Original Article

Criterion Validation of the Edinburgh Postnatal Depression Scale as a Screening Tool for Depression in Patients with Advanced Metastatic Cancer

Mari Lloyd-Williams, MD, MRCGP, Trevor Friedman, MRC Psych, and Nicky Rudd, MRCP
Leicestershire Hospice (M.L.-W., N.R.), and Department of Liaison Psychiatry (T.F.), Leicester General Hospital, Leicester, United Kingdom

Abstract
It is estimated that 25% of palliative care patients will have identifiable symptoms of depression. Near the end of life, the distinction between what can be called “appropriate sadness” and depression may be difficult. Many palliative care units use rating scales to help identify patients who may be depressed. It is believed that symptoms such as guilt, worthlessness, and hopelessness may be more discriminating than other symptoms for depression within this population. The Edinburgh postnatal depression scale (EPDS) was devised for use in women in the postnatal period and does not contain any somatic-type symptoms. It consists of 10 items, each rated on a four-point scale, and includes items on guilt, thoughts of self-harm, and hopelessness. It has not previously been used for screening in cancer patients. In a study of 100 inpatients receiving palliative care, a cutoff of 13 on the EPDS had a sensitivity of 81% and a specificity of 79% for detecting cases of depression. There was a low rate of misclassifications. This study suggests that the EPDS may be useful as a screening instrument for palliative care patients. J Pain Symptom Manage 2000;20:259–265. © U.S. Cancer Pain Relief Committee, 2000.

Key Words
Depression, screening, terminal illness, palliative care, EPDS

Introduction
Depression is a common psychiatric illness in patients with advanced metastatic cancer. It has been estimated that 25% of patients admitted to a hospice have identifiable symptoms of depression.¹ The presence of depression not only reduces quality of life for patients but can also impact on the effective palliation of physical symptoms. Additionally, psychological and psychiatric morbidity can itself be a major source of distress to terminally ill patients and to their relative and friends.

The treatment of depression in terminally ill patients should be the same as treatment in any other patient with depression. It should consist of good psychological support, together with antidepressant medication at an appropriate dosage for an adequate length of time. There are many antidepressants available with acceptable side effect profiles, and pa-
patients identified as depressed even within the last 4 to 6 weeks of life may still benefit from treatment.

Many depressed patients near the end of life are not treated due to the difficulties in assessing psychiatric symptoms in patients with terminal illness.\textsuperscript{2,3} Patients may be reluctant to disclose their feelings of depression with medical and nursing staff for fear of being perceived as a “bad” or “weak” patient. There may also be a lack of awareness of the symptoms of depression in palliative care. One of the main difficulties in establishing a diagnosis of an illness where there are no biological markers, physical signs, or effective diagnostic test is deciding what can be called “appropriate sadness” as patients approach the end of life with what is a depressive illness.

It has been suggested that specific criteria should be used to diagnose depression in patients with advanced cancer.\textsuperscript{4,5} These criteria should replace somatic symptoms with nonsomatic alternatives. Recently, Chochiov et al. found that this approach only affected prevalence rates if a low threshold approach was used by the investigators.\textsuperscript{6}

The use of screening tools to assess for psychiatric morbidity has been advocated.\textsuperscript{7–13} These tools range from structured questionnaires to single questions. Snaith emphasized that “a good rating scale should consist of not too many items and not too few and they should have been selected by both the intuition of clinical experience and by the statistical process of item analysis, but unfortunately these desiderata apply to very few scales.”\textsuperscript{14} Silverstone\textsuperscript{15} suggested that none of the instruments currently available could be recommended as a “gold standard” for measuring or screening for depression in the medically ill. Psychiatry differs from any other specialty in that diagnosis can only be made on the basis of a very comprehensive and thorough history and, with very few exceptions, diagnostic tests have no role to play. Psychiatric assessments are difficult and time-consuming to undertake and are seldom carried out by palliative care doctors.\textsuperscript{16}

Several instruments have been developed with a particular score or cutoff threshold assigned to predict a “case” of depression. The majority of rating scales contain a number of symptoms or feelings on which the patient indicates their own response, and the scores are calculated by the person administering the scale. The symptoms of depression are not unique and can occur in other psychiatric illness. Depression scales are therefore nonspecific and cannot and should not be used for diagnostic purposes.\textsuperscript{17} After screening, the patient who scores at a predefined cutoff threshold or above could either receive treatment for depression empirically or be referred for further assessment.

Depression comprises different constructs, and scales have been devised that measure different elements (e.g., cognitions, behavior, somatic symptoms, etc.). Different scales measure different facets of a depressive illness.\textsuperscript{18} Many instruments have been developed and validated on physically healthy patients and those that have been studied in cancer populations have been focused on patients with early disease or patients undergoing active treatment. In the terminally ill, symptoms such as guilt, suicidal ideation, and worthlessness may be more useful discriminating symptoms for depression than vegetative symptoms (e.g., poor appetite, weight loss, and sleep disturbance), which are common manifestations of advanced cancer.

The Edinburgh postnatal depression scale\textsuperscript{19} (EPDS) was developed and validated to assess for depression in women in the postnatal period. It excludes the somatic symptoms of depression and was devised to that it could be administered by health care workers with no prior knowledge of psychiatry. It contains questions on worthlessness, subjective sadness, and suicidal ideation (see Appendix 1). The scale contains 10 items selected from the Hospital Anxiety and Depression Scale (HAD), The Irritability, Depression and Anxiety Scale, and the Anxiety and Depression Scale. It is scored in a similar format to the HAD scale (i.e., each question has four possible responses scored from 0–3; the most negative response is given the highest score). The original scale was validated on 84 mothers in the postnatal period and found to have a sensitivity of 86% and specificity of 78% using a cutoff threshold of 12/13.

The EPDS has been further evaluated using a large community sample in the U.K.\textsuperscript{20} and in a North American population\textsuperscript{21} and found to have acceptable sensitivity, specificity, and positive predictive rates for depression in the postnatal population. It has also been validated in
females outside the postnatal period and found to have acceptable sensitivity and specificity.\textsuperscript{22} The authors of the scale believed this scale could be used for other populations, but that the scale required validation prior to use. To our knowledge the EPDS has not been used as a screening tool for patients with physical illness.

The use of validated screening tools could improve the detection and treatment of depression in the terminally ill. The aim of this study was to validate the EPDS, which does not contain somatic symptoms but includes symptoms thought to be important in detecting depression in patients with advanced metastatic cancer.

\textbf{Method}

Patients 18–70 years old who had a prognosis of 6 months or less, as determined by the consultant in charge of their care, were considered eligible to take part in the study. All were in patients in a Hospice or oncology unit for at least 48 hours prior to participating in the study. Patients were required to be able to understand both written and spoken English and be capable of completing self-assessment scales unaided and respond to a semistructured interview. The exclusion criteria included patients who were currently prescribed antidepressant medication, had suspected or confirmed cerebral metastases, or a prognosis of 1 week or less. The aims of the study were to establish the use of the EPDS as a screening tool, and it was felt that patients already diagnosed as being depressed and prescribed antidepressant medication would not normally be included within a screening population. Patients with cerebral metastases and those with a very short prognosis of 1 week or less. The aims of the study were to establish the use of the EPDS as a screening tool, and it was felt that patients already diagnosed as being depressed and prescribed antidepressant medication would not normally be included within a screening population.

Patients were asked to complete the EPDS and were interviewed using the Present State Examination (PSE) by the first author. The PSE is a semistructured psychiatric interview based on establishing if a range of symptoms are present and, if so, the degree of severity.\textsuperscript{23} It is a clinical instrument based on a method of cross-examination used by psychiatrists to decide if specific symptoms are present. The PSE takes into account somatic and nonsomatic features of depression by the nature of questioning and the score allocated. It has been widely used for research purposes within psychiatry, and it has been found to be a valid instrument when interviewers had been adequately trained and supervised.\textsuperscript{24} Both the first and second authors of this paper are experienced in PSE interviewing and are tutors on the intensive PSE course. The interviewer was blind to the EPDS score when carrying out the interview.

According to the protocol, the medical staff responsible for the patient's care was informed of any patients who were found to have a psychiatric diagnosis according to the PSE interview.

To ensure a scale can be useful within a population, it is important that a cut-off threshold is calculated. A number of cutoff thresholds were analyzed for sensitivity: the number of patients scoring at this threshold or above who are actually true cases; specificity, the number scoring below a threshold who were true non-cases; the positive predictive value, the probability that a score at the threshold or higher would be a true case; and the negative predictive value, the probability that a score below the threshold would not be a case. The internal reliability of the scale was calculated by using Cronbach alpha coefficient\textsuperscript{25} and factor analysis was also carried out. All patient's demographic details were coded and the date of death recorded.

\textbf{Results}

One hundred patients who fulfilled the inclusion criteria participated in the study. A further 30 patients were either unable to complete the questionnaire or the PSE due to physical illness. All patients had metastatic carcinoma at the time of interview and were receiving palliative care only, with emphasis on pain relief and symptom control. Of the participating patients, 56 were females and 44 were males, and the mean age of patients was 57.3 years (range 25–69 years).

Cancer of the lung accounted for 26% of all diagnoses, breast cancer 24%, and cancer of the colon 11%. Of the 100 patients, the majority (60%) was married, 21% were separated or divorced, 8% were widowed, 8% had never married, and 3% were cohabiting.
The prevalence of depression identified by the PSE interview according to International Classification of Diseases (10th edition) criteria was 22%. A further 5% of patients were identified as having other psychiatric morbidity, which included anxiety disorders.

The mean age of patients identified as cases of depression was 56.6 years (95% C.I., 52.5–60.7). All patients who were interviewed survived a mean of 32.3 days (95% C.I., 25.8–38.8), with a range of 2 days to 180 days and a median survival of 23 days. The mean survival time for patients who were diagnosed as being depressed was 38 days (95% C.I., 26–50 days) and the median survival time was 26 days. This difference was not significant ($P = 0.3401$).

**The Validity of the EPDS**

The mean score for the complete 10-item Edinburgh scale was 10.9 (95% C.I., 10.01–11.96; range 2–25). At a cutoff of 13, 27% scored above this threshold. Scoring high on the EPDS did not influence survival: patients scoring above 13 on the 10-item scale survived a mean of 36 days compared to a mean survival of 32 days for all patients in the study.

The sensitivity, specificity, positive predictive value, and negative predictive value at each threshold can be seen in Tables 1 and 2. Although developed for identifying depression, the Edinburgh scale had acceptable sensitivity and specificity in identifying all cases of psychiatric morbidity. A cutoff threshold of 13 gave a sensitivity of 78% and specificity of 82% for identifying all cases of psychiatric morbidity. Across all thresholds, the EPDS performed better in identifying cases of depression. The optimum threshold of the EPDS for identifying cases of depression was 13, which had a sensitivity of 81% and specificity of 79%. When a lower threshold of nine or ten is considered for identifying PSE cases of depression, the sensitivity is 100% but the specificity is lower.

At high cutoff thresholds, the Edinburgh scale has high positive predictive values and negative predictive values.

**Factor Analysis of the EPDS**

The factor analysis was performed using an oblique rotation and revealed three factors that accounted for 52% of the variance of the scale. Four items (items three, five, six, and seven) load onto factor one, and three items (items one, two, and four) load onto factor two. The three items of the EPDS that are derived from the Hospital Anxiety and Depression scale load negatively onto a separate factor, suggesting that these items may be measuring a different construct of depression. (See Table 3.)

The internal reliability of the EPDS was calculated using Cronbach’s alpha coefficient. The scale has an alpha coefficient of 0.7805.

**Discussion**

The aim of this study was to validate the use of a rating scale to identify depression in a population of terminally ill patients. Using the International Classification of Diseases criteria from the PSE interview, this study of 100 terminally ill patients found a prevalence of 22% for depression. This is consistent with published rates. The EPDS was able to correctly identify cases of depression, and a cutoff threshold of 13 gave the optimal sensitivity and specificity. The majority of cases were depression of moderate severity. Depressed patients had a mean age of 56.6 years, compared to the mean age of all patients recruited of 57.2 years.

The sample size of 100 patients is larger than many samples of terminally ill patients used for research studies and allowed statistical analysis to be carried out. Recruiting patients who are terminally ill is difficult, as many patients are too frail to undergo any form of extensive interviewing.

<table>
<thead>
<tr>
<th>Threshold scores</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>95</td>
<td>86</td>
<td>81</td>
<td>68</td>
<td>50</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>31</td>
<td>37</td>
<td>52</td>
<td>61</td>
<td>66</td>
<td>79</td>
<td>85</td>
<td>89</td>
</tr>
<tr>
<td>PPV (%)</td>
<td>29</td>
<td>31</td>
<td>36</td>
<td>42</td>
<td>42</td>
<td>53</td>
<td>55</td>
<td>58</td>
</tr>
<tr>
<td>NPV (%)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>97</td>
<td>94</td>
<td>94</td>
<td>91</td>
<td>86</td>
</tr>
</tbody>
</table>

**Table 1**

Threshold Score of EPDS for Identifying Cases of Depression

PPV = Positive Predictive Value, NPV = Negative Predictive Value.
A prognosis of 6 months or less was required for participation in the study, but the majority of patients survived for a much shorter period (a median survival for all patients of 23 days).

Depression fluctuates during a terminal illness, and this study does not address the issue of when is the most appropriate time to screen for depression. Although the median survival was low, eight patients who were detected as being depressed as part of this study were commenced on medication by their oncologist and derived therapeutic benefit.

In devising the EPDS, Cox et al. suggested that for the 10-item scale, a score of 12 or 13 was most likely to predict mothers in the postnatal period who may be suffering from a major depressive illness requiring further assessment. In terminally ill patients, a cutoff threshold of 13 gave the optimal sensitivity of 81% and specificity of 79%. The positive predictive value at this threshold was 53%. At lower cutoff thresholds, higher sensitivities were obtained but at the expense of lower specificity. If the scale is to be used as a crude screening instrument with all possible cases to be identified, a cutoff threshold of 10 could be advocated. All patients “screened in” at this level would require further evaluation by a clinician to determine whether depression was present at such a level to require treatment. Although the EPDS appeared to be equally sensitive in detecting other psychiatric morbidity, it has not been developed for this purpose. Further research, which we are undertaking, may reveal a threshold at which “psychological distress” could be identified.

The sensitivity and specificity of the EPDS in this study is higher than previous studies using other scales, namely the Hospital Anxiety and Depression Scale, for the assessment of depression in patients with advanced metastatic cancer. This suggests that the EPDS has acceptable validity and could be used as a screening tool in terminally ill patients.

The factor analysis revealed three factors. Items relating to subjective sadness, hopelessness, and thoughts of self-harm loaded positively onto factor one, and items relating to blame and anxiety loaded positively onto factor two. There are no other published studies of factor analysis of the EPDS with which to compare these findings. The internal consistency of the Edinburgh scale as measured by Cronbach’s alpha for this group of patients was 0.78. Although it is recommended that the alpha coefficient should be above 0.8 for a screening instrument, the internal reliability of a scale reflects the number of items. The lower internal reliability of the EPDS in this study did not influence its efficacy. For a scale to be acceptable

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>-0.00</td>
<td>-0.04</td>
<td>-0.88</td>
</tr>
<tr>
<td>Two</td>
<td>0.10</td>
<td>0.31</td>
<td>-0.54</td>
</tr>
<tr>
<td>Three</td>
<td>0.10</td>
<td>0.68</td>
<td>0.20</td>
</tr>
<tr>
<td>Four</td>
<td>-0.12</td>
<td>0.71</td>
<td>-0.15</td>
</tr>
<tr>
<td>Five</td>
<td>0.07</td>
<td>-0.02</td>
<td>-0.77</td>
</tr>
<tr>
<td>Six</td>
<td>0.85</td>
<td>-0.64</td>
<td>0.02</td>
</tr>
<tr>
<td>Seven</td>
<td>0.04</td>
<td>0.65</td>
<td>-0.09</td>
</tr>
<tr>
<td>Eight</td>
<td>0.66</td>
<td>0.24</td>
<td>0.01</td>
</tr>
<tr>
<td>Nine</td>
<td>0.31</td>
<td>0.01</td>
<td>-0.31</td>
</tr>
<tr>
<td>Ten</td>
<td>0.70</td>
<td>-0.06</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

PPV = Positive Predictive Value, NPV = Negative Predictive Value.
as a screening instrument in patients who are weak and unwell, a small number of items is preferable, and this together with other features of the EPDS support its use. The test-retest reliability of the EPDS remains to be established.

The scores of the EPDS were evenly distributed, with no questions eliciting a high number of positive responses. It is interesting that an instrument designed for the postnatal population should have good efficacy with the terminally ill population. The use of questions relating to subjective sadness, guilt, hopelessness, and thoughts of self-harm appear to be independent of physical disability, and a terminally ill patient’s positive response is likely to reflect low mood.

What is the next step for the use of rating scales in the terminally ill patient? In a survey of rating instruments, Meakin suggested that their increased use should be accompanied by a validation of treatment to establish which patients should be screened to benefit from treatment. This concept is equally applicable in terminal illness as in other physical illnesses. The use of rating scales therefore is still subject to debate, but together with effective clinical evaluation and adequate treatment, rating scales have a role in the assessment of depression in the terminally ill.

It is acknowledged that the distress of terminally ill patients encompasses far more than the presence or absence of a psychiatric illness. Indeed, some aspects of psychological distress in terminally ill patients (e.g., spiritual distress), are distinct, rarely included in assessment, and may indeed be impossible to measure. However, depression is a treatable symptom that can greatly influence the quality of life of a terminally ill patient and measures should be taken to improve its detection and treatment. The Edinburgh scale, which includes items on guilt, subjective sadness, and thoughts of self-harm, appears able to identify terminally ill patients who may be depressed, and it could be a useful instrument for screening for depression in this population.

This initial study should be followed by additional studies of the construct validity and the test-retest reliability of the scale. We plan further research on its efficacy in the terminally ill population using a larger sample of patients.

References
Appendix 1: The Edinburgh Postnatal Depression Scale

In the past 7 days:

1. I have been able to laugh and see the funny side of things
   as much as I always could.
   not quite so much now.
   definitely not so much now.
   not at all.

2. I have looked forward with enjoyment to things
   as much as I ever did.
   rather less than I used to.
   definitely less than I used to.
   hardly at all.

3. I have blamed myself unnecessarily when things went wrong

4. I have been anxious or worried for no good reason
   not at all.
   hardly ever.
   sometimes.
   very often.

5. I have felt scared or panicked for no very good reason
   quite a lot.
   sometimes.
   not much.
   not at all.

6. Things have been getting on top of me
   most of the time, and I haven’t been able to cope at all.
   sometimes, and I haven’t been coping as well as usual.
   rarely, and I have coped quite well.
   almost never, and I have been coping as well as ever.

7. I have been so unhappy, I have had difficult sleeping
   most of the time.
   sometimes.
   very often.
   not at all.

8. I have felt sad or miserable
   most of the time.
   quite often.
   not at all.

9. I have been so unhappy, I have been crying
   most of the time.
   quite often.
   only occasionally.
   never.

10. The thought of harming myself has occurred to me
    quite often.
    sometimes.
    hardly ever.
    never.