The Edmonton Symptom Assessment System (ESAS) – poor performance as screener for major depression in patients with incurable cancer

Elisabeth Brenne 1,2, Jon H Loge 1,4,5, Hanne Lie 3,4, Marianne J Hjermstad 1,5, Peter M Fayers 1,6 and Stein Kaasa 1,2

On behalf of the EPCRC, European Palliative Care Research Collaborative

---

1European Palliative Care Research Centre, Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Norway
2Cancer Clinic, St. Olav’s Hospital, University Hospital of Trondheim, Norway
3Regional Centre for excellence in Palliative Care, Oslo University Hospital, Norway
4Department of Behavioural Sciences in Medicine, Institute of Basic Medical Sciences, University of Oslo, Norway
5Regional Centre for Excellence in Palliative Care, South Eastern Norway, Oslo University Hospital, Norway
6Department of Public Health, Aberdeen University Medical School, UK

Corresponding author:
Elisabeth Brenne
PRC, Department of Cancer Research and Molecular Medicine. Norwegian University of Science and Technology, NTNU, N-7491 Trondheim, Norway
Email: elisabeth.brenne@ntnu.no
Abstract

Background: Depressive symptoms are prevalent in patients with advanced cancer, sometimes of a severity that fulfil the criteria for a Major Depressive Episode (MDE).

Aim: The aim of the present study was to investigate how the item on depression in the Edmonton Symptom Assessment System (ESAS-Depression) with a 0-10 numerical rating scale (NRS) performed as a screener for MDE. A possible improved performance by adding the ESAS-anxiety item was also examined for.

Design: An international cross-sectional study including patients with incurable cancer was conducted. The ESAS score was compared against MDE as assessed by the Patient Health Questionnaire (PHQ-9). Screening performance was examined by sensitivity, specificity and the Kappa coefficient.

Setting: Patients with incurable cancer (n=969), median age 63 years and from eight nationalities, provided report. Median Karnofsky Performance Status was 70. Median survival was 229 (205-255) days.

Results: PHQ-9-MDE was present in 133 of 969 patients (13.7%). ESAS-Depression screening ability for PHQ-9-MDE was limited. Area under the receiver operating characteristic curve (AUC) was 0.71 (0.66 - 0.76). Valid detection or exclusion of PHQ-9-MDE could not be concluded at any ESAS-Depression cut-off; by the cut-off NRS ≥2, sensitivity was 0.69 and specificity was 0.60. By the cut-off NRS ≥4, sensitivity was 0.51 and specificity was
0.82. Combined mean ratings by ESAS-Depression and ESAS-Anxiety revealed similar limited screening ability.

**Conclusions:** The depression and anxiety items of the ESAS, a frequently used assessment tool in palliative care settings, seem to measure a construct other than MDE as assessed by the PHQ-9 instrument.

**Keywords**

Palliative care, neoplasms, depression, depressive disorder, symptom assessment
**Introduction**

*Depression in patients with advanced cancer*

Patients with advanced cancer experience multiple symptoms such as pain, fatigue and anxiety, with depressive symptoms being among the commonest (1, 2). The term depression expresses a spectre from the transient feeling of sadness to major depression disorder (3-5). Assessed by different questionnaires, depression estimates vary around 30-50% of patients with advanced cancer (1, 6). Major Depressive Disorder is defined by the DSM and ICD psychiatric classification systems as one or more Major Depressive Episodes (MDEs) (7, 8). According to the DSM-5, a MDE is characterized by the persistent presence of at least five of nine depressive symptoms that cause significant distress or functional impairment for two weeks or longer. A systematic review estimated the point prevalence of MDE in palliative care cancer populations to 16.5% (13.1 – 20.3) (2) as opposed to estimations of 0.9% - 4.6% in general populations (9). The high prevalence underlines depression as an important concern in palliative care. This is further accentuated by the fact that major depression can be relieved by treatment; making identification of major depression an integral part of palliative care (10-12).

*Depression in symptom assessment*

In palliative care cancer programs, symptom assessment tools are used to assess symptom intensity, prevalence and course, to guide clinical consultations and to monitor treatment effects (13-15). Patients’ self-report of symptoms, also known as Patient Reported Outcomes (PROs), is a core assessment method in palliative care given this care’s central focus on
symptoms and quality of life (15-17). A plethora of instruments exists, ranging from single-item assessment tools to comprehensive assessment tools for multiple symptoms each often assessed by multiple items per dimension. Brief methods are advocated for monitoring the many symptoms in progressively diseased patients (18-20). Most tools assessing multiple symptoms include one or more items on depression symptoms. The second most applied assessment of depression in palliative care research is the Edmonton Symptom Assessment System (ESAS) that includes nine simple items on frequent symptoms (21, 22).

**Screening for depression**

Screening for depression by single items has been investigated in patients with advanced cancer. Screening by use of the MDE main criterion on lowered mood was originally promising (23). The single item “Are you depressed?” performed poorly in three later studies among patients with advanced cancer (24-26). In contrast to these findings, Taylor (27) found good screening capabilities by use of the main criterion on lowered mood for MDE. The question “Have you felt depressed most of the day, nearly every day, for two or more weeks?” identified MDD assessed by a structured interview with a sensitivity 0.80 and a specificity of 0.85 (27). This finding indicates that simple items are possible screeners for MDE in patients with advanced cancer.

**The ESAS-Depression item**

The Edmonton Symptom Assessment System (ESAS) assesses nine symptoms that are common among patients with advanced cancer, on 0-10 numerical rating scales (NRS). The ESAS includes one item of the symptoms pain, fatigue, nausea, depression, anxiety, drowsiness, appetite, general well-being and shortness of breath (21). The ESAS was
developed for a pragmatic day to day assessment of symptoms (21, 28, 29). The ESAS has successfully been implemented in daily symptom monitoring in palliative care worldwide (28, 30) and is frequently applied in palliative care research assessment (22). A recently developed extended version of the ESAS, the European Association for Palliative Care (EAPC) Basic Dataset, adds items on sleep, constipation and vomiting (31). There is no definitive guide how to interpret patient scorings of the emotional items ESAS-Depression and ESAS-Anxiety (28, 29). In validation studies, the ESAS-Depression item has been compared to the Hospital Anxiety and Depression rating Scale (HADS)(32-37), the Rotterdam Symptom Checklist (38-40), and later the Patient Health Questionnaire numerical sumscore (PHQ-9)(41, 42); PHQ-9 can be analysed by sumscore or by the MDE diagnostic concept. Palliative care guidelines recommend psychiatric classification as the standard reference for depression (12, 14). Interpretation of ESAS-Depression scorings into clinical decisions would be facilitated if ESAS-Depression could be applied as a screening instrument for MDE.

**ESAS-Depression – a valid indicator of depression?**

Previous studies have examined the screening abilities of ESAS-Depression. Vignaroli et al. examined a mixed cancer population (33) and compared ESAS-Depression with HADS (HADS-D ≥11). With acceptable sensitivity (0.83) and low specificity (0.47), Vignaroli proposed ESAS-Depression score ≥2 as a cut-off and proposed further research on cut-off values for severe depression. Bagha et al. examined cancer outpatients (42). They found that ESAS-Depression performed well compared to PHQ-9 sumscore (≥10 of 27) and proposed ESAS-Depression as an initial screening instrument to exclude non-depressed patients before screening with a more extended instrument. An ESAS-Depression score ≥2 was proposed as cut-off (sensitivity 0.86, specificity 0.72). Ripamonti et al (37) found good screening performance of
the ESAS Depression item compared to HADS (HADS cutoff ≥11) in patients with non-advanced cancer. They proposed a score ≥4 as an optimal cut-point.

Documentation of the screening capabilities of the ESAS-Depression item in the palliative care context is limited. Studies have been small and depression has been conceptualised differently. Still, ESAS has been claimed to be a valid screener (34, 35, 40).

**ESAS-Depression and ESAS-Anxiety**

Anxiety and depression frequently co-occur in patients with advanced cancer (2, 34, 43-45). The combined assessment of these symptoms is integrated in many instruments as a measure of psychological distress. The ESAS-anxiety item also tapping an underlying depression construct is therefore a reasonable hypothesis that has not yet been investigated.

**Aims of the study**

The aims of this study were to test screening capabilities of ESAS-Depression alone and in combination with ESAS-Anxiety when compared to Major Depressive Episode as assessed by the PHQ-9 in patients with incurable cancer.

The specific research questions were:

1. Does the ESAS-Depression item have adequate screening ability for a Major Depressive Episode (MDE) assessed by the PHQ-9 instrument (PHQ-9-MDE) in patients with incurable cancer?
2. Does the additional assessment ESAS-Anxiety improve screening ability for MDE assessed by the PHQ-9 instrument (PHQ-9-MDE) in patients with incurable cancer?
Materials and methods

Design and sampling

An observational cross-sectional study was conducted as part of an international multicenter study; the Computer Symptom Assessment (CSA) study run by the European Palliative Care Collaborative (EPCRC) (16, 46). The EPCRC project aimed to improve assessment and classification of depression, pain and cachexia in patients with advanced cancer (22, 45, 47-49). Inclusion criteria were incurable cancer, aged 18 years or older. Exclusion criteria were obvious cognitive impairment, language problems or physical disability preventing participation. Patients were included in palliative care inpatient- and outpatient units, hospices, and in general oncology, surgical and medical wards.

The study was approved by the appropriate ethical authorities at each study site. All patients provided written informed consent.

Data collection

A total of seventeen medical centres in eight countries participated from October 2008 to December 2009. The centres were: Australia: Braeside Hospital NS, West Australian Centre for Cancer and Palliative Care, Curtin University, Southern Adelaide Palliative Services; Austria: Medical University of Graz, Department of Internal Medicine, Division of Oncology; Canada: University of Alberta. Grey Nuns Community Hospital Division of Palliative Care Medicine; UK: University of Bristol, Department of Palliative Medicine. Bristol Haematology & Oncology Centre, St George's Hospital Medical School, St Georges University of London; Germany: University Hospital RWTH Aachen, Department of Palliative Medicine Aachen University; Italy: Rehabilitation and Palliative Care, National Cancer Institute of Milan, Unità Cure Palliative, Liguria, Genova; Norway: Palliative Medicine
Touch screen sensitive laptops were used for data collection; one section was filled in by patients and one section by health care personnel. The native languages; Norwegian, English, German or Italian were used in patient assessment. Each question had to be completed to move to the next question. A research nurse or study coordinator provided assistance if necessary.

**Measurements**

Health care personnel provided demographic and medical data (Table 1). Several PROs (46), among them the Edmonton Symptom Assessment System (ESAS)(21) and the Patient Health Questionnaire (PHQ-9)(41), were rated by the patients.

**ESAS.** The ESAS assesses nine symptoms; pain, tiredness, nausea, depression, anxiety, drowsiness, appetite, feeling of well-being and shortness of breath (21, 50). The items were introduced by “Please mark the number that best describes your situation right now:” “Depression” and “Anxiety”. The anchors of the scales were “Not depressed”, “Worst possible depression”, “Not anxious, and “Worst possible anxiety”. The 0-10 NRS scoring by the ESAS-Depression item and the mean ESAS-Depression and ESAS-Anxiety scoring for each patient, were examined.

**PHQ-9.** The PHQ-9 instrument is composed of 10 questions (51). The nine symptom criteria of MDE (DSM-5) are assessed: 1) little interest or pleasure in doing things, 2) feeling down, depressed or hopeless, 3) disrupted sleep, 4) feeling tired or having little energy, 5) appetite changes, 6) feeling bad about oneself or as a failure, 7) trouble concentrating, 8) slowness or
fidgety, restlessness and 9) thoughts of being better off dead or hurting oneself. The items are introduced by “Over the last two weeks, how often have you been bothered by...”. Each item is rated using four response options: 0, Not at all; 1, Several days; 2, More than half the days; 3, Nearly every day. Each symptom is regarded as present by ratings 2 or 3. A tenth question asks for the symptoms’ influence on functioning, but is not included in the standard analysis. The original English and the authorized Norwegian, German and Italian translations were used (41, 51). Standard scoring according to the DSM-5 diagnostic algorithm for MDE was applied as the main assessment (PHQ-9-MDE) (41, 51); at least five of the nine symptoms should be present, and at least one of these must be a main symptom. The PHQ-9 numeric sumscore (0-27) was also used as a comparator.

**Statistical analysis**

*Descriptive statistics.* Standard descriptive statistics were applied with frequencies, mean (sd), median (range, quartiles) of patient characteristics. Survival was calculated by the Kaplan Meyer method (52).

*Sensitivity and Specificity.* Indication for clinical support for detection of major depression was evaluated according to suggestions by Löwe (53); a minimum specificity of 0.75 and a maximum sensitivity above the specificity value (53).

In a two-step screening procedure, sensitivity in the first step would be most important to not overlook depressed subjects; sensitivity of 0.85 was considered putative for this purpose (42).

*Area under the ROC Curve (AUC).* Combined sensitivities and specificities were visualized in a Receiver Operating Characteristic (ROC) curve. The AUC provides an estimate of overall
discrimination, and for evaluations of an appropriate cut-off value of the screening item. AUC of 0.5 - 0.7 indicates low accuracy, 0.7 - 0.9 indicates moderate accuracy and 0.9 - 1.0 indicates high accuracy (54, 55).

**Coefficient Kappa χ.** Cohen’s coefficient χ estimated strength of agreement between ESAS dichotomized by the different cut-offs and PHQ-9-MDE (56). Clinical usefulness was evaluated according to suggestions by Landis (57): < 0.00, Poor; 0.00 - 0.20, Slight; 0.21 - 0.40, Fair; 0.41 - 0.60, Moderate; 0.61 - 0.80, Substantial; 0.81 - 1.0, Almost perfect.

The PASW 21 statistical package (IBM SPSS, Armonk, NY, USA. 2012) and an online statistical calculator ([http://statpages.org/ctab2x2.html](http://statpages.org/ctab2x2.html) 2014.03.18) were used with the statistical analyses.

**Results**

Altogether 1070 patients completed the study (Figure 1). Information on 15 patients was incomplete because of technical failure, and they were omitted from analyses. Four patients withdrew their informed consent. Of the remaining 1051 patients, 26 did not complete ESAS-Depression and ESAS-Anxiety, and 56 did not complete the PHQ-9, leaving a sample of 969 patients (92.2%) for the present study, 48% female and 52% male. Mean Karnofsky Performance Status (KPS) was 70.9 (SD 16.4). Median age was 63 (range 18-91). Median survival was 229 days (95% confidence interval 205-255). Gastrointestinal cancer was the commonest type of cancer (26.6%), 17.3% had breast cancer, 16.7% had cancer in the respiratory organs and 16.5% had urological cancer.
Figure 1. Patient recruitment

1070 patients registered

- 4 patients
  - Written consent retracted

1051 patients included

- 15 patients
  - Technical failure

969 patients
  - ESAS and PHQ-9 completed

Patient sample
Table 1. Patient characteristics

<table>
<thead>
<tr>
<th></th>
<th>Sample patients</th>
<th>Non-sample patients*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subjects</strong></td>
<td>969 (100%)</td>
<td>82 (100%)</td>
</tr>
<tr>
<td><strong>SOCIODEMOGRAPHIC DATA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>465 (48%)</td>
<td>38 (46 %)</td>
</tr>
<tr>
<td>Male</td>
<td>504 (52%)</td>
<td>44 (54 %)</td>
</tr>
<tr>
<td>Age (Years)</td>
<td>63</td>
<td>68</td>
</tr>
<tr>
<td>Median</td>
<td>63</td>
<td>68</td>
</tr>
<tr>
<td>Range</td>
<td>18-91</td>
<td>28-98</td>
</tr>
<tr>
<td>Living situation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse</td>
<td>643 (65.4%)</td>
<td>52 (66.7%)</td>
</tr>
<tr>
<td>Living alone</td>
<td>254 (26.2%)</td>
<td>21 (25.9%)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤9 years</td>
<td>333 (34.5%)</td>
<td>34 (41%)</td>
</tr>
<tr>
<td>10-12 years</td>
<td>338 (35.0%)</td>
<td>32 (40%)</td>
</tr>
<tr>
<td>&gt;12 years</td>
<td>295 (30.6%)</td>
<td>5 (6.3%)</td>
</tr>
<tr>
<td>Nationality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian</td>
<td>448 (50.4%)</td>
<td>32 (39%)</td>
</tr>
<tr>
<td>Austrian</td>
<td>99 (10.2%)</td>
<td>1 (1.2%)</td>
</tr>
<tr>
<td>Swiss</td>
<td>92 (9.5%)</td>
<td>18 (22%)</td>
</tr>
<tr>
<td>Italian</td>
<td>88 (9.1%)</td>
<td>3 (3.7%)</td>
</tr>
<tr>
<td>British</td>
<td>81 (8.4%)</td>
<td>4 (4.9%)</td>
</tr>
<tr>
<td>Australian</td>
<td>61 (6.3%)</td>
<td>9 (11%)</td>
</tr>
<tr>
<td>Canadian</td>
<td>31 (3.2%)</td>
<td>3 (3.7%)</td>
</tr>
<tr>
<td>German</td>
<td>29 (3.0%)</td>
<td>12 (14.6%)</td>
</tr>
<tr>
<td><strong>MEDICAL DATA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatients</td>
<td>420 (43.4%)</td>
<td>31 (37.8%)</td>
</tr>
<tr>
<td>Inpatients</td>
<td>547 (56.6%)</td>
<td>51 (62.2%)</td>
</tr>
<tr>
<td>Survival**</td>
<td>229</td>
<td>52</td>
</tr>
<tr>
<td>Days. Median</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>205 - 255</td>
<td>29 - 76</td>
</tr>
<tr>
<td>Mean</td>
<td>70.9</td>
<td>58.6</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>16.4</td>
<td>20.2</td>
</tr>
<tr>
<td>&lt;80</td>
<td>542 (56.6%)</td>
<td>60 (73.2%)</td>
</tr>
<tr>
<td>≤80</td>
<td>748 (78.2%)</td>
<td>72 (87.9%)</td>
</tr>
<tr>
<td>Major depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-9-MDE***</td>
<td>133 (13.7%)</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>307 (31.7%)</td>
<td></td>
</tr>
<tr>
<td>PHQ-9 sumscore ≥10 of 27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressants****</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>131 (13.5%)</td>
<td>10 (12.3%)</td>
</tr>
<tr>
<td>With PHQ-9-MDE***</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Without PHQ-9-MDE***</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Cancer diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>257 (26.6%)</td>
<td>17 (20.7%)</td>
</tr>
<tr>
<td>Breast</td>
<td>167 (17.3%)</td>
<td>10 (12.2%)</td>
</tr>
<tr>
<td>Respiratory organs</td>
<td>161 (16.7%)</td>
<td>16 (19.5%)</td>
</tr>
<tr>
<td>Urinary/male genital</td>
<td>159 (16.5%)</td>
<td>15 (18.3%)</td>
</tr>
<tr>
<td>Gynecological</td>
<td>26 (2.7%)</td>
<td>6 (7.3%)</td>
</tr>
<tr>
<td>Other</td>
<td>199 (20.2%)</td>
<td>18 (22.0%)</td>
</tr>
<tr>
<td><strong>Mean ESAS reports</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(standard deviation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>2.2 (2.3)</td>
<td></td>
</tr>
<tr>
<td>Tiredness</td>
<td>3.7 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Nausea</td>
<td>1.1 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>1.9 (2.3)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>2.1 (2.3)</td>
<td></td>
</tr>
<tr>
<td>Drowsiness</td>
<td>3.4 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Reduced appetite</td>
<td>3.3 (3.0)</td>
<td></td>
</tr>
<tr>
<td>Feeling of well-being</td>
<td>3.4 (2.4)</td>
<td></td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>1.9 (2.4)</td>
<td></td>
</tr>
</tbody>
</table>
Half of the patient sample was recruited in Norway, the remaining patients being uniformly spread among the other nationalities (Table 1). The prevalence of a Major Depressive Episode (MDE) as assessed by PHQ-9 (PHQ-9-MDE) was 13.7% (Table 1). Out of these 133 patients, 33 patients received antidepressant medication for other than pain. Ninety-eight patients receiving antidepressants did not fulfil the PHQ-9-criteria for MDE. The mean score for ESAS-Depression was 1.9 (standard deviation 2.3), and mean ESAS-Anxiety was 2.1 (standard deviation 2.3). Patients who did not complete the PROs (non-sample patients) had lower KPS and shorter survival than those who did. About the same proportion of patients received antidepressants in the sample (13.5%) and the non-sample (12.3%) groups.

Figure 2. ROC curve for ESAS Depression, ESAS-Anxiety and for mean ESAS-Depression and ESAS-Anxiety combined. Reference standard: PHQ-9-MDE (see text)
Sensitivity and specificity. The ROC curves (Figure 2) showed relatively low overall accuracies. The AUC was about 0.70 for all measures (Table 2) indicating that the ESAS-Depression item is a poor discriminator of MDE as defined by the PHQ-9-MDE. Similarly, the use of the combined ESAS-Depression and ESAS-Anxiety score added little. Also by use of a cut off on the PHQ-9 sumscore (≥10 of 27) as an alternative to MDE, AUC was about the same (Table 2).

Table 2. Area under the Receiver Operating Curve (AUC) with 95% Confidence interval (CI95)

<table>
<thead>
<tr>
<th>Items</th>
<th>AUC</th>
<th>CI95</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A specificity of 0.75 combined with a sensitivity of 0.75 or higher was not reached by any cut-off (Tables 3 and 4).

An adequate sensitivity of 0.85 was reached for mean ESAS-Depression and ESAS-Anxiety at the cut-off NRS ≥0.5. Specificity at this cut-off was low 0.34, which means that 66% of patients without MDE would be further evaluated for MDE (false positives).

At the cut-off of NRS ≥2 both for the ESAS-Depression item and for mean ESAS-Depression and ESAS-Anxiety, the present sensitivity was 0.69, specificity was 0.60 (Tables 3 and 4). At the cut-off of NRS ≥4, sensitivity was 0.51, specificity was 0.82 for ESAS-Depression; for mean ESAS-Depression and ESAS-Anxiety sensitivity was 0.48 and specificity was 0.83 (Tables 3 and 4).

The highest sensitivity reached by the ESAS-Depression item was 0.79 with a specificity of 0.46 at the cut-off NRS ≥1. At this most liberal cut-off, 21% of the patients with PHQ-9-MDE

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESAS-Depression</td>
<td>0.71</td>
<td>0.66 – 0.76</td>
</tr>
<tr>
<td>Mean ESAS-Depression and ESAS-Anxiety</td>
<td>0.71</td>
<td>0.65 – 0.76</td>
</tr>
<tr>
<td>ESAS-Anxiety</td>
<td>0.67</td>
<td>0.62 – 0.73</td>
</tr>
<tr>
<td>ESAS-Depression when compared to PHQ-9 sumscore (≥10 of 27)</td>
<td>0.72</td>
<td>0.68 – 0.76</td>
</tr>
</tbody>
</table>
would be overlooked and 54% of the patients without PHQ-9-MDE would be further evaluated for MDE.

Coefficient Kappa $\kappa$: Coefficient $\kappa$ was in the range 0.03 - 0.26 for all cut-offs of the ESAS items (Tables 3 and 4). This represents poor to slight agreement between the ESAS items and PHQ-9-MDE, for all cut-off points.

**Discussion**

ESAS-Depression had limited screening ability for Major Depressive Episode (MDE) when MDE was assessed by the PHQ-9 instrument in patients with incurable cancer. AUC was 0.71, and either sensitivity or specificity was inadequate depending on the 0-10 NRS cut-off point. Combining an anxiety item with the depression item made no improvement. Using ROC curves, there were no satisfactory cut-off values for either detection or exclusion of PHQ-9-MDE. Assuming the PHQ-9 instrument assesses MDE adequately in the population, the ESAS-Depression item appears insufficient to screen for MDE.

Former suggestions of applying ESAS-Depression with a cut-off point of NRS $\geq$2 or NRS $\geq$4, provide little help for screening MDE according to this study. Applying the ESAS-Depression cut-off point NRS $\geq$2 as advocated by Vignaroli (33) and Bagha (42), the present study revealed a sensitivity of 0.69 and a specificity of 0.60 which are inadequate both for exclusion and inclusion of MDE. The suggested cut-off point NRS $\geq$4 by Ripamonti (37) provided an inadequate sensitivity of 0.51 and a specificity of 0.82 in this study (Table 3).

Some other studies on patients with advanced cancer have been published. Teunissen (34) (n=54) found limited screening ability comparing ESAS-Depression with HADS (HADS-D $\geq$11). Delgado-Guay (35) (n=216) explored the association between ESAS-Depression and
HADS-Depression and found a low Spearman correlation of 0.39. Carvajal (40)(n=90), found a weighted coefficient Kappa of 0.32 between ESAS-Depression and RSCL. The authors though concluded that the ESAS is valid for screening purposes.

Patients with advanced cancer describe the ESAS-Depression and ESAS-Anxiety items difficult to rate (58, 59); the terms “depression” and “anxiety” were perceived unspecific. An unclear perception of the term depression might indicate more underlying concepts and different understandings in patients with incurable cancer. Resistance against reporting on psychiatric disorders is a general finding (60) and may influence scorings (59). Selby (61) found a floor effect with discrepancy between low ratings of ESAS-Depression and ESAS-Anxiety and high emotional impact and burden. The ESAS-Depression and ESAS-Anxiety items were low rated in the actual study; mean ESAS-Depression was 1.9, mean ESAS-Anxiety was 2.1. Noguera (36)(n=100) found a deflated floor effect and better screening performance by use of the more colloquial term “discouraged” compared to “depressed” in translating ESAS into Spanish. Refinement of wordings might have screening potentials not yet inquired in patients with advanced cancer.

Content of the terms “depression” and “anxiety” seem differently loaded in cultures (14, 36, 62). The actual study mainly included European patients; half of them were Norwegian. National differences cannot be concluded from the actual study due to insufficient sample sizes; however, the study by Noguera (36) points to a need of thorough considerations of further cultural adjustments of ESAS-Depression.

Adding the ESAS-Anxiety scoring to ESAS-Depression gave about the same limited screening ability. With the limited screening properties of ESAS-Depression, an eventual increased validity of adding anxiety measurement cannot be evaluated, and should be further
investigated. Another solution to increase accuracy of simple screening instruments might be the combined assessment with other key symptoms than anxiety. Payne (63) increased accuracy significantly by questioning the second main depressive symptom “Loss of interest” to the simple question “Are you depressed?” in a study in the palliative care context (sensitivity 0.91, specificity 0.68, reference criterion MDD by DSM-IV psychiatric interview). Questioning the two main depressive symptoms by the DSM complete symptom descriptions with a 0-10 NRS response, has not been examined in patients with advanced cancer.

There are several adjusted versions of ESAS (28). Inference from the current study is limited to the ESAS version applied in the study. Regarding symptom time frame, the patients were asked about depression “just now”; the term “Depression” was used, not “Depression = feeling sad” as in the ESAS revised version (64).

A prerequisite for the inferences of the screening performance explored in this study is the adequacy of the PHQ-9 instrument as a reference standard for MDE. The revised DSM classification, DSM-5, promotes PHQ-9 to be investigated as an added severity measurement of MDE (7). The instrument is increasingly used in clinical studies, also in patients with cancer (12, 42, 65, 66). PHQ-9 is applied in clinical assessment throughout the cancer trajectory (42). Among the applied translated versions of PHQ-9, validity is only examined for the German version (51, 53) and not the Norwegian and Spanish version. The PHQ-9 patients’ report has shown strong agreement with MDE assessed by psychiatric interviews as the gold standard of diagnosing in several patient populations (53, 67-69); however some studies found limited detection performance of PHQ-9 (70, 71). Heterogeneity of populations is one proposed condition to variations in performance of the PHQ-9 (72). PHQ-9 is not validated in
patients with advanced cancer. Luckett in a review, commented on five of the nine PHQ-9 questions as problematic in a palliative care population; sleep, fatigue, appetite, concentration and restlessness (73). The problem with symptom overlap between depression and cancer affects the psychiatric classification system itself applied in patients with advanced cancer (7, 45, 74). Palliative care guidelines define the psychiatric classification systems as the clinical standard (11, 14). The diagnostic algorithm might however reduce overestimation of MDE by the overlapping somatic depressive and cancer symptoms (75). The prevalence rate of MDE measured by PHQ-9 in the present study was 13.7%. This is in line with prevalence rates of major depression in patients with advanced cancer (2, 6) and may indicate that the use of PHQ-9-MDE is acceptable in the study.

Strengths of the study include international multinational collaboration, inclusion of patients across countries and cultures, a large sample size of patients with advanced cancer and blinded reports due to computer-based assessments.

**Limitations**

Inference from this study can only be drawn for patients with incurable cancer. The inclusion was performed conveniently and not consecutively which might have skewed the sample and introduced a healthy bias. The frailest patients did not fill in the PROs supporting this interpretation. Still, the lowered performance status and the survival estimates indicate the sample is representative for patients with incurable cancer in an early palliative phase. About fifty percent of the patients were Norwegian, and the results might therefore not be generalizable for all countries. The wording of the item in Norwegian might not be optimal and need reconsideration. Another limitation is the use of PHQ-9 as the external criterion for MDE and not a diagnosis based upon a psychiatric diagnostic interview. This was done for
Conclusion

ESAS-Depression seems to measure a construct other than PHQ-9-MDE in patients with advanced cancer. The underlying construct of ESAS-Depression should be further investigated in patients with advanced cancer as should the validity of the PHQ-9 as a determinant of MDE in this patient group. Cultural and translational considerations should be addressed.

Acknowledgements

The authors wish to thank Even Fyllingen for important assistance in the development of the computer programme and Turid Follestad for statistical assistance. We thank all participating centres for important contributions.

Funding

This study was supported by the EU’s 6th framework, contract no. 037777, PI Stein Kaasa, and funding from Trondheim University Hospital, Norway. Elisabeth Brenne received a grant from the Norwegian Cancer Society.

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

APPENDIX

About the EPCRC
The European Palliative Care Research Collaborative is funded by the European Commission’s Sixth Framework Programme (contract no LSHC-CT-2006-037777) with the overall aim to improve treatment of pain, depression and fatigue through translation research. Core scientific group / work package leaders: Stein Kaasa (project coordinator), Frank Skorpen, Marianne Jensen Hjermstad, and Jon Håvards Loge, Norwegian University of Science and Technology (NTNU); Geoffrey Hanks, University of Bristol; Augusto Caraceni and Franco De Conno, Fondazione IRCCS Istituto Nationale dei Tumori, Milan; Irene Higginson, King’s College London; Florian Strasser, Cantonal Hospital St. Gallen; Lukas Radbruch, RWTH Aachen University; Kenneth Fearon, University of Edinburgh; Hellmut Samonigg, Medical University of Graz; Ketil Bø, Trollhetta AS, Norway; Michael Schade, Bender MedSystems GMbH, Austria; Odd Erik Gundersen, Verdande Technology AS, Norway. Scientific advisory group: Neil Aaronson, The Netherlands Cancer Institute; Vickie Baracos and Robin Fainsinger, University of Alberta; Patrick C. Stone, St. George’s University of London; Mari Lloyd Williams, University of Liverpool. Project management: Stein Kaasa, Ola Dale, and Dagny F. Haugen, NTNU.

References

37. Ripamonti CI, Bandieri E, Pessi MA, Maruelli A, Buonaccorso L, Miccinesi G. The Edmonton Symptom Assessment System (ESAS) as a screening tool for depression and anxiety in non-advanced patients with solid or haematological malignancies on cure or follow-up. Support Care Cancer. 2014;22(3):783-93.