

“Are You Depressed?” Screening for Depression in the Terminally Ill

Harvey Max Chochinov, M.D., Keith G. Wilson, Ph.D.,
Murray Enns, M.D., and Sheila Lander, R.N.

***Objective:** This study compared the performance of four brief screening measures for depression in a group of terminally ill patients. The methods included 1) a single-item interview assessing depressed mood, 2) a two-item interview assessing depressed mood and loss of interest in activities, 3) a visual analog scale for depressed mood, and 4) the Beck Depression Inventory—Short Form. **Method:** Semistructured diagnostic interviews for depression were administered to 197 patients receiving palliative care for advanced cancer. The interview diagnoses served as the standards against which the screening performance of the four brief screening methods was assessed. **Results:** Single-item interview screening correctly identified the eventual diagnostic outcome of every patient, substantially outperforming the questionnaire and visual analog measures. **Conclusions:** Brief screening measures for depression are important clinical tools for terminally ill patients. For diagnostic purposes, however, they do not approach the validity of a single-item interview that asks, in effect, “Are you depressed?” (Am J Psychiatry 1997; 154:674–676)*

Depression is a common problem in older adults and medically ill patients that often goes unrecognized by treating clinicians. For this reason, several efforts have been made to develop self-report screening inventories that would improve the accuracy of detection of depressive symptoms and yet be brief enough for routine administration to the medically ill. Although these measures have generally been found to correlate highly with full-length depression inventories (1–4), in some cases their concordance with a standard from clinician-administered interviews has been less than optimal (5). Recently, Mahoney et al. (6) reported preliminary evidence that a single-item interview for depression was as accurate as the 30-item Geriatric Depression Scale (7) in identifying depressive disorders in older adults. Because single-item interview screening is an approach that should be applicable even to terminally ill patients, further investigation of this method

represents an important direction for research into brief screening alternatives. In this study of patients receiving palliative care for advanced cancer, we compared the performance of four brief screening measures for depression: 1) a single-item interview assessing depressed mood, 2) a two-item interview assessing depressed mood and loss of interest or pleasure in activities, 3) a visual analog scale, and 4) the Beck Depression Inventory—Short Form (2).

METHOD

This study was approved by the Ethical Review Committee of the University of Manitoba, Faculty of Medicine. Before their participation and after the protocol had been fully explained, all patients gave written informed consent. As part of a larger study investigating the prevalence and correlates of depression in the medically ill, we interviewed 200 inpatients receiving palliative care for advanced terminal cancer. For 197 of these subjects (94 male and 103 female; mean age=71.0 years, SD=10.7), complete data were available from a semistructured diagnostic interview for depression, a visual analog scale measure of depressed mood, and the Beck Depression Inventory—Short Form.

The diagnostic interview was adapted from the Schedule for Affective Disorders and Schizophrenia (SADS) (8). In the interests of maintaining a short protocol for this group of severely ill patients, only items pertaining to the diagnoses of major and minor depressive episodes according to the Research Diagnostic Criteria (RDC) (9) were included. Most interviews were conducted by psychiatric nurses, whose ratings have been shown to have good concordance with psychiatrists' ratings in structured clinical judgments of this type (10). As discussed elsewhere (11), reliability was established by having an observer (a psychiatrist or psychologist) attend a random sample of 27 interviews.

Received June 20, 1996; revision received Oct. 3, 1996; accepted Oct. 25, 1996. From the Departments of Psychiatry and Medicine, University of Manitoba; the Manitoba Cancer Treatment and Research Foundation, Winnipeg; the Department of Medicine and School of Psychology, University of Ottawa; and the Rehabilitation Centre, Ottawa. Address reprint requests to Dr. Chochinov, Department of Psychiatry, University of Manitoba, 771 Bannatyne Ave., Winnipeg, Man. R3E 3N4, Canada.

Supported by grants from the Manitoba Cancer Treatment and Research Foundation and the Manitoba Mental Health Research Foundation and by the National Cancer Institute of Canada with funds from the Canadian Cancer Society.

The authors thank Mr. Kuldip Maini for statistical assistance.

TABLE 1. Ability of Four Brief Screening Measures to Identify Depressed Mood in 197 Terminally Ill Cancer Patients Diagnosed as Depressed by a Full Interview

Measure	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	False Positive Rate	False Negative Rate
Single-item interview screening	1.00	1.00	1.00	1.00	0.00	0.00
Two-item interview screening	1.00	0.98	0.86	1.00	0.02	0.00
Beck Depression Inventory—Short Form score ≥ 8	0.79	0.71	0.27	0.96	0.29	0.21
Visual analog scale score ≤ 55	0.72	0.50	0.17	0.92	0.50	0.29

In the RDC, the diagnosis of a major depressive episode requires the presence of either a depressed mood "most of the day, nearly every day" or a loss of interest or pleasure in "all or almost all activities," along with at least four other criterion symptoms. Minor depression requires the presence of a depressed mood and at least two other symptoms. Our diagnoses are based on strict adherence to these severity thresholds, which identify only relatively severe presentations of depression as meeting the diagnostic criteria (11). For the present purpose, single-item screening refers to the diagnostic efficiency of the SADS depressed mood item by itself, when held against the standard of the full interview. Two-item screening refers to the diagnostic performance of both the depressed mood and the loss of interest items of the SADS considered together.

The Beck Depression Inventory—Short Form is a 13-item version of the standard 21-item Beck Depression Inventory, with which it correlates at 0.96 (2). The short form was developed as a rapid screening technique for use with medical patients. Although it is appropriate for use as a self-administered measure, most patients in our study group required that it be presented orally by the interviewer.

The visual analog scale was taken from the Memorial Pain Assessment Card (12). It consisted of a 100-mm line anchored at the ends with the referents "worst possible mood" and "best possible mood"; lower scores indicate greater levels of depression. In general, the measurement of mood state by visual analog scale is a common assessment practice in palliative care, where it is considered to provide a simple and valid screen for depression (13).

RESULTS

The reliability checks confirmed that there was good concordance between the interviewers and observers in the overall identification of depressive disorders ($\kappa=0.76$). A total of 24 patients (12.2%) met the RDC for either major ($N=15$) or minor ($N=9$) depressive disorder. Table 1 shows the concordance of the four different brief screening measures with the diagnoses based on the full interview.

The most salient aspect of table 1 is the fact that the single item assessing depressed mood correctly identified the eventual diagnostic outcome of every patient in the study group. In effect, whenever a patient acknowledged the presence of a depressed mood at a threshold level of severity, the probability of having at least two more of the RDC symptoms (thereby meeting the RDC for, at minimum, a minor depressive episode) was 100%. Patients who did not meet this criterion, on the other hand, were never diagnosed with a depressive syndrome on the basis of other criteria. In fact, extending the screening interview to two items had the effect of identifying four patients as depressed who were classified as not depressed on the basis of the full interview.

With the Beck inventory short form, the sensitivity

and specificity values converged around a score of 8, which is the recommended cutoff for the identification of moderate depression (2). At this score, the sensitivity and specificity values correspond well with those reported by other investigators (14), but they reflect the misclassification of five (21%) of the 24 depressed patients and a high false positive rate (29%, $N=51$ of 173).

Similar problems were evident with the visual analog scale measure. A cutoff score of 55 mm provided a reasonable trade-off between sensitivity and specificity but resulted in the misclassification of seven depressed patients (29%) and 87 nondepressed patients (50%).

Statistical comparisons between the four screening methods confirmed that they were significantly different (Cochran's $Q=151.4$, $df=3$, $p<0.0001$). Pairwise comparisons using McNemar's chi-square ($df=1$) indicated that the single-item and two-item interviews yielded comparable rates of diagnosis ($p>0.10$), which were lower than those associated with the Beck short form and the visual analog scale measures (all p values <0.0001). The last two procedures also differed significantly from each other, with the visual analog scale resulting in the highest classification rates ($p<0.0001$).

DISCUSSION

Those who work with the terminally ill recognize the need for brief screening measures of depression, and a number of alternatives are now available. However, the present findings, together with similar results reported by Mahoney et al. (6), suggest that caution should be exercised in recommending to clinicians that the routine use of such measures would markedly improve their identification of clinical depressive disorders in an efficient way. Our findings indicate that a single-item interview for depressed mood provides a reliable and remarkably accurate screen, which, we submit, is brief enough for inclusion in most clinical contacts that permit direct patient interview. In this context, the failure to identify depression may not be due to a lack of sufficiently sensitive brief screening measures but, rather, to a failure to (in effect) make the simple inquiry, "Are you depressed?"

In addition to depressed mood, contemporary diagnostic criteria for depression include a loss of interest or pleasure in activities. Empirically, this item did not improve our diagnostic accuracy, nor could it have done so, but conceptually it merits consideration for

inclusion in a brief screening interview, because it provides for complete coverage of the core criterion symptoms of depression and thereby removes the possibility of false negative results. Indeed, this logic is embedded in most structured interviews, which commonly allow skipping inquiry into the remaining criterion symptoms if both depressed mood and loss of interest are absent. Although the two-item interview was found to have a slightly higher rate of false positive identifications when compared with diagnoses based on the full interview, its performance in this respect was still far superior to the self-report and visual analog scale measures. These last two techniques certainly have a role to play in providing gross assessments when direct interviews are not feasible, in quantifying the severity of a depressive syndrome once it is identified, and in monitoring change over time. However, their efficiency as diagnostic aids in palliative care settings is unlikely to approach that of a brief personal interview.

REFERENCES

- Allen N, Ames D, Ashby D, Bennetts K, Tuckwell B, West C: A brief sensitive screening instrument for depression in late life. *Age Ageing* 1994; 23:213-218
- Beck AT, Beck RW: Screening depressed patients in family practice: a rapid technic. *Postgrad Med* 1972; 52:81-85
- Burke WJ, Roccaforte WH, Wengel SP: The short form of the Geriatric Depression Scale: a comparison with the 30-item version. *J Geriatr Psychiatry Neurol* 1991; 4:173-178
- Koenig HG, Cohen HJ, Blazer DG: A brief depression scale for use in the medically ill. *Int J Psychiatr Med* 1992; 22:183-195
- Gerety MB, Williams JW Jr, Mulrow CD, Cornell JE, Kadri AA, Rosenberg J, Chiodo LK, Long M: Performance of case-finding tools for depression in the nursing home: influence of clinical and functional characteristics and selection of optimal threshold scores. *J Am Geriatr Soc* 1994; 42:1103-1109
- Mahoney J, Drinka TJK, Abler R, Gunter-Hunt G, Matthews C, Gravenstein S, Carnes M: Screening for depression: single question versus GDS. *J Am Geriatr Soc* 1994; 42:1006-1008
- Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO: Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982-1983; 17:37-49
- Endicott J, Spitzer RL: A diagnostic interview: the Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry* 1978; 35:837-844
- Spitzer RL, Endicott J, Robins E: Research Diagnostic Criteria: rationale and reliability. *Arch Gen Psychiatry* 1978; 35:773-782
- Lesage AD, Cyr M, Toupin J: Reliable use of the Present State Examination by psychiatric nurses for clinical studies of psychotic and nonpsychotic patients. *Acta Psychiatr Scand* 1991; 83:121-124
- Chochinov HM, Wilson KG, Enns M, Lander S: Prevalence of depression in the terminally ill: effects of diagnostic criteria and symptom threshold judgments. *Am J Psychiatry* 1994; 151:537-540
- Fishman B, Pasternak S, Wallenstein SI, Houde RW, Holland JC, Foley KM: The Memorial Pain Assessment Card: a valid instrument for the evaluation of cancer pain. *Cancer* 1987; 60:1151-1158
- Ahles TA, Ruckdeschel JC, Blanchard EB: Cancer-related pain, II: assessment with visual analogue scales. *J Psychosom Res* 1984; 28:121-124
- Volk RJ, Pace TM, Parchman ML: Screening for depression in primary care patients: dimensionality of the short form of the Beck Depression Inventory. *Psychol Assessment* 1993; 5:173-181