Oncologists’ perception of depressive symptoms in patients with advanced cancer:
Accuracy and relational correlates

Lucie Gouveia¹,²
*Corresponding author
Email: lucie.gouveia@umontreal.ca

Angélique Bonnaud-Antignac³
Email: angelique.bonnaud@univ-nantes.fr

Sophie Lelorain³
Email: sophie.lelorain@univ-lille3.fr

Florence Cousson-Gélie⁶
Email:florence.cousson-gelie@univ-montp3.fr

Anne Brédart⁴
Email: anne.bredart@curie.fr

Serge Sultan¹,²
Email: serge.sultan@umontreal.ca

Sylvie Dolbeault⁴
Email: sylvie.dolbeault@curie.fr

¹Département de psychologie, Université de Montréal
Pavillon Marie-Victorin
90, ave Vincent d’Indy
H2V 2S9
Montreal, Qc, Canada

²Centre de recherche, CHU Sainte-Justine
3175, Chemin de la Côte-Sainte-Catherine
H3T 1C5
Montreal, Qc, Canada

³Unité de Recherche en Sciences Cognitives et Affectives, URECA, EA 1059
Université Lille 3, UFR de Psychologie
Rue du Barreau, BP 60149
F-59653 Villeneuve d'Ascq cedex

⁴Psycho-Oncology Unit, Institut Curie
26 rue d’Ulm, 75248
Paris Cedex, France

⁵Équipe de recherche en sciences humaines et sociales appliquées à la cancérologie
Université de Nantes, EA4275 SPHERE
Nantes St-Herblain, France

⁶Institut régional du cancer, Pôle prévention Epidaure
208 Avenue des Apothicaires
34298 Montpellier Cedex 5
Université Montpellier 3
Montpellier, France
Abstract

Background: Health care providers often inaccurately perceive depression in cancer patients. The principal aim of this study was to examine oncologist-patient agreement on specific depressive symptoms, and to identify potential predictors of accurate detection.

Methods: 201 adult advanced cancer patients (recruited across four French oncology units) and their oncologists (N=28) reported depressive symptoms with eight core symptoms from the BDI-SF. Various indices of agreement, as well as logistic regression analyses were employed to analyse data. Results: For individual symptoms, medians for sensitivity and specificity were 33% and 71%, respectively. Sensitivity was lowest for suicidal ideation, self-dislike, guilt, and sense of failure, while specificity was lowest for negative body image, pessimism, and sadness. Indices independent of base rate indicated poor general agreement (median DOR = 1.80; median ICC = .30). This was especially true for symptoms that are more difficult to recognise such as sense of failure, self-dislike and guilt. Depression was detected with a sensitivity of 52% and a specificity of 69%. Distress was detected with a sensitivity of 64% and a specificity of 65%. Logistic regressions identified compassionate care, quality of relationship, and oncologist self-efficacy as predictors of patient-physician agreement, mainly on the less recognisable symptoms. Conclusions: The results suggest that oncologists have difficulty accurately detecting depressive symptoms. Low levels of accuracy are problematic, considering that oncologists act as an important liaison to psychosocial services. This underlines the importance of using validated screening tests. Simple training focused on psychoeducation and relational skills would also allow for better detection of key depressive symptoms that are difficult to perceive.
Oncologists’ perception of depression in patients with advanced cancer

Keywords: cancer; oncology; depression; symptom assessment; physician-patient relations; patient-centered care
Background

Depression is a common emotional experience in people with advanced cancer. A review of the literature (Mitchell et al. 2011) suggests that many patients in palliative care suffer from adjustment disorders (~15.4%), minor depressive disorders (~9.6%), or major depression (~16.5%). Indeed, patients with brain metastases have been found to report more emotional symptoms than physical complaints (Cordes et al. 2014).

Stromgren et al. (2001) found that, amongst 102 patients with advanced cancer, more than half reported significant levels of depression. However, less than a third of these cases were reported in medical records. Similar findings have repeatedly been reported in the general cancer population, suggesting that physicians and other health care providers (HCPs) may inaccurately perceive patient distress, particularly depression (Lampic and Sjödén 2000; Werner et al. 2012; Keller et al. 2004; Trask et al. 2002). This is problematic considering that HCPs serve as the first line to psychosocial services. In addition to disrupting resource allocation, failing to understand the patient’s personal experience can hinder the collaborative process on which important medical decisions rest. Few studies have examined this issue amongst individuals with late-stage cancer.

The aim of this study was to better understand detection of depression in advanced care patients by measuring patient-oncologist agreement on specific depressive symptoms and by examining relational skills as predictors of accurate detection.

Physician accuracy on patient depression

Depression is defined by the World Health Organisation “as a common mental disorder, characterized by sadness, loss of interest or pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, feelings of tiredness and poor concentration”
In the context of cancer care, it can be understood as a type of distress, defined by the National Comprehensive Cancer Network (NCCN) as an “unpleasant emotional experience” that varies in magnitude and may interfere with coping abilities (Holland et al. 2013). Although depression may be referred to as a psychiatric diagnosis, the term is also used to describe subclinical levels of the disorder, as in the present research. The definition also varies according to the method of measurement. Over the past few decades, it has consistently been reported that HCPs often fail to detect depression in cancer patients (e.g. Lamic and Sjödén 2000; Okuyama et al. 2011; Werner et al. 2012). Although diverse statistical indices have been employed to assess HCP accuracy on patient depression, findings generally converge.

Patient ratings of their own depression are typically used as the reference point against which HCP ratings are compared. While some studies use standardised tools for patients and HCPs, others only do so for patients. Most commonly reported is sensitivity (number of cases detected by HCPs/total number of cases) and specificity (number of non-cases detected by HCPs/total number of non-cases). Low sensitivity values of 12.2 to 30.4% suggest that physicians have difficulty detecting depression when it is present. Specificity (74 to 97%) is generally higher, which may reflect a tendency to prematurely rule out depression (Passik et al. 1998; Werner et al. 2012; Okuyama et al. 2011).

Kappa statistics evaluating agreement between patient and physician ratings of patient distress range from .04 to .17 (Keller et al. 2004; Passik et al. 1998; Werner et al. 2012; Fukui et al. 2009; Sollner et al. 2001; Chidambaram et al. 2014), indicating poor accuracy (Landis and Koch 1977). Despite rare contradicting reports, most recent studies
support the idea that oncologists struggle to discriminate between cases and non-cases of depression.

Although several studies deal with recognition of depression in cancer patients, almost none have detailed their results at the symptom level. This represents a major gap in the literature considering that detection of depression is contingent on the recognition of specific signs. To our knowledge, only one research team has taken a symptomatic approach. Passik et al. (1998) reported findings suggesting that physicians' perception of symptoms associated with obvious signs might be more accurate than that of other less recognisable ones. No additional studies have further pursued this hypothesis.

Another issue is the use of inappropriate indices of accuracy (Passik et al. 1998; Trask et al. 2002; Werner et al. 2012) where other indices are recommended (Peat and Barton 2005; Glas et al. 2003). A simple product-moment correlation, for example does not reflect the absolute agreement between two ratings, but rather their similarity in ranking. The intraclass correlation coefficient (ICC) is preferable, as it accounts for the distance between physician and patient scores (Peat and Barton 2005). For the analysis of dichotomous variables, an index of agreement that is much less dependent on prevalence than the kappa is the diagnostic odds ratio\(^2\) (DOR), which represents the odds of caseness in ‘test positives’ (i.e. patients rated as distressed by oncologists) relative to the odds of caseness in ‘test negatives’ (Glas et al. 2003).

Key symptoms of depression in adult oncology

There has been much discussion around distinctive symptoms of depression in the medically ill (Trask 2004). Various screening instruments exclude somatic symptoms, which typically overlap with the side effects of physical illness. In accordance with this,
research suggests that affective and cognitive symptoms are optimal for identifying depression in this population (Sultan et al. 2010), as they lower the rate of false negatives. Studies in cancer care support this idea (Reuter et al. 2004; Warmenhoven et al. 2012). Key symptoms may differ according to cancer stage, due to changes in somatic symptoms and patient status (Mitchell et al. 2012). This has yet to be verified, as there is little research on detection of depression amongst patients with advanced cancer, possibly due to recruitment and attrition difficulties.

Potential predictors of accurate detection

Based on preliminary research, many factors seem to influence oncologists’ ability to accurately detect depressive symptoms in their patients. For example, a number of studies indicate that physicians’ empathic attitude and skills have an important impact on how accurately they perceive distress in cancer patients as well as the extent to which patients feel understood (Razavi et al. 2003; Merckaert et al. 2008; Fukui et al. 2009). According to Neumann et al. (2009)’s model, an empathic style of communication increases the accuracy of caregivers’ perceptions and diagnoses by encouraging patient disclosure. More generally, it is thought that the quality of the patient-physician relationship allows for better detection of distress (Newell et al. 1998; Ryan et al. 2005).

Another potential element which may enhance perception of patient depression is oncologists’ self-efficacy in detecting distress. In fact, confidence in personal skills appears to be one of the main barriers to successful screening (Mitchell et al. 2008). However, this idea deserves to be nuanced, as the construct of self-efficacy is easily confounded with overconfidence, a characteristic which may harm rather than enhance performance (Moores and Chang 2009).
Study objectives

Our first objective was to estimate oncologists’ ability to accurately detect individual depressive symptoms amongst advanced cancer patients, in addition to depression and psychological distress, and to compare the results across symptoms. It was hypothesized that patient-oncologist agreement would be lower for less obvious symptoms (sense of failure, guilt, self-dislike, suicidal ideation), compared to more recognisable ones (sadness, pessimism, negative body image). Unlike the former, the latter are associated with specific cues, such as crying/droopy facial expression (sadness), reactions to negative prognoses (pessimism) and hair loss (negative body image). We also wanted to identify key symptoms that contribute to accurate detection of depression and distress. The second main objective was to examine relational variables as predictors of oncologist accuracy for each symptom (i.e. physician-reported empathy, self-efficacy in detecting distress, and quality of relationship with patients).

Methods

Procedure

A cross-sectional design involving patient-physician dyads was elaborated. Oncologists at the ‘Institut Curie’ (Paris and Saint-Cloud), the ‘Institut de Cancérologie de l’Ouest’ (Nantes), the ‘Hôpital Nord Laennec’ (Nantes), and the ‘Polyclinique Bordeaux Nord Aquitaine’ (Bordeaux) were invited to participate. Those interested completed questionnaires examining professional characteristics and empathic skills. Each physician was asked to choose ten of their own patients meeting a set of selection criteria (see below). In consultation, they introduced the study to these patients, and handed them a consent form with depression and distress questionnaires. Patients who
agreed to participate had one week to complete the documents and mail them back to the coordinating center in a pre-paid envelope. The physicians completed an analogous set of questionnaires in a perspective taking task (Sultan et al. 2011), in which they provided the answers which they thought their patient had given. This paradigm allowed the assessment of patient-physician agreement. The protocol was approved by the institutional review board of the Institut Curie (DR-2011-318) and by the French national advisory committee for the processing of information in health research (11.202).

Participants

Oncologists

Sixty-four oncologists were contacted. Of these, 14 refused to participate, 11 had ineligible patients, and 11 accepted but did not follow through for reasons related to time and/or motivation. Twenty-eight oncologists (10 male) participated in the study. Differences between these participants and those who dropped out are unknown. The age of participating oncologists ranged from 31 to 64 years (Table 1).

Patients

The sample of patients for the present study consisted of 201 advanced cancer patients (146 female). To participate, patients needed to meet the following criteria: age 18+ years, metastatic cancer from and beyond the 4th line of chemotherapy for primary breast cancer, or from and beyond the 2nd line of chemotherapy for any other type of primary cancer. Patients had to have already consulted the physician at least 3 times before their inclusion, so that they had a minimum knowledge of each other (Lelorain et al. 2014). Exclusion criteria were confirmed psychiatric pathology and hematological cancers. The age of patients ranged from 27 to 89 years old. Diagnoses included breast

[Tapez un texte]
Oncologists’ perception of depression in patients with advanced cancer

Cancer (45.3%), colorectal cancer (20.9%), lung cancer (14.9%), and others (18.9%; Table 1).

Measures

**Depression and depressive symptoms**

A short form of the Beck Depression Inventory (BDI-SF) was used to measure Depression and depressive symptoms (Collet and Cottraux 1986). Each item refers to one cognitive or affective symptom (Self-Dislike, sense of Failure, Guilt, Negative Body Image, Pessimism, Suicidal Ideation, Sadness, and Dissatisfaction with Life), and was selected for medical settings (Beck and Beck 1972; Sultan et al. 2010). For each item, the responder chooses one of four statements of varying intensity (0-3), according to his/her present state. A cutoff of 3 yields the best trade-off between sensitivity and specificity when screening for depression in patients with chronic illnesses (Sultan et al. 2010). The internal consistency for this sample was very good (α = .81). Convergent and predictive validity have also been supported (Furlanetto et al. 2005). In a population of women with metastatic breast cancer, the BDI-SF performed better than the Hospital Anxiety and Depression Scale in screening for DSM-IV depressive disorders (Love et al. 2004). It has been shown to recognize 88% of clinical cases amongst diabetes patients (Sultan et al. 2010). In this study, individual items served as measures of symptoms. A cutoff of 1 was used, discriminating between presence and absence of any given symptom.

**Distress**

Distress was assessed via the Distress Thermometer (DT; Dolbeault et al. 2008), originally developed by Roth et al. (1998). This visual analogue scale ranges from ‘no distress’ to ‘extreme distress’. The DT is recommended by the NCCN (Holland et al.)
A cutoff score of 4/10 is recommended, and has been identified as optimal for research purposes in a sample of cancer survivors (Boyes et al. 2013). As a screening test, the DT rarely misses clinical cases of distress, though it does not reliably exclude sub-clinical ones (e.g. Mitchell 2007). A more thorough evaluation is needed when looking to identify purely clinical cases.

**Potential predictors of patient-physician agreement**

Four variables relating to relational skills were assessed. Physicians completed the Jefferson Scale of Physician Empathy (JSPE; Hojat et al. 2002). Confirmatory analyses of the French version have failed to support the existence of an over-arching global factor (Zenasni et al. 2012). However, support was found for two factors within the questionnaire: Compassionate Care (CC) and Perspective Taking (PT). While the latter measures a cognitive aspect of empathy, the former concerns emotional processes (Hojat et al. 2002). The PT and CC scores consist of ten and eight items, respectively. In the present database, Cronbach’s alphas were .57 (CC), .64 (PT), and .74 (total). Despite support for the questionnaire’s construct validity (Glaser et al. 2007), it is undermined by low internal consistency.

Physicians also rated their sense of self-efficacy in detecting patient distress on a self-developed Likert scale: “In general, I feel competent to detect my patients’ emotional distress and needs (1 = strongly disagree; 7 = strongly agree)”. Post-consultation, they rated the quality of the patient-physician relationship using a similar scale: “What is the quality of your relationship with this patient? (1 = very difficult relationship; 7 = very easy relationship)”. Statistical analysis
The DOR and the ICC\textsuperscript{b} were used to calculate agreement between patients’ and physicians’ scores on patient Depression, depressive symptoms, and Distress. Patient ratings on the BDI-SF and the DT were used as reference points against which physician ratings were compared. To allow for inter-study comparisons, we also calculated other indices typically seen in the literature, such as the kappa statistic.

To identify which symptoms best contributed to patient-physician agreement on Depression and Distress, two stepwise logistic regressions were performed. Agreement (versus disagreement) on Depression (1\textsuperscript{st} model) or Distress (2\textsuperscript{nd}) was entered as the dependent variable. Eight predictor variables (patient-physician agreement/disagreement on each symptom) were then entered in both models, using the forward Likelihood Ratio method. Agreement versus disagreement was determined for each dyad according to the established cutoffs (i.e. 3 for Depression, 1 for depressive symptoms and 4 for Distress).

Next, a hierarchical logistic regression model was constructed, entering control variables in the first block and then adding the four predictor variables in a second block. This model was run to predict agreement on each of the eight symptoms, as well as Depression and Distress. Due to lack of research, the confounding factors are unclear. Control variables were thus identified from the study’s large dataset. Correlation analyses were performed on sociodemographic and clinical variables, to determine their relationship with patient-physician agreement on Depression, individual depressive symptoms, and Distress. Significant correlations were retained as control variables (Cohen 1988).

Analyses were performed through IBM SPSS Statistics 20 and an alpha level of .05 was set for statistical significance.
Results

Preliminary analyses

The mean Depression score was 3.94 (SD = 3.33), with a 51.5% rate of significant depression. Pessimism (51.8%) and Sadness (42.6%) were the most prevalent depressive symptoms. Guilt (14.0%) and Suicidal Ideation (17.0%) were the rarest. The mean Distress score was 1.80 (SD = 1.60), with a 25.9% rate of significant distress.

Mean level comparisons indicate moderate differences between physician and patient scores on Distress (d = -.76; 49.3% overestimation). Small differences were found for Suicidal Ideation (d = .33; 13.4% underestimation) and Negative Body Image (d = -.30; 39.8% overestimation). Weak differences were found for Sadness (d = -.22; 32.8% overestimation) and Pessimism (d = -.20; 36.3% overestimation). No significant differences were found on the remaining symptoms and Depression scores (Table 2).

Patient-physician agreement

Sensitivity was only slightly higher for Depression (68.9%) than for Distress (64.3%; Table 3). Specificity was higher for Distress (64.7%) than for Depression (52.0%). Regarding symptoms, sensitivity was highest for Pessimism (73.5%), Negative Body Image (68.4%), and Dissatisfaction (49.2%). Specificity was highest for Suicidal Ideation (94.6%), Self-Dislike (85.1%), and Guilt (84.9%).

Percent agreement and the kappa coefficient were not coherent. All kappa values indicated only slight agreement, except that of depression which indicated fair patient-physician agreement (κ = .21).

The DOR obtained for Depression was small (2.41; Rosenthal 1996), although near moderate (the odds that a patient reporting distress be judged as distressed was 2.41
times that of a patient who did not report distress). A moderate value (3.31) was obtained for Distress. All symptom DORs were small, except for Suicidal Ideation (4.52).

Similarly, no good or excellent ICCs were obtained (Landis and Koch 1977). Values for Distress (.52), Sadness (.48), Depression (.42), and Suicidal Ideation (.40) indicated fair agreement. The next three highest were Pessimism (.36), Negative Body Image (.30), and Dissatisfaction (.30). Agreement was poor on Self-Dislike (.17), Guilt (.15), and Sense of Failure (.14). With the exception of Suicidal Ideation (due to high specificity), this order of symptoms provides some support for the idea that less obvious symptoms are particularly difficult to detect. However, overlapping confidence intervals indicate minimal differences.

Key symptoms in accurate detection of depression and distress

In decreasing order of odds ratios (OR), patient-physician agreement on Pessimism (OR = 6.27; 95% confidence interval (CI) = 2.94-13.36; \( p < .000 \)), Negative Body Image (OR = 4.27; 95% CI = 2.01-9.07; \( p < .000 \)), Sadness (OR = 3.72; 95% CI = 1.77-7.82; \( p < .000 \)), and Dissatisfaction (OR = 3.20; 95% CI = 1.51-6.78; \( p < .002 \)), were retained in the first model, as the most significant predictors of agreement on Depression.

This led to an overall model characterised by a correct classification power of 76.8%. A test of the model against the constant-only model was significant, \( \chi^2 \) (df = 4, N = 190) = 76.36, \( p < .001 \), Nagelkerke \( R^2 = .45 \), indicating that the model statistically distinguished between agreement and non-agreement on Depression.

In decreasing order of ORs, patient-physician agreement on Guilt (OR = 4.65; 95% CI = 2.18-9.94; \( p < .000 \)) and Dissatisfaction (OR = 3.91; 95% CI = 2.02-7.58; \( p < .000 \)) were retained in the second model, as the most significant predictors of agreement on Distress.
This led to an overall model characterised by a correct classification power of 71.1%. A test of the model against the constant-only model was significant, $\chi^2 (df = 2, N = 190) = 34.20, p < .001$, Nagelkerke $R^2 = .23$, indicating that the model statistically distinguished between agreement and non-agreement on Distress.

Relational variables predictive of patient-physician agreement

Correlation analyses revealed that patient status, cancer site, patient gender and age showed significant relationships to at least one of the dependent variables. These variables were integrated as control variables. Physician age and gender were also retained, given their similarity to the patient variables. As expected, the control variables significantly predicted patient-physician agreement in the regression analyses (data available upon request).

Agreement on Depression was not significantly associated with any of the predictor variables, beyond the effect of controls (Table 4). Agreement on Distress was associated with higher-quality relationships (OR_1.81; 95% CI_1.28-2.56; $p_=.001$). Agreement on several symptoms was significantly related to higher CC, perception of higher-quality patient-physician relationships and higher self-efficacy in detecting distress. Agreement on Sense of Failure (OR_1.54; 95% CI_1.03-2.32; $p_=.037$) was associated with higher CC. Results approached significance for Guilt (OR_1.61; 95% CI_1.00-2.56; $p_=.050$). Agreement on sense of Failure (OR_1.41; 95% CI_1.02-1.95; $p_=.040$), Dissatisfaction with life (OR_1.95; 95% CI_1.40-2.73; $p_=.000$), Guilt (OR_1.55; 95% CI_1.10-2.18; $p_=.013$), and Self-Dislike (OR_1.56; 95% CI_1.11-2.19; $p_=.010$) were associated with higher-quality relationships, although the ORs are small. Agreement on Sadness (OR_1.92; 95% CI_1.27-2.91; $p_=.002$) was associated with self-
efficacy. Contrary to predictions, however, agreement on sense of Failure (OR_0.62; 95% CI_0.39,-0.97; p_0.037) and Self-Dislike (OR_0.59; 95% CI_0.36-0.97; p_0.039) were associated with lower PT.

**Discussion**

The present study demonstrates poor oncologist accuracy on patient depressive symptoms, particularly those that are more subtle in nature. Accuracy on pessimism, sadness, dissatisfaction with life, and negative body image emerged as key elements when exploring factors predicting accuracy on depression and distress as a whole. Additionally, physicians who reported higher levels of compassionate care, relationship quality and self-efficacy in detecting distress tended to be more accurate on individual depressive symptoms.

Patient-physician agreement on all symptoms was low. Still, agreement on the intensity of easily recognisable symptoms (sadness, pessimism, negative body image, and dissatisfaction with life) was consistently (though insignificantly) higher than that of less obvious symptoms (self-dislike, guilt, sense of failure). This is in line with the findings reported by Passik et al. (1998). Interestingly, however, is that overestimation was highest for the former. This may be explained by a tendency to amplify symptoms that are easier to perceive. Indeed, appearances can be misleading; a female patient who has lost her hair will not necessarily hold a negative body image. In this study, negative body image was the most overestimated symptom at 39.8%, indicating that oncologists relied too heavily on appearances when rating this symptom. Similarly, Holmes and Eburn (1989) found that nurses were better able to detect distress symptoms such as appearance and tiredness, although these were generally overestimated. Pessimism was the second
most overestimated symptom in this study at 36.3%. This corresponds to the findings by Faller et al. (1995), who reported that professional caregivers tended to underestimate the amount of hope held by cancer patients. An exception was suicidal ideation which, although difficult to detect as indicated by a low sensitivity score, received the highest accuracy scores. This can be explained by an almost-perfect specificity (94.6%).

Recognition of cases was slightly higher for depression than it was for distress, while recognition of non-cases was higher for distress. These results contradict the literature, as the opposite is most commonly found. Still, overestimation was far more frequent for distress. This may be explained by physicians’ tendency to rate the DT in a polarized manner (low distress vs. high distress) – a trend which was not observed on the psychometrically more reliable BDI-SF. Overall though, accuracy was better on distress than it was on depression and symptoms.

Results suggest that both affective and cognitive symptoms are involved in accurate detection of depression and distress. Accurate detection of pessimism, sadness, dissatisfaction with life, and negative body image accounted for nearly half of the variation in accurate detection of depression. Accurate detection of dissatisfaction with life and guilt contributed the most to accurate detection of distress, although they accounted for less (23%). These may be key symptoms involved in identification of depression and distress amongst adults with advanced cancer. These analyses, however, are still exploratory and should be pursued further.

Support was also found for the hypothesis predicting that oncologists’ relational skills would be associated with patient-oncologist agreement on depressive symptoms. In
accordance with Neumann et al. (2009)’s model of empathic communication, the quality of the patient-oncologist relationship and compassionate care were predictive of agreement on several symptoms. Interestingly, these results were found for the symptoms with the lowest levels of patient-physician agreement as measured by the ICC, suggesting that relational skills are especially important for evaluating symptoms that are harder to perceive.

Moreover, the results suggest that self-efficacy in detecting patient distress may also play a part, namely in detecting sadness. However, this result only surfaced for one symptom out of eight. One explanation for this is that the scale used may be a better measure of overconfidence than of healthy self-efficacy. A multi-item questionnaire would most likely be needed to reliably measure this construct.

Unexpectedly, perspective taking predicted inaccuracy on patient sense of failure and self-dislike. Again, this may be due to a gap between the construct which the scale is meant to measure and that which it actually taps into. Whereas compassionate care captures open-mindedness toward empathy, perspective taking is centered on self-evaluation of empathic skills. The latter scale may inadvertently be measuring overconfidence in one’s own empathic skills. Such a phenomenon has been observed amongst pharmacy students; those with poor empathy skills were found to largely overestimate their personal abilities (Austin and Gregory 2007). A performance task would most probably have been a more valid measure.

The present study has several limitations. First, it must be noted that the situation in which oncologists were placed is unnatural and may therefore limit the applicability of the results. Perhaps physicians tended to overestimate symptoms simply because the
perspective-taking task attracted their attention to them. Secondly, the results may be
affected by a selection bias, as less than 50% of the contacted physicians participated in
the study. Perhaps interest in empathy is related to accuracy on patient distress. Thirdly,
the limited sample size combined with the high number of variables likely led to
underpowered analyses. The findings should therefore be considered as exploratory in
nature. Fourthly, many of the measures have limited reliability due to either low internal
consistency (JSPE) or a one-item structure (depressive symptoms, self-efficacy, quality
of relationship). Fifthly, some of the predictor variables are not independent and thus may
violate the logistic regression assumptions. Consequently, results involving the
perspective-taking and compassionate care scores from the JSPE should be considered
with caution. Sixthly, it may be argued that between-physician differences explain part of
the results. To explore this avenue, we compared agreement rates between physicians and
found no significant differences (Figures 1 and 2). Multilevel analyses with larger
samples would be recommended in future studies.

Despite its limitations, this work enriches research on detection of distress in quite
a few ways. For one, it points to the importance of using standardised tests to screen for
depression, as patient-physician agreement is low on all symptoms. In addition, this study
sheds light on the relational and psychological evaluation skills necessary for accurate
detection of depression and distress in cancer patients. Teaching these to HCPs could
help them decide whether they should refer patients to psychosocial services when test
scores are at a borderline level or unavailable. Once a profile of key symptoms is well
delineated, training could be made a lot simpler by focusing on those signs that allow for
most efficient detection of depression (and other forms of distress). Moreover, this study
adds to current literature on patient-HCP agreement by examining individual symptoms. Previous studies have not offered this level of analysis, and have often presented inappropriate statistical indices. Finally, this study adds to the existing literature by focusing on homogeneous samples that are difficult to recruit, patients and oncologists included. Such properties eliminate potential confounding variables and increase the study’s internal validity.

**Conclusion**

The use of robust indices clearly illustrated oncologists’ lack of accuracy on depressive symptoms, especially covert ones. Although the cross-sectional design of this study prevents us from establishing directionality of associations, the findings clearly emphasize the role of *relational skills* in detecting these symptoms. They demonstrate the value of using structured screening instruments and of training physicians in relational and key-symptom assessment skills. Such measures could significantly enhance the detection and handling of patient depression.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

LG elaborated hypotheses, conducted statistical analyses and drafted the manuscript. SL helped conceive the study, collected the data and revised the manuscript. AB helped conceive the study, supervised data collection and revised the manuscript. SD co-supervised data collection and discussed earlier versions of the study. ABA and FCG participated in data collection. SS supervised the whole project, contributing conceptual, theoretical and methodological suggestions, and revised the manuscript.
Acknowledgements

This project was funded by the French National Cancer Institute (SHS SPE 2010) and supported by the CHU Sainte-Justine Foundation, the Larry and Cookie Rossy Foundation, and Industrial Alliance. These finding bodies did not participate in design, collection, analysis, or interpretation of data.

Endnotes

a. DOR = (sensitivity X specificity)/[(1 – sensitivity)X(1 – specificity)]; 1.5 = small, 2.5 = medium, 4 = large, 10 = very large (Rosenthal 1996). b. < .40 = poor agreement, .40 - .59 = fair agreement, .60 - .74 = good agreement, ≥ .75 = excellent agreement (Landis and Koch 1977).
References


Feasibility and performance of the adapted distress thermometer. *Palliative & Supportive Care, 6*(02), 107-117, doi:10.1017/S1478951508000187.


[Tapez un texte]


Table 1

Sample Description.

<table>
<thead>
<tr>
<th>Variables</th>
<th>201 Patients</th>
<th>28 Oncologists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>M</td>
</tr>
<tr>
<td>Age</td>
<td>61.97</td>
<td>11.49</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>55 (27.4)</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>146 (72.6)</td>
<td></td>
</tr>
<tr>
<td>Years of education / practice</td>
<td>2.64</td>
<td>.91</td>
</tr>
<tr>
<td>Cancer site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>91 (45.3)</td>
<td></td>
</tr>
<tr>
<td>Colorectal</td>
<td>42 (20.9)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>30 (14.9)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>38 (18.9)</td>
<td></td>
</tr>
<tr>
<td>Patient status&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.08</td>
<td>.91</td>
</tr>
<tr>
<td>Physician specialty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical oncology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td>1 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Palliative care</td>
<td>5 (17.9)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3 (10.7)</td>
<td></td>
</tr>
<tr>
<td>Patient Depression (BDI-SF, 0–24)</td>
<td>3.46</td>
<td>3.33</td>
</tr>
<tr>
<td>Patient Distress (DT, 0–10)</td>
<td>1.80</td>
<td>1.60</td>
</tr>
</tbody>
</table>

Note. <sup>a</sup>0 = normal activity; 1 = some symptoms, but still near fully ambulatory; 2 = < 50% of daytime in bed; 3 = > 50%; 4 = completely bedridden.
Table 2

Comparisons between Oncologist and Patient Ratings.

<table>
<thead>
<tr>
<th>Measure</th>
<th>M (SD)</th>
<th></th>
<th>r</th>
<th>t (d)</th>
<th>Underestimation (%)</th>
<th>Acceptable Estimation (%)</th>
<th>Overestimation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient</td>
<td>Oncologist</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>3.46 (.33)</td>
<td>3.94 (3.50)</td>
<td>.29***</td>
<td>1.67 (-.14)</td>
<td>15.9</td>
<td>62.7^a</td>
<td>21.4</td>
</tr>
<tr>
<td>A) Sadness</td>
<td>.54 (.72)</td>
<td>.70 (.73)</td>
<td>.31***</td>
<td>2.66** (-.22)</td>
<td>18.4</td>
<td>48.8^b</td>
<td>32.8</td>
</tr>
<tr>
<td>B) Pessimism</td>
<td>.77 (.88)</td>
<td>.95 (.91)</td>
<td>.22**</td>
<td>2.27* (-.20)</td>
<td>2.2</td>
<td>41.3</td>
<td>36.3</td>
</tr>
<tr>
<td>C) Failure</td>
<td>.34 (.69)</td>
<td>.30 (.53)</td>
<td>.08</td>
<td>-63 (.07)</td>
<td>18.4</td>
<td>63.2</td>
<td>18.4</td>
</tr>
<tr>
<td>D) Dissatisfact.</td>
<td>.35 (.57)</td>
<td>.47 (.67)</td>
<td>.18*</td>
<td>2.16 (-.19)</td>
<td>17.9</td>
<td>57.2</td>
<td>24.9</td>
</tr>
<tr>
<td>E) Guilt</td>
<td>.25 (.66)</td>
<td>.24 (.57)</td>
<td>.08</td>
<td>-13 (.02)</td>
<td>11.9</td>
<td>74.6</td>
<td>13.4</td>
</tr>
<tr>
<td>F) Self-Dislike</td>
<td>.21 (.47)</td>
<td>.17 (.42)</td>
<td>.09</td>
<td>-.95 (.09)</td>
<td>14.9</td>
<td>73.1</td>
<td>11.9</td>
</tr>
<tr>
<td>G) Suicidal Idea</td>
<td>.26 (.63)</td>
<td>.09 (.35)</td>
<td>.29***</td>
<td>3.65*** (.33)</td>
<td>13.4</td>
<td>82.1</td>
<td>4.5</td>
</tr>
<tr>
<td>H) Body Image</td>
<td>.74 (.90)</td>
<td>1.01 (.90)</td>
<td>.18*</td>
<td>3.27** (-.30)</td>
<td>21.9</td>
<td>38.3</td>
<td>39.8</td>
</tr>
<tr>
<td>Distress</td>
<td>1.80 (1.60)</td>
<td>3.07 (1.73)</td>
<td>.35***</td>
<td>9.47*** (-.76)</td>
<td>8.5</td>
<td>42.3^c</td>
<td>49.3</td>
</tr>
</tbody>
</table>

Note. ^a Evaluations of depression were considered acceptable when situated within 17 points away from the patient’s score. This margin is based on an α of .81, calculated for the patient BDI-SF; ^b Evaluations on BDI-SF items were considered acceptable when they exactly matched the patient’s score; ^c Evaluations of distress were considered acceptable when situated within 6.3 points away from the patient’s score. This margin is based on a test-retest r of .80, reported in a recent validation study of the DT (Tang et al. 2011).

*p < .05, **p < .01, ***p < .001.
Table 3

Accuracy of Oncologists’ Ratings.

<table>
<thead>
<tr>
<th>Measure (base rate %)</th>
<th>Cutoff</th>
<th>Agreement (%)</th>
<th>Se (%)</th>
<th>Sp (%)</th>
<th>(\kappa)</th>
<th>DOR</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression (51.5)</td>
<td>≥3</td>
<td>60.7</td>
<td>68.9 (59.5-77.1)</td>
<td>52.0 (42.3-61.7)</td>
<td>.21 (.14-.34)</td>
<td>2.41 (1.35-4.28)</td>
<td>.42 (.24-.56)</td>
</tr>
<tr>
<td></td>
<td>≥1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A) Sadness (42.6)</td>
<td>41.0</td>
<td>32.5 (.23-.43)</td>
<td>47.3 (38.3-56.5)</td>
<td>.19 (.08-.32)</td>
<td>0.43 (.24-.78)</td>
<td>.48 (.31-.61)</td>
<td></td>
</tr>
<tr>
<td>B) Pessimism (51.8)</td>
<td>39.1</td>
<td>73.5 (64.2-81.1)</td>
<td>44.2 (34.6-54.2)</td>
<td>.18 (.05-.31)</td>
<td>2.20 (1.21-4.00)</td>
<td>.36 (.15-.51)</td>
<td></td>
</tr>
<tr>
<td>C) Failure (25.0)</td>
<td>65.0</td>
<td>34.0 (22.4-47.9)</td>
<td>75.3 (67.9-81.6)</td>
<td>.09 (-.05-.24)</td>
<td>1.57 (.79-3.15)</td>
<td>.14 (-.14-.35)</td>
<td></td>
</tr>
<tr>
<td>D) Dissatisfaction (30.5)</td>
<td>62.0</td>
<td>49.2 (37.1-61.4)</td>
<td>67.6 (59.5-74.8)</td>
<td>.16 (.02-.30)</td>
<td>2.02 (1.09-3.74)</td>
<td>.30 (.07-.47)</td>
<td></td>
</tr>
<tr>
<td>E) Guilt (14.0)</td>
<td>77.0</td>
<td>28.6 (15.3-47.1)</td>
<td>84.9 (78.8-89.5)</td>
<td>.12 (-.04-.28)</td>
<td>2.25 (.90-5.64)</td>
<td>.15 (-.12-.36)</td>
<td></td>
</tr>
<tr>
<td>F) Self-Dislike (19.1)</td>
<td>72.9</td>
<td>21.1 (11.1-36.4)</td>
<td>85.1 (78.8-89.8)</td>
<td>.07 (-.08-.21)</td>
<td>1.52 (.62-3.72)</td>
<td>.17 (-.10-.37)</td>
<td></td>
</tr>
<tr>
<td>G) Suicide Ideas (17.0)</td>
<td>82.0</td>
<td>20.6 (10.4-36.8)</td>
<td>94.6 (90.0-97.1)</td>
<td>.19 (.02-.36)</td>
<td>4.52 (1.55-13.20)</td>
<td>.40 (.21-.55)</td>
<td></td>
</tr>
<tr>
<td>H) Body Image (47.5)</td>
<td>53.5</td>
<td>68.4 (58.5-76.9)</td>
<td>40.0 (31.1-49.6)</td>
<td>.08 (-.05-.21)</td>
<td>1.44 (.81-2.59)</td>
<td>.30 (.08-.47)</td>
<td></td>
</tr>
<tr>
<td>Distress (25.9)</td>
<td>≥4</td>
<td>64.7</td>
<td>64.3 (45.8-79.3)</td>
<td>64.7 (57.4-71.5)</td>
<td>.17 (.05-.28)</td>
<td>3.31 (1.44-7.61)</td>
<td>.52 (.36-.63)</td>
</tr>
</tbody>
</table>

Note. 95% confidence interval in parentheses; Se = Sensitivity; Sp = Specificity; \(\kappa\) = Kappa statistic. Full statistical information is available upon request.
Table 4  

*Logistic Regression Analysis of Patient-Physician Agreement on Depressive Symptoms as a Function of Relational Variables.*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Sadness</th>
<th>Pessimism</th>
<th>Failure</th>
<th>Dissatisfaction</th>
<th>Guilt</th>
<th>Self-Dislike</th>
<th>Suicidal Ideas</th>
<th>Negative Body Image</th>
<th>Global Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of relationship</td>
<td>1.90 (.66-1.23)</td>
<td>1.20 (.88-1.63)</td>
<td>1.41* (1.02-1.95)</td>
<td>1.95*** (1.40-2.73)</td>
<td>1.55* (1.10-2.18)</td>
<td>1.56* (1.11-2.19)</td>
<td>.97 (.66-1.42)</td>
<td>1.05 (.78-1.41)</td>
<td>1.345 (.983-1.840)</td>
</tr>
<tr>
<td>Compassionate Care</td>
<td>.76 (.52-1.12)</td>
<td>.91 (.63-1.34)</td>
<td>1.54* (1.03-2.32)</td>
<td>.90 (.61-1.33)</td>
<td>1.61* (1.0-2.56)</td>
<td>1.13 (.73-1.74)</td>
<td>1.10 (.68-1.78)</td>
<td>1.25 (.86-1.82)</td>
<td>.887 (.604-1.303)</td>
</tr>
<tr>
<td>Perspective Taking</td>
<td>.70 (.45-1.09)</td>
<td>.86 (.56-1.31)</td>
<td>1.10 (.71-1.70)</td>
<td>.62 (.37-1.06)</td>
<td>.59* (.36-97)</td>
<td>.68 (.39-1.20)</td>
<td>.93 (.62-1.40)</td>
<td>.865 (.561-1.334)</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>1.92** (1.27-2.91)</td>
<td>1.35 (90-2.01)</td>
<td>.87 (58-1.32)</td>
<td>1.40 (.93-2.12)</td>
<td>.92 (.58-1.49)</td>
<td>1.56 (1.11-2.19)</td>
<td>1.04 (.64-1.68)</td>
<td>1.06 (.72-1.55)</td>
<td>1.412 (.936-2.128)</td>
</tr>
</tbody>
</table>

**Model characteristics**

| Correct classification (%)   | 65.1     | 64.5     | 70.0     | 68.5     | 80.0     | 72.4     | 82.0     | 61.5     | 67.2     |
| Nagelkerke R$^2$             | .13      | .09      | .16      | .18      | .17      | .18      | .09      | .07      | .14      |

*Note.* ORs adjusted for site of cancer, patient status, gender and age of physicians and patients.  
\( \text{p} < .06, \text{*p} < .05, \text{**p} < .01, \text{***p} < .001. \)
Figures

Figure 1. Percent frequency of patient-oncologist agreement on depression.
Agreement/disagreement was determined according to the BDI-SF cutoff score (3). The figure only features the oncologists who saw ten patients (n = 12). Values are displayed with 95% confidence intervals. Physician #6 was in agreement with all of his patients.

Figure 2. Percent frequency of patient-oncologist agreement on distress.
Description of data: Agreement/disagreement was determined according to the DT cutoff score (4). The figure only features the oncologists who saw ten patients (n = 12). Values are displayed with 95% confidence intervals.
Figure 1
Figure 2