

# Symptoms of Major Depression in People With Spinal Cord Injury: Implications for Screening

Charles H. Bombardier, PhD, J. Scott Richards, PhD, James S. Krause, PhD, David Tulsy, PhD, Denise G. Tate, PhD

**ABSTRACT.** Bombardier CH, Richards JS, Krause JS, Tulsy D, Tate DG. Symptoms of major depression in people with spinal cord injury: implication for screening. *Arch Phys Med Rehabil* 2004;85:1749-56.

**Objectives:** To provide psychometric data on a self-report measure of major depressive disorder (MDD) and to determine whether somatic symptoms are nonspecific or count toward the diagnosis.

**Design:** Survey.

**Setting:** Data from the National Spinal Cord Injury Statistical Center representing 16 Model Spinal Cord Injury Systems.

**Participants:** Eight hundred forty-nine people with spinal cord injury who completed a standardized follow-up evaluation 1 year after injury.

**Interventions:** Not applicable.

**Main Outcome Measures:** The Patient Health Questionnaire-9 (PHQ-9), a measure of MDD as defined by the *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*. We computed descriptive statistics on rates of depressive symptoms and probable MDD, evaluated internal consistency and construct validity, and analyzed the accuracy of individual items as predictors of MDD.

**Results:** Exactly 11.4% of participants met criteria for probable MDD. Probable MDD was associated with poorer subjective health, lower satisfaction with life, and more difficulty in daily role functioning. Probable MDD was not related to most demographic or injury-related variables. Both somatic and psychologic symptoms predicted probable MDD.

**Conclusions:** The PHQ-9 has promise as a tool with which to identify probable MDD in people with SCI. Somatic symptoms should be counted toward the diagnosis and should alert health care providers to the likelihood of MDD. More efficient screening is only one of the quality improvement efforts needed to enhance management of MDD.

**Key Words:** Depression; Rehabilitation; Spinal cord injuries.

© 2004 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

**S**YMPTOMS OF DEPRESSION have been studied extensively among people with spinal cord injury<sup>1,2</sup> (SCI). From these reports we know that depressive symptoms, as well as major depressive disorder<sup>3</sup> (MDD), are highly prevalent and associated with negative outcomes. As a result, considerable emphasis has been placed on screening, diagnosing and treating MDD more effectively in people with SCI.<sup>4</sup> Despite these efforts, there is little evidence that detection or treatment of MDD among people with SCI has improved. This situation is not unique to SCI rehabilitation; clinicians and researchers in primary care settings continue to struggle with underdiagnosis and undertreatment of MDD.<sup>5</sup> Specialists in SCI rehabilitation can learn from innovations in primary care that aim to improve screening measures, to integrate screening into routine clinical care, and to develop models of care that increase the probability that efficacious treatments will be delivered to patients with MDD in sufficient dosages and of adequate duration to be clinically effective (for a review, see Pignone et al<sup>5</sup>; for a critique, see Coyne et al<sup>6</sup>).

There are numerous barriers to efficient and effective screening, diagnosis, and treatment of MDD, as recommended in the Consortium for Spinal Cord Medicine's clinical practice guideline on depression.<sup>4</sup> One barrier is that current screening measures are considered too long and cumbersome to be used routinely in typical medical practice.<sup>7-9</sup> Most commonly used measures of depressive symptomatology—such as the Beck Depression Inventory, Center for Epidemiological Studies–Depression Scale, Older Adult Health and Mood Questionnaire, Inventory to Diagnose Depression, and the Zung Depression Rating Scale—have 20 to 22 items and require on average 2 to 5 minutes for able-bodied patients to complete.<sup>9</sup> Another barrier is that, with few exceptions,<sup>10</sup> most “depression” screening instruments have item content, question format, time frame, or cutoff scores that make it difficult to relate self-report data to a *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*<sup>3</sup> (DSM-IV) diagnosis of MDD.<sup>2</sup> Use of non-diagnostic measures also creates an additional step in the treatment planning process, because clinicians must conduct a separate evaluation to diagnose MDD.<sup>4</sup> Delayed or absent diagnostic evaluations and a failure to distinguish MDD from more general psychologic distress can affect important decisions about treatment, especially whether antidepressant medications should be recommended. A third barrier pertains to whether clinicians attribute somatic symptoms (ie, weight change, appetite change, low energy, psychomotor changes) to MDD or to the primary effects of SCI.<sup>2</sup> How they answer this question is a factor in arriving at a diagnosis of MDD. Existing data present a conflicting picture, with 1 study suggesting that psychologic symptoms are “core elements” of depression in SCI,<sup>11</sup> whereas another shows that symptoms uniquely predictive of major depression include psychomotor changes, changes in appetite, and altered sleep.<sup>10</sup>

This study presents preliminary data on a relatively new screening measure devised to identify MDD among primary care patients. The Patient Health Questionnaire-9<sup>7,12</sup> (PHQ-9) attempts to reduce barriers to effective diagnosis and treatment in several

From Rehabilitation Medicine, University of Washington School of Medicine, Seattle, WA (Bombardier); Spain Rehabilitation Center, University of Alabama, Birmingham, AL (Richards); College of Health Professions, Medical University of South Carolina, Charleston, SC (Krause); Kessler Medical Rehabilitation Research & Education Corp and University of Medicine and Dentistry of New Jersey, West Orange, NJ (Tulsy); and University of Michigan Model Spinal Cord Injury Care System, University of Michigan, Ann Arbor, MI (Tate).

Supported by funds from the National Institute on Disability and Rehabilitation Research in the Office of Special Education and Rehabilitation Services in the US Department of Education, Spinal Cord Injury Model Systems program.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the author(s) or upon any organization with which the author(s) is/are associated.

Correspondence to Charles H. Bombardier, PhD, Dept of Rehabilitation Medicine, Box 359740, Harborview Medical Center, 325 9th Ave, Seattle, WA 98104, e-mail: chb@u.washington.edu. Reprints are not available from the author.

0003-9993/04/8511-9245\$30.00/0

doi:10.1016/j.apmr.2004.07.348

ways. First, with just 9 items, the PHQ is substantially shorter than most depression screening measures. Shorter measures may be more acceptable to patients and more likely to be used for routine screening in primary care settings.<sup>8</sup> Next, the items of the PHQ-9 exactly parallel the 9 diagnostic criteria that make up the DSM-IV criteria for major depression. The format and temporal framework of the items also corresponds to the DSM-IV criteria in that subjects are asked whether each symptom bothered them “not at all,” “several days,” “more than half the days,” or “nearly every day” over the past 2 weeks. Basing the item content and format on DSM-IV criteria potentially obviates the need for a separate follow-up diagnostic assessment. The clinician determines the diagnosis by reviewing patient responses to each of the 9 diagnostic criteria, clarifying responses if needed, and also ruling out physical causes of depression, normal bereavement, and history of mania before offering pharmacologic treatment or psychotherapy.<sup>7</sup> Unlike most other measures, the PHQ-9 was developed, tested, and refined for use with medical patients, not psychiatric patients or community residents. This is important because the criterion validity had to be established in patients with high rates of nonspecific physical symptoms that may confound the diagnosis of MDD. The instrument also has to show acceptability among nonpsychiatric patients, plus its ease of use by busy primary care practitioners. Each of these issues is relevant to screening for depression among people with SCI.

Our primary goal in this study was to provide preliminary data on the clinical utility of the PHQ-9 with people with SCI, including answering the question of the validity of somatic symptoms in diagnosing MDD. Our first hypothesis was that the PHQ-9 would show good psychometric properties and produce results consistent with prior studies of MDD in people with SCI. We measured internal consistency (Cronbach  $\alpha$ ) and tested several a priori hypotheses about the PHQ-9 that, if supported, would provide evidence for its construct validity in an SCI population. We predicted that scores on the PHQ-9 would correlate inversely with satisfaction with life and subjective health as well as correlate positively with greater interference with daily role functioning. Next, we predicted that rates of probable MDD found in this sample would be consistent with those of similar studies. We examined the relationship of PHQ-9–diagnosed MDD to demographic factors and injury-related variables to determine whether the measure is biased by any of these variables. Previous studies have shown that depressive symptoms are generally unrelated to level of injury, injury severity, or disability. However, based on epidemiologic studies, we anticipated higher rates of MDD might be found among women with SCI and younger age cohorts.<sup>13,14</sup> Our final hypothesis was that somatic symptoms (sleep changes, psychomotor changes, appetite changes, poor energy) would be as predictive of probable MDD as psychological symptoms.

## METHODS

### Participants

Participants were recruited from 16 Model Spinal Cord Injury Systems (MSCIS) centers throughout the United States with both urban and rural catchment areas. The study sample was comprised of 849 persons who sustained traumatic SCI between August 30, 2000, and April 1, 2003. Four sites also provided data on satisfaction with life and subjective health. Data from this subsample of 144 persons were used in analyses related to the construct validity of the PHQ-9. Participants were 17 years old or older and met the eligibility criteria of the MSCIS. That is, they received acute medical or surgical care, inpatient rehabilitation, outpatient rehabilitation, or day-hospital rehabilitation within one of the MSCIS programs. Exclusion

criteria were completion of an organized rehabilitation program before system admission, non-US citizen, or lacking a clear neurologic impairment after SCI.

All participants gave informed consent for data collection. Each MSCIS center obtained approval of its individualized research protocols in accordance with the ethical standards of its local institutional review boards. We facilitated appropriate standards of care concerning safety and confidentiality of people discovered to be depressed and/or suicidal by distributing a protocol for handling potentially suicidal subjects for inclusion in human subjects applications and by reviewing the protocol with data collectors. The protocol outlined plans for professional back-up for depressed persons needing further evaluation, referral procedures, and procedures for identifying emergency conditions under which confidentiality could be broken to prevent imminent harm to the subject.

### Materials and Methods

This cross-sectional study was based on data collected during the 1-year follow-up assessment. Data were obtained in person (9.5%), by telephone (79.8%), self-administered questionnaire (8.5%), or by using a combination of these methods (2.2%) with eligible subjects who consented to participate. Initial demographic and injury-related characteristics were obtained from interviews, physical examinations, and hospital records near the time of a person's initial rehabilitation hospitalization. Neurologic injury was classified using a combination of the *International Standards for Neurological Classification of Spinal Cord Injury* developed by the American Spinal Injury Association (ASIA) and the International Medical Society of Paraplegia.<sup>15</sup> Subjects were categorized into 2 injury levels, tetraplegia and paraplegia. The ASIA Impairment Scale was used to classify injury severity into complete (ASIA grade A) versus incomplete (ASIA grades B–E).

**Functional independence.** The motor component of the FIM instrument was used to measure injury-related impairment at the time of discharge from inpatient rehabilitation. The motor FIM has 13 items that assess degree of independence, from 1 (complete dependence) to 7 (complete independence), in 4 areas of function: self-care activities, sphincter control, mobility, and locomotion.<sup>16</sup> Motor FIM scores range from 13 (maximal dependence) to 91 (maximal independence). Rehabilitation therapists completed motor FIM scores for all patients at discharge.

**Probable MDD.** As noted earlier, the PHQ-9 was developed to facilitate identification and diagnosis of DSM-IV MDD in medical samples. To be consistent with DSM-IV MDD diagnostic criteria, each of the 9 depression items is rated according to how persistent the symptom has been in the past 2 weeks: 0 (not at all), 1 (several days), 2 (more than half the days), or 3 (nearly every day) (table 1). MDD can be diagnosed by first determining whether a patient reported at least 1 of the 2 essential criteria, depressed mood or anhedonia, on more than half of the last 14 days. If so, the remaining items are reviewed to determine if a total of 5 or more symptoms were present more than half the days in that time period. However, the item on thoughts of death or suicide was counted if reported at all. In this article, we considered those who met these categorical criteria to have “probable MDD.” The PHQ-9 may also be used as a continuous measure of depression severity. The values for each of the 9 item scores can be summed, resulting in a total score ranging from 0 to 27.

The PHQ-9 showed good internal consistency (Cronbach  $\alpha=.89$ ) and test-retest reliability ( $r=.84$ ) in a study of 3000 primary care patients and 3000 obstetrics and gynecology patients.<sup>7</sup> Kroenke et al<sup>7</sup> also evaluated the PHQ-9's diagnostic

Table 1: PHQ-9 Items

Over the last 2 weeks, how often have you been bothered by any of the following problems?
Response options: (0) not at all, (1) several days, (2) more than half the days, (3) nearly every day
1. Little interest or pleasure in doing things.
2. Feeling down, depressed, or hopeless.
3. Trouble falling or staying asleep, or sleeping too much.
4. Feeling tired or having little energy.
5. Poor appetite or overeating.
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down.
7. Trouble concentrating on things, such as reading the newspaper or watching television.
8. Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual.
9. Thoughts that you would be better off dead or of hurting yourself in some way.

NOTE. Reprinted from the PRIME-MD® Patient Health Questionnaire. ©1999, Pfizer Inc.

validity by comparing its total score in diagnosed MDD to an independent, structured diagnostic assessment of DSM-IV MDD made by a mental health professional. When used as a categorical diagnostic measure, the PHQ-9 had a sensitivity of 73% and a specificity of 98% for independently diagnosed MDD. Construct validity was documented through significant correlations between increasing levels of depressive symptoms and poorer health-related quality of life (HRQOL), greater disability days, and more physician visits.

**Daily role functioning.** The DSM-IV criterion C for major depression has to do with whether the symptoms cause “clinically significant distress or impairment in social, occupational, or other important areas of functioning.”<sup>3</sup> Therefore, an item in the PHQ-9 assesses the extent to which symptoms of depression have made it difficult for “you to do your work, take care of things at home, or get along with other people?” Response options include not at all, somewhat, very, or extremely difficult. Only subjects who reported at least 1 symptom of depression were asked this question.

**Subjective health.** This is a single item from the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) that assesses a patient’s perception of his/her overall health. The item reads: “In general, would you say that your health is: excellent, very good, good, fair, or poor.” For analyses, answers are assigned a score ranging from 1 (excellent) to 5 (poor). This question alone has been highly predictive of HRQOL and accounts for a significant proportion of the variance in the SF-36.<sup>17</sup>

**Satisfaction with life.** The Satisfaction With Life Scale is a validated measure of subjective satisfaction with life; it has 5 items, all of which load on a single factor of well-being.<sup>18</sup> Each item is a statement of satisfaction with life and is rated on a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) to produce a total score between 5 and 35.

### Statistical Analyses

To examine internal consistency, we calculated item-total correlations and the Cronbach  $\alpha$  on the PHQ-9 items. We used Spearman correlations and chi-square tests to measure the association between the PHQ-9 and measures of quality of life, subjective health, and difficulty in role functioning. Next we calculated descriptive statistics, including proportions of the

sample that met criteria for mild, moderate, and severe depressive symptoms, as well as the proportion that met criteria for probable MDD. We computed chi-square tests to determine whether rates of probable MDD were biased by demographic or injury-related factors. Last, to evaluate the predictive utility of somatic and psychologic items, we created 2×2 contingency tables for each PHQ-9 symptom and the criterion measure, which was whether the person met criteria for probable MDD. Various measures of clinical usefulness were computed for each item, including sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV), and positive and negative likelihood ratios.<sup>19</sup> These indices were used because they can characterize the predictive characteristics of each PHQ-9 symptom. Based on the study by Clay et al,<sup>10</sup> we calculated a measure of predictive efficiency for each symptom. According to Widiger et al,<sup>20</sup> an indicator is an efficient predictor if the ratio of the base rate to 1 minus the base rate is greater than the ratio of the false positive rate to the true positive rate.

As applied in this study, the sensitivity of a symptom refers to the proportion of subjects with probable MDD who report that symptom, or how certain it is that reporting a particular symptom identifies everyone who has MDD. Specificity of a symptom is the proportion of subjects without probable MDD who do not report the symptom, or how certain it is that not reporting a symptom identifies everyone without probable MDD. The PPV of a symptom refers to the fraction of people who report the symptom and also have probable MDD or our confidence that those with a symptom have probable MDD. The NPV of a symptom corresponds to the fraction of patients who do not report a symptom and do not have probable MDD, or our confidence that the absence of a symptom means the absence of probable MDD. A positive likelihood ratio refers to the odds that a given symptom would be reported by a patient with probable MDD, as opposed to without. A negative likelihood ratio refers to the odds that a given symptom would not be reported by a patient with MDD versus without.

## RESULTS

### Sample Characteristics

Our sample included 849 participants who completed evaluations at 1 year post-SCI. Sample characteristics were 645 (76.0%) male; 616 (72.9%) white; 186 (21.9%) African American; 2 (0.2%) Native American, Eskimo, or Aleut; 14 (1.6%) Asian or Pacific Islanders; 18 (2.1%) other/unclassified; and 88 (10.4%) Hispanic. Nearly half (47.6%) had ASIA grade A “complete” injuries and 45.5% had paraplegia. Mean age at the

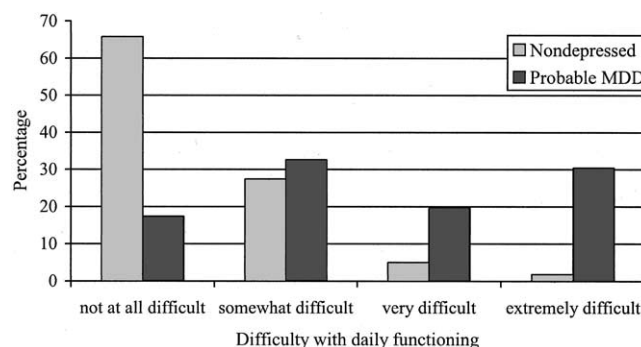


Fig 1. Relation between probable major depression and difficulty in daily role functioning (n=638).

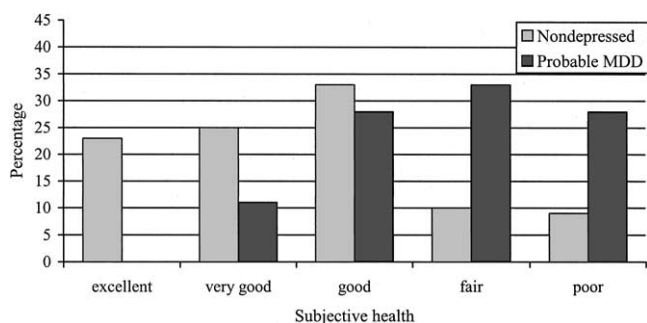


Fig 2. Relation between probable major depression and subjective health (n=144).

time of injury ± standard deviation was 36.9±15.0 years (range, 17–88y). Of the sample, 22.7% had less than a high school education, 58.1% had completed high school, and 19.1% reported having more than a high school education.

**Internal Consistency and Construct Validity**

To assess the internal consistency of the PHQ-9, we calculated the overall Cronbach α (.87). The corrected item total correlations ranged from .72 (depressed mood) and .69 (failure) to .45 (psychomotor agitation/retardation) and .48 (suicidal ideation). In terms of construct validity, greater depressive symptom severity as measured by the PHQ-9 correlated inversely with satisfaction with life (Spearman  $\rho_{144} = -.51, P < .001$ ) and subjective health (Spearman  $\rho_{144} = .50, P < .001$ ; higher numbers reflect poorer health). Greater depressive symptom severity correlated positively with greater difficulty in daily role functioning (Spearman  $\rho_{638} = .62, P < .001$ ). The relationships between probable MDD and subjective health and difficulty with daily role functioning are shown in figures 1 and 2. Fifty percent of patients with probable MDD reported that their depressive symptoms made it very difficult or extremely difficult to do work, take care of things at home, or get along with people, whereas only 6.8% of those with less severe depressive symptoms had this much difficulty ( $\chi^2_{1} \text{ test} = 129.4, P < .001$ ). In addition, 61.1% of those with probable MDD reported their health as fair to poor, whereas only 19% of those without probable MDD reported their health as fair to poor ( $\chi^2_{1} \text{ test} = 15.1, P < .001$ ). No one with probable MDD reported their health as excellent, whereas 23% of those without MDD reported excellent health.

**Depression Symptom Severity and Rate of Probable Major Depression**

The average PHQ-9 total score for the entire sample was 5.48 (95% confidence interval, 5.07–5.88). Based on previously pub-

lished severity ranges,<sup>7</sup> the sample was divided into 6 categories of depression symptom severity (table 2). Fifty-eight percent of the sample reported no symptoms or only minimal symptoms of depression. Ninety-six (11.4%) people met criteria for probable MDD, with a mean PHQ-9 score of 18.1±3.9, whereas the nondepressed group had an overall mean score of 3.8±3.9. We dichotomized the sample into probable MDD and nondepressed subjects for the subsequent analyses.

**Relationship of Probable Major Depression to Clinical and Injury-Related Variables**

Rates of probable MDD did not differ with respect to injury level (paraplegia vs tetraplegia), injury severity (complete vs incomplete), FIM discharge scores, sex, race (white vs non-white), or level of education at the time of injury (<high school vs high school vs >high school). Rates of probable MDD did not differ significantly by method of data collection; however, about 20% of the 20 people who completed the measure by themselves had probable MDD versus 10% to 11.6% of those interviewed by other methods. The rate of probable MDD was significantly higher among people aged 25 to 49 years (15.0%), compared with older (8.7%) or younger (6.5%) age groups (overall  $\chi^2_{2} \text{ test} = 11.6, P = .003$ ). The proportion with MDD was significantly lower among those who were single when injured, compared with those who were married or divorced, widowed, or separated (overall  $\chi^2_{2} \text{ test} = 15.6, P < .001$ ).

**Individual Symptoms as Predictors of Probable Major Depression**

Figure 3 displays the entire range of individual symptom reports for the total sample. The second column of table 3 shows the percentage of the sample that reported each symptom was present more than half the days of the week. The symptoms most frequently reported to this degree were sleep disturbance (26.0%), feeling tired or having little energy (24.3%), anhedonia (19.6%), and depressed mood (18.7%). Nearly 1 in 6 (15.4%) reported thinking it would be better to be dead or of injuring themselves in some way within the past 2 weeks.

Indexes of the predictive utility for each of the PHQ-9 symptoms are also provided in table 3. Three psychologic and 2 somatic symptoms are highly sensitive (>80%) indicators of probable MDD: depressed mood, disturbed sleep, decreased energy, anhedonia, and feelings of failure. That is, more than 80% of all those with probable MDD reported each of these symptoms. All PHQ-9 symptoms are relatively specific. Two somatic symptoms (psychomotor changes, appetite change) as well as having trouble concentrating are among the 5 items with specificities above 90%. All symptoms had relatively low PPV, implying that a large proportion of those who report a particular item will not have MDD. In contrast, the NPV of all items was quite high. The

Table 2: Rates of Depression Symptom Severity

Label	Definition	Rate, n (%)
No depressive symptoms	PHQ-9=0	199 (23.4)
Minimal depressive symptoms	PHQ-9=1 to 4	294 (34.6)
Mild depressive symptoms	PHQ-9=5 to 9	170 (20.0)
Moderate depressive symptoms	PHQ-9=10 to 14	101 (11.9)
Moderate/severe depressive symptoms	PHQ-9=15 to 19	48 (5.7)
Severe depressive symptoms	PHQ-9=20 to 27	37 (4.4)
Probable MDD	Reports at least 5/9 symptoms (must include either depressed mood or anhedonia)	96 (11.4)

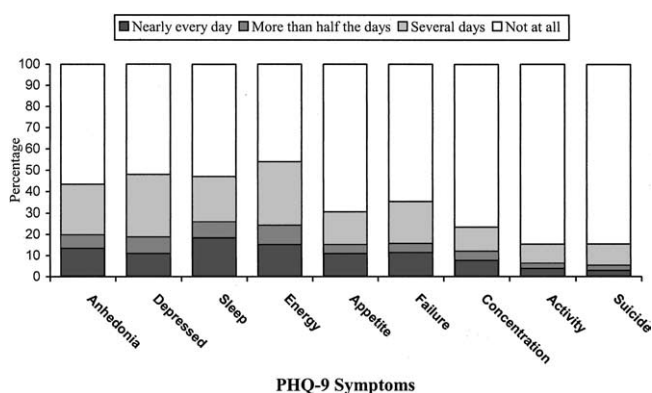


Fig 3. Reports of PHQ-9 depression symptoms for the entire sample (N=849).

probability of not having MDD was very high when none of the items were reported. Likelihood ratios for a positive test were uniformly high. Depending on the symptom, the odds that reporting a given item would occur in a participant having probable MDD versus a participant not having MDD ranged from 5:1 (sleep disturbance) to 18:1 (psychomotor changes). The likelihood ratios for a negative test tended to be less robust. Trouble concentrating, psychomotor changes, and suicidal ideation had only about one-half to two-thirds odds of being denied by someone with probable MDD. However, some symptoms, such as depressed mood, sleep disturbance, and poor energy, had low odds of being denied by someone who met probable MDD criteria. All symptoms met the criteria for efficient predictors,<sup>20</sup> although psychomotor changes and thoughts of death did so only marginally.

## DISCUSSION

The data provide preliminary evidence that the PHQ-9 may be a useful screen for probable MDD in people with SCI. The measure has shown high internal consistency and generally strong item-total correlations. The measure also shows moderately strong correlations with conceptually related variables such as satisfaction with life, subjective health, and difficulty performing daily activities. Half of the subjects with probable MDD reported that depressive symptoms make it difficult or extremely difficult to function at work, at home, or socially. Sixty-one percent of those with probable MDD rated their health as fair to poor. Satisfaction with life and subjective health were significantly and inversely correlated with increasing depressive symptomatology. Our results are consistent with

previous research that showed that more severe depressive symptoms are associated with diminished health and well-being in people with SCI,<sup>2</sup> as well as among people with other medical disorders.<sup>21</sup>

The rate of probable MDD detected in this sample is consistent with the rates found in previous studies of MDD among people with SCI. Major depression rates found in studies that used DSM diagnostic criteria range widely, from about 9.8% to 37.5%, although most estimates fall in the 15% to 23% range (table 4). The rate reported in this study is also consistent with the base rate of MDD (10%) among 3000 primary care patients in the PHQ validation study.<sup>12</sup> In contrast, epidemiologic studies show the point prevalence of MDD in primary care settings is slightly lower, between 4.8% and 8.6%.<sup>22</sup> The original validity study showed that PHQ-9 diagnosed MDD had a sensitivity of 73% and a specificity of 98% and the PHQ was not found to have a systematic tendency to overdiagnose or underdiagnose MDD.<sup>12</sup> Subsequent research showed that scoring the PHQ-9 as a continuous measure and using a cutoff score of 10 or more resulted in the most accurate predictions (88% sensitivity, 88% specificity), compared with an independent diagnostic assessment by a mental health professional.<sup>7</sup> If we use a cutoff of 10 or more to determine presence or absence of probable MDD in our sample, a prevalence of 22% is obtained. Therefore, depending on how the PHQ-9 is scored, rates of probable MDD may differ by a factor of 2, although with both estimates falling within the range we would expect based on prior research.

Probable MDD was not related to SCI-related variables such as injury severity, level of injury, or magnitude of functional impairment. Demographic factors such as gender, level of education, and race were also unrelated. However, age and marital status seemed to influence rates. Contrary to predictions, the middle-age group (age range, 25–49y), not the youngest group, reported probable MDD significantly more frequently than either the older or younger age groups. This contrasts in part with data showing that rates of depression are higher among younger birth cohorts in the general population.<sup>13</sup> Next, single people met criteria for probable MDD significantly less frequently than did those who were married or divorced, widowed, or separated. In contrast, divorce or separation are often found to be risk factors for MDD among people with primary psychiatric disorders.<sup>23</sup>

Together with previous studies of the PHQ-9, these data suggest the measure is worthy of further study as a diagnostic screening measure based on self-report by people with SCI. It represents important advances in screening for MDD through its brevity, fidelity to DSM-IV diagnostic criteria, good diag-

Table 3: Indexes of Clinical Usefulness for Individual Items Predicting MDD Diagnosed by PHQ-9

Symptom	Base Rate* (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Likelihood Ratio for a Positive Test	Likelihood Ratio for a Negative Test
Anhedonia	19.6	84.4	88.7	49.1	97.8	7.5	.18
Depressed mood	18.7	93.8	90.9	56.9	99.1	10.3	.07
Sleep disturbance	26.0	89.5	82.2	39.1	98.4	5.0	.13
Poor energy	24.3	87.5	83.8	40.8	98.1	5.4	.15
Appetite change	15.0	72.6	92.2	54.3	96.4	9.3	.30
Feelings of failure	15.5	80.2	92.8	58.8	97.3	11.1	.21
Trouble concentrating	11.8	56.2	93.8	54.0	94.3	9.1	.47
Psychomotor changes	6.3	37.9	97.7	67.9	92.5	18.0	.64
Suicidal ideation	15.4	56.4	89.8	41.1	94.2	5.1	.49

\*The percentage of subjects reporting they had this symptom more than half the days over the past 2 weeks. However, in the case of suicidal ideation, having this symptom several days or more per week is considered clinically significant.

Table 4: Summary of Studies on Rates of MDD After SCI

Study	Population Studied	Sample Size	Diagnostic Method	Rates of MDD or Probable MDD, n (%)
Clay et al <sup>10</sup>	Inpatients	133	IDD, DSM-III-R	17 (12.8)
Federoff et al <sup>31</sup>	Inpatients	55	DSM-III	12 (21.8)
Frank et al <sup>32</sup>	Mixed	32	DSM-III	14 (37.5)
Frank et al <sup>11</sup>	Inpatients	132	IDD, DSM-III-R	13 (9.8)
Fullerton et al <sup>33</sup>	Inpatients	30	SADS/RDC	9 (30.0)
Judd et al <sup>34</sup>	Inpatients	84	DSM-III, HDRS	9 (10.7)
Judd and Brown <sup>35</sup>	Inpatients	227	DSM-III	31 (13.7)
MacDonald et al <sup>36</sup>	Community	53	CDM (RDC)	8 (15)
Tate et al <sup>37</sup>	Inpatients	30	DSM-III-R	7 (23.3)

Abbreviations: CDM, Clinical Depression Measure; HDRS, Hamilton Depression Rating Scale; IDD, Inventory to Diagnose Depression; RDC, research diagnostic criteria; SADS, Schedule for Affective Disorders and Schizophrenia.

nostic sensitivity, and excellent specificity in medical samples and by potentially obviating the need for separate diagnostic assessments. However, before firm conclusions can be made about its diagnostic validity with people with SCI, studies are needed that evaluate its sensitivity and specificity compared with a criterion standard depression diagnostic measure such as the Structured Clinical Interview for the DSM-IV.<sup>24</sup> Research is needed that compares the diagnostic accuracy of the PHQ-9 as a DSM-IV categorical measure versus using a cutoff score and comparing self-administration of the PHQ-9 to interview-based administration among people with SCI. Finally, to promote more widespread screening for major depression, clinical researchers should continue testing the potential clinical utility of even shorter measures such as the 2-item screen suggested by Whooley et al.<sup>8</sup>

This study confirms and extends previous research by Clay et al<sup>10</sup> that suggests somatic symptoms of depression such as appetite change, sleep disturbance, and poor energy are sensitive and specific to people with probable MDD. Within this large SCI sample assessed 1 year after injury, all depression symptoms had high specificity and high NPV in predicting who would or would not meet criteria for probable MDD. Moreover, all symptoms met criteria as efficient predictors of probable MDD according to the formula suggested by Widiger et al.<sup>20</sup>

Therefore, these results lend credence to the "inclusive" approach for diagnosing MDD in people with SCI.<sup>9</sup> The inclusive approach includes depressive symptoms in diagnosing MDD regardless of the suspected etiology of the symptoms. In contrast, the etiologic approach follows DSM-IV rules and includes symptoms in a depression diagnosis unless the symptom is "the direct physiological consequence of the general medical condition."<sup>3(p325)</sup> A third diagnostic strategy, the substitutive approach, replaces etiologically ambiguous somatic symptoms with purely psychologic symptoms of depression. The inclusive approach is thought to be the most reliable and sensitive but may increase the risk of false-positive diagnoses.<sup>9</sup> The substitutive approach represents movement away from universally accepted diagnostic criteria and conflicts with the fact that neither national nor international nomenclature groups have recommended altering depression diagnostic criteria for other specific diseases.<sup>25</sup> With regard to the etiologic approach, the sensitivity and specificity data presented here suggest somatic symptoms of depression are not nonspecific symptoms of SCI but are predictive of probable MDD and should be "counted" toward the diagnosis.

### Limitations

As noted earlier, this study lacks a criterion standard diagnostic measure of MDD to which the PHQ-9 can be compared. Therefore, rates of probable MDD reported here remain only estimates of DSM-IV MDD. However, because of the stringent criteria used to define caseness, our data most likely underestimate MDD in this sample. And, we do not have data on the test-retest or interrater reliability of the PHQ-9. Prior data on test-retest reliability showed good temporal stability ( $r=.84$ ) over a 48-hour period and included a methodology that compared 2 of the data collection methods (self-report vs telephone administration) used in this study.

Although our sample was large and geographically and racially diverse, participants were recruited from the MSCIS centers of care and were not population based.<sup>26</sup> Therefore, the data have limited usefulness in helping to determine important epidemiologic information such as the prevalence of MDD. Other limitations that may have influenced the representativeness and validity of our findings are potential selection biases, subject retention biases, and under- or overreporting of symptoms on self-report measures. We were unable to determine the extent to which probable major depression had already been identified and treated because we did not have information on antidepressant medication use by the sample or about how many subjects were receiving psychotherapy. However, general information on postdischarge services suggests that 18.8% of patients received outpatient psychologic or vocational counseling at some time between discharge and their first anniversary of their injury.<sup>27</sup>

Another limitation is that these data are limited to 1 point in time—1 year after injury. The predictive validity of the PHQ-9 somatic items may vary as a function of time since injury. For example, soon after injury and especially during acute rehabilitation, there may be more nonspecific somatic symptoms such as sleep disturbance and weight loss. This may detract from the diagnostic sensitivity and specificity of PHQ-9 somatic symptoms as they relate to MDD. Rates of probable MDD also may vary over time. Some research has found a general improvement in depressive symptoms over time<sup>28</sup>; however, other research suggests that a subgroup of people who show high levels of anxiety and depressive symptoms early continue to experience high levels of distress 1 to 2 years after injury.<sup>29</sup> Longitudinal studies are needed of major depressive disorder in unselected cases. Optimally, these studies would examine symptoms of MDD soon after injury and at multiple time points thereafter. It may be particularly interesting to note not

only changes in the proportion of probable MDD over time but also longitudinal changes in rates of somatic versus psychologic symptoms.

### CONCLUSIONS

This study supports a shorter and more streamlined approach to screening for depression after SCI. A 9-item self-report measure may be able to diagnose MDD with sufficient accuracy to reduce the need for a separate diagnostic interview. The brevity of this measure, the fidelity to DSM-IV criteria, and the straightforward diagnostic rules may encourage more widespread adoption of routine depression screening. These data also encourage an inclusive approach to diagnosing MDD in which clinicians count all DSM-IV symptoms toward the diagnosis rather than attributing the symptoms to nonspecific effects of SCI. For health care providers who do not adopt a universal screening approach, symptoms such as psychomotor changes, feelings of failure, depressed mood, appetite change, and difficulty concentrating are 9 to 18 times more likely to be reported by people with probable MDD and can serve as reliable red flags to trigger further assessment of depressive symptoms.

Screening is not an end in itself but should be instituted only as a first step toward effective treatment for MDD. Although systematic screening and feedback can significantly improve recognition of MDD, it is also important to acknowledge that simple screening and feedback to the health care provider is likely to improve management of MDD only modestly.<sup>5</sup> A meta-analysis<sup>5</sup> of primary care studies showed that programs that incorporate other quality improvement strategies produce more robust positive effects on long-term patient outcomes. To improve depression care significantly, proactive multidisciplinary efforts are needed to educate patients about major depression, to ensure evidence-based treatment, to promote adherence to treatment, and to use data management systems that provide feedback to the health care provider about needed alterations in the treatment plan.<sup>30</sup> Moreover, in SCI, there is a dearth of evidence regarding effective treatment for MDD, side effects of commonly used antidepressant medications, and patient preferences regarding treatment. The absence of such basic information probably inhibits more widespread adoption of quality improvement programs. Studying ways potentially to reduce barriers to screening and diagnosis is a useful step toward achieving this longer-term goal.

**Acknowledgments:** We thank Dr. Steve Williams and 2 anonymous reviewers for their helpful comments on earlier drafts of this article.

### References

1. Frank RG, Elliott TR, Corcoran JR, Wonderlich SA. Depression after spinal cord injury: is it necessary? *Clin Psychol Rev* 1987; 7:611-30.
2. Elliott TR, Frank RG. Depression following spinal cord injury. *Arch Phys Med Rehabil* 1996;77:816-23.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington (DC): APA; 1994.
4. Consortium for Spinal Cord Medicine. Depression following spinal cord injury: a clinical practice guideline for primary care physicians. Washington (DC): Paralyzed Veterans of America; 1998.
5. Pignone M, Gaynes B, Rushton J, et al. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;136:765-76.
6. Coyne JC, Palmer SC, Thompson R. Questionnaires for depression and anxiety. Routine screening entails additional pitfalls [letter]. *BMJ* 2001;323:168.
7. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression symptom severity measure. *J Gen Intern Med* 2001;16:606-13.
8. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med* 1997;12:439-45.
9. Williams JW Jr, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? *JAMA* 2002;287:1160-70.
10. Clay D, Hagglund K, Frank R, Elliott T, Chaney J. Enhancing the accuracy of depression diagnosis in patients with spinal cord injury using Bayesian analysis. *Rehabil Psychol* 1995;40:171-80.
11. Frank RG, Chaney JM, Clay DL, et al. Dysphoria: a major symptom factor in persons with disability or chronic illness. *Psychiatry Res* 1992;43:231-41.
12. Spitzer RL, Kroenke K, Williams J. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999;282:1737-44.
13. Weissman M, Bruce M, Leaf P, Holzer C. Affective disorders. In: Robins L, Regier D, editors. *Psychiatric disorders in America*. New York: Free Pr; 1991. p 52-80.
14. Krause JS, Kemp B, Coker J. Depression after spinal cord injury: relation to gender, ethnicity, aging, and socioeconomic indicators. *Arch Phys Med Rehabil* 2000;81:1099-109.
15. American Spinal Injury Association and International Medical Society of Paraplegia. International standards for neurological classification of spinal cord injury. 6th ed. Chicago; ASIA; 2000.
16. Linacre J, Heinemann A, Wright B, Granger C, Hamilton B. The structure and stability of the Functional Independence Measure. *Arch Phys Med Rehabil* 1994;75:127-32.
17. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;31:247-63.
18. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess* 1985;49:71-4.
19. Sackett DL. A primer on the precision and accuracy of the clinical examination. *JAMA* 1992;267:2638-44.
20. Widiger TA, Hurt SW, Frances A, Clarkin JF, Gilmore M. Diagnostic efficiency of DSM-III. *Arch Gen Psychiatry* 1984;41:1005-12.
21. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients. Results from the Medical Outcomes Study. *JAMA* 1989;262:914-9.
22. Association for Healthcare Policy Research. Depression in primary care: treatment of major depression. In: Panel DG, editor. Rockville: US Department of Health and Human Services, US Public Health Services, Agency for Health Care Policy and Research; 1993. Publication No. 93-0552.
23. Fava M, Kendler K. Major depressive disorder. *Neuron* 2000;28:335-41.
24. First M, Gibbon M, Spitzer R, Williams J. User's guide for the Structured Clinical Interview for DSM IV Axis I disorders. New York: Biometrics Res Dept, New York State Psychiatric Institute; 1996.
25. Gill D, Hatcher S. Antidepressants for depression in people with physical illness. *Cochrane Database Syst Rev* 2000;(4): CD001312.
26. DeVivo MJ, Go BK, Jackson AB. Overview of the national spinal cord injury statistical center database. *J Spinal Cord Med* 2002; 25:335-8.

27. National Spinal Cord Injury Statistical Center. Annual report. Birmingham: Univ Alabama; 2003.
28. Richards JS. Psychologic adjustment to spinal cord injury during first postdischarge year. *Arch Phys Med Rehabil* 1986;67:362-5.
29. Craig A, Hancock KM, Dickson HG. A longitudinal investigation into anxiety and depression in the first 2 years following a spinal cord injury. *Paraplegia* 1994;32:675-9.
30. Katon W, VonKorff M, Lin E, Simon G. Rethinking practitioner roles in chronic illness: the specialist, primary care physician and the practice nurse. *Gen Hosp Psychiatry* 2001;23:138-44.
31. Federoff J, Lipsey J, Starkstein S, Forrester A, Price T, Robinson R. Phenomenological comparisons of major depression following stroke, myocardial infarction or spinal cord injury. *J Affect Disord* 1991;22:83-9.
32. Frank RG, Kashani JH, Wonderlich S. Depression and adrenal function in spinal cord injury. *Am J Psychiatry* 1985;142:252-3.
33. Fullerton DT, Harvey R, Klein M. Psychiatric disorders in patients with spinal cord injury. *Arch Gen Psychiatry* 1981;32:369-71.
34. Judd FK, Burrows GD, Brown DJ. Depression following acute spinal cord injury. *Paraplegia* 1986;24:358-63.
35. Judd FK, Brown DJ. Psychiatric consultation in a spinal injuries unit. *Aust N Z J Psychiatry* 1992;26:218-22.
36. MacDonald M, Nielson W, Cameron M. Depression and activity patterns of spinal cord injury persons living in the community. *Arch Phys Med Rehabil* 1987;68:339-43.
37. Tate D, Forchheimer M, Maynard F, Davidoff G, Dijkers M. Comparing two measures of depression in spinal cord injury. *Rehabil Psychol* 1994;38:53-61.