001$ level.

patients, pharmacologic treatment was initiated at the intake evaluation. Because the PDSQ measures current psychiatric state, and this might change after several days or a week of treatment, we examined the association between the time-2 PDSQ scores and the other questionnaires rather than using the time-1 PDSQ. The data in Table 5 show that the PDSQ subscale scores were significantly correlated with other measures of the same symptom domain. Every PDSQ subscale was most highly correlated with the validity scale assessing the same construct. Seven of these 12 correlations were greater than 0.70 (see Table 5), and the mean of the correlations between the PDSQ subscales and the concordant validation scales was 0.72. In contrast, the mean of the 132 correlations between the PDSQ subscales and the nonconcordant measures of psychopathology was 0.17. Thus, the PDSQ subscales demonstrated very high levels of discriminant and convergent validity.

Association With Psychiatric Diagnosis

For each of the disorders assessed by the PDSQ, we compared the mean diagnosis-specific subscale score in patients with and without that DSM-IV diagnosis. The data in Table 6 show that, for patients with each diagnosis, the corresponding PDSQ subscale score was significantly elevated.

DISCUSSION

To our knowledge, the PDSQ is the first self-administered questionnaire explicitly developed to screen for several DSM-IV Axis I disorders in psychiatric patients. Longer, multidimensional questionnaires such as the Minnesota Multiphasic Personality Inventory-2[29] and the Millon Clinical Multiaxial Inventory-II[30] have been used as diagnostic aids; however, they were not designed to be congruent with the current diagnostic nomenclature. Moreover, these inventories are too long and their scoring too time consuming to be routinely completed and scored in an office waiting area before the initial evaluation. Scales have been developed to detect specific DSM Axis I disorders such as depression,[1] PTSD,[16] and bulimia,[2] but they are limited to only one type of pathology. Recent attempts to develop tools to assist primary care providers with the recognition and diagnosis of psychiatric disorders in primary care patients have yielded mixed results. The Symptom Driven Diagnostic System[31] is a brief self-report questionnaire that screens for 5 Axis I disorders. The scale has poor screening properties (e.g., in the authors' cross-validation study, average sensitivity for the 5 disorders covered = 56%; average specificity = 81.2%), and its psychometric properties have not been established. The Primary Care Evaluation of Mental Disorders[32] includes a brief self-report Patient Questionnaire (PQ) and an interview guide for the clinician. Eighty percent of primary care patients screen positive on the PQ, indicating that the measure has low specificity and poor positive predictive value.

Our preliminary findings suggest that the PDSQ is a reliable and valid measure that can be incorporated into routine clinical outpatient practice without disruption. The subscales achieved high levels of internal consistency and test-retest reliability. The subscale scores were significantly associated with clinicians' diagnoses that were

made blind to PDSQ information, and the items correlated much more highly with their own parent subscale than with the other subscales. The PDSQ diagnosis-specific subscales demonstrated excellent discriminant validity, correlating much more highly with established measures of the same constructs than with measures of other types of psychopathology. While the mean correlation between the PDSQ diagnosis-specific subscales and the validation scales was high (.72), 2 of these correlations were below .60. The lowest correlation was between the PDSQ drug abuse scale and the Drug Abuse Screening Test (DAST). Closer inspection of the data revealed that only 2 of the 49 patients who completed the DAST were diagnosed with a drug use disorder, and on the PDSQ drug abuse subscale, only 4 patients scored above zero. Thus, the study needs to be replicated in a sample with a higher frequency of drug use disorders to evaluate this subscale more fully. The only other correlation below .60 was between the PDSQ panic subscale and the Beck Anxiety Inventory (BAI). It is likely that this relatively low correlation is due to the nonspecificity of the BAI for panic anxiety. The PDSQ directly asks about panic attacks, whereas the BAI evaluates general cognitive and somatic anxiety symptoms that might occur in other anxiety disorders. Unfortunately, the current state of the field offers no more specific self-report assessment of panic disorder symptomatology.

The PDSQ assessed most of the Axis I disorders typically evaluated by semistructured research interviews. Noteworthy exceptions are anorexia nervosa, psychotic disorders, and mania. Anorexia was included in the earlier versions of the PDSQ tested in primary care settings. However, we found upwards of 50% of primary care patients responded yes to screening questions about preoccupation about weight, belief of being fat despite what others say, and intense fear of becoming fat. Moreover, the psychometric properties of the anorexia subscale (internal consistency, discriminant and convergent validity) were consistently the poorest of all of the PDSQ subscales. Consequently, we dropped this subscale from the PDSQ. We did not assess mania or psychosis on the PDSQ because these disorders are more frequently present in inpatient than outpatient settings, and we were skeptical that it was possible to briefly screen for these symptoms on a self-administered questionnaire. We are currently pilot testing new 5- and 7-question subscales that screen for psychosis and mania, respectively. We have also developed a 7-item dysthymia subscale, and we will examine if it can distinguish dysthymic disorder from major depressive disorder.

The psychiatric diagnoses made in this study were determined by unstructured clinical evaluations. These diagnoses were then used to validate the PDSQ subscales. It would have been preferable to use semistructured research interviews to make patients' diagnoses, because these research instruments improve diagnostic reliability, and their structured and comprehensive format ensures inquiry for all diagnoses.[33,34] Because of the potentially greater diagnostic error associated with clinical diagnosis, we did not determine cutoff points on the PDSQ subscales to identify cases (i.e., individuals who screened positive for a disorder). At the present time, we would recommend that clinicians follow up positive responses on the scale. Patients who endorse symptoms of the syndromes assessed by the measure should be questioned about those symptoms. In the future, we will develop thresholds to distinguish individuals who do and do not screen positive on a particular subscale. The development of such thresholds involves striking a balance between diagnostic sensitivity and specificity; consequently, even after case-defining thresholds are empirically established, we would recommend that clinicians review patients' item responses on the questionnaire.

A potentially important clinical application of the PDSQ is the detection of psychopathology that otherwise would not have been ascertained during a routine clinical evaluation. Although there was evidence of considerable diagnostic comorbidity in our sample, the rate of comorbidity was lower in this study than in research studies using semistructured interviews.[35–37] For example, only 34.5% of the patients diagnosed with major depression were diagnosed with a comorbid disorder. On the PDSQ, however, more than half of these patients reported many symptoms of nondepressive disorders. Because comorbidity predicts poorer outcome,[38,39] it is important to study whether this "undiagnosed" pathology has prognostic significance.

The delivery of mental health services is rapidly evolving, and 2 changes in particular may compromise the accuracy of psychiatric diagnoses in the clinical setting. Management and fiscal pressures have forced clinicians to reduce the time devoted to evaluating their patients; consequently, it is often not possible to conduct a thorough anamnesis and diagnose potentially important disorders comorbid with the primary disorder. A second financially driven change in the delivery of mental health services is the more frequent use of paraprofessionals such as social workers and nurse clinicians to conduct intake evaluations. These professionals do not receive as much training in psychiatric diagnosis as do psychologists and psychiatrists, and the risk of diagnostic errors may be increased.

Can an inexpensive and valid diagnostic screening test help reduce clinical diagnostic errors? This can be tested by comparing the agreement between clinicians' diagnoses made in the usual clinical manner or made after reviewing their patients' responses to a screening questionnaire such as the PDSQ with diagnoses based on a comprehensive research diagnostic interview. If the screening test is a useful diagnostic aid, then the clini-

cians' rate of agreement with the research interview should increase.

Although the results of this study are encouraging, replication and extension to samples with different demographic and clinical characteristics are necessary. The present sample comprised insured, acutely ill psychiatric outpatients who were predominantly white. Few patients had chronic psychotic disorders, and relatively few had substance use disorders. The performance of the PDSQ in settings that service more severe populations (e.g., community health centers, day hospital programs) needs to be demonstrated. Similarly, its utility in primary care settings, where the severity of disorders is milder, warrants study.

REFERENCES

1. Zimmerman M, Coryell W, Corenthal C, et al. A self-report scale to diagnose major depressive disorder. Arch Gen Psychiatry 1986;43:1076–1081
2. Thelen MH, Farmer J, Wonderlich S, et al. A revision of the Bulimia Test: the BULIT-R. Psychol Assess 1991;3:119–124
3. Hyler SE, Reider RD, Williams JBW, et al. The Personality Diagnostic Questionnaire: development and preliminary results. J Pers Disord 1988; 2:229–237
4. Burgess JW. The Personality Inventory Scales: a self-rating clinical instrument for diagnosis of personality disorder. Psychol Rep 1991;69: 1235–1246
5. Coolidge FL, Merwin MM. Reliability and validity of the Coolidge Axis II Inventory: a new inventory for the assessment of personality disorders. J Pers Assess 1992;59:223–238
6. Morey LC. Personality Assessment Inventory. Lutz, Fla: Psychological Assessment Resources; 1991
7. Klein MH, Benjamin LS, Rosenfeld R, et al. The Wisconsin Personality Disorders Inventory I: development, reliability, and validity. J Pers Disord 1993;7:285–303
8. Clark LA. Manual for the Schedule for Normal and Abnormal Personality. Minneapolis, Minn: University of Minnesota Press; 1992
9. Zimmerman M, Coryell W. The validity of a self-report questionnaire for diagnosing major depressive disorder. Arch Gen Psychiatry 1988;45: 738–740
10. Duncan-Jones P, Henderson S. The use of a two-phase design in a prevalence study. Soc Psychiatry 1978;13:231–237
11. Wing JK, Nixon JM, Mann SA, et al. Reliability of the PSE (9th ed) used in a population study. Psychol Med 1977;1:505–516
12. Garner DM, Olmstead MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. Int J Eat Disord 1983;2:15–34
13. Leary MR. A brief version of the Fear of Negative Evaluation Scale. Pers Soc Psychol Bull 1983;9:371–375
14. Marks IM, Mathews AM. Brief standard self-rating for phobic patients. Behav Res Ther 1979;17:263–267
15. Turner SM, Beidel DC, Dancu CV, et al. An empirically derived inventory to measure social fears and anxiety: The Social Phobia and Anxiety Inventory psychological assessment. Psychol Asses J Consult Clin Psychol 1989;1:35–40
16. Foa EB, Riggs DS, Dancu CV, et al. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. J Trauma Stress 1993;6:459–473
17. Coles M, Bogert KV, Krause MS, et al. The Obsessive-Compulsive MCP-Oxford Scale. Presented at the 30th annual convention of the Association for the Advancement of Behavior Therapy; November 1996; New York, NY
18. Meyer TJ, Miller ML, Metzger RL, et al. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990;28:487–495
19. Beck AT, Brown G, Epstein N, et al. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56: 893–897
20. Selzer ML. The Michigan Alcoholism Screening Test: the quest for a new diagnostic instrument. Am J Psychiatry 1971;127:1653–1658
21. Skinner HA. The Drug Abuse Screening Test. Addict Behav 1982;7: 363–371
22. Pilowsky I. Dimensions of hypochondriasis. Br J Psychiatry 1967;113: 89–93
23. Othmer E, DeSouza C. A screening test for somatization disorder (hysteria). Am J Psychiatry 1985;142:1146–1149
24. Swartz M, Hughes D, George L, et al. Developing a screening index for community studies of somatization disorder. J Psychiatr Res 1986;20: 335–343
25. Zimmerman M, Farber NJ, Hartung J, et al. Screening for psychiatric disorders in medical patients: a feasibility and patient acceptance study. Med Care 1994;2:603–608
26. Zimmerman M, Lish JD, Farber NJ, et al. Screening for depression in medical patients: is the focus too narrow? Gen Hosp Psychiatry 1994;16: 388–396
27. Zimmerman M, Lish JD, Lush DT, et al. Suicidal ideation among urban medical outpatients. J Gen Intern Med 1995;10:573–576
28. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait multi-method matrix. Psychol Bull 1959;56:81–105
29. Butcher JN, Graham JR, Williams CL, et al. Development and use of the MMPI-2 Content Scales. Minneapolis, Minn: University of Minnesota Press; 1990
30. Millon T. Manual for the Millon Clinical Multiaxial Inventory-II (MCMI-II). Minneapolis, Minn: National Computer Systems; 1987
31. Broadhead WE, Leon AC, Weissman MM, et al. Development and validation of the SDDS-PC screen for multiple mental disorders in primary care. Arch Fam Med 1995;4:211–219
32. Spitzer RL, Williams JBW, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. JAMA 1994;272:1749–1756
33. Endicott J, Spitzer RL. A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia (SADS). Arch Gen Psychiatry 1978;35: 837–844
34. Spitzer RL, Williams JBW, Gibbon M, et al. Structured Clinical Interview for DSM-III-R (SCID). Washington, DC: American Psychiatric Association; 1990
35. Goisman RM, Goldenberg I, Vasile RG, et al. Comorbidity of anxiety disorders in a multicenter anxiety study. Compr Psychiatry 1995;36:303–311
36. Sanderson WC, DiNardo PA, Rapee RM, et al. Syndrome comorbidity in patients diagnosed with a DSM-III-R anxiety disorder. J Abnorm Psychol 1990;3:308–312
37. Sanderson WC, Beck AT, Beck J. Syndrome comorbidity in patients with major depression or dysthymia: prevalence and temporal relationships. Am J Psychiatry 1990;147:1025–1028
38. Brown TA, Barlow DH. Comorbidity among anxiety disorders: implications for treatment and DSM-IV. J Consult Clin Psychol 1992;6:835–844
39. Keller MB, Klerman GL, Lavori PW, et al. Long-term outcomes of episodes of major depression. JAMA 1984;252:788–792