

The Depression Interview and Structured Hamilton (DISH): Rationale, Development, Characteristics, and Clinical Validity

KENNETH E. FREEDLAND, PhD, JUDITH A. SKALA, PhD, RN, ROBERT M. CARNEY, PhD, JAMES M. RACZYNSKI, PhD, C. BARR TAYLOR, MD, CARLOS F. MENDES DE LEON, PhD, GAIL IRONSON, MD, PhD, MARSTON E. YOUNGBLOOD, MA, MPH, K. RANGA RAMA KRISHNAN, MD, AND RICHARD C. VEITH, MD, FOR THE ENRICHD INVESTIGATORS

Objective: The Depression Interview and Structured Hamilton (DISH) is a semistructured interview developed for the Enhancing Recovery in Coronary Heart Disease (ENRICHD) study, a multicenter clinical trial of treatment for depression and low perceived social support after acute myocardial infarction. The DISH is designed to diagnose depression in medically ill patients and to assess its severity on an embedded version of Williams' Structured Interview Guide for the Hamilton Depression scale (SIGH-D). This article describes the development and characteristics of the DISH and presents a validity study and data on its use in ENRICHD. **Methods:** In the validity study, the DISH and the Structured Clinical Interview for DSM-IV (SCID) were administered in randomized order to 57 patients. Trained interviewers administered the DISH, and clinicians administered the SCID. In ENRICHD, trained research nurses administered the DISH and recorded a diagnosis. Clinicians reviewed 42% of the interviews and recorded their own diagnosis. The Beck Depression Inventory (BDI) was administered in both studies. **Results:** In the validity study, the SCID diagnosis agreed with the DISH on 88% of the interviews (weighted $\kappa = 0.86$). In ENRICHD, the clinicians agreed with 93% of the research nurses' diagnoses. The BDI and the Hamilton depression scores derived from the DISH in the two studies correlated 0.76 ($p < .0001$) in the validity study and 0.64 ($p < .0001$) in ENRICHD. **Conclusions:** These findings support the validity of the DISH as a semistructured interview to assess depression in medically ill patients. The DISH is efficient in yielding both a DSM-IV depression diagnosis and a 17-item Hamilton depression score. **Key words:** depressive disorder, psychiatric status rating scales, psychological tests, coronary disease, myocardial infarction.

BDI = Beck Depression Inventory; DISH = Diagnostic Interview and Structured Hamilton; DSM-IV = *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition; ENRICHD = Enhancing Recovery in Coronary Heart Disease; HRSD = Hamilton Rating Scale for Depression; MI = myocardial infarction; SIGH-D = Structured Interview Guide for the Hamilton Depression Scale.

INTRODUCTION

The Depression Interview and Structured Hamilton (DISH) is a semistructured interview that was devel-

From the Department of Psychiatry (K.E.F., J.A.S., R.M.C.), Washington University School of Medicine, St. Louis, Missouri; Behavioral Medicine Unit and Department of Health Behavior (J.M.R.), University of Alabama at Birmingham, Birmingham, Alabama; Department of Psychiatry (C.B.T.), Stanford University School of Medicine, Palo Alto, California; Department of Preventive Medicine (C.F.MDL.), Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois; Department of Psychology (G.I.), University of Miami, Coral Gables, Florida; Department of Biostatistics (M.E.Y.), University of North Carolina, Chapel Hill, North Carolina; Department of Psychiatry (K.R.R.K.), Duke University Medical Center, Durham, North Carolina; and Department of Psychiatry and Behavioral Sciences (R.C.V.), University of Washington, Seattle, Washington.

Address reprint requests to: James D. Hosking, PhD, Collaborative Studies Coordinating Center, Department of Biostatistics CB 8030, University of North Carolina, 137 E. Franklin Street, Suite 203, Chapel Hill, NC 27514-4145.

Received for publication May 22, 2001; revision received December 14, 2001.

DOI: 10.1097/01.PSY.0000028826.64279.29

oped for the Enhancing Recovery in Coronary Heart Disease (ENRICHD) study, a multicenter clinical trial of treatment for depression and/or low perceived social support after acute myocardial infarction (MI) (1, 2). The DISH is used to 1) screen medical patients for depressive disorders, 2) diagnose major and minor depression and dysthymia according to DSM-IV criteria (3), 3) rate the severity of depression on a structured version of the Hamilton Rating Scale for Depression (4, 5), and 4) document the history and course of depression. It is designed to minimize respondent burden without sacrificing the thoroughness or accuracy of the interview. This article describes the rationale, development, and characteristics of the DISH and presents data on its validity.

Rationale

Eligibility for ENRICHD required a unipolar depressive disorder and/or low social support (1, 2). Enrollment soon after the index MI was essential because a key aim was to reduce the risk of early reinfarction or death (6). Thus, patients had to be interviewed and diagnosed at bedside by research nurses. The nurses also had to assess the severity of the depressive episode. Most patients were too ill to tolerate lengthy assessments and had little unencumbered time during their brief hospital stay, so a highly efficient interview was needed to minimize respondent burden. Many patients were unaccustomed to discussing their emotional problems and had symptoms that are hard to assess in hospitalized medical patients, so it was es-

essential to conduct the interviews in a flexible, sensitive manner that would foster trust, rapport, and disclosure rather than in the neutral, tightly structured manner of an epidemiological study.

To identify a suitable interview, we conducted literature reviews, sought the advice of experts in psychiatric interviewing, diagnosis, and epidemiology, and examined a variety of widely used instruments, including the Diagnostic Interview Schedule (DIS) (7–11), modified versions of the DIS used in previous studies of depression in cardiac patients (12, 13), the Composite International Diagnostic Interview (CIDI) (14–17), the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (18, 19), and the Structured Clinical Interview for DSM (SCID) (20–22).

Unfortunately, none of these instruments completely met the requirements discussed above. The DIS was developed for the Epidemiologic Catchment Area Study (23), and the CIDI was used in the National Comorbidity Survey (24, 25) and the World Health Organization study of Psychological Problems in General Health Care (26, 27). Both are suitable for lay interviewers, but they are inflexible and insensitive to change, and they emphasize lifetime rather than current psychopathology. The modified DIS (12) assesses the presence and duration of depression and anxiety symptoms in cardiac patients. It emphasizes current psychopathology, and the symptoms are probed in a different order than on the standard DIS: somatic symptoms are assessed first. Although the modified DIS is for lay interviewers, it is inflexible and insensitive to change, and the diagnosis is determined by clinicians. The World Health Organization SCAN (18, 19) is a complex, minimally structured interview for experienced mental health clinicians. Finally, the SCID (20–22) is a widely used, semistructured diagnostic interview. The nonpatient version is used with nonpsychiatric populations, including medical patients, but it is administered by experienced mental health clinicians. Also, if the SCID is administered shortly after an MI, its “past month” duration probes are insensitive to depression symptoms with a post-MI onset.

Even if these instruments had been suitable, they could be used only to diagnose depression, not to measure its current severity. Despite its shortcomings, the Hamilton Rating Scale for Depression (HRSD) (4) is used in most treatment trials to measure depression severity. Its reliability is low even in the hands of experienced clinicians. Williams developed the Structured Interview Guide for the Hamilton Depression scale (SIGH-D) (5) to improve the reliability of the HRSD and to facilitate its use by lay interviewers. Because some of its probes overlap with ones needed

to diagnose depressive disorders, sequential administration of the SIGH-D and a diagnostic interview would increase respondent burden. We realized, however, that this could be prevented by integrating the SIGH-D and an interview for diagnosing DSM-IV depressive disorders in medical patients.

Development of the Interview

The DISH incorporates material from the 17-item version of the HRSD (4, 28, 29), the SIGH-D (5), the DIS (7–11), the modified DIS (12), and DSM-IV (3). Most of the Hamilton items were taken verbatim from Williams’ SIGH-D (5); others were modified slightly to facilitate integration with the DSM-IV probes. A preliminary version was field tested in 1996 on 214 patients with coronary heart disease in the ENRICH pilot study. Data analyses and feedback from the interviewers were used to refine the instrument. The revised version was used throughout the ENRICH clinical trial. A detailed, 60-page guide to the use of the DISH (30) was written.

Characteristics of the DISH

The DISH is divided into three sections. The “Optional Opening Questions” section is comprised of open-ended questions to build rapport and encourage disclosure. Examples include “Would you mind telling me about your heart attack?” and “What’s it been like for you here in the hospital?”

The next section is entitled “Current Depression Symptoms.” As shown in Table 1 and Figure 1, it includes the probes needed to diagnose major and minor depression and dysthymia and to rate the severity of depression for the past week on the 17-item Hamilton scale. The items eliminate redundancies between the SIGH-D and DSM-IV probes. DSM-IV symptoms are coded absent, subthreshold, present, or present but due entirely to direct physiological effects of medical illness or its treatment. Symptom duration in weeks is specified in an adjacent column. There is a separate “days” field for symptoms present less than 2 weeks (eg, symptoms with post-MI onset).

Some probes are mandatory; others are optional. Mandatory probes are administered verbatim and are required for determining whether a DSM-IV depressive disorder is present and for obtaining an HRSD score. The wording of some probes depends on the patient’s preferred terms for symptoms. For example, some depressed patients deny feeling “sad” or “depressed” but admit to feeling “down” or “blue.” The burden is on the interviewer to judge whether the patient’s terms are synonymous with the DSM-IV cri-

DEPRESSION INTERVIEW AND STRUCTURED HAMILTON (DISH)

TABLE 1. Current Depression Symptoms Section of the DISH

DISH Items	Symptoms	DSM-IV		HRSD	Other ^a
		Major/Minor Depression	Dysthymia		
2-4	Anhedonia, inactivity, social withdrawal	x		x	x
5-8	Dysphoric mood	x	x	x	
9	Irritability, anger				x
10-12	Changes in appetite and/or weight	x	x	x	
13-17	Insomnia or hypersomnia	x	x	x	
18-19	Fatigue, loss of energy	x	x	x	
20-22	Guilt, worthlessness, low self-esteem	x	x	x	
23	Hopelessness		x		
24-26	Thoughts of death, suicidal features	x		x	
27	Impaired concentration, indecision	x	x		
28	Cognitive anxiety			x	
29	Somatic anxiety			x	
30	Hypochondriacal concerns			x	
31	Loss of interest in sex			x	
32-35	Psychomotor retardation, agitation	x		x	
36	Insight			x	
37	Distress or functional impairment ^b	x	x		
38	Exclusionary symptoms ^c				x

^a "Other" items (social withdrawal, anger and irritability, and exclusionary symptoms) do not count toward the depression diagnosis or HRSD score.

^b Indicates whether "the symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning," as per DSM-IV criteria.

^c This item is a brief screen for psychotic and other symptoms that would necessitate exclusion from a treatment trial such as ENRICH (hallucinations, paranoia, aphasia, etc.).

teria. Some of the mandatory probes assess atypical features (increased appetite, weight gain, hypersomnia) and bereavement. Optional probes clarify the context, frequency, severity, duration, and qualitative features of symptoms. The interviewer may rephrase them if necessary and may improvise follow-up questions if further clarification is needed. There is ample space on the interview form to quote or paraphrase the patient's response to each probe and to note behavioral observations. Interviewers are trained to write extensive notes with particular emphasis on information that clarifies the basis for their symptom coding decisions.

The threshold is very high in DSM-IV for attributing symptoms to medical illness, medication, or substance abuse rather than to depression. In DSM-IV and on the DISH, symptoms count toward depression except if clearly and completely accounted for by the direct physiological effects of a medical condition or medication (eg, weight gain due solely to edema). Symptoms such as fatigue or dysphoric mood rarely meet this stringent test, even in acute MI patients.

The Current Depression items can be administered in the order they appear on the interview form. The first items assess the cardinal symptoms of depression (dysphoric mood and loss of interest or pleasure in usual activities). This facilitates rapid screening of

nondepressed patients, but many medical patients are more comfortable if somatic symptoms are assessed first. The order may be varied as needed to facilitate disclosure and rapport.

At the end of the Current Depression section, there is a brief assessment of signs or symptoms of major psychiatric disorders (eg, paranoia, delusions, hallucinations, hypomania, confusion). This is not a comprehensive evaluation but rather a brief screen for severe psychiatric comorbid conditions, such as schizophrenia or bipolar disorders, that are rarely encountered in studies of post-MI patients but that would require exclusion from the trial if present.

The Psychiatric History section concludes the interview. As shown in Table 2, most of the items address past history and treatment of major depression. The number of past episodes, the age at first onset, and the age at onset of the last prior episode are elicited, along with family history of depression. Other items screen for history of bipolar disorder, alcoholism, and other disorders that might necessitate exclusion from the study. This section is not repeated when the DISH is readministered at follow-up. Instead, the Longitudinal Course Chart is used to document the interim course of the depressive disorder (if any) since the baseline interview. The approximate dates of any exacerbations,

CURRENT DEPRESSION SYMPTOMS						
#	DYSPHORIC MOOD	CODE	HRSD	DSM-IV	DURATION	
5	<p>[REQ] WHAT'S YOUR MOOD BEEN LIKE THIS WEEK? [OPT] HAVE YOU BEEN FEELING (SAD, DOWN, DEPRESSED, UNHAPPY, etc.)? [OPT] HAVE YOU FELT LIKE CRYING, OR HAVE YOU ACTUALLY BEEN CRYING OR HAD CRYING SPELLS SOMETIME IN THE LAST WEEK ?</p>					
	<p>[REQ] [RATE MOOD THIS WEEK, BASED ON PATIENT'S ANSWERS TO THIS ITEM AND YOUR OBSERVATIONS.] 0---- MOOD IS NOT DYSPHORIC 1---- MOOD IS DYSPHORIC, BUT THIS IS APPARENT ONLY IN PT.'s ANSWERS TO QUESTIONS 2---- MOOD IS DYSPHORIC, AND IS SEVERE ENOUGH THAT THE PT. TALKS SPONTANEOUSLY (WITHOUT BEING ASKED) ABOUT IT. 3---- MOOD IS DYSPHORIC, AND IS SEVERE ENOUGH THAT YOU CAN TELL NOT ONLY FROM THE PT.'s ANSWERS BUT ALSO FROM HIS/HER FACIAL EXPRESSION, VOICE, POSTURE, CRYING, ETC. 4---- MOOD IS DYSPHORIC, AND IS SO SEVERE THAT IT IS OBVIOUS IN VIRTUALLY EVERYTHING THE PT. SAYS AND DOES; PT. LOOKS AND SOUNDS VERY DEPRESSED</p>		HRSD			
6	<p>[REQ] [RATE CURRENT MOOD, BASED ON ITEM #5, PROBES FOR FREQUENCY & DURATION, AND YOUR OBSERVATIONS. IF MOOD IS DYSPHORIC, ASK]: HAVE YOU BEEN FEELING (SAD, DEPRESSED, EMPTY, etc.) MOST OF THE TIME? HOW LONG HAVE YOU BEEN FEELING LIKE THAT? 0---- NOT DYSPHORIC (Duration=N) 1---- DYSPHORIC SOME DAYS (Duration=weeks) 2---- DYSPHORIC MOST DAYS (Duration=weeks) M--- MEDICAL Sx (Duration=weeks) R--- REFUSED (Duration=R) U--- UNABLE TO ASSESS (Duration=U) DURATION IN DAYS IF <2 WEEKS: _____</p>			DSM-IV MAJOR: 2 MINOR: 2 DYSTH: 2		

Fig. 1. Sample page from the Depression Interview and Structured Hamilton (DISH) form. The page has been compressed to reduce the size of the figure. On the actual DISH form, there is more room to write notes, and there are spaces between the items to improve the readability of the form. The Hamilton dysphoric mood item is rated in the HRSD column in item 5, and the DSM-IV dysphoric mood symptom is coded in the DSM-IV column in item 6. The duration (in weeks) of the symptom is entered in the four columns to the right of the DSM-IV code. The "Code" column is not used on this page of the interview, but it is included on every page in the Current Symptoms section to maintain a visually consistent format from one page to the next.

remissions, relapses, recurrences, or new depressive episodes are recorded.

Diagnostic Forms

Several forms are used in conjunction with the DISH. The DSM-IV Diagnosis Guide enables the interviewer to quickly determine whether the diagnostic criteria for major or minor depression or dysthymia have been met. The guide form includes an outline of the criteria for each disorder and a grid with a row for each criterion symptom and a column for each depressive disorder. Symptom codes and durations from the DISH are copied into the cells of the grid. The diagnosis, episode duration, and HRSD score are coded on the Diagnostic Summary Form at baseline and the Diagnostic Follow-Up Form at follow-up.

STUDY 1: CLINICAL VALIDITY ANCILLARY STUDY

Methods

An ENRICHD ancillary study was conducted at the St. Louis site to evaluate the concurrent validity of the DISH vis-à-vis the SCID. The depression sections of the nonpatient version of the SCID were used because this version is recommended for use with medical patients.

The ancillary study and ENRICHD samples were mutually exclusive. The participants were cardiology patients at Washington University Medical Center who responded to recruitment flyers or who were approached by study recruiters and who signed an informed consent document approved by the Washington University Medical Center Human Studies Committee. Medical records were reviewed to document heart disease, including 13 (23%) patients with stable coronary disease, 15 (26%) with a history of MI, 9 (16%) with unstable angina, 15 (26%) with cardiomyopathy, 17 (30%) with congestive heart failure, 9 (16%) with an arrhythmia, 6 (11%) with a history of coronary artery bypass graft surgery, and 9 (16%) with a history of coronary angioplasty. The aim was to evaluate the validity

DEPRESSION INTERVIEW AND STRUCTURED HAMILTON (DISH)

TABLE 2. Psychiatric History Section of the DISH

DISH Item	Psychiatric History
39	Prior history of depression
40	Number of prior major depressive episodes
41	Age at onset of first prior major depressive episode
42	Age at onset of last prior major depressive episode
43	Prior treatment for depression ^a
44	Current treatment for depression ^a
45	Diagnosed by a psychiatrist as having "manic depression"
46	Family history of depression
47	Treatment for any other psychiatric disorder or emotional problem

^a Psychotherapy or counseling, antidepressant medication, electroconvulsive therapy, and/or psychiatric hospitalization are noted if treatment for depression is endorsed.

of the DISH for patients with various forms of heart disease, not only patients recovering from an acute MI.

Unlike the ENRICHD trial, ancillary study participants were medically stable, had been out of the hospital for at least 2 weeks, and were interviewed in a quiet, secluded interview room. These were more favorable conditions than typical of ENRICHD baseline interviews.

The BDI was administered approximately 10 minutes before the first interview. The order of interview administration was randomized (DISH first in 28 cases; SCID first in 29), and the patients had a 10-minute break between the interviews. The second interviewer was blind to the results of the first interview, and participants were asked to reveal nothing to the second interviewer about the first interview. The DISH interviews were administered by five research nurses and lay interviewers who had studied the DISH manual, completed a 6-hour DISH training course, and had experience with a minimum of three DISH administrations with supervisory feedback before interviewing any study participants. Two interviewers conducted 84% of the interviews. The SCID was administered by a clinical social worker and two clinical psychologists, all with formal training and experience in diagnosing and treating depression in cardiac patients. They also completed the SCID videotape training course (31) and completed at least three practice SCID administrations under supervision of the first author before conducting the study interviews. One interviewer conducted 74% of the interviews. There was no relationship between the DISH and SCID interviewer pairings ($p = .61$).

Results

The participants ($N = 57$) were 57 ± 10 years old; 26 (46%) were female and 17 (30%) were African American. The diagnoses were combined into three categories: nondepressed vs. minor depression or dysthymia vs. major depression. The DISH and SCID diagnoses agreed in 88% of the cases (weighted $\kappa = 0.81$, 95% CI = 0.68–0.94; $\chi^2 = 73.6$, $p < .0001$). Out of seven discordant cases, three were between minor depression and no depression, and four were between minor and major depression (ie, all disagreements were between adjacent categories). The former occurred at

lower BDI scores (mean = 9.7) than the latter (mean = 20.0). The diagnoses were then collapsed into two categories, depressed vs. nondepressed, corresponding to how the DISH was used to determine eligibility for ENRICHD. In this analysis, $\kappa = 0.89$ (95% CI = 0.78–1.00; $\chi^2 = 45.6$, $p < .0001$). With the SCID diagnosis as the criterion and the DISH as the test, sensitivity = 0.94, specificity = 0.96, positive predictive value = 0.97, and negative predictive value = 0.93.

The Pearson correlation coefficient between the BDI and the 17-item HRSD embedded in the DISH was 0.76 ($p < .0001$). Analysis of variance (ANOVA) was used to compare the BDI and HRSD scores of the DISH diagnosis groups. The BDI scores in the major depression ($N = 17$), minor depression or dysthymia ($N = 13$), and nondepressed groups ($N = 27$) were 24.3 ± 6.9 , 19.7 ± 6.6 , and 6.1 ± 4.4 , respectively ($F = 58.7$, $p < .0001$). Post hoc tests showed significant differences among all three groups. The HRSD scores were 21.6 ± 6.6 , 13.7 ± 4.1 , and 5.6 ± 5.2 , respectively ($F = 11.7$, $p < .0001$). Subsequent tests again identified significant differences among all three groups.

ANOVA was also used to compare the groups as to the time required to administer the DISH. Including instructions to the patient and the Optional Opening Questions, the DISH required 38 ± 18 minutes for nondepressed patients, 58 ± 18 minutes for patients with minor depression or dysthymia, and 64 ± 20 minutes for patients with major depression ($F = 11.7$, $p < .0001$). Subsequent tests showed significant differences between the nondepressed and both depressed groups, but not between the major depression and minor depression or dysthymia groups.

STUDY 2: ENRICHD TRIAL

Methods

The methods and baseline findings of ENRICHD have been reported elsewhere (1, 2). All participants signed informed consent documents approved by the clinical site's institutional review board. We report here on baseline data from ENRICHD concerning the validity of the DISH. Research nurses from eight ENRICHD clinical sites studied the DISH manual, completed two intensive DISH training workshops, and administered at least three DISH interviews under supervision before interviewing study patients. The DISH and the BDI (32) were administered within 28 days of the patient's index MI. The nurses were instructed to consult with a clinical supervisor or the first author only if unable to reach an independent diagnosis in especially difficult cases. Consults were requested on fewer than 1% of the baseline interviews. Seventy-eight percent of the interviews were conducted at bedside during the hospitalization; the others were conducted after discharge at the patient's residence (9%), an outpatient facility (5%), or elsewhere (8%). Participants were enrolled for depression alone, low perceived social support alone (as assessed by the ENRICHD Social Support Instrument; 1), or both. Because all nondepressed participants had

TABLE 3. Beck Depression Inventory and Hamilton Depression Scores By Depression Diagnosis and Low Perceived Social Support (LPSS) Status

Category	Beck Depression Score			Hamilton Depression Score		
	N	Mean	SD	N	Mean	SD
Overall (N = 2481)	2456	15.8	8.4	2404	15.5	7.4
Nondepressed with LPSS (N = 647)	635	9.7	6.4	612	8.8	6.0
Depressed without LPSS (N = 976)	969	16.7	6.8	957	17.6	6.3
Dysthymia only (N = 31)	31	15.8	6.1	31	14.1	5.3
Minor depression (N = 475)	472	13.9	4.8	468	14.7	4.9
Major depression (N = 470)	466	19.5	7.4	458	20.8	6.2
Depressed with LPSS (N = 858)	852	19.4	8.8	835	18.0	6.5
Dysthymia only (N = 35)	35	18.5	7.6	34	14.9	5.7
Minor depression (N = 336)	334	14.9	6.3	333	14.3	4.9
Major depression (N = 487)	483	22.6	9.0	468	20.9	6.1

low social support, they are unrepresentative of the nondepressed post-MI population as a whole.

During the first year of recruitment, almost 100% of the interviews were independently reviewed by designated study clinical psychologists or psychiatrists at each clinical site. The purpose of this procedure was to determine the extent to which mental health clinicians with advanced training and experience in psychiatric diagnosis would agree with the nurse-interviewers' diagnostic conclusions. The reviewers were given DISH forms that had been completed by nurse-interviewers and asked to read the qualitative notes, inspect the symptom and psychiatric history codes, form diagnostic impressions based solely on the interview material, and code their diagnoses on standardized forms. The reviewers were asked to determine their diagnoses before being unblinded to the nurse-interviewers' diagnostic conclusions. Individual clinicians were excused from reviewing any interview on which they been consulted by the nurse-interviewer; in these cases, the review was reassigned to a different clinician. The review requirement was relaxed in the second year because by then most of the nurses had extensive experience with the interview and diagnostic disagreements had become rare. Nurses who joined the staff in year 2 were still required to submit their interviews for review until there was diagnostic agreement in 10 successive cases.

Results

All 2481 ENRICHD participants completed the baseline DISH. BDI scores were obtained from 2456 (99%) and 17-item HRSD scores were extracted from the DISH from 2404 (97%) of these patients.¹ The mean age was 61 ± 13 years; 44% were female, 34% were racial minorities, and 26% had less than 12 years of education. The most common comorbid medical conditions were hypercholesterolemia (63%), hypertension (61%), cardiovascular disease (45%), and diabetes (33%); 65% had a history of cigarette smoking, and 27% had at least one previous MI.

One thousand thirty-seven (42%) of the interviews were independently reviewed by clinicians. The specific

¹ The HRSD score was not computed if there were any missing items.

diagnoses agreed in 969 (93%) of the cases, and there was 98% agreement as to whether the patient was depressed ($\kappa = 0.95$). Table 3 shows that 647 (26%) of the patients were nondepressed and enrolled solely on the basis of low social support. Of the remaining patients, 957 (39%) had major depression, 811 (33%) had minor depression, and 66 (3%) had dysthymia.²

Table 3 also displays BDI and HRSD depression scores by diagnostic category and social support status. In every diagnostic category, the scores are within the expected ranges. Type III multivariate analysis of covariance was used to determine the effects of depression diagnosis (major vs. minor or dysthymia vs. not depressed) on the BDI and HRSD scores, controlling for low social support (present vs. absent). The multivariate effect of depression diagnosis was significant (Wilks' λ , $F = 405.2$, $p < .0001$), as were the univariate effects of diagnosis on the BDI ($F = 539.8$, $p < .0001$) and HRSD ($F = 817.5$, $p < .0001$). The covariate (low social support) had a significant effect on the BDI ($F = 77.3$, $p < .0001$) but not on the HRSD ($F = 3.3$, $p = .07$). On post hoc tests, all groups differed significantly from all other groups on both measures ($p < .0001$).

The HRSD and BDI scores were correlated to test the concurrent validity of the SIGH-D version of the HRSD embedded in the DISH. In previous studies, correlations between BDI and HRSD scores have ranged from 0.61 to 0.87 across a variety of clinical populations (33). Among the 2379 ENRICHD patients with both scores, the correlation was 0.64 ($p < .0001$.)

² These figures are not interpretable as prevalence estimates because patients who screened negative both for depression and inadequate social support were excluded from participation.

DISCUSSION

Clinical Validity and Uses of the DISH

The concurrent validity of the DISH was evaluated in an ENRICHD ancillary study and in the baseline phase of the ENRICHD trial. Both studies show that with well-trained interviewers, the DISH yields clinically valid depression diagnoses and 17-item Hamilton depression scores. The clinical reviewers agreed with the nurses on 93% of the interviews in the ENRICHD trial, and there was 88% agreement between the DISH and SCID diagnoses in the ancillary study. The disagreements between the DISH and the SCID were, in every instance, between adjacent categories (ie, major vs. minor depression, or minor depression vs. no depression). In both studies, the DISH diagnoses were consistent with the BDI scores, and the correlations were moderately high between the BDI and the HRSD scores that were extracted from the DISH. Taken together, these findings support the use of the DISH as an interview for diagnosing depression and assessing its severity in post-MI patients.

None of the challenges that necessitated the development of the DISH are unique to the ENRICHD trial. Furthermore, similar challenges will confront investigators of comorbid depressive disorders in other medical patients, such as those with diabetes, cancer, or AIDS. With relatively minor modifications (in particular, with a different set of optional opening questions), the DISH can be adapted for use with these patients.

Limitations of the Studies

Although the ancillary validity study (study 1) was specifically designed to evaluate the concurrent validity of the DISH, the same cannot be said of the ENRICHD trial itself (study 2). The primary purpose of the trial's clinical review procedure was to determine whether given the same set of interview data, psychologists and psychiatrists with expertise in psychiatric diagnosis would agree with the research nurse-interviewers' depression diagnoses. The procedure was included in the study protocol because of the importance of depression as an eligibility criterion and as an outcome variable. Evaluation of the clinical validity of the DISH was only a secondary aim. Consequently, the review procedure lacked the methodological rigor of studies designed specifically to evaluate the concurrent validity of diagnostic interviews.

The discriminant validity of the DISH (ie, its ability to differentiate between depressive disorders and other psychiatric disorders) and its sensitivity to change have also not yet been studied. The sensitivity of the HRSD to change is well established, but the

sensitivity of the DISH to remission of DSM-IV depressive disorders deserves further investigation.

Neither the interrater reliability nor the test-retest reliability of the DISH were evaluated in either study. The reliability of instruments such as the DIS that are used in epidemiological studies is maximized by eliminating as much variance as possible in the administration of the interview. Nearly all questions are mandatory, and they are presented in an invariant order. These techniques could have been used to maximize the reliability of the DISH, but only at the expense of its clinical validity and of the willingness of some patients to be interviewed at all. Because the DISH is a flexible, semistructured interview, reliability analysis is a high priority for further research.

Limitations of the DISH

It is often quite difficult for interviewers to distinguish between threshold and subthreshold symptoms, especially for features that are mild or that fluctuate from day to day. This is a source of error variance at the item level on the DISH that can be minimized by rigorous training and quality assurance efforts, but it probably cannot be eliminated.

Symptoms that seem to be due to the medical illness also pose a dilemma. A symptom counts toward a depression diagnosis unless there is clear evidence that it is due entirely to the direct physiological effects of the medical condition. Many symptoms, however, cannot be so easily disqualified. Interviewers are often reluctant to count ambiguous symptoms such as fatigue unless there is some affirmative evidence that they actually *are* due, at least in part, to depression. The interviewer's decision about whether to count such symptoms may be context-dependent: If cognitive or affective symptoms such as dysphoria are present, it may be easier for the interviewer to believe that somatic symptoms such as fatigue might also be due to depression. This may reflect good clinical judgment, but it is technically at odds with the rules of DSM-IV. It is another source of error variance at the item level. Uncertainty about whether to count ambiguous symptoms may also contribute to inaccurate diagnoses.

Some patients can recall the durations of their depressive symptoms with great precision, especially if they were precipitated by a salient, discrete event. Many patients, however, provide only vague descriptions of symptom durations or onsets, and some give very confusing, inconsistent estimates. Furthermore, the temporal relationship between acute stressors and depressive episodes varies considerably across individuals. An acute MI can precipitate the onset of de-

pression in patients who were not depressed before their MI. If such patients are interviewed during the index hospitalization, as in ENRICHHD, the duration of depressive symptoms may be less than the 2-week minimum required by DSM-IV, and the stability of the diagnosis becomes a matter of concern. In other cases, the onset of some or all of the depressive symptoms precede the MI by weeks, months, or even years. For these patients, the MI might have exacerbated existing depressive symptoms, masked them, or had no discernible effect on them. In short, it may be even more difficult to determine symptom and episode durations in studies of comorbid depression in hospitalized medical patients than it is in studies of stable medical outpatients, psychiatric patients, or community populations. The reliability of symptom and episode duration data on the DISH should be studied.

CONCLUSION

The Depression Interview and Structured Hamilton is a semistructured interview to assess comorbid depressive disorders in post-MI and other medically ill patients. It is suitable for use by well-trained research interviewers and mental health clinicians. It is efficient in yielding both a DSM-IV depression diagnosis and a 17-item Hamilton depression score while eliminating redundant probes. Data on its performance in ENRICHHD and in an ancillary study comparing it with the SCID support the concurrent validity of the DISH. Additional studies are needed to evaluate its test-retest reliability, discriminant validity, sensitivity to change, and performance in other patient populations.

Supported by Contracts N01-HC-55140, N01-HC-55141, N01-HC-55142, N01-HC-55143, N01-HC-55144, N01-HC-55145, N01-HC-55146, N01-HC-55147, and N01-HC-55148 from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland. The authors thank the ENRICHHD research nurses who administered the DISH and the study clinicians who overread them at clinical centers across the United States. We thank Wilson M. Compton III, MD, the late Samuel B. Guze, MD, Lee N. Robins, PhD, and Janet B. W. Williams, DSW, for their invaluable advice. We also thank Iris Csik, Christianne Meier, Teresa Benoist, Patricia Hoffman, Steven Berger, Angela Misuraco, and Angela Tanner for their assistance with the clinical validity ancillary study.

REFERENCES

1. ENRICHHD Investigators. Enhancing Recovery in Coronary Heart Disease Patients (ENRICHHD): study design and methods. *Am Heart J* 2000;139:1-9.
2. ENRICHHD Investigators. Enhancing Recovery in Coronary Heart Disease Patients (ENRICHHD): baseline characteristics. *Am J Cardiol* 2001;88:316-22.
3. DSM-IV. Diagnostic and statistical manual of mental disorders. 4th ed. Washington DC: American Psychiatric Association; 1994.
4. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;12:56-62.
5. Williams JBW. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiatry* 1988;45:742-7.
6. McGovern PG, Pankow JS, Shahar E, Doliszny KM, Folsom AR, Blackburn H, Luepker RV. Recent trends in acute coronary heart disease—mortality, morbidity, medical care, and risk factors. *N Engl J Med* 1996;334:884-90.
7. Robins LN, Helzer JE, Croughan J, Ratcliff K. National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Arch Gen Psychiatry* 1981;38:381-9.
8. Singerman B, Stoltzman RK, Robins LN, Helzer JE, Croughan JL. Diagnostic concordance between DSM-III, Feighner, and RDC. *J Clin Psychiatry* 1981;42:422-6.
9. Robins LN, Helzer JE, Ratcliff KS, Seyfried W. Validity of the Diagnostic Interview Schedule, version II: DSM-III diagnoses. *Psychol Med* 1982;12:855-70.
10. Helzer JE, Robins LN, McEvoy LT, Spitznagel EL, Stoltzman RK, Farmer A, Brockington IF. A comparison of clinical and diagnostic interview schedule diagnoses: physician reexamination of lay-interviewed cases in the general population. *Arch Gen Psychiatry* 1985;42:657-66.
11. Helzer JE, Robins LN. The Diagnostic Interview Schedule: its development, evolution, and use. *Soc Psychiatry Psychiatr Epidemiol* 1988;23:6-16.
12. Carney RM, Rich MW, Freedland KE, Saini J, teVelde A, Simeone C, Clark K. Major depressive disorder predicts cardiac events in patients with coronary artery disease. *Psychosom Med* 1988;50:627-33.
13. Frasure-Smith N, Lesperance F, Talajic M. Depression following myocardial infarction: impact on 6-month survival. *JAMA* 1993;270:1819-25.
14. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, Farmer A, Jablenski A, Pickens R, Regier DA. The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988;45:1069-77.
15. Wittchen HU, Robins LN, Cottler LB, Sartorius N, Burke JD, Regier D. Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). The Multicentre WHO/ADAMHA Field Trials. *Br J Psychiatry* 1991;159:645-53.
16. Janca A, Robins LN, Bucholz KK, Early TS, Shayka JJ. Comparison of Composite International Diagnostic Interview and clinical DSM-III-R criteria checklist diagnoses. *Acta Psychiatr Scand* 1992;85:440-3.
17. Janca A, Robins LN, Cottler LB, Early TS. Clinical observation of assessment using the Composite International Diagnostic Interview (CIDI): an analysis of the CIDI field trials—wave II at the St Louis site. *Br J Psychiatry* 1992;160:815-8.
18. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, Jablenski A, Regier D, Sartorius N. SCAN: Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry* 1990;47:589-93.
19. Janca A, Ustun TB, Sartorius N. New versions of World Health Organization instruments for the assessment of mental disorders. *Acta Psychiatr Scand* 1994;90:73-83.

DEPRESSION INTERVIEW AND STRUCTURED HAMILTON (DISH)

20. Spitzer RL, Williams JBW, Gibbon M, First MB. User's guide for the Structured Clinical Interview for DSM-III-R: SCID. Washington DC: American Psychiatric Press; 1997.
21. Spitzer RL, Williams JB, Gibbon M, First MB. The Structured Clinical Interview for DSM-III- R (SCID). I. History, rationale, and description. *Arch Gen Psychiatry* 1992;49:624-9.
22. Williams JB, Gibbon M, First MB, Spitzer RL. The Structured Clinical Interview for DSM-III- R (SCID). II. Multisite test-retest reliability. *Arch Gen Psychiatry* 1992;49:630-6.
23. Robins LN, Regier DA, editors. *Psychiatric disorders in America: the Epidemiologic Catchment Area Study*. New York: The Free Press; 1991.
24. Kessler RC, McGonagle KA, Zhao S, Nelson CB. Lifetime and 12-month prevalence of DSM- III-R psychiatric disorders in the United States: results from the National Comorbidity Study. *Arch Gen Psychiatry* 1994;51:8-19.
25. Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994;151:979-86.
26. Sartorius N, Uestuen B, Costa e Silva JA, Goldberg D. An international study of psychological problems in primary care: preliminary report from the World Health Organization collaborative project on psychological problems in general health care. *Arch Gen Psychiatry* 1993;50:819-24.
27. Ormel J, VonKorff M, Ustun T. Common mental disorders and disability across cultures: results from the WHO Collaborative Study on Psychological Problems in General Health Care. *JAMA* 1994;272:1741-8.
28. Elkin I, Parloff MB, Hadley SW, Autry JH. NIMH Treatment of Depression Collaborative Research Program: background and research plan. *Arch Gen Psychiatry* 1985;42:305-16.
29. Elkin I, Shea MT, Watkins JT, Imber SD. National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. *Arch Gen Psychiatry* 1989;46:971-82.
30. Freedland KE. DSM-IV Depression Interview and Structured Hamilton (DISH). In: ENRICH Investigators. *The National Heart, Lung, and Blood Institute Enhancing Recovery in Coronary Heart Disease (ENRICH) Study manual of operations*. Chapel Hill (NC): Collaborative Studies Coordinating Center, Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill; 1997.
31. SCID-101 for DSM-IV: training videotapes for the Structured Clinical Interview for DSM-IV axis I disorders (SCID). New York: Biometrics Research Department, New York State Psychiatric Institute. 1996.
32. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4: 561-71.
33. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev* 1988;8:77-100.