Agreement for depression diagnosis between DSM-IV-TR criteria, three validated scales, oncologist assessment, and psychiatric clinical interview in elderly patients with advanced ovarian cancer

Wadhih Rhondali1
Gilles Freyer2
Virginia Adam3
Marilène Filbet4
Martine Derzelle5
Gaelle Abgrall-Barbry6
Sophie Bourcelet7
Jean-Louis Machaviole8
Muriel Chomat-Neufread9
Olivier Gisserot10
Rémi Largillier11
Annick Le Rol12
Frank Priou13
Pierre Saltel14
Claire Falandry15

1 Clinique Mon Repos, Clinea, Marseille, France; 2 Medical Oncology Unit, Centre Hospitalier Lyon Sud, Université Lyon 1, Pierre-Bénite, France; 3 Institut de Cancérologie de Lorraine Alexis Vautrin, Vandoeuvre-lès-Nancy, France; 4 Palliative Unit, Centre Hospitalier Lyon Sud, Université Lyon 1, Pierre-Bénite, France; 5 Institut Jean Godinot, Reims, France; 6 Tenon Hospital, Assistance Publique Hôpitaux de Paris, Paris, France; 7 Centre Léon Bérard, Lyon, France; 8 Centre François Baclesse, Caen, France; 9 Centre Hospitalier de la région d’Annecy, Pringy, France; 10 Hôpital d’Instruction des Armées Sainte-Anne, Toulon, France; 11 Centre Azuèren de Cancérologie, Moungins, France; 12 Medical Oncology, Hôpital Purpétuel Secours, Levallois-Perret, France; 13 Medical Oncology, Centre Hospitalier Départemental Les Ouates, La Roche-sur-Yon, France; 14 Supportive Care Department, Centre Léon Bérard, Lyon, France; 15 Geriatrics and Oncology Unit, Centre Hospitalier Lyon Sud, Université Lyon 1, Pierre-Bénite, France

Background: Depression, a major outcome in cancer patients, is often evaluated by physicians relying on their clinical impressions rather than patient self-report. Our aim was to assess agreement between patient self-reported depression, oncologist assessment (OA), and psychiatric clinical interview (PCI) in elderly patients with advanced ovarian cancer (AOC).

Methods: This analysis was a secondary endpoint of the Elderly Women AOC Trial 3 (EWOT3), designed to assess the impact of geriatric covariates, notably depression, on survival in patients older than 70 years of age. Depression was assessed using the Geriatric Depression Scale-30 (GDS), the Hospital Anxiety Depression Scale, the distress thermometer, the mood thermometer, and OA. The interview guide for PCI was constructed from three validated scales: the GDS, the Hamilton Depression Rating Scale, and the Montgomery Asberg Depression Rating Scale (MADRS). The Diagnostic and Statistical Manual of Mental Disorders, fourth edition, revised (DSM) criteria for depression were used as a gold standard.

Results: Out of 109 patients enrolled at 21 centers, 99 (91%) completed all the assessments. Patient characteristics were: mean age 78, performance status ≥2: 47 (47%). Thirty six patients (36%) were identified as depressed by the PCI versus 15 (15%) identified by DSM. We found moderate agreement for depression identification between DSM and GDS (k=0.508) and PCI (k=0.431) and high agreement with MADRS (k=0.663). We found low or no agreement between DSM with the other assessment strategies, including OA (k=0.043). Identification according to OA (yes/no) resulted in a false-negative rate of 87%. As a screening tool, GDS had the best sensitivity and specificity (94% and 80%, respectively).

Conclusion: The use of validated tools, such as GDS, and collaboration between psychologists and oncologists are warranted to better identify emotional disorders in elderly women with AOC.

Keywords: depression, elderly, cancer, screening, geriatric assessment

Introduction
Depression is a major outcome in cancer patients, with an estimated frequency of 16%, and this frequency is reported to be higher for elderly patients (45%). Depression is the most common psychiatric illness in patients with cancer and it is known to reduce patients’ quality of life and to decrease their adherence to medical treatments. Depressive symptoms (asthenia, sleep disturbances, anorexia, etc) have been reported to significantly affect patients and their families. Moreover, recent studies have...
shown that depression is an independent predictive factor for cancer-related mortality,7–9 even though depression usually responds well to treatment (nonpharmacologic and pharmacologic).10–12

The evaluation of depression has been a challenge because symptoms cover a broad spectrum, from sadness to major depressive disorder, and because mood change is often difficult to assess when patients are confronted by repeated threats to life or are experiencing pain.6

For a depression diagnosis, clinicians typically rely on their clinical impression rather than on patient self-report or structured assessment.13,14 A systematic routine screening for depression has been implemented in some cancer care settings. Nevertheless, a substantial amount of data shows that depression remains underdiagnosed and undertreated in older patients, especially in the context of cancer,15–19 with low agreement between physicians and patients about the prevalence of mild and moderate/severe depression (33% and 13%, respectively).20–22 Similar results have been reported for nurses.14,21 Several reasons have been highlighted to explain those barriers to diagnosis and thus to adequate treatment.

First, many cancer patients are not treated for depression due to assessment difficulties, especially with older patients.6,19 This population often develops symptoms that differ from those experienced by younger patients, with more somatic complaints and fewer emotional symptoms (sadness, guilt, anhedonia).18,19 Patients in later adulthood are more reluctant to report depressive symptoms such as sadness or suicidal ideations.12 Furthermore, they might consider these symptoms as belonging to aging processes.12

Those difficulties may be due to nondisclosure by the patients themselves, who may feel they are wasting the doctor’s time or that they are in some way to blame for their distress and therefore choose to hide their feelings.24 There is also the wrong belief among some professionals that terminal illness invariably causes depression,25 and there are disagreements about the meaning of the term “depression” within psychiatric classification.26 We also know that some symptoms of depression (eg, fatigue, loss of appetite) might be expected consequences of physical illness.27,28 It is possible that some of the difficulties faced when attempting to identify depression might be decreased by the use of appropriate screening scales (adapted to the population and to the clinical setting).12

This study aimed to assess agreement between patient self-reported depression using different scales,18 oncologist assessment (OA) based on the oncologist’s clinical impressions, and a psychiatric clinical interview (PCI) conducted by a psychologist using the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, revised (DSM)29 among elderly patients with advanced ovarian cancer (AOC).

Methods

This analysis was a secondary endpoint of the Groupe d’Investigateurs Nationaux pour l’Etude des Cancers Ovariens (GINECO) Elderly Women AOC Trial 3 (EWOT3) Phase II multicenter trial, designed to assess the impact of geriatric covariates, notably depression, on survival in patients older than 70 years of age receiving six courses of carboplatin as first-line chemotherapy for stage 3–4 ovarian cancer. The local ethics committee and the Institutional Review Board at the Hospices Civils de Lyon approved this study, and all patients gave written informed consent.

Participants

To be included, patients had to have advanced (International Federation of Gynecology and Obstetrics stage 3 or 4) ovarian cancer and be older than 70 years of age. They could not have received chemotherapy previously and should have had a life expectancy of at least 3 months. Inclusion and exclusion criteria and patient characteristics have been previously reported.30

Assessments

Emotional disorders – and particularly depression – were assessed using patient self-report, OA, and PCI. Chronologically, oncologists were asked to provide their clinical impression before the self-reports. The PCI was performed within 10 days of enrollment according to the trial design, and the interviewer was blinded to the OA and to the self-report results.

Self-reports

All self-reports used were validated in their French form and included the Geriatric Depression Scale-30 (GDS),31 the Hospital Anxiety Depression Scale (HADS),32,33 the distress thermometer (DT), and the mood thermometer (MT).34 The GDS is a 30-item scale. This scale has been developed and validated especially for depression screening among elderly patients35,36 and has also been used in cancer patients.37 Scores of 0–9 are considered normal, 10–19 indicate mild depression, and 20–30 indicate severe depression.35 The HADS is a 14-item scale with separate subscales for anxiety (HADS-A) and depression (HADS-D). It has been validated in cancer patients for depression and anxiety diagnosis38 and can be administered or used as a self-assessment tool. The DT is a
one-item self-report screening tool for assessing psychological distress in cancer patients. The DT asks patients to rate the intensity of their psychological distress using an eleven-point visual analog scale from 0 (no distress) to 10 (extreme distress), with a score of 4 used to indicate a significant level of distress. The MT is a one-item self-report screening tool for measuring mood in cancer patients. Patients are asked to answer the question “How depressed have you been today and over the last week?” by rating their answer on a 0–10 scale (0 = normal mood; 10 = extreme depression). A score of 4 is used to indicate a significant level of depression. All of these self-report assessment tools were transformed into binary variables with the following cutoff scores: ≥10/30 for GDS-30, >14/42 for HADS, >4/10 for MT, and >4/10 for DT.

**Oncologist assessment**

Oncologists were invited to give their clinical impression regarding depression (yes/no). This item was previously shown to be correlated to survival in two previous GINECO studies of elderly patients with AOC.

**Psychiatric clinical interview**

A PCI was conducted by psychologists within 10 days of enrollment. The interview guide for the PCI was constructed and adapted from the DSM criteria and three validated scales: GDS, Hamilton Depression Rating Scale (HDRS), and Montgomery Asberg Depression Rating Scale (MADRS), according to a structured interview guide for evaluating depression in elderly patients developed by Tison. Psychologists were to assess depression using DSM criteria (DSM major depressive disorder, yes/no) and to give their clinical impression (PCI, yes/no). The DSM was considered to be the gold standard.

**Procedure and outcomes**

Demographic information (ie, age) and clinical data (ie, performance status, albumin level, weight, cancer diagnosis, and metastatic status) were collected.

Patient functioning was assessed using two validated tools: basal Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL).

On the day of enrollment, patient assessments included a clinical evaluation, laboratory results, and psychological self-reports (HADS, DT, MT, and GDS-30). In the 10 days after enrollment, a PCI based on MADRS and HDRS questionnaires was conducted.

Survival data were collected from the medical chart, starting from the time of enrollment to the patient’s death. Patients alive at the time of data collection were included in the analysis as censored cases using the date of last follow-up.

**Statistical considerations**

The number of patients to include was defined in the first aim of the GINECO EWOT3 study, a secondary endpoint of which is examined in the current study, using previous findings from our group. A single analysis requires a sample of 105 patients, assuming that depression triples the risk of death (two-tailed test), with a significance level of α=5% and a power of 90% (β=10%). This calculation was based on the assumption, on the basis of results from the EWOT1 and EWOT2 studies, that a death rate of 85% is expected at 2 years (86.4%, according to EWOT1 and EWOT2).

Categorical variables were presented by their frequencies and percentages and continuous variables by their means and standard deviations if they were normally distributed. When not normally distributed, data were reported by their median and interquartile range (Q1–Q3) and analyzed using nonparametric methods. Mann–Whitney U-tests and chi-square tests were performed to determine factors associated with depression according to the DSM. Agreements between depression diagnosis and DSM, PCI, and self-reported assessments were evaluated using Cohen’s kappa coefficient (κ).

P-values <0.05 were considered statistically significant. All statistical analyses were performed using SPSS v17 (SPSS, Inc., Chicago, IL, USA).

**Results**

**Patient characteristics**

One hundred and nine patients had a psychological assessment; of these, 99 had a complete set of psychological data and were included in the analysis. Patients’ characteristics are presented in Table 1.

**Depression assessments**

Out of the 99 patients included, 15 (15%) were depressed according to DSM and 39 (39%) were depressed according to PCI. We found different levels of agreement between DSM – considered as the gold standard – and other assessments (Table 2):

- no agreement (kappa coefficient between 0.0 and 0.2): with the DT (κ=−0.060), MT (κ=0.058), HADS (κ=0.127), HDRS (κ=0.151), and OA (κ=−0.043),
- moderate agreement (kappa coefficient between 0.4 and 0.6): with GDS (κ=0.508) and with PCI (κ=0.431),
- high agreement (kappa coefficient between 0.6 and 0.8): with MADRS (κ=0.663).
Table 1 Patient characteristics, N=99

<table>
<thead>
<tr>
<th>Variables</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td></td>
</tr>
<tr>
<td>≥80</td>
<td>78 (5 )</td>
</tr>
<tr>
<td>≥85</td>
<td>38 (38)</td>
</tr>
<tr>
<td>PS, N (%)</td>
<td></td>
</tr>
<tr>
<td>0–1</td>
<td>52 (53)</td>
</tr>
<tr>
<td>2–3</td>
<td>47 (47)</td>
</tr>
<tr>
<td>ADL, median (Q1–Q3)</td>
<td></td>
</tr>
<tr>
<td>Number of patients with an ADL ≤4</td>
<td>5.5 (5–6)</td>
</tr>
<tr>
<td>IADL, median (Q1–Q3)</td>
<td>22 (17–26)</td>
</tr>
<tr>
<td>Number of patients with an IADL ≤24, N (%)</td>
<td>68 (69)</td>
</tr>
<tr>
<td>Depression according to DSM criteria, N (%)</td>
<td>15 (15)</td>
</tr>
<tr>
<td>Depression according to PCI, N (%)</td>
<td>39 (39)</td>
</tr>
<tr>
<td>Overall survival in months, mean (SD)</td>
<td>15.7 (9.2)</td>
</tr>
</tbody>
</table>

Abbreviations: ADL, Activities of Daily Living scale; DSM, Diagnostic and Statistical Manual of Mental Disorders, fourth edition, revised; IADL, Instrumental Activities of Daily Living scale; N, number; PCI, psychiatric clinical interview; PS, performance status; Q1–Q3, first through third quartiles; SD, standard deviation.

Considering the agreement between CPI and other assessments, results are the following:

- no agreement (kappa coefficient between 0.0 and 0.2): with the DT (κ=−0.060), MT (κ=0.193), HADS (κ=−0.051), and OA (κ=−0.061),
- low agreement (kappa coefficient between 0.2 and 0.4): with HDRS (κ=0.389),
- moderate agreement (kappa coefficient between 0.4 and 0.6): with DSM (κ=0.431), with GDS (κ=0.517), and with MADRS (κ=0.494).

The tool giving the best combination of sensitivity/specificity for depression screening among elderly patients was the GDS (cutoff score ≥10), with a sensitivity of 0.94 and a specificity of 0.80 as compared to DSM (Table 3). The DT, MT, and HADS had very poor sensitivity/specificity for depression identification (0.50/0.40, 0.50/0.57, and 0.50/0.67, respectively). Among clinical assessments, a good sensitivity/specificity (1.00/0.71) was obtained with the PCI, whereas OA resulted in a very low sensitivity (0.13).

### Discussion

The main objective of this study was to assess the level of agreement between patient self-reported depression rated on different scales, oncologist clinical impression, and

Table 2 Comparison of depression diagnosis using different strategies between depressed and nondepressed patients according to DSM diagnosis, N=99

<table>
<thead>
<tr>
<th>Variables</th>
<th>Depressed patients</th>
<th>Nondepressed patients</th>
<th>κ coefficient</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT ≥4</td>
<td>50 (60)</td>
<td>8 (53)</td>
<td>−0.060 (−0.189; 0.069)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MT &gt;4</td>
<td>36 (43)</td>
<td>8 (53)</td>
<td>0.058 (0.096; 0.212)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GDS ≥10</td>
<td>17 (20)</td>
<td>15 (100)</td>
<td>0.508 (0.326; 0.690)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HADS &gt;14</td>
<td>28 (33)</td>
<td>8 (53)</td>
<td>0.127 (−0.051; 0.305)</td>
<td>0.01</td>
</tr>
<tr>
<td>HDRS ≥7</td>
<td>49 (58)</td>
<td>15 (100)</td>
<td>0.151 (0.051; 0.251)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MADRS ≥16</td>
<td>8 (10)</td>
<td>14 (94)</td>
<td>0.663 (0.473; 0.853)</td>
<td>0.109</td>
</tr>
<tr>
<td>PCI</td>
<td>24 (29)</td>
<td>15 (100)</td>
<td>0.431 (0.268; 0.594)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>OA</td>
<td>15 (18)</td>
<td>2 (13)</td>
<td>−0.043 (−0.228; 0.223)</td>
<td>0.851</td>
</tr>
</tbody>
</table>

Note: *McNemar’s test.

Abbreviations: κ, Cohen’s kappa coefficient; DSM, Diagnostic and Statistical Manual of Mental Disorders, fourth edition, revised; DT, Distress Thermometer; GDS, Geriatric Depression Scale; HADS, Hospital Anxiety Depression Scale; HDRS, Hamilton Depression Rating Scale; MADRS, Montgomery Asberg Depression Rating Scale; MT, Mood Thermometer; N, number; OA, oncologist assessment; PCI, psychiatric clinical interview.

### Depression and overall survival

Mean survival was 15.7 months (standard deviation = 9.2 months). We did not find any significant difference regarding patients’ characteristics between depressed and nondepressed patients according to DSM (Table 4). Patients with major depressive disorder according to the DSM tended to have a shorter overall survival than did nondepressed patients, at 14.9 months (3.0–18.3 months) versus 15.2 months (9.4–22.8 months); however, this difference was not statistically significant (P=0.281).
a PCI conducted by a psychologist using the DSM criteria among elderly patients with AOC. We found moderate agreement for depression identification between DSM and GDS (κ=0.508) and PCI (κ=0.431) and high agreement with MADRS (κ=0.663). We found low or no agreement between DSM and the other assessments strategies, including OA (κ≈0.043). One explanation may be the difficulty, according to the DSM classification, to connect somatic symptoms with depression, considering that ovarian cancer may lead to appetite change, fatigue, and libido or sleep disruption.

This result confirms the importance of the National Comprehensive Cancer Network guidelines and the work of Jacobsen regarding the implementation of systematic strategies for distress screening among cancer patients.\(^48\text{–}50\) Indeed, OA results in a false-negative rate of more than 80%, leading to the recommendation that health care professionals without specific training for depression diagnosis should be encouraged to use validated tools in their daily clinical practice.

The screening tool with the best combination of sensitivity/specificity for depression among elderly patients with AOC was the GDS, with a sensitivity of 0.94 and a specificity of 0.80. These results suggest that the GDS is a good tool, as part of a screening “routine” among elderly patients with AOC, and should be preferred to more classic and unspecific tools such as the DT or the HADS. Our results might be partially explained by the fact that most of the scales for depression screening/diagnosis were validated among middle-aged patients.\(^51\)

The systematic use of this type of tool (eg, GDS) has been widely advocated to improve the identification of depression in this frail population.\(^2,20\text{–}52\) However, the vast majority of cancer centers do not have a strategy for the identification of depression in patients with late-stage cancer. Our results suggest that some existing tools are consistent with the DSM diagnosis of depression that is still considered as the gold standard. Nonetheless, our findings suggest that we should distinguish first-line tools (screening) from second-line tools (diagnostic).\(^11,20,50,53\) Indeed, even if these tools allow easy identification of depression, in case of treatment failure and/or severe risk factors, these patients should be referred to a mental health specialist to confirm the diagnosis.\(^54\)

Our group has recently published a report proposing a geriatric vulnerability score in elderly patients with AOC.\(^30\) In this previous work, we used the HADS for the development of the vulnerability score. Our findings suggest that this score is less accurate than the GDS for the identification of depression. More research is necessary on the best way to screen for depression among this specific population, so as to allow simple screening in daily practice. It is important to recognize that in this population of patients with severe physical and/or psychosocial distress, only very brief assessments are likely to be adopted in clinical practice.

We did not find significant difference in overall survival between depressed patients and nondepressed patients (Table 4, Figure S1). These results are in contradiction with our previous results, which reported a significant impact of OA and HADS on survival.\(^45,55\) These results were recently confirmed in a pooled analysis of EWOT1, EWOT2, and EWOT3 (submitted) data, leading to the idea that emotional disorders in general, including cancer distress, might have a more significant impact on survival than authentic depression as defined by DSM. Such results are in line with the current debate on HADS performance, considered as poor in depression identification\(^56,57\) but powerful in predicting various medical outcomes.\(^58\) More research is necessary to better characterize the association between both depression and survival among elderly cancer patients.

The main limitation of this study was that this analysis was a secondary endpoint of the GINECO EWOT3 Phase II multicenter trial, designed to assess the impact of geriatric covariates, notably depression, on survival in elderly patients with AOC. As a result, the sample size was not calculated.
regarding sensitivity and specificity of the different tools as compared to DSM, limiting generalization of our results.

Another limitation is that all the threshold scores used for this study come from existing literature and it is known that sensitivity and specificity of the instrument depends on cutoff points, which might need to be changed in different patient populations (eg, inpatients, outpatients, and elderly patients, or by stage and type of disease).

This is probably true for all these tools, and future research should try to explore this issue using the same tool among different populations. Finally, our results suggest that depression systematic screening using specific tool (eg, GDS) is feasible and efficient, and that these tools should be selected according to the population and the clinical setting. More research is also necessary to better understand what should be done in case of positive screening for depression. Research that includes prospective longitudinal studies will help to characterize depression and develop optimal treatment strategies.

Acknowledgments

The authors acknowledge Douglas Micheau-Bonnier, Nicolas Gane, and Bénédicte Votan from the GINECO study office, the Délégation à la Recherche Clinique et à l’Innovation of the Hospices Civils de Lyon and Raymonde Maraval-Gaget. We also thank the following investigators and psychologists who participated in the trial: Drs B Weber, MC Kaminsky, and E Luporsi (Centre Alexiis Vautrin-Brabois, Vandoeuvre-lès-Nancy), Prof H Curé, Drs AM Savoye and G Yazbek (Institut Jean Godinot, Reims), Drs I Ray-Coquard, JP Guastalla, and O Tredan (Centre Léon Bérard, Lyon), Dr E Sevin (Centre François Baclesse, Caen), Drs L Stefani and J Provençal (Centre Hospitalier de la région d’Annecy, Pringy), Drs S Kalla, F Savinelli, and G Deplanque (Groupe Hospitalier Saint-Joseph, Paris), Drs M Combe and M Atlassi (Centre Hospitalier du Mans, Le Mans), Dr J Salvat (Hôpitaux du Léman, Thonon-les-Bains), Prof E Pujade-Lauraine and J Alexandre, and Drs L Chauvet and JM Tigaud (Hôpital Hôtel-Dieu, Paris), Dr J Meunier and D Bosquet (Centre Hospitalier Régional d’Orléans, Orléans), Dr J Cretin and E Roux (Clinique Bonnefon, Alès), Dr M Fabbro and P Champoiral (Institut du Cancer de Montpellier, Montpellier), Dr MN Certain and M Gillon (Centre Hospitalier d’Aix-en-Provence, Aix-en-Provence), Drs E Legouffe and MA Armand (Clinique de Valdegour, Nîmes), Mrs I Gourmoud (Hôpital d’Instruction des Armées Sainte-Anne, Toulon), Drs C Ligeza-Poisson and P Deguiral, A Harel (Clinique Mutualiste de l’Estuaire, Saint-Nazaire), Dr C Tep and K Pietrain (Centre Hospitalier Départemental Les Oudairies, La Roche-sur-Yon), Dr F Rousseau and P Bensoussan (Institut Paoli Calmettes, Marseille), N Darchy (Hôpital Perpétuel Secours, Levallois-Perret), and G Marini (Centre Azuréen de Cancérologie, Mougins).

This work was supported by research grants from the French Ministry of Health (Programme Hospitalier de Recherche Clinique 2006, 27–39) and La Fondation de France (grant #200601589).

This work was presented as a poster at the 48th Annual Meeting of the American Society of Clinical Oncology, June 1–5, 2012.

Disclosure

The authors report no conflicts of interest in this work.

References


Supplementary material

Figure S1 Kaplan–Meier estimates of overall survival according to the presence or the absence of depression using DSM criteria.

Abbreviation: DSM, Diagnostic and Statistical Manual of Mental Disorders, fourth edition, revised.