Screening for Psychological Distress in Palliative Care: A Systematic Review

Parvez Thekkumpurath, MRCPsych, Chitra Venkateswaran, MD, Manoj Kumar, MRCPsych, MD, and Michael I. Bennett, FRCP, MD
Cancer Research UK Clinical Centre, St. James’s University Hospital, University of Leeds (P.T., C.V); Leeds Mental Health Trust (M.K.); St. Gemma’s Hospice (M.B), Leeds, United Kingdom

Abstract
Psychological distress is common in the terminally ill. It is often undetected and undertreated and has significant impact on the individual and family. There is a growing consensus on a broader concept of psychological suffering conceptualized as “distress” in the palliative care setting. Psychological screening programs play an important role in improving detection and management of distress. National and international guidelines recommend routine screening. This systematic review summarizes the evidence for screening for psychological distress in a palliative care setting. The review includes studies that compare screening questionnaires against a gold standard criterion of semistructured or structured psychiatric interview. Eight studies were identified which examined the performance of 10 screening questionnaires, ranging from single items to multidimensional questionnaires with up to 33 items. The performances of these questionnaires are described in terms of their sensitivity, specificity, and positive and negative predictive values. There are very few studies that examine the validity of questionnaires against credible criteria such as psychiatric interview and most studies have so far focused on depression. Unidimensional scales appear to perform equally well compared to the longer versions. This review summarizes the evidence, the quality of this evidence, and future challenges to improve identification and management of distress in palliative care.

Key Words
Distress, screening, palliative care, terminally ill, depression, anxiety, adjustment disorder, review, validation, questionnaires

Introduction
Psychological Distress as a Concept
Medicine has always dealt with traditional disease categories, which are mutually exclusive (e.g., cancer, bronchitis), and psychiatry has mirrored this approach. Psychological distress can be conceptualized as a categorical construct following this tradition or as a unitary dimensional construct along a continuum of mild to severe.
The term “distress” was first introduced in this context as a working concept based on expert consensus to refer to emotional or psychological problems in people with cancer. The National Comprehensive Cancer Network defines distress as:

“...a multifactorial unpleasant emotional experience of a psychological (cognitive, behavioral, emotional), social, and/or spiritual nature... Distress extends along a continuum ... from common normal feelings of vulnerability, sadness ... to problems that can become disabling, such as depression, anxiety, panic, social isolation, and existential and spiritual crisis.”

The term encompasses a wide range of emotions from sadness to more complex psychological syndromes such as depression and anxiety disorders. This unitary concept of distress is very appealing in its simplicity, pragmatism, and appeal to all.

Conventional categories of psychiatric diagnoses have been used to describe, understand, identify, and manage distress. Although this attracts criticisms of stigma and questionable validity in a medically unwell population, the translation to clinical usefulness such as prognostication and management is undoubted. However, while using such a framework, the focus so far has been on identification and management of syndromes such as depression. There is increasing evidence of much broader and mixed forms of psychiatric morbidity, such as adjustment disorder, in the terminally ill.

Prevalence of Distress

Prevalence of distress in the terminally ill varies depending on the way it is measured and the diagnostic criteria applied. Self-report measures, such as the Hospital Anxiety and Depression Scale (HADS), typically tend to show higher estimates as they try to be over inclusive and sensitive rather than specific.

Estimates of distress also vary depending on the patient population, for example, between day hospice and inpatient settings. The most consistent rates are derived from structured or semistructured interviews by using operationalized diagnostic criteria, such as ICD-10 or DSM-IV. Most of these studies report psychological morbidity between 20% and 30%. Higher rates are seen in people with advanced and metastatic disease and in patients with a poor prognosis. At least 10%–15% of these people warrant specialist psychological or psychiatric help.

In a palliative care population, estimates of specific diagnostic categories, such as depression, range from 6.7% to 30%. In contrast, anxiety disorders are conspicuous by their absence or very low rates. Only one study reported on adjustment disorders and found a prevalence of 15.8%.

There is a high degree of correlation between the presence of anxiety and depressive disorders in this population.

Importance of Detecting Distress

Palliative care aims to improve the quality of life of patients through prevention and relief of all forms of suffering, including psychological, by means of early identification, assessment, and treatment. The majority of patients receiving palliative care services have a diagnosis of cancer and up to one-third of these patients suffer psychological distress. The prevalence increases to more than 50% in advanced cancer and people with poor prognosis. This is not surprising in a group of patients who are coming to terms with failing health and death.

Psychological problems are not routinely picked up in medical settings, where more than 50% of these problems may go undetected, and the scenario is not different in cancer patients. Normalizing of distress by patients and a lack of awareness and skill by clinicians in identifying and differentiating distress from appropriate sadness, along with limited management options, contribute to this underdetection.

Assessment and diagnosis of psychological problems in palliative care is challenging. Physical symptoms such as fatigue, changes in sleep and appetite, concentration, and energy are common in palliative care and are also used to diagnose traditional categories, such as depression. Substitution by nonsomatic symptoms and exclusion of these items have been tried, but there is growing consensus that physical symptoms are useful in supporting the diagnosis, and including them affects the diagnostic process only in “low-threshold approaches.”
The cost of untreated distress is significant in terms of the impact on a patient’s quality of life and alleviation of physical symptoms, such as pain, physical functioning, role and social functioning, and impact on their families. Once identified, psychological distress in the terminally ill is amenable to treatment through good psychological support and medications as in any other setting. Even patients with a life expectancy of only four to six weeks benefit from the use of appropriate intervention. It is in this context that a concerted effort is required to improve identification and management strategies of psychological distress in palliative care.

Screening Tools Used to Detect Distress
Scope of Review
Screening questionnaires have been used in cancer patients since the 1980s, and there are currently a number of such instruments used in a wide variety of cancer settings, including palliative care. Over the years, questionnaires have become more sophisticated, briefer, and easier to use than earlier versions. This has occurred as a result of the application of statistical techniques, such as factor analysis and principal component analysis. Screening questionnaires need to fulfill certain criteria before they can be used in a particular population. These include psychometric soundness, such as validity and reliability, along with feasibility, acceptability, and ease of use.

It is well known that questionnaires identify different subgroups based on the clinical setting and culture. Establishing the validity of a screening questionnaire is a key and primary step prior to its use in a specific population. This is done by comparing the performance of the questionnaires against “gold standard” criteria for measuring the construct in question. Some gold standards are too specific in that they measure only certain dimensions of a construct, such as depression. Often, one questionnaire is validated against another questionnaire and it is hard to draw any meaningful clinical conclusion when it is not clear what is being measured. The most appropriate gold standard is a psychiatric interview that generates validated psychiatric diagnostic categories. This approach is not without flaws, but is the best available method in the absence of any valid markers, physical signs, or tests for psychological distress.

This review reports on validation studies reporting the use of screening questionnaires to detect psychological distress in palliative care populations. We included studies that examined the validity of screening questionnaires by comparing them against DSM-IV or ICD-10 diagnosis generated by structured or semistructured psychiatric interviews. A previous review in 2003 reported on depression screening tools in this population and identified three studies.

Search Methodology
The review set out to identify studies examining criterion validity of screening questionnaires to detect psychological distress by comparing against a formal psychiatric interview as gold standard in a terminally ill population.

We searched the following databases: Medline 1966—, Psychinfo1967—, and Embase 1980—.

Search terms were defined under three main headings. MeSH terms were augmented with keyword searches where appropriate:

- Palliative care (the population)—palliative care, palliative$, terminally ill, terminal$, cancer, and neoplasms.
- Psychological distress (the problem)—depression, anxiety, adjustment disorders, psychological distress, and distress.
- Assessment—questionnaires, sensitivity and specificity, diagnosis, mass screening, and receiver operating characteristics curve.

Studies were included if they met the following criteria:

1. Sample was from a palliative care or terminally ill population.
2. Validation of questionnaires was the primary aim.
3. Questionnaires were compared against a formal psychiatric interview.

Studies that had a heterogeneous cancer population at various stages of the illness trajectory were excluded. Studies were included only if the screening questionnaires were compared against a formal psychiatric interview, which was set as the gold standard in the a priori
question. This is in line with recommendations by the Expert Working Group of the European Association for Palliative Care. Thus, only those studies where the diagnoses were verified by independent clinical criteria were included in the review.

A large number of validation studies were excluded, as they were being compared against other questionnaires. These were not considered credible criteria as set out in the a priori clinical question.

The a priori question, search terms, and methodology and the criterion for critical review of selected process were agreed after independent selection and then discussion among all four researchers.

This broad search yielded 3,300 hits. By a process of visual screening, 178 potential articles were identified and abstracts obtained. This was supplemented by hand searches and following up references from identified articles. Eight articles were identified which met the inclusion criteria.

Identified papers were read and data were extracted on the questionnaire, characteristics of validation population, and questionnaire psychometric properties (sensitivity, specificity, and positive and negative predictive values). We summarized the performance of screening questionnaires examined in these studies (Table 1) and commented on the strengths and weaknesses of the studies.

Results

Unidimensional Scales

Brief measures have particular appeal in palliative care due their ease of administration in a physically frail population.

Single-Item Interviews. The most studied measures are the single items adapted from structured interview schedules: “Are you depressed?” and “Have you lost interest or pleasure?” Chochinov et al. report 100% sensitivity and specificity for the single item “Are you depressed?” and a slightly lower rate when combined with another item about loss of interest. However, when Lloyd Williams et al. examined the validity of a single item “Are you depressed?” in a UK palliative care population, the question performed modestly, with a sensitivity of 54% and specificity of 74% in correctly identifying those with depression. A recent study from Ireland reports a sensitivity of 90% and specificity of 67% when the two items were combined. Individually the items were less sensitive but more specific. Similarly, in a small study of the single-item interview question in 45 patients, only a moderate agreement (kappa 0.49) was found. The study used a different methodology and does not report sensitivity and specificity.

Akechi and colleagues examined the validity of two single-item questions “Are you depressed?” and “Have you lost interest or pleasure?” alone and in combination in a Japanese palliative care population. In detecting major depression and adjustment disorder together, the screening items showed a high specificity (>0.95), but their sensitivity remained low (<0.68) either alone or in combination. In contrast, they performed well in detecting major depression alone (sensitivity 0.79 to 1.00, specificity 0.86–0.92), demonstrating the focus of these items on the core diagnostic criteria for depression.

Visual Analog Scale. Chochinov et al. examined the efficacy of a 100 mm visual analog scale, anchored at “worst possible mood” and “best possible mood” at the ends, adapted from the Memorial Pain Assessment Card. The scale’s performance was poor compared to other questionnaires, with a sensitivity of 72% and specificity of 50%.

Verbal Mood Rating Scale. A verbal mood rating scale of 0–10, indicating severe low mood and high mood at the ends, showed a reasonable sensitivity of 80%, but had poor specificity of 43% in identifying cases of depression as diagnosed by DSM-IV criteria. The authors comment that the experience of rating one’s mood in this way was a novel experience for many patients, affecting its performance.

Multidimensional Scales

These scales have more than one item and, as the name suggests, assess more than one dimension of distress, for example, cognitive and behavioral aspects. They take longer to complete and theoretically are thought to be more specific.

Edinburgh Depression Scale. The 10-item Edinburgh Depression Scale (EDS), adapted from the HADS, was originally developed for use
<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Author, Year, and Country of Study</th>
<th>Population</th>
<th>Interview, Target Disorder</th>
<th>Sample</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edinburgh Depression Scale, EDS, 10 items</td>
<td>Lloyd-Williams, 2000, UK</td>
<td>Inpatients, advanced cancer</td>
<td>PSE, All psychiatric diagnoses</td>
<td>100</td>
<td>0.71</td>
<td>0.82</td>
<td>0.62</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Lloyd-Williams, 2004, UK</td>
<td>Day unit</td>
<td>Clinical interview, Depression</td>
<td>74</td>
<td>0.70</td>
<td>0.80</td>
<td>0.56</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Lloyd-Williams, 2007, UK</td>
<td>Day unit</td>
<td>Clinical interview, Depression</td>
<td>246</td>
<td>0.72</td>
<td>0.74</td>
<td>0.55</td>
<td>0.85</td>
</tr>
<tr>
<td>Brief Edinburgh Depression Scale, BEDS, six items</td>
<td>Lloyd-Williams, 2007, UK</td>
<td>Day unit, advanced cancer</td>
<td>PSE, Depression</td>
<td>246</td>
<td>0.72</td>
<td>0.83</td>
<td>0.65</td>
<td>0.87</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale, HADS, 14 items</td>
<td>Le Fevre, 1999, UK</td>
<td>Inpatients, advanced cancer</td>
<td>CIS-R, All psychiatric diagnoses</td>
<td>79</td>
<td>0.74</td>
<td>0.71</td>
<td>0.52</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Akechi, 2006, Japan</td>
<td>Outpatients, advanced cancer</td>
<td>SCID, Depression and adjustment disorders</td>
<td>209</td>
<td>0.80</td>
<td>0.67</td>
<td>0.41</td>
<td>0.92</td>
</tr>
<tr>
<td>General Health Questionnaire, 12 items</td>
<td>Le Fevre, 1999, UK</td>
<td>Inpatients, advanced cancer</td>
<td>CIS-R, All psychiatric diagnoses</td>
<td>79</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Single item: “Are you depressed?”</td>
<td>Chochinov, 1997, Canada</td>
<td>Inpatients, advanced cancer</td>
<td>SADS, Major and minor depression</td>
<td>200</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Akechi, 2006, Japan</td>
<td>Outpatients, advanced cancer</td>
<td>SCID, Depression and adjustment disorders</td>
<td>209</td>
<td>0.47</td>
<td>0.97</td>
<td>0.81</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Lloyd-Williams, 2004, UK</td>
<td>Day unit</td>
<td>Clinical interview, Depression</td>
<td>74</td>
<td>0.55</td>
<td>0.74</td>
<td>0.44</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Meyer, 2003, UK</td>
<td>Inpatients, advanced cancer</td>
<td>SCID, Major and mild depression</td>
<td>45</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Payne, 2007, Ireland</td>
<td>Inpatients, advanced cancer</td>
<td>Clinical interview, Depression</td>
<td>167</td>
<td>0.70</td>
<td>0.81</td>
<td>0.57</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>Chochinov, 1997, Canada</td>
<td>Inpatients, advanced cancer</td>
<td>SADS, Major and minor depression</td>
<td>200</td>
<td>1.00</td>
<td>0.98</td>
<td>0.86</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Akechi, 2006, Japan</td>
<td>Outpatients, advanced cancer</td>
<td>SCID, Depression and adjustment disorders</td>
<td>209</td>
<td>0.68</td>
<td>0.94</td>
<td>0.76</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Payne, 2007, Ireland</td>
<td>Inpatients, advanced cancer</td>
<td>Clinical interview, Depression</td>
<td>167</td>
<td>0.91</td>
<td>0.68</td>
<td>0.49</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Chochinov, 1997, Canada</td>
<td>Inpatients, advanced cancer</td>
<td>SADS, Major and minor depression</td>
<td>200</td>
<td>0.72</td>
<td>0.50</td>
<td>0.17</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>Meyer, 2003, UK</td>
<td>Inpatients, advanced cancer</td>
<td>SCID, Major and mild depression</td>
<td>45</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mood Evaluation Questionnaire, 33 items</td>
<td>Chochinov, 1997, Canada</td>
<td>Inpatients, advanced cancer</td>
<td>SADS, Major and minor depression</td>
<td>200</td>
<td>0.79</td>
<td>0.71</td>
<td>0.27</td>
<td>0.96</td>
</tr>
<tr>
<td>Beck Depression Inventory -Short Form, 13 items</td>
<td>Lloyd Williams, 2004, UK</td>
<td>Day unit</td>
<td>Clinical interview, Depression</td>
<td>74</td>
<td>0.80</td>
<td>0.43</td>
<td>0.34</td>
<td>0.85</td>
</tr>
</tbody>
</table>

PSE = Present State Examination; CIS-R = Clinical Interview Schedule-Revised; SCID = Structured Clinical Interview for DSM-III-R; SADS = Schedule for Affective Disorders and Schizophrenia; PPV = Positive predictive value; NPV = Negative predictive value.
in mothers in the postnatal period. Although the absence of somatic items suits a palliative care population, the focus on core depressive symptoms, such as anhedonia, attracts similar criticisms to the HADS.

When evaluated by comparing to a semi-structured psychiatric interview, the Present State Examination, a sensitivity of 71% and specificity of 82% was found. It is not clear whether the study looked actively for minor mood disorders, such as adjustment disorders. In a later study, when the same author compared EDS along with two other screening items to a semistructured interview based on DSM-IV criteria for depression, a sensitivity of 70% and specificity of 80% in identifying depression was achieved. These values, although lower than the initial results, are still comparable to performance of screening questionnaires.

In a study primarily concerned with the development of the briefer version of the scale, EDS showed a sensitivity and specificity of 71% and 74%, respectively, when evaluated against the Present State Examination. Brief Edinburgh Depression Scale (BEDS). This brief version of the 10-item EDS was developed by item reduction following factor analysis. Six items were identified that loaded on a single factor and accounted for 36% of the explained variance. In this study, this six-item version, the BEDS, scored similarly to its parent version, with a sensitivity of 72% and specificity of 83% at a cut off score of six. This showed promise as screening tool for depression.

Hospital Anxiety and Depression Scale. The HADS was first developed to assess the presence of depression and anxiety in the hospitalized medically ill population. The 14-item scale, with subscales for depression and anxiety, was developed with focus on core depressive symptoms, such as anhedonia. Although HADS is one of the most widely used valid screening questionnaires for anxiety and depression in the medical setting, there is controversy surrounding its use in palliative care. In Japanese palliative care patients, the HADS, along with two single-item questions, were compared to the Structured Clinical Interview for DSM-III-R. At the optimal cutoff score of 13, a sensitivity of 80% and specificity of 67% were achieved. The cutoff score of 13 in this Asian population is significantly lower than for the previous study in a European population. They found the HADS to a better screening instrument in this population when targeting adjustment disorder and major depression together than the single item questionnaire.

General Health Questionnaire-12. The 12-item, briefer version of the General Health Questionnaire, with focus on global distress rather than specific diagnostic categories, makes it a reasonable choice. Only one study reports the use of the General Health Questionnaire-12. The authors report only the area under the curve of 0.81 on a receiver operating characteristics analysis in identifying depression. Other parameters are not reported, which makes comparison difficult.

Mood Evaluation Questionnaire. This is a 33-item self-report measure developed specifically for assessment of depression in palliative care and rehabilitation patients. It has only psychological and cognitive items. Against the Structured Clinical Interview for DSM-III-R, the authors report moderate agreement (kappa 0.52). The reporting of results as agreements makes it difficult to compare with other studies.

Beck Depression Inventory-Short Form. This 13-item short version of the Beck Depression Inventory was developed as a rapid screening questionnaire for use in medically ill patients. At the recommended cutoff of 8, it showed a sensitivity of 79% and specificity of 75%.

Discussion

There are very few studies that compare the criterion validity of screening questionnaires in palliative care populations with standardized psychiatric interview. Many studies show promising results, but few have been replicated by other researchers or in sufficiently
large numbers. This is undoubtedly a reflection of the difficulties in the recruitment of patients into palliative care research studies.

Most of the studies focused on depression and only a few included other diagnostic categories. More recent studies examined a broader and wider spectrum of psychological problems, with increasing evidence for the prevalence of such a spectrum of disorders in this population.

Unidimensional scales attract immediate attention from clinicians because of their simplicity and ease of use. One example is the Distress Thermometer, which is becoming increasingly used in cancer settings. The validity of a unidimensional measurement of complex psychological constructs remains to be proven. Many of these questionnaires rely on a degree of self-evaluation and familiarity with the construct that it measures. These may pose problems in non-Western populations less familiar with such constructs.3

Multidimensional scales vary in their length from 6 to 30 items. They focus on various dimensions of distress, such as cognitive, behavioral, and somatic, and may theoretically capture the construct better. Excessive focus on one domain, for example, somatic aspects, makes them less desirable in palliative care settings. The lengths of the questionnaires are an obvious disadvantage.

The single-item questionnaire “Are you depressed?” show variable results across different cultures. Its usefulness as a screening aid for psychological distress remains to be proved.

The EDS fairs well in identifying depression in this population, but needs evaluation against gold standards that encompass a wider diagnostic spectrum reflecting psychological distress. The six-item, briefer version, BEDS, shows promise as a screening tool for depression.

The HADS anxiety subscale fared at least equally well to the 14-item full questionnaire in identifying cases of depression, shedding light into the psychopathology in this population, where emotionally distressing symptoms, such as anxiety and low mood, arising in the context of a stressor or illness, co-exist. Distinct diagnostic categorization may be inappropriate, and if one must for research or other purposes, then categories such as adjustment disorders, which capture the essence of these emotionally distressing states, should be included as well.

Screening for psychological distress is only one aspect of improving the management of this clinical issue. Of equal importance are interventions based on the outcomes of such screening. The increased use of screening instruments in a population should be accompanied by the validation of treatments to establish benefits from screening-led detection and treatment.13

Summation and immediate feedback of the screening results through user-friendly technology also have a role in improving effectiveness of the overall screening process. Computer-aided, touch-screen technology has been used successfully in quality-of-life screening, with good clinician and patient acceptance.32 Whether such technology can be adapted to psychological screening programs and in a patient population with poor performance status, such as palliative care, remains to be established.

Research and development of integrated screening programs must remain a priority. Screening tools, with their cutoff thresholds, only provide guidance. To rely on these as diagnostic aids would be a mistake. Screening should not be confused with assessment.33 Assessment is a more comprehensive process involving formulation of the problems, building a therapeutic relationship, and initiation of management strategies. Screening, on the other hand, is a prospective method of identifying distress that ideally leads to assessment in identified patients. Equally relevant is the judicious use of these scales. They should be part of an integrated program, where appropriate assessment and treatment is offered following screening.2

The usefulness of psychiatric questionnaires depends on changes in professional behavior, such as using them in routine practice to produce action on the results by clinicians. Without this, the whole exercise becomes a costly and cumbersome affair.34 Screening programs in these settings are most effective when coupled with an educational program to raise awareness among nonpsychiatric staff and the presence of a dedicated and responsible specialist psychiatric input.34

In applying this review to a consideration of these issues, it must be emphasized that it
focuses on the palliative care population and the findings may not be applicable to the whole cancer population. Further studies of screening tools also are needed.

References


28. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale—a