

ASSESSING DEPRESSIVE SYMPTOMS IN FIVE PSYCHIATRIC POPULATIONS: A VALIDATION STUDY

MYRNA M. WEISSMAN, DIANE SHOLOMSKAS, MARGARET POTTENGER,
BRIGITTE A. PRUSOFF AND BEN Z. LOCKE¹

Weissman, M. M. (Yale University School of Medicine, Department of Psychiatry, Depression Research Unit, Connecticut Mental Health Center, New Haven, CT 06519), D. Sholomskas, M. Pottenger, B. A. Prusoff and B. Z. Locke. Assessing depressive symptoms in five psychiatric populations: A validation study. *Am J Epidemiol* 106:203-214, 1977.

Data from five psychiatric populations and a community sample are presented on the CES-D, a 20-item self-report depression symptom scale developed by the Center for Epidemiologic Studies. Results show that the scale is a sensitive tool for detecting depressive symptoms and change in symptoms over time in psychiatric populations, and that it agrees quite well with more lengthy self-report scales used in clinical studies and with clinician interview ratings. Although a symptom scale cannot differentiate between diagnostic groups, the CES-D has demonstrated its validity as a screening tool for detecting depressive symptoms in psychiatric populations.

depression; psychiatric status rating scales; psychiatry

In 1971 a series of studies were initiated by the Center for Epidemiologic Studies (CES) to develop epidemiologic techniques for the continuous measurement of psychiatric impairment, particularly depressive symptomatology, in the community (1). As part of the procedures, the CES developed a 20-item self-report symptom rating scale (CES-D) to measure depressive mood in the community (2, 3). This scale was used in household surveys sponsored by the

CES in Kansas City, Missouri, from October, 1971, through January, 1973, and in Washington County, Maryland, from December, 1971, through March, 1973; over 3800 individuals participated (4-7).

Self-administered questionnaires to determine frequency and severity of depressive symptoms in the community have been used recently by other investigators (8-11). Self-reports are attractive because they are economical, they do not require extensive rater training, and they avoid the problem of observer bias.

However, in any survey assessing psychiatric symptoms in the community in which psychiatric interviews and diagnostic classification are not used, one of the principal problems is translating from the self-report symptom scales to what the psychiatrist considers a case. Instruments used in community surveys frequently treat psychiatric morbidity as a unitary phenomenon, quantitative in terms of severity but without distinction among discrete syndromes (12). This concept is different from the one used by clinical psychiatrists where weights given to symptoms

Received for publication January 20, 1977, and in final form May 24, 1977.

Abbreviations: CES, Center for Epidemiologic Studies; CES-D, 20-item self-report depression symptom scale developed by CES; DSM-II, Diagnostic and Statistical Manual II, American Psychiatric Association, NHI, New Haven Schizophrenia Index; SCL-90, Symptom Check List, 90-item self-report questionnaire.

From the Yale University School of Medicine, Department of Psychiatry, Depression Research Unit, Connecticut Mental Health Center, 904 Howard Avenue, Suite 2A, New Haven, CT 06519 (reprint requests to Dr. Weissman at this address).

¹ Center for Epidemiologic Studies, National Institute of Mental Health, Rockville, MD.

Supported by ADAMHA Contract #ADM-42-74-83 (DBE) from the Center for Epidemiologic Studies, National Institute of Mental Health.

are not necessarily linear and symptoms are classified into diagnostic groups or dimensions.

As a result of these differences in determining a case, there has been a gap between epidemiologic and clinical studies in psychiatry and a problem in translating epidemiologic findings into clinical terms. Studies distinguishing among diagnoses or examining rates for a single diagnosis have for the most part relied upon hospital or clinical populations and ignored community populations. Alternately, studies which have used quantitative measures of community psychiatric impairment have been ignored in ascertaining rates of psychiatric illness.

The Kansas City and Washington County community surveys provided a rich source of epidemiologic data on the self-reported symptoms of depression in a large heterogeneous, randomly-selected population. While the CES-D, the main method of assessing depressive symptoms, was a composite of several well-known depression scales, it was a new scale, it was a subject self-report, and it was not a diagnostic assessment (2). Therefore, it was unclear how the scale scores on the CES-D related to scales used in clinical studies of depression, or how patients who were diagnosed and under treatment for depression or other psychiatric disorders in clinical studies would score on it. Since there are no specific laboratory tests to confirm the diagnosis of depression, the validity of the scale depends ultimately on clinicians' judgment. Therefore, it was important to determine if subjects who were considered as having depressive symptoms on the CES-D (i.e., reaching over a certain score) would be judged as clinically depressed by a clinician. Moreover, was the scale specific to depression? Would patients diagnosed as depressed score higher on the symptom scale than patients with other diagnoses, or did the scale merely measure overall psychiatric impairment? Would

psychiatric patients with other diagnoses also score high on the CES-D?

The data presented in this paper will partially answer these questions. We will compare results on the CES-D obtained in a community sample with those obtained in five psychiatric populations, including patients diagnosed and treated for acute depression. We will also compare the results of the CES-D in these populations with those obtained with other standard depression scales.

Specifically, we are interested in the validity of the CES-D as well as its utility as a screening instrument for case finding. Concurrent validity will be determined by its ability to discriminate between diagnosed psychiatric patients and a community sample; to differentiate among different types of psychiatric patients; and by agreement with a variety of other types of symptom scales. Discriminant validity will be demonstrated by lack of agreement with variables for which differences should be evident. Utility as a screening instrument will be tested by comparing cases as defined by a certain cut-off score against cases defined by a clinical criterion and determining false-positive and negative rates.

MATERIALS AND METHODS

Community sample. The community sample was comprised of 3845 randomly-selected adults, aged 18 years and over, drawn from Kansas City, Missouri, and Washington County, Maryland. Full details of the sampling scheme and method of procedure have been described elsewhere (5).

Psychiatric patients. All psychiatric patients came from treatment facilities of the Connecticut Mental Health Center, affiliated with the Yale University Department of Psychiatry. There were 406 psychiatric patients divided into five groups. These included: acutely depressed patients; recovered formerly depressed patients; drug-

addicted patients; alcoholics; and schizophrenics. Consecutive male and female subjects between the ages of 18 and 65 who were receiving outpatient treatment were included. All patients were interviewed after the study had been explained to them and they had agreed to participate by signing a consent form.

Acutely depressed patients. One hundred forty-eight acutely depressed patients coming to the Depression Research Clinic for outpatient treatment of acute depression of at least moderate severity were included. The criteria for admission for treatment included a depression of sufficient intensity to reach a total score of 7 or more on the Raskin Three Area Depression Scale (range 3–15), based on the patient's verbal report, behavior and secondary symptoms of depression (13). Patients were excluded if the depression was secondary to another predominant syndrome, such as schizophrenia, alcoholism or drug addiction, either currently or by history. Persons of subnormal intelligence, and persons having serious physical illness or other conditions which precluded taking of psychotropic medication were also excluded.

Recovered depressed patients. Recovered depressives were 87 patients who had had a documented acute depressive episode which had been treated at the Depression Research Clinic and from which they had recovered. For inclusion as recovered depressives, they must have been currently asymptomatic as assessed by a Raskin score of less than 7. The original inclusion criteria for treatment of depression was the same as described above for acute depressives. Since these patients had been part of the Depression Research Clinic studies, considerable data on their symptoms and social functioning during an acute depressive illness were available and the acute depressive episode was well documented.

Drug addicts. The drug addicted sub-

jects were 60 patients currently maintained on methadone as part of the treatment program of the Drug Dependence Unit. All patients had a history of prior serious opiate abuse, most commonly heroin.

Alcoholics. The alcoholics included in this study were 61 patients coming for outpatient treatment specifically for alcoholism. Eligibility for treatment included the overt recognition by the patient, and agreement of the physician, that the patient's dysfunction was primarily a drinking problem, and the patient was motivated to withdraw from alcohol. Patients were selected and interviewed at the point they were determined eligible for admission to the treatment program.

Schizophrenics. The schizophrenic subjects were 50 patients with an unequivocal diagnosis of schizophrenia, according to the American Psychiatric Association DSM-II, and the New Haven Schizophrenia Index (NHSD) (14). Therefore, they must have had at least one documented episode of hallucinations or delusions. They were currently in outpatient treatment receiving a phenothiazine and group therapy for their schizophrenic illness at the Community Support Clinic.

Assessments

Clinical symptoms were assessed by clinician ratings and by patient self-report on the following scales:

Clinician rating scales

Hamilton Rating Scale (15). The Hamilton Rating Scale is a widely-used 17-item scale completed by a clinician and based on information elicited from the patient during an interview. The items are measured on a 3- or a 5-point scale. A total score is obtained by summing the scores of the individual items. Therefore, the total score range is 0 to 62, higher score indicating more impairment. Mean total scores in ambulatory acutely depressed patients are

usually about 20. Five factor scores have been derived which cover the dimensions of depression.

Raskin Depression Scale (13, 14). The Raskin Depression Scale is the clinician's assessment of the patient made during an interview with the patient and covering three areas: the patient's verbal report; behavior; and secondary symptoms of depression. Each area is rated on a 5-point scale and the scores are summed to yield a total score of 3-15. A score of 7 or higher is considered a depression of sufficient severity to be treated with psychopharmacologic agents. Mean total pretreatment scores in acutely depressed ambulatory patients are about 10 (16).

Self-report scales

Symptom Checklist (SCL-90) (17, 18). The Symptom Checklist (SCL-90) is a self-report rating scale oriented towards the symptomatic behavior of psychiatric outpatients. It is derived from the Hopkins Symptoms Checklist (18). It is comprised of 90 items rated on a 5-point scale. The items reflect nine primary symptom dimensions that are believed to underlie the majority of symptom behaviors observed in these patients (17).

CES-D. The CES-D scale consists of 20 items from previously developed scales. The items were selected to represent the major symptoms in the clinical syndrome of depression, as identified by clinical judgment, frequency of use in other questionnaires for depression, and factor analytic studies. Reliability and discriminatory power were also taken into account wherever such data were available. Major contributors to the CES-D scale were Zung's depression scale (19); a depression inventory developed by Beck (20); a self-report rating of depression by Raskin (13); the depression scale of the Minnesota Multiphasia Inventory (21); and a scale developed by Gardner (22).

Of the component scales, only the Minnesota Multiphasia Inventory has been

widely used among general populations. The scales by Zung and Beck have been validated in clinical groups. The CES-D scale as a whole has now been validated in several different clinical populations, including a comparison with a general population (23-25).

The instrument is completed by the patient and oriented around symptoms of depression. It asks for feelings during the week preceding the interview. The scale range of answers is from 0 to 3; in all cases, but for four questions, higher score indicates more impairment. For those four questions, the scores are reversed. A total score for the scale is made by summing all items for each patient. The total score has a possible range of 0 to 60, and this single total score is used as an estimate of the degree of depressive symptomatology. Based on findings from the Kansas City and Washington County community surveys, scores of 16 or more (the lower bound of the approximate upper quintile of scores for the combined population) were considered "cases" of depression (5). Persons with a score of 16 or more must have had at least six of the 20 symptoms in the CES-D with persistence for most of the previous week or a majority of the symptoms on the scale for shorter periods of time.

Table 1 shows the 20-item scale. The full details of its development, and reliability and pretesting have been described elsewhere (2, 7).

Raters. The assessments were made by two social workers with considerable clinical experience, and previous training and experience in the use of these scales. Both social workers underwent conjoint training in the use of the scales prior to beginning the study until they achieved high interrater agreement. Agreement between raters was periodically checked.

Method of administration. In order to standardize ratings, the scales were always administered in the order listed above. The research assistant remained with the patient while the self-report

TABLE 1

Twenty-item self-report depression symptom scale developed by the Center for Epidemiologic Studies (CES-D)

INSTRUCTIONS FOR QUESTIONS: Below is a list of the ways you might have felt or behaved. Please tell me how often you have felt this way during the past week.

- Rarely or none of the time (less than 1 day)
- Some or a little of the time (1-2 days)
- Occasionally or a moderate amount of time (3-4 days)
- Most or all of the time (5-7 days)

During the past week:

- 1 I was bothered by things that usually don't bother me
- 2 I did not feel like eating; my appetite was poor
- 3 I felt that I could not shake off the blues even with help from my family or friends.
- 4 I felt that I was just as good as other people.
- 5 I had trouble keeping my mind on what I was doing.
- 6 I felt depressed
- 7 I felt that everything I did was an effort.
- 8 I felt hopeful about the future.
- 9 I thought my life had been a failure.
- 10 I felt fearful
- 11 My sleep was restless.
- 12 I was happy.
- 13 I talked less than usual
- 14 I felt lonely.
- 15 People were unfriendly.
- 16 I enjoyed life.
- 17 I had crying spells
- 18 I felt sad
- 19 I felt that people dislike me
- 20 I could not get "going."

forms were being completed, in order to insure completion and to answer questions. This order of administration and procedures is the standard practice in our clinic.

RESULTS

Characteristics of the samples

Table 2 presents the sociodemographic characteristics of the five psychiatric samples. The individual samples are representative of the clinic population from which they derive. Considerable differences in age, sex and social class are found between samples, as expected. Because of

these differences, data from the five populations will not be pooled.

Comparison of CES-D scores in different populations

Table 3 shows the mean total CES-D scores for males and females within the five psychiatric populations and in the community samples. Looking first at the total CES-D scores, the psychiatric populations score considerably higher (more depressive symptoms) than the community sample. The acutely depressed patients score the highest (38.10), considerably over the cut-off score of 16. The recovered depressives and schizophrenic patients score below the cut-off range and within the same range (14.85 and 12.98, respectively), which is still slightly higher than the community sample. The scores of the drug-addicted and alcoholic populations lie somewhere between the acutely depressed and the other psychiatric populations. The higher scores of the acutely depressed as compared to the other psychiatric populations indicate that the scale is not merely measuring overall psychiatric impairment. The scores for males and females in 4 of the 6 samples are similar. The two exceptions are: the drug-addicted and the community sample females who score higher than the males in their respective samples.

CES-D scores for nondepressed and depressed patients within the five samples

The various populations were divided, according to their Raskin Depression Scores, into not depressed (a Raskin Depression Score of 3-6) and clinically depressed (a Raskin Depression Score of 7 or higher). As expected, the acute depressives all fell into the depressed category and the recovered depressives into the not depressed category (table 4).

Using the Raskin Depression Score as criteria, 19 out of 60 (32 per cent) drug-addicted patients, 36 out of 61 (59 per cent) alcoholic patients, and 14 out of 50 (28 per

TABLE 2
Sociodemographic characteristics of the five psychiatric populations

Characteristics	Acute depressives (N = 148)		Recovered depressives (N = 87)		Drug addicts (N = 60)		Alcoholics (N = 61)		Schizophrenics (N = 50)	
	No	%	No	%	No.	%	No	%	No	%
<i>Sex</i>										
Male	28	19	19	22	47	78	36	59	13	26
Female	120	81	68	78	13	22	25	41	37	74
<i>Age (years)</i>										
18-24	31	21	19	22	30	50	10	16	4	8
25-44	87	59	50	57	30	50	40	66	18	36
45-64	30	20	18	21			11	18	28	56
<i>Race</i>										
White	132	89	77	89	30	50	33	54	40	80
Non-white	16	11	10	11	30	50	28	46	10	20
<i>Marital status</i>										
Never married	20	14	12	14	26	43	14	23	15	30
Currently married/Divorced remarried	90	60	53	61	26	44	25	41	21	42
Currently divorced/Separated	35	24	20	23	8	13	20	33	14	28
Widowed	3	2	2	2			2	3		
<i>Employment status</i>										
Employed full time	46	31	27	31	15	24	12	20	11	22
Employed part-time	26	18	18	21	4	7	6	10	6	12
Unemployed	76	51	42	48	41	68	43	70	33	66
<i>Social class*</i>										
One	7	5	5	6			1	2		
Two	15	10	12	14	2	4	2	3		
Three	47	32	24	28	7	13	4	7	2	4
Four	54	36	33	37	16	31	26	43	19	38
Five	25	17	13	15	27	52	28	45	29	58
<i>Religion</i>										
Catholic	87	59	47	54	33	55	26	43	33	66
Protestant	41	28	27	31	24	40	32	52	17	34
Jewish	12	8	6	7	2	3	1	2		
Other or none	8	5	7	8	1	2	2	3		

* Based on the Hollingshead Two Factor Index of Social Position, A Hollingshead, Copyright, Yale University, 1957. Social class data are missing on eight subjects in the drug addict population.

cent) schizophrenic patients were also clinically depressed.

When the sample is divided by patients diagnosed as depressed on the Raskin Depression Scale, most of the not depressed sample scored close to the community sample score of 9 and considerably below the cut-off score of 16. The recovered depressives had the highest score (14.85) because

there were some patients in this group who had borderline depressive symptoms (i.e., Raskin Depression Scores of 5 and 6).

All of the depressed subgroups within each population scored considerably higher than the not depressed subjects within the particular population and higher than the community sample. The scores of the depressed subgroups were in

the range of 24.84 for the drug-addicted subjects to 31.11 for the depressed alcoholics. None of the depressed subgroups scored as high as the acute depressives (38.10). These findings again indicated that the CES-D was sensitive to depressive symptomatology within the various populations and was not merely measuring overall impairment. It was the depressed subgroups within each population who were contributing most to the overall mean score of the population. The not depressed subgroups were closer to the range of the community sample on depressive symptoms.

Correlationship between CES-D and other scales.

Correlations were examined between the CES-D mean total scores and the mean total scores of the clinician ratings, the Hamilton and the Raskin Depression Scale, and the self-report scale (SCL-90)

and with age, social class, and sex in the different populations (table 5).

Correlations between the CES-D total scores and other scales in all the populations were highly significant. They were highest for all subjects between the CES-D and the SCL-90, both of which were symptom self-report scales and are, therefore, most alike in content and method. However, for the alcoholics and schizophrenics, correlations are high between CES-D and other scales whether they are clinician ratings or self-reports.

There were no significant correlations between the total CES-D scores and the patient's age, social class and sex, with two exceptions. There were modest correlations between social class and CES-D total scores in schizophrenics (schizophrenic patients of a higher social class had higher symptom scores on the CES-D) and between sex and CES-D total scores in drug-addicted patients (female addicts had higher CES-D scores).

TABLE 3
Mean CES-D scores for males and females in different populations

Populations		Mean CES-D scores			p-value
		Total	Male	Female	
Acute depressives	\bar{X}	38.10	37.14	38.33	N.S.
	S.D.	9.01	8.31	9.15	
	N	148	28	120	
Recovered depressives	\bar{X}	14.85	18.58	13.81	N.S.
	S.D.	10.06	9.94	9.84	
	N	87	19	68	
Drug addicts	\bar{X}	17.05	15.57	22.39	<.05
	S.D.	10.69	9.99	11.39	
	N	60	47	13	
Alcoholics	\bar{X}	22.97	21.94	24.44	N.S.
	S.D.	13.58	14.98	11.11	
	N	61	36	25	
Schizophrenics	\bar{X}	12.98	13.08	12.95	N.S.
	S.D.	12.94	11.30	13.47	
	N	50	13	37	
Community sample	\bar{X}	9.10	7.90	9.93	<.001
	S.D.	8.60	7.60	9.14	
	N	3932	1614	2318	

TABLE 4
Mean CES-D scores for nondepressed and depressed patients in five psychiatric populations

Populations		Mean CES-D scores		
		Total	Nondepressed (Raskin 3-6)	Depressed (Raskin 7+)
Acute depressives	\bar{X}	38.10		38.10
	S.D.	9.01		9.01
	N	148		148
Recovered depressives	\bar{X}	14.85	14.85	
	S.D.	10.06	10.06	
	N	87	87	
Drug addicts	\bar{X}	17.05	13.44	24.84
	S.D.	10.69	8.58	10.62
	N	60	41	19
Alcoholics	\bar{X}	22.97	11.24	31.11
	S.D.	13.58	9.91	9.09
	N	61	25	36
Schizophrenics	\bar{X}	12.98	7.25	27.71
	S.D.	12.94	7.67	12.04
	N	50	36	14

TABLE 5
Correlations of the CES-D with other depression scores and with age, social class, and sex in five populations

Populations	Correlations with CES-D total scores					
	Clinician Ratings		Self-Report	Sociodemographic		
	Hamilton	Raskin	SCL-90	Age	Social Class	Sex
Acute depressives (N = 148)	.49†	.28†	.72†	-.10	-.03	.05
Recovered depressives (N = 87)	.65†	.64†	.73†	-.02	-.03	-.20
Drug addicts (N = 60)	.70†	.49†	.76†	-.15	.04	.26*
Alcoholics (N = 61)	.82†	.76†	.87†	-.20	-.10	.09
Schizophrenics (N = 50)	.85†	.79†	.84†	-.22	-.32*	.00

* $p < .05$

† $p < .001$

Correlations between the CES-D total score and the SCL-90 factors, which includes a depression factor, were examined (table 6). While the correlations between all factors and total scores were significant, they were highest with the depression factor (.73 to .89). This was consistent in the five psychiatric populations.

Comparing cut-off scores for the CES-D using the Raskin Score as criterion

As noted previously, in the community survey a score of 16 or higher was consid-

ered a case. Using a Raskin Depression Score of 7 or higher as the criterion for depression, the CES-D score of 16 or higher was tested as a screening criterion for depressive symptoms. Sensitivity and specificity were calculated (table 7). The acute depressives had the highest sensitivity (99 per cent). For the recovered depressives, the specificity was moderately high (56 per cent), indicating that some of the borderline cases (Raskin Depression Score of 5 or 6) were rating themselves as 16 or higher on the CES-D. Sensitivity and spec-

ificity were quite high for alcoholics and schizophrenics (94 per cent and 93 per cent, respectively). Sensitivity for the drug-addicted patients was acceptable (74 per cent), but not as high. This indicated that 16 or higher was satisfactory in most of the patients with the exception of the drug-addicted subjects where a high false positive rate was obtained.

Change in symptoms over time

Table 8 shows the mean CES-D total scores and standard deviation on 35 acutely depressed patients at admission for the treatment of the acute episode, and after one and four weeks of treatment with psychotropic medication. All patients were judged by the clinician to have shown

some improvement over the four weeks, and over 70 per cent were asymptomatic. This improvement is reflected in the CES-D scores over the four weeks.

Not shown here is the magnitude of the change for patients by degree of clinical recovery. Patients who were considered asymptomatic by the clinician (Raskin Score of less than 7) had a mean decrease in the CES-D score of 20 points whereas those patients who were judged clinically as still mildly symptomatic (Raskin score over 6) only showed a decrease of 12 points.

DISCUSSION

These results point out the validity and utility of the CES-D, as well as its limitations.

TABLE 6
Correlations of CES-D total scores with factors of the SCL-90 in five psychiatric populations

SCL-90 Factors	Correlations with CES-D total score*				
	Depressives		Drug Addicts	Alcoholics	Schizophrenics
	Acute	Recovered			
Depression	.73	.79	.79	.89	.86
Somatization	.47	.44	.63	.54	.61
Obsessive compulsive	.56	.62	.65	.77	.74
Interpersonal sensitivity	.61	.52	.65	.76	.81
Anxiety	.63	.57	.66	.78	.82
Hostility	.32	.44	.58	.64	.79
Phobic anxiety	.50	.46	.57	.61	.68
Paranoid ideation	.47	.45	.51	.67	.70
Psychoticism	.50	.58	.69	.73	.69

* All correlations are significant, $p < .001$

TABLE 7
Classification of patients as depressed and not depressed using a cut-off score of 16 on the CES-D and the Raskin score as criterion

CES-D	Raskin Depression Score*									
	Acute depressives (N = 148)		Recovered depressives (N = 87)		Drug addicts (N = 60)		Alcoholics (N = 61)		Schizophrenics (N = 50)	
	7+	3-6	7+	3-6	7+	3-6	7+	3-6	7+	3-6
16+ (depressed)	147	0	0	38	14	17	34	4	13	5
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
0-15 (not depressed)	1	0	0	49	5	24	2	21	1	31
	<i>c</i>	<i>d</i>	<i>c</i>	<i>d</i>	<i>c</i>	<i>d</i>	<i>c</i>	<i>d</i>	<i>c</i>	<i>d</i>
Sensitivity $a/a + c$	99%		-		74%		94%		93%	
Specificity $d/b + d$	-		56%		59%		84%		86%	

* Raskin Score 7+ = depressed, 3-6 = not depressed.

TABLE 8
Depressive symptoms CES-D scores in depressed patients over four weeks of treatment

	CES-D total scores (N = 36)	
	Mean	S D
Admission for acute episode	39.11	9.49
After one week of treatment	29.29	10.82
After four weeks of treatment	20.91	12.48

Validity

Evidence for the content, concurrent, and discriminant validity of the CES-D was demonstrated as follows:

Content validity. Content validity refers to the extent to which items in a scale are a representative sample from a universe generally accepted as defining the behaviors of interest (26). The items in the CES-D were specifically selected because of their content validity. Representative items from widely-used and previously validated scales assessing depression were selected and these items were chosen because they represented major components of depressive symptomatology as identified in the clinical literature and factor analytic studies.

Concurrent validity. Concurrent validity refers to the relationship between the test results and some external criterion measured about the same time. When applied to rating scales this type of validity can involve demonstrating that groups of patients differing from each other in some known characteristic, external to the rating scale, are discriminated from each other by the rating scale; that scores on the rating scale agree with data from other concurrent measures or that an instrument can measure change (27). Evidence for the concurrent validity of the CES-D was demonstrated by the following results:

- 1) The CES-D differentiated psychiatric patients from community normals.
- 2) Acutely depressed patients scored higher (more symptomatic) than other psychiatric patients.
- 3) Depressed subgroups within each of the

three psychiatric populations (alcoholics, drug addicts, and schizophrenics) scored higher than not depressed patients within each of these populations.

- 4) Acutely depressed patients scored higher than recovered depressives.
- 5) Correlations between the CES-D and other depression scales obtained by either self-report or by clinician interview were high.

Discriminant validity. One way to demonstrate discriminant validity is to show that a test does not correlate significantly with variables from which it should differ (28). Discriminant validation was demonstrated by low correlation of the CES-D with such variables as age, sex, and social class, and by the higher correlations of the CES-D with the depression factor of the SCL-90 as contrasted with the other factors of that scale.

Utility for case finding

Symptom scales can be valuable as screening measures in identifying high risk groups. When used for these purposes, cut-off scores are developed to distinguish possible cases from non-cases. Subjects' scores using these cut-off points are then compared to an accepted criterion of a "case" and agreement between the cases defined by the cut-off scores and against the criterion is tested in terms of numbers of false positives and negatives. Based on previous research a cut-off score of 16 was developed for the CES-D.

When this score was applied to the various psychiatric populations, low false negative and positive rates and high sensitivity and specificity was demonstrated in the acutely depressed, the alcoholic, and the schizophrenic populations. The cut-off score was less useful with the recovered depressives where a large number of borderline cases (i.e., patients who had mild symptoms of depression) probably were included, or where issues of response set may be operating. It was also less useful with drug addicted patients where high

false positive rates were obtained, suggesting that higher cut-off scores may be required in this population. These results suggest the utility of the scale as a screening instrument and that some adjustments in cut-off scores may be necessary for certain populations.

Limitations

The CES-D is quite sensitive to detecting symptoms in a variety of psychiatric populations and can identify high risk groups who have depressive symptoms independent of diagnosis. However, as with any symptom scale, it has certain limitations. It cannot differentiate between subjects who have acute depressive symptoms in the absence of other psychiatric diagnosis (primary depression) and those who have depressive symptoms in association with other psychiatric and/or medical conditions (secondary depression) (29). The distinction between primary and secondary depression may be important since there is evidence that these disorders may have different epidemiologies, clinical course, family history and response to treatment (30).

Moreover, the CES-D could not be used to ascertain rates of psychiatric disorders as defined and diagnosed in treatment settings. Diagnostic scales which give weights and classify symptoms into groups or dimensions would be required to make the distinction between primary and secondary depression or to obtain epidemiologic rates of specific disorders as diagnosed in clinical settings. The well known unreliability of psychiatric diagnosis has been a major deterrent to their use in epidemiologic studies. However, progress in increasing their precision and reliability has recently been made (31).

Conclusion

These results show that the CES-D is a sensitive tool in studies of psychiatric populations for detecting depressive symptoms and for measuring change in symp-

toms over time (32). In epidemiologic studies it may be a useful index for the study of association between depressive symptoms and factors of interest. It has applicability in clinical studies for detecting depressive symptoms in a wide variety of psychiatric patients and as a measure of change in psychiatric patients. Results obtained on the CES-D agree quite well with other more widely used and more lengthy self-report scales and clinician ratings.

While the CES-D, as any other symptom scale, cannot differentiate between diagnostic groups, it has demonstrated validity as a screening tool for case finding in psychiatric populations and for detecting groups at high risk for depression.

REFERENCES

1. Roth A, Locke B Continuous mental health assessment. Presented at the 101st Meeting of the American Public Health Association, San Francisco, CA, 1973
2. Markush RE, Favero RV. Epidemiologic assessment of stressful life events, depressed mood, and psychophysiological symptoms: A preliminary report. In *Stressful Life Events: Their Nature and Effects* Edited by BS Dohrenwend, BP Dohrenwend. New York, John Wiley & Sons, 1973, pp 171-190
3. Hornstra R, Klassen D The course of depression. *Compr Psychiatry*, in press
4. Comstock GW, Helsing KJ: Characteristics of respondents and nonrespondents to a questionnaire for estimating community mood. *Am J Public Health* 97 233-239, 1973
5. Comstock GW, Helsing KJ: Symptoms of depression in two communities. *Psychol Med* 6:551-564, 1976
6. Radloff LS Sex differences in depression: The effects of occupation and marital status. *Sex Roles* 1:249-265, 1975
7. Radloff LS. The CES-D Scale. A self-report depression scale for research in the general population. *Appl Psychol Meas*, in press
8. Mellinger GD, Balter MB, Parry JK, et al. An overview of psychotherapeutic drug use in the United States. In *Drug Use: Epidemiological and Sociological Approaches*. Edited by E Josephson, EE Carroll. New York, Hemisphere Publishing Corp., 1974, pp 333-336
9. Blumenthal MD: Measuring depressive symptomatology in a general population. *Arch Gen Psychiatry* 32:971-978, 1975
10. Warheit GJ, Holzer CE III, Schwab JJ: An analysis of social class and racial differences in depressive symptomatology: A community study. *J Health Soc Behav* 14:291-299, 1973
11. Martin FM, Brotherston JHF, Chave SPW: Inci-

- dence of neurosis in a new housing estate. *Br J Prev Soc Med* 11:196-202, 1957
12. Weissman MM, Myers JK: The New Haven community survey 1967-75: Depressive symptoms and diagnosis. Presented at the Society for Life History Study of Psychopathology Meeting, Fort Worth, TX, October 6-8, 1976
 13. Raskin A, Schulterbrandt J, Reatig N, et al: Factors of psychopathology in interview, ward behavior, and self-report ratings of hospitalized depressives. *J Consult Psychol* 31:270-278, 1967
 14. Raskin AJ, Schulterbrandt J, Reatig N, et al: Replication of factors of psychopathology in interview, ward behavior, and self-report ratings of hospitalized depressives. *J Nerv Ment Dis* 148:87-98, 1969
 15. Hamilton M A rating scale for depression. *J Neurol Neurosurg Psychiatry* 23:56-62, 1960
 16. Weissman MM, Lieb J, Prusoff B, et al. A double-blind trial of maprotiline (Ludiomil) and amitriptyline in depressed outpatients. *Acta Psychiatr Scand* 52:225-236, 1975
 17. Derogatis LR, Lipman RS, Covi L: "SCL-90" An outpatient psychiatric scale: Preliminary report. *Psychopharmacol Bull* 9:13-27, 1973
 18. Derogatis LR, Rickels K, Rock AF: The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *Br J Psychiatry* 128:280-289, 1976
 19. Zung WWK A self-rating depression scale. *Arch Gen Psychiatry* 12:63-70, 1965
 20. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. *Arch Gen Psychiatry* 4:561-571, 1961
 21. Dahlstrom WG, Welsh GS: An MMPI Handbook. Minneapolis, University of Minnesota Press, 1960
 22. Gardner EA: Development of a symptom checklist for the measurement of depression in a population. Unpublished manuscript, 1968
 23. Craig TJ, Van Natta P: Validation of the community mental health assessment interview instrument among psychiatric inpatients. Working Paper #B-27a for the Center for Epidemiologic Studies. Rockville, MD, National Institute of Mental Health, November 1973
 24. Craig TJ, Van Natta P: Prevalence and persistence of depressive symptoms in patient and community populations. *Am J Psychiatry*, in press
 25. Craig TJ, Van Natta P. Recognition of depressed affect in hospitalized psychiatric patients. Staff and patient perceptions. *Dis Nerv Syst*, in press
 26. Dohrenwend BP, Dohrenwend BS. The problems of validity in field studies of psychological disorder. *J Abnorm Psychol* 70:52-59, 1965
 27. Spitzer R, Endicott J: Psychiatric rating scales. In *Comprehensive Textbook of Psychiatry II*. Edited by A Friedman, H Kaplan, B Sadock. Baltimore, Williams & Wilkins Co., 1975, pp 2026-2031
 28. Cronback LJ, Meehl PC Construct validity in psychological tests. *Psychol Bull* 52:281-302, 1955
 29. Robins E, Guze S Classification of affective disorders The primary-secondary, the endogenous and the neurotic-psychotic concepts. In *Recent Advances in the Psychobiology of the Depressive Illness* DHEW Publication # (HSM) 70-9053 Washington, DC, USGPO, 1972
 30. Weissman MM, Pottenger M, Kleber H, et al: Symptom patterns in primary and secondary depression: A comparison of primary depressives with depressed opiate addicts, alcoholics, and schizophrenics. *Arch Gen Psychiatry*, in press
 31. Spitzer R, Endicott J, Robins E: Clinical criteria for psychiatric diagnosis and the DSM-III. *Am J Psychiatry* 132:1187-1199, 1975
 32. Weissman MM, Prusoff B, Newberry P. Comparison of the CES-D with standardized depression rating scales at three points in time. Technical Report from Yale University under Contract ASH 74-166, Center for Epidemiologic Studies, National Institute of Mental Health, Rockville, MD, 1975