A Comparison of Depression Screening Instruments in Hepatitis C and the Impact of Depression on Somatic Symptoms


Objective: Treatment of hepatitis C (HCV) with pegylated interferon-alpha (IFNα) can cause depression in approximately 30% of patients and underscores the need for effective detection of depression prior to and during IFNα treatment. Elevated rates of depression in untreated HCV can be a barrier to initiating HCV therapy and can impact fatigue and physical symptoms. In this preliminary study, we examined the accuracy of the seven-item Hamilton Depression Rating Scale (HAM-7) and Patient Health Questionnaire-9 (PHQ-9) in detecting depression in HCV-infected patients and determined the effect of major depression on somatic symptoms.

Methods: We conducted a preliminary comparison of operating characteristics of the PHQ-9 and HAM-7 to the MINI International Neuropsychiatric Interview for major depression in 116 individuals with chronic HCV assessed in an ambulatory office setting.

Results: Currently depressed chronic hepatitis C patients had significantly higher scores on all the scales compared with nondepressed patients. HAM-7 and PHQ-9 scores were significantly correlated with somatic and physical symptoms scales. Both the PHQ-9 and HAM-7 demonstrated comparable accuracy in detecting depression in comparison to the MINI. Conclusions: Our results suggest that the HAM-7 and PHQ-9 both have good operating characteristics compared with a criterion standard measure. Given that depression was associated with fatigue and increased somatic complaints, improved detection and treatment of depression could reduce disability and facilitate treatment for depressed HCV-infected patients.

Received January 22, 2011; revised February 8, 2011; accepted February 9, 2011. From University Health Network, Toronto General Hospital (SS, SEA); Medical Psychiatry Program, University Health Network (SS, SEA); Department of Psychiatry, University of Toronto (SS, DB, SEA); Department of Psychiatry, University of Toronto; King Saud University, Riyadh, Saudi Arabia (AAJ, FA); Department of Medicine, University of Toronto (GH); Liver Centre, Toronto Western Hospital (GH). Send correspondence and reprint requests to Sanjeev Sockalingam, M.D., FRCP, University Health Network, Toronto General Hospital, 200 Elizabeth Street, SEN-228, Toronto, ON M5G 2C4, Canada. e-mail: sockalingam@uhn.on.ca

© 2011 The Academy of Psychosomatic Medicine. Published by Elsevier Inc. All rights reserved.
Detecting Depression in Hepatitis C

scales. In addition, patients may be susceptible to IFN-MDE given the increased rates of depression in untreated patients with HCV. Rates of depression in untreated CHC patients have ranged from 7% to 62% and mood disorders may go undiagnosed in many patients. Moreover, baseline depressive symptoms are the most robust risk factor for IFNα-MDE and reinforce the importance of accurate identification of depression pre-IFNα treatment for HCV. Severe IFNα-MDE may precipitate serious psychiatric complications, such as suicidal ideation or psychosis, and may result in discontinuation of IFNα therapy.

Similar to other medical illnesses, depression may increase the severity of physical symptoms in patients with untreated CHC. In a retrospective study of 800 CHC patients, HCV infection was associated with higher rates of chronic fatigue and depression. Most patients with CHC report moderate to severe physical difficulties, specifically fatigue and flu-like symptoms, during IFNα therapy. Research suggests that the development of IFNα-MDE is also associated with a change in sleep quality, which can contribute to fatigue. Although several studies have shown higher rates of somatic symptoms and fatigue in depressed patients, limited data are available on the impact of depression on physical and fatigue symptoms in CHC patients.

Despite the importance of depression on the physical and mental health of patients with CHC, a paucity of studies have examined the effectiveness of depression screening tools in untreated patients with CHC and patients receiving IFNα for CHC. Using a cut-off score of 16 or more on the self-report scale, the Center for Epidemiologic Studies-Depression Scale (CES-D), Clark et al. determined that the CES-D had internal consistency reliability, construct validity, and predictive validity for depression in CHC patients. However, studies have failed to agree on optimal CES-D cut-off scores for the screening for major depression in HCV-infected patients. Dbouk et al. evaluated the performance of the CES-D, Beck Depression Inventory (BDI), and the Patient Health Questionnaire-9 (PHQ-9) as self-report screening tools for depression in an HCV patient population. This study demonstrated a strong correlation between PHQ-9, BDI, and CES-D scores for depression. In a second study of injection drug users infected with CHC, the CES-D and BDI had moderate agreement in screening for depression.

Recent studies have raised questions about the effectiveness of self-report depression measures in HCV compared with routine or standardized clinical assessment in specialty clinics. Leutscher et al. found benefit in using the Major Depression Inventory (MDI), a self-report measure, in detecting depression during IFNα therapy. However, this study did not evaluate the psychometric properties of the MDI in this patient population. They concluded that major depression in patients undergoing HCV therapy is often overlooked and advocated for the use of self-report scales, such as the MDI, in clinical practice. However, Philips et al. compared the CES-D to a single standardized question by a care provider. These studies were limited by the lack of comparison to a standardized psychiatric interview and clinician-rated depression scale.

Furthermore, no studies have examined the optimal cut-off scores for clinician-rated scales for CHC patients. Abbreviated clinician-rated scales, such as the seven-item Hamilton Depression Rating scale (HAM-7) (Toronto version), may have benefit in detecting depression in patients with CHC and can be used to determine response to pharmacologic and psychological treatments. In primary care settings, the HAM-7 is comparable to the 17-item Hamilton Depression Rating Scale (HAM-17) in measuring remission for patients on antidepressant treatment. Therefore, the HAM-7 may provide an alternative depression screening measure in CHC patients and may provide additional utility in determining remission of depression symptoms with treatment.

In this study of patients with CHC, we conducted a preliminary study comparing the performance of a self-administered depression measure, the PHQ-9, and a validated clinician-administered depression scale, the HAM-7, to a criterion standard interview for depression. The study was the first step to determining the operating characteristics of the HAM-7 and PHQ-9 specifically in a CHC patient population. Furthermore, we hoped to determine if previous associations between depression and increased somatic and fatigue symptoms observed in other patient populations applied to CHC patients.

METHODS

Study Sample

Patients were identified from consecutive assessments between July 1, 2008 and January 1, 2010 at the Liver Centre, Toronto Western Hospital, which is an ambulatory liver disease program with a mixed secondary and tertiary referral base. The clinic includes an integrated psychiatric consultation service that evaluates and manages psychiat-
A total of 144 patients were referred for psychiatric assessment within the defined study period, for whom complete psychiatric data was available for 116. Twenty-eight patients had incomplete self-report data due to language barriers (2), visual impairment (2) or patient refusal to complete the self-report questionnaires (22). For the purpose of this study, we analyzed the data collected for the 116 patients.

All patients had a confirmed diagnosis of HCV with a positive HCV quantitative RNA and all patients underwent HCV genotyping. Liver biopsy results within 12 months before or after the psychiatric assessment were used to determine the severity of liver disease.

Patients were assessed by the psychiatrist for the presence of a past or current substance use disorder as per Diagnostic Statistical Manual of Mental Disorders-IV (DSM-IV) criteria. Due to the potential impact of illicit drug use and alcohol use during HCV treatment, the psychiatric assessment focused on these specific substances and permitted an estimation of alcohol abuse/dependence and drug abuse/dependence rates in the sample.

Depression Measures

The MINI International Neuropsychiatric Interview (MINI) is a structured diagnostic interview with good reliability and validity and includes a depression module assessing patients for a current major depressive episode (MDE).\textsuperscript{20} The MINI was administered by a psychiatrist in the clinic and was used as the reference standard for depression. In addition, the MINI was used to determine a lifetime diagnosis of major depressive disorder or bipolar disorder using the respective modules.

We used the Patient Health Questionnaire-9 (PHQ-9) as our self-report measure of depression.\textsuperscript{21} We selected the
Detecting Depression in Hepatitis C

PHQ-9 over other self-report measures, such as the BDI or CES-D, due to its brevity, correspondence with DSM-IV criteria, and its widespread use in medically ill patient populations including HCV-infection.\textsuperscript{16,22–25} Scores on the PHQ-9 can range from 0 to 27 and mild, moderate, moderately severe, and severe levels of depressive symptoms correspond to cut-off scores of 5, 10, 15, and 20, respectively.

An abbreviated clinician-rated depression scale, the Toronto seven-item Hamilton Depression Rating Scale (HAM-7) offers comparable psychometric properties as the longer 17-item Hamilton Depression Rating Scale,\textsuperscript{26,27} and its brevity is an advantage\textsuperscript{(18)}. The HAM-7 incorporates seven specific domains of the HAM-17: depression, anhedonia, guilt, two anxiety domains, energy, and suicidal ideation.\textsuperscript{28} The Toronto HAM-17 overlaps with the seven-items on the HAM-17 identified by Santen and colleagues as the most sensitive items for detecting response to antidepressant therapy.\textsuperscript{29} Scores on the HAM-7 can range from 0 to 26 with scores of 3 or less indicating remission of depressive symptoms.

Fatigue and Physical Symptom Measures

Although the PHQ-15 is a measure of somatic symptoms in patients with somatoform disorders,\textsuperscript{30} the PHQ-15 has also been used to estimate overall somatic symptom burden in elderly and military samples.\textsuperscript{31,32} Scores on the PHQ-15 range from 0 to 30 and cut-off scores for mild, moderate, and severe somatic symptom severity are 5, 10, and 15, respectively.

The Fatigue Severity Scale (FSS) is a nine-item self-report scale for measuring fatigue and each item is rated on a seven-point Likert type scale.\textsuperscript{33} Higher scores on the FSS represent a higher impact of fatigue on everyday life. It has been used in a range of medical conditions, including HIV-infection,\textsuperscript{34} Parkinson’s disease,\textsuperscript{35} and multiple sclerosis.\textsuperscript{36} In a study examining the psychometric qualities of the FSS in patients with HCV, the FSS demonstrated good internal consistency, reliability, and test-retest reliability.\textsuperscript{37}

The Epworth Sleepiness Scale (ESS) is a self-administered measure of daytime sleepiness and can be a harbinger of underlying sleep disorders, such as obstructive sleep apnea, which can contribute to fatigue.\textsuperscript{38} Respondents are asked to rate their likelihood of dozing off in eight situations, using a four-point Likert scale, and total scores can range from 0 to 24. A score of 0 to 10 is normal, 10 to 12 is borderline, and greater than 12 is abnormal.

Demographics

Demographic data, specifically age, gender, human immunodeficiency virus (HIV), and liver biopsy results were collected retrospectively from patient charts. The most recent liver biopsy results were analyzed, and liver biopsies were conducted within 12 months of psychiatric assessment. A total of 21 patients who were genotypes 2 and 3 did not have a liver biopsy and three patients who were genotype 1 did not have a liver biopsy within the 12-month time frame. None of these 24 patients had clinical evidence of liver cirrhosis as per clinical examination and liver function tests.

Procedures

Psychiatric assessment consisted of a comprehensive self-report questionnaire with several psychosocial measures to determine current psychiatric risk and current psychopathology. The self-report questionnaire included the PHQ-9, the PHQ-15, and the FSS. Psychiatric assessment used the MINI to assess current MDE and the HAM-7 to determine depression severity.

Statistical Analysis

All statistical analyses were performed using PASW (IBM SPSS, Chicago, IL) 18.0. Descriptive statistics with means and standard deviations were reported for continuous variables. Categorical variables were compared using the $\chi^2$ test. We used $t$-tests to compare means of continuous variables. Person’s correlation coefficient was used to assess for correlation between scores on the different depression measures. $P < 0.05$ was set as the level of statistical significance.

To determine the cut-off scores on depression measures for identifying depression in patients with HCV-infection, we used a receiver operating analysis. A receiver operating curve (ROC) demonstrates the sensitivity vs. one minus the specificity for possible cut-off point. We determined the optimal cut-off scores by examining the score that combined the maximum sensitivity and specificity. An area under the curve (AUC) was used to determine the ability of depression scales to differentiate depressed and nondepressed patients as per the MINI.
RESULTS

Characteristics of Study Sample

The characteristics of 116 patients are summarized in Table 1. A total of 67 (58%) patients had a current MDE diagnoses. Approximately 66% of patients were male with a mean age of 49.8 years. A majority of the sample had HCV genotype 1 (82%), and nearly a third (31%) suffered from advanced fibrosis. Only two patients were HIV-HCV co-infected. Rates of lifetime alcohol and drug use were 54% and 60%, respectively, in this sample. Patients had lower rates of current drug and alcohol use; specifically, 19% were diagnosed with a current alcohol use disorder and 13% were diagnosed with a current drug use disorder. Moreover, 50% of the sample had a history of a major depressive disorder and 3% had a pre-existing diagnosis of bipolar disorder.

With respect to current depressive symptoms, 55% had moderate depressive symptoms as per the PHQ-9. Mean fatigue scores on the FSS were 41.7 ± 15.7, indicating a high degree of fatigue in the sample. Approximately 17% of patients had an abnormal result for daytime sleepiness as per the ESS.

Depressed vs. Nondepressed Patients

Table 1 provides a comparison of demographic data between depressed and nondepressed patients as defined by the MINI. There was no significant difference between the severity of liver fibrosis, gender, proportion of patients on IFNα therapy, patients with genotype 1 HCV, and age. Furthermore, depressed and nondepressed patients did not significantly differ on current or lifetime alcohol or drug use and past mood disorders. Currently depressed patients had significantly higher scores on the PHQ-15 (P < 0.001), FSS (P < 0.001), ESS (P < 0.05), PHQ-9 (P < 0.001), and HAM-7 (P < .001) scores (see Table 2).

### TABLE 1. Patient Characteristics (Total Sample = 116)

<table>
<thead>
<tr>
<th></th>
<th>All (116)</th>
<th>Depressed (67)</th>
<th>Nondepressed (49)</th>
<th>Comparison of Depressed and Nondepressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male), n (%)</td>
<td>76 (66%)</td>
<td>43 (37%)</td>
<td>33 (28%)</td>
<td>NS</td>
</tr>
<tr>
<td>Age, years (mean ± SD)</td>
<td>49.8 ± 8.8</td>
<td>50.9 ± 8.09</td>
<td>48.2 ± 9.62</td>
<td>NS</td>
</tr>
<tr>
<td>Genotype 1, n (%)</td>
<td>95 (82%)</td>
<td>55 (47%)</td>
<td>40 (34%)</td>
<td>NS</td>
</tr>
<tr>
<td>On IFNα Therapy, n (%)</td>
<td>30 (26%)</td>
<td>21 (18%)</td>
<td>9 (8%)</td>
<td>NS</td>
</tr>
<tr>
<td>HIV Positive, n (%)</td>
<td>2 (2%)</td>
<td>2 (2%)</td>
<td>0 (0%)</td>
<td>NS</td>
</tr>
<tr>
<td>Past history of alcohol use disorder</td>
<td>63 (54%)</td>
<td>26 (18%)</td>
<td>37 (13%)</td>
<td>NS</td>
</tr>
<tr>
<td>Current alcohol use disorder</td>
<td>22 (19%)</td>
<td>10 (9%)</td>
<td>12 (10%)</td>
<td>NS</td>
</tr>
<tr>
<td>Past history of drug use disorder</td>
<td>70 (60%)</td>
<td>28 (24%)</td>
<td>42 (36%)</td>
<td>NS</td>
</tr>
<tr>
<td>Current drug use disorder</td>
<td>15 (13%)</td>
<td>6 (5%)</td>
<td>9 (8%)</td>
<td>NS</td>
</tr>
<tr>
<td>Past history of mood disorder major depressive disorder</td>
<td>55 (47%)</td>
<td>18 (16%)</td>
<td>37 (32%)</td>
<td>NS</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>3 (3%)</td>
<td>0 (0%)</td>
<td>3 (3%)</td>
<td>NS</td>
</tr>
<tr>
<td>Liver fibrosis ≥ grade 3, n (%)</td>
<td>36 (31%)</td>
<td>21 (18%)</td>
<td>15 (13%)</td>
<td>NS</td>
</tr>
</tbody>
</table>

All tests used Chi Square statistics.
NS = not statistically significant.
Depression determined by the MINI.
* Drug use disorders included any substance except nicotine, caffeine, cannabis, and alcohol.

### TABLE 2. Comparison of Fatigue, Physical Symptom, and Depression Measures

<table>
<thead>
<tr>
<th></th>
<th>All (116)</th>
<th>Depressed (67)</th>
<th>Nondepressed (49)</th>
<th>Comparison of Depressed and Nondepressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSS (mean ± SD)</td>
<td>41.7 ± 15.68</td>
<td>46.2 ± 13.13</td>
<td>35.3 ± 16.89</td>
<td>P&lt;0.001*</td>
</tr>
<tr>
<td>ESS (mean ± SD)</td>
<td>7.4 ± 5.46</td>
<td>8.6 ± 5.12</td>
<td>5.7 ± 5.55</td>
<td>P&lt;0.05*</td>
</tr>
<tr>
<td>PHQ-9 (mean ± SD)</td>
<td>11.0 ± 6.47</td>
<td>13.4 ± 6.24</td>
<td>7.51 ± 5.11</td>
<td>P&lt;0.001*</td>
</tr>
<tr>
<td>PHQ-15 (mean ± SD)</td>
<td>10.8 ± 5.88</td>
<td>12.8 ± 5.64</td>
<td>8.04 ± 5.08</td>
<td>P&lt;0.001*</td>
</tr>
<tr>
<td>HAM-7 (mean ± SD)</td>
<td>4.7 ± 3.43</td>
<td>6.4 ± 3.20</td>
<td>2.4 ± 2.12</td>
<td>P&lt;0.001*</td>
</tr>
</tbody>
</table>

NS = not statistically significant.
Depression determined by the MINI.
* All tests used t-tests.
Both PHQ-9 and HAM-7 depressive symptoms were significantly correlated with PHQ-15, FSS, and ESS scores (see Table 3). PHQ-9 and HAM-7 scores were moderately correlated with each other ($r = 0.53$, $P < 0.001$).

### Depression Measures

The ROC curves for the HAM-17, HAM-7, and PHQ-9 are shown in Figure 2. The AUCs were higher for the HAM-7 (0.85, 95% CI 0.77–0.92) compared with the PHQ-9 (0.76, 95% CI 0.68–0.85). The optimal dichotomy cut-off point for the HAM-7 was ≥5 on the HAM-7 (sensitivity 74%, specificity 87%) and the PHQ-9 was ≥10 (sensitivity 73%, specificity 70%).

### DISCUSSION

Patients with CHC are at an increased risk of psychiatric co-morbidity and, using a gold standard evaluation, we found that 58% of our patients suffered from a current major depressive episode. Despite recent studies reporting major depression rates as low as 7% in untreated HCV-infected patients, the high rates of depression in our sample can be understood by the clinic’s depression screening protocol and threshold for referral. Patients who underwent assessment for depression had multiple risk factors for current major depressive. Our results highlight the utility of both the PHQ-9 and the HAM-7 to detect depression in this high-risk patient group.

In addition, a current MDE were significantly, although moderately, associated with physical symptoms, fatigue and daytime sleepiness, which was confirmed by significant correlations between these measures and both PHQ-9 and HAM-7 scores. This is congruent with previously established associations between depression and fatigue, and reinforces the need for considering depression in patients with prominent fatigue and physical complaints in the context of HCV-infection. Treatment of major depression could improve symptoms related to fatigue and physical symptoms. Further, a diagnosis of a current MDE corresponded to significantly higher scores on the PHQ-9 and HAM-7.

Although the administration of both the HAM-7 and MINI by the same clinician precludes our ability to discern the superiority of the HAM-7 in our patient sample, both the HAM-7 and PHQ-9 are brief measures and could theoretically be used in hepatology clinics to assess patients with CHC for depression. The use of one measure over the other should be based on the need for patient administered tools, due to resource limitations or the possibility of integrating a clinician rated scale into routine assessments. The HAM-7 has been studied by primary care physicians and was determined to be effective in assessing depressive symptom severity and response and has been used in this manner. Despite its brevity, the HAM-7 requires training of hepatology nurses and clinicians prior to regular use. In contrast, the PHQ-9 has been shown to be equal or superior to other depression measures in primary care settings and does not require clinician training to administer as it is a self-report measure. It is a brief self-report tool, easy to score, and administered in a variety of medical settings. As a result, both measures have their strengths and should be considered as a screening tool for depression in CHC.

### TABLE 3. Correlation Data for Depressive and Somatic Symptoms

<table>
<thead>
<tr>
<th></th>
<th>PHQ-9</th>
<th>HAM-7</th>
<th>PHQ-15</th>
<th>FSS</th>
<th>ESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ-9</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAM-7</td>
<td>0.53**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-15</td>
<td>0.65**</td>
<td>0.47**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSS</td>
<td>0.41**</td>
<td>0.42**</td>
<td>0.49**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>ESS</td>
<td>0.45**</td>
<td>0.37**</td>
<td>0.21*</td>
<td>0.34**</td>
<td>–</td>
</tr>
</tbody>
</table>

Correlation reported as r-values.
* $P < 0.05$.
** $P < 0.01$.  

![FIGURE 2. Receive Operating Curve for PHQ-9 & HAMD7](image-url)
patient populations based upon the clinic setting, resources, and availability of training.

Our findings should nevertheless be interpreted in the context of our study limitations. First, both the HAM-7 and MINI were completed by the same clinician, whereas the PHQ-9 was completed by the patient. Thus, the clinician administering the HAM-7 and MINI could be biased in the interpretation, and scoring of data and could falsely increase the accuracy of the HAM-7. This precludes us from drawing any significant conclusions on the accuracy of the HAM-7 vs. the PHQ-9 in detecting depression in CHC patients.

Second, we analyzed both untreated HCV-infected patients and those patients receiving IFNα therapy, which could increase heterogeneity of our sample and make it difficult to interpret study results. However, 30 of 116 patients were receiving IFNα therapy when assessed for depression, and there was no significant difference in the number of patients receiving IFNα therapy between depressed and nondepressed patients. Third, our study was limited to patients who were “high-risk” for psychiatric illness, which limits generalizability to all patients with CHC. However, our inclusion criteria were broad and would likely capture most patients with HCV referred to hepatologists and gastroenterologists. Lastly, our rates of current alcohol (19%) and drug (13%) use disorders may be additional confounders to the accurate detection of depression in our sample.

In summary, unrecognized IFN-MDE can lead to serious psychiatric complications and can compromise completion and, thus, efficacy of IFNα based therapy. Given the HCV treatment implications of depression in untreated CHC patients, routine screening for depressive symptoms is warranted in order to mitigate the risk of IFNα-MDE. Our results indicate that both self-report and clinician-rated scales are suitable to identify clinical depressive symptoms in patients with CHC, and a cut-off of 5 and 10 on the HAM-7 and PHQ-9, respectively, can be used to detect depression in HCV-infected patients. Both tools are easy to implement in hepatology clinics, and their accuracy may assist in identifying patients who are at greatest risk of neuropsychiatric complications on IFNα therapy. Our study also underscores the need to consider a diagnosis of depression in patients with significant fatigue and physical symptoms. Future studies will be needed to determine the effect of depression treatment on overall physical symptom burden and fatigue and optimal psychiatric interventions. Moreover, additional studies are needed that compare the HAM-7 and PHQ-9 to an independent criterion standard administered by an interviewer blinded to the results of both measures in order to determine whether measures have different or similar operating characteristics.

Disclosure: Dr. Sanjeev Sockalingam is a speaker for Hoffman-Roche Canada. Dr. Abbey is on the advisory boards for Pfizer Canada and Eli Lilly Canada.

References

14. Clark CH, Mahoney JS, Clark DJ, et al: Screening for depression in hepatitis C population: the reliability and validity of the
Detecting Depression in Hepatitis C


