Impact of severity and type of depression on quality of life in cases identified in the community

Authors: Roberto Nuevo\textsuperscript{1,2}, Caroline Leighton\textsuperscript{1}, Graham Dunn\textsuperscript{3}, Christopher Dowrick\textsuperscript{4}, Ville Lehtinen\textsuperscript{5}, Odd Steffen Dalgard\textsuperscript{6}, José Luis Vázquez-Barquero\textsuperscript{7}, Patricia Casey\textsuperscript{8}, José Luis Ayuso-Mateos\textsuperscript{1,2,*}

\textsuperscript{1}Department of Psychiatry, Hospital Universitario de la Princesa, Madrid, Spain

\textsuperscript{2}Instituto de Salud Carlos III, Centro de Investigación en Red de Salud Mental, CIBERSAM, Spain

\textsuperscript{3}Biostatistics group, Division of Epidemiology & Health Sciences, The Medical School, University of Manchester, UK

\textsuperscript{4}Division of Primary Care, University of Liverpool, UK

\textsuperscript{5}National Research and Development Centre for Welfare and Health, Turku, Finland

\textsuperscript{6}National Institute of Public Health, Oslo, Norway

\textsuperscript{7}Department of Psychiatry, Universidad de Cantabria, Hospital Universitario Marqués de Valdecilla, Santander, Spain

\textsuperscript{8}Department of Psychiatry, University College Dublin, Mater Misericordiae Hospital, Dublin, Ireland

* Corresponding author: J.L. Ayuso-Mateos, Hospital Universitario de La Princesa, Servicio de Psiquiatría. C/ Diego de León 62, 28006 Madrid, Spain. Telephone: +34-91-5202366; Fax: +34-91-5202284; e-mail: joseluis.ayuso@uam.es
Abstract:
The present work aimed to test the relationship between depression severity and several dimensions of functioning as assessed by the SF-36 at the population level. A sample of 551 participants from the second phase of the ODIN study (228 controls without depression, and 313 persons endorsing ICD criteria for major depression) were selected for a further assessment of several variables, including quality of life (QoL) related to physical and mental health as measured with the SF-36. Statistically significant differences between controls and the major depression group were found in both physical and mental markers of health, regardless of the level of depression severity; however, there were very few differences in QoL between levels of depression as defined by ICD-10. Likewise, there were no differences according to the type of depression (single episode vs. recurrent). These results suggest doubts about the adequacy of the current ICD classification of depression into three levels of severity.

Keywords: major depression; disability; quality of life; severity; SF-36
Introduction

Depressive disorders are an important public health problem because they are highly prevalent and produce worse quality of life (QoL) and important levels of disability (Ayuso-Mateos et al., 2001b; Bijl et al., 1998; Kessler and et al., 1994; Offord and et al., 1996). It has been projected that in the future they will be the second most frequent cause of years lost to disability in the world (World Health Organization, 2001). The impact of depression on patients’ wellbeing and quality of life have been shown to be equal to, or greater than, that of several other major chronic medical conditions (Wells and Sherbourne, 1999; Sprangers et al., 2000; Bonicatto et al., 2001).

Cases of depression are classified into subgroups according to the criteria of the International Classification of Disorders (ICD) to facilitate treatment planning in the clinical situation. The ICD-10 (World Health Organization, 1992) classifies episodes of major depression into two types: “single” and “recurrent”. Each episode is, in turn, classified according to level of severity: “mild”, “moderate”, “severe”, “severe with psychotic features”, and “unspecified severity”. The level of severity is defined according to the ICD-10 by the number of depressive symptoms. The sum of these symptoms, which produces the total score used to determine the level of depression, does not take into consideration the type of symptoms that the patient presents. In other words, according to the current categorical system, patients who present symptoms such as plans to commit suicide, depressed mood for two weeks, increased fatigue and reduced concentration and attention, but who do not present disturbed sleep, diminished appetite, decreased interest and ideas of guilt and unworthiness, would have a “mild” level of depression. On the contrary, patients endorsing eight of the nine symptoms, but without suicidal thoughts, would have a diagnosis of “severe” depression. The only
exception is the presence of psychotic symptoms, which changes the level of severity to “severe with psychotic features” regardless of the number of depressive symptoms.

It is assumed that the more severe the level of the depressive episode, the greater likelihood the episode will be associated with lower QoL, even though little empirical evidence exists to support this. The distribution of QoL in the depressed population has been relatively little studied (Kruijshaar et al, 2003c). Several studies have provided evidence for a continuum in the impact on QoL (e.g., da Silva Lima and de Almeida Fleck, 2007; Ravindran et al, 2002) as the severity of depression increases but they have generally used limited samples given its size the procedure of recruitment, or the diagnostic method, and above all they have not been focused in the distinction of depression severity according to the criteria of the World Health Organization.

It has been suggested that “mild” major depression (MD) could be an unnecessary category, because it is limited in time, the impact on QoL that it produces are low, and it would be diagnosed with low frequency in clinical samples, compared with population samples (Dowrick et al, 1998b; Keller et al, 1995). Hence, the interest of studying whether subjects with mild MD constitute a different sample from individuals without a depressive disorder, according to the quality of life (QoL) that they present.

The aim of this study was to analyze whether the type of MD (single or recurrent) and the different levels of severity (mild, moderate, severe), as defined by ICD-10, correlate with the impact on patients’ QoL. For this study it is assumed that recurrent episodes and higher level of severity of MD would have a higher negative impact on QoL.

**Material and Methods**
Subjects

The sample for this study comprised 551 persons: 238 persons without MD diagnosis, used as control group and 313 persons endorsing a diagnosis of MD. Data were obtained from the Outcome of Depression International Network (ODIN) project. This project is a multicentre epidemiological study designed to provide reliable and valid data on the prevalence, risk factors and outcome of depressive disorders within the European Union, based on an epidemiological sampling framework, as well as to assess the impact of psychological interventions on the outcome of depression. A detailed description of the aims and methods of the study is provided elsewhere (Ayuso-Mateos et al, 2001a; Dowrick et al, 1998a).

The study was undertaken in five countries and nine geographical areas, and included a total sample of 8764 individuals. Each centre identified one rural and one urban setting in which to conduct its research. Community study samples were identified through census registers or lists of patients registered with primary care physicians. The different research teams involved made the choice to use one sampling procedure or the other based on their previous experience in community surveys. The overall response rate for the first phase of the survey was 65%, which included variations by study site, ranging from 54% (rural Ireland) to 90% (urban Spain).

Procedure

A two-phase sampling procedure (Dunn et al, 1999) was adopted in the ODIN study, with phase one consisting of the use of mental health screening instruments and collection of sociodemographic and medical data. This screening phase of the ODIN study included the Beck Depression Inventory (BDI) (Beck et al, 1961), with a cut-off score above 12 (Lasa et al, 2000) and General Health Questionnaire (GHQ-12).
Depression and quality of life
(Goldberg et al, 1997). The second stage included offering all of those scoring at or above the BDI threshold, and a random 5% of responders below threshold, a detailed interview with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) Version 2.0 (World Health Organization, 1994). The SCAN was used to generate diagnoses of depressive disorders on the basis of ICD-10 and DSM-IV categories.

Health related QoL was assessed using the short form of the Medical Outcome Study (MOS) General Health Survey (SF-36) (Ayuso-Mateos et al, 1999;Stewart et al, 1988). The SF-36 has been shown to be a valid and reliable measure of QoL in depressed patients, with a high internal consistency for most subscales (0.7 in the pain dimension to 0.9 in the physical functioning) (Ayuso-Mateos, J. L., Lasa, L., Vazquez-Barquero, J. L., Oviedo, A, and Diez-Manrique, J. F., 1999). The translations of the SF-36 in several different languages have been shown to be culturally appropriate and comparable in their content (Vilagut et al, 2005). In the SF-36 higher scores indicate better QoL.

The cases were the patients with a diagnostic of Major Depression for ICD-10 after SCAN interview. These included single and recurrent depressive episodes (F 32, F 33) that were mild, moderate or severe on the SCAN (n=313), and they completed at least 6 of the 8 subscales of the SF-36 questionnaire. Controls were the individuals that had scores under 12 in the BDI, no diagnostic at the SCAN and completed at least six of the eight subscales on the SF-36 questionnaire (n=238). The severe level included the subjects diagnosed as having severe MD with psychotic features (n=6; 1.1% of the total sample). Final sample sizes for depression severity groups were: Mild (n = 132); Moderate (n = 124); Severe (n = 57).

Statistical analysis
First, statistical comparisons (chi-square for categorical variables and one-way analyses of variance for continuous) were performed in order to test differences across groups in the main characteristics (age, gender, distribution by country, marital status, rural vs. urban setting, presence vs. absence of long term illness or disability, and number of days which has been sick during the last six months) of the groups of depression severity. Then, a series of ANOVAs were performed to study the differences between levels of severity in each of the eight subscales on the SF-36. Post-hoc analyses (Scheffé when homoscedasticity could be assumed, Games-Howell when it could not) were performed to test for differences between specific groups. Effect sizes (Hedges’s \( g \)) were calculated for each paired comparison. These analyses (ANOVAs and post-hoc tests) were then repeated, controlling for age, gender, country, marital status, setting, presence of long term illness or disability and number of days sick during the last 6 months, as it has been suggested that these variables affect the level of QoL, and could be confounders of the effect of depression severity (Bijl and Ravelli, 2000; Kruijshaar et al, 2003b; Jorngarden et al, 2006) (Buist-Bouwman et al, 2004). Age and number of days sick were entered as covariate, and the remaining potential confounders were included as independent variables in the model, and main effects of severity of depression were checked.

To test whether the level of severity of depression was associated with a greater number of symptoms, measured by the BDI scores, a univariate ANOVA was performed comparing the four groups of severity with the BDI score as dependent variable. Moreover, to test the interaction of BDI scores and level of depression with extent of QoL, correlation analyses were separately performed for each group of severity between the BDI score and each of the eight SF-36 subscales.
Finally, *t*-tests for independent samples were carried out to compare mean SF-36 scores between persons with recurrent depression and persons with a single episode. All analyses were performed with the statistical package SPSS, release 14.0.

**Results**

The comparison of demographics characteristics across levels of depression severity indicated statistically significant differences in age (F (3,547) = 6.6; *p* < .001), marital status (χ² (9) = 56.3; *p* < .001), distribution across countries (χ² (12) = 127.9; *p* < .001), and setting (χ² (3) = 8.3; *p* = .040); there were no statistically significant differences according to gender (χ² (3) = 2.2; *p* = .539). Concretely, the depressed groups were older (mean of age ranging from 43.9 (SD = 10.5) for the severe depression group to 44.9 (SD = 11.1) for the moderate depression group) than the control group (mean = 40.1; SD = 12.6), with no differences across levels of severity; a higher proportion of individuals who were separated or divorced was related to higher levels of depression; people with severe or moderate depression were more prevalent in the UK and Ireland and less prevalent in Spain; and there were more people with severe or moderate depression living in an urban setting. Likewise, there were statistically significant differences across severity groups in the percentage of persons with long-term illness or disability (χ² (3) = 91.3; *p* < .001), with higher percentages for mild and moderate depression groups (30.2 % and 34.6 % respectively) and lower for control (15.9%) and severe depression (19.2 %) groups. There were statistically significant differences across severity groups in the number of days sick during the last six months (F(3,288) = 11.0; *p* < .001). According to the post-hoc analyses, the control group reported less days sick (mean = 1.46; SD = .84) than the depression groups, without differences between them.
Depression and quality of life

The one-way ANOVAs with depression severity as independent variable and each of the eight subscales of the SF-36 as dependent variable resulted in significant differences (p < .001) in all cases (see Table 1), with effect sizes (eta-squared) ranging between .07 and .26 for the physical health variables and between .36 and .58 for the mental health variables. Post-hoc analyses (with p set to .00625 (.05 / 8) controlling for type I error associated with multiple comparisons) indicated that the control group had statistically significantly higher scores (that is, better QoL related with health status) on all of the eight subscales compared with the rest of the groups, with large effect sizes (for instance, ranging between .94 and 2.57 for the paired comparisons between controls and mild depression groups)\(^1\); with no differences between the three levels of severity in people with depression, except for the subscale of mental health between mild and moderate depression, with a lower score for the moderate depression group (p < .001; g = .53, 95% CI: .60 / 1.26), and significantly lower scores for the severe group compared with the mild group in emotional role limitations (p < .001; g = .79; 95% CI: .46 / 1.11), social functioning (p < .001; g = .64, 95% CI: .32 / 96), vitality/energy (p < .001; g = .79, 95% CI: .47 / 1.11), and mental health (p < .001; g = .95, 95% CI: .62 / 1.28). There were not statistically significant differences between moderate and severe depression groups.

Allowance for confounders in the Analyses of Variance led to qualitatively similar conclusions. When these analyses were carried out introducing all the confounders into the model, the main effect of severity remained significant (p < .001) for the eight SF-36 subscales, and the same pattern of specific differences between groups in post-hoc comparisons was found.

\(^1\) Specific results of each paired comparison available upon requests to the corresponding author.
ANOVA comparing BDI scores between groups of depression severity indicated a significant effect (F(3,546) = 450.6, p < .001). Post-hoc analyses showed that the control group had a significantly (p < .001) lower score (BDI mean 3.62, SD = 3.4) than the remaining groups; the mild MD group (BDI mean = 21.4, SD = 7.7) had significantly lower scores than the severe group (p < .001), but without differences compared with the moderate MD (p = .365; BDI mean = 22.8; SD = 7.3); and the severe group (BDI mean = 29.3; SD = 2.9) had significantly higher scores than the rest of the groups (p < .001).

Finally, comparisons of means according to the type of depression (single MD episode vs. recurrent MD) indicated that there were no statistically significant differences between these two groups for any of the eight SF-36 subscale scores (see Table 2). The distribution of cases of mild, moderate or severe MD did not differ significantly between single- and recurrent-episode MD.

**Discussion**

The main goal of the present study was to analyze the relationship between depression severity level and the QoL related with health status. It seems clear with the present results that depression is negatively affecting them, although it is not a very marked effect. There were practically no differences between each step from one to the next level of severity of depression, and the main differences appeared at the step from normal mood state to depression. At this step, the QoL deteriorates, and from there mostly remains, regardless of the level of severity according to ICD-10 classification of depression or whether there is a single episode or recurrent depressive episodes.
Depression and quality of life

In the sampled studied, the classification of depression into three levels of severity proposed by the ICD-10 is not clearly validated in terms of its impact on QoL that they provoke. The current system for distinguishing among mild, moderate or severe depression is categorical, involving a simple addition of symptoms, with no consideration of the weight according to its importance or intensity. It could be that different symptoms have different weight in the impact on QoL of depressive subjects, and the inclusion of additional symptoms to those with the highest weights is not indicating an increase in severity of depression.

Recurrent MD was not found to be associated with worse QoL compared with a single episode of MD. This lack of difference in degree of QoL between recurrent and single episodes could be explained, hypothetically, by the fact that individuals with recurrent MD adapt to their situation, and therefore they give more positive answers to the SF-36, than subjects in their first episode. Another hypothesis would be that individuals with recurrent depression “know” their disorder, and seek treatment earlier (Kruijshaar et al, 2003a; Merikangas et al, 1994; Paykel, 2003) or that their healthcare professionals “know” their disorder better and therefore treat them earlier and more successfully, so their depressive episodes are shorter, which could increase their QoL, diminishing differences with single MD. The present sample is from a community setting, and could present differences with a clinical setting, where the observation that recurrent depression is more disabling than a single depressive episode was originally made (O'Leary et al, 2000).

The comparisons between the categories of the depressive episode and severity in the BDI scores, also indicated that the threshold between mild and moderate depression could be arbitrary: the control group had lower scores than the rest of
Depression and quality of life

groups, severe group had higher scores, but there were no differences between mild and moderate.

**Validation of the current classifications of depressive episodes**

The current model for classification of mood disorders takes the depressive episode as a categorical entity; the model is unitary, with disorders differentiated largely on the basis of severity (Kessing, 2007). The way ICD-10 handles severity has been praised, and recommended for preservation in future editions (Paykel, 2002).

There have been very few studies that validate this classification of severity levels of depressive episodes with other clinical correlates, such as biological markers, prognosis, outcomes, and treatment responses. This classification has only been validated by Kessing (Kessing, 2004c), in a sample comprising all psychiatric inpatients in Denmark discharged between 1994 and 1999 with a diagnosis of a single depressive episode, testing its power to predict risk of relapse and suicide—with the result that patients with increasing levels of severity were significantly associated with a higher risk. Hämäläinen J, (Hamalainen et al, 2004) in a population study in Finland, showed that more severe depressive episodes are associated with a predictably larger use of health services.

The current classification has been questioned by some researchers. Parker argues (Parker G, 2005b;Parker G, 2005a;Parker, 2006) that the ICD-10 describes major depression as a single condition varying by severity, and that this model has not generated any replicable biological changes or correlates at a satisfactory level, and has not been informative in identifying treatment-specific effects.

In our study, the subjects with mild depressive episodes presented a significant difference in QoL compared with subjects without depression. This difference is greater
Depression and quality of life

than the one seen between the different severity groups. Moreover, in follow-up studies, it has been detected that mild MD patients are five times more likely to have major depression after one year (Kessing, 2004b; Lyness et al, 2006). Mild MD may occur independently of a lifetime history of a major depressive disorder (Rapaport et al, 2002). Our findings support the necessity of having specific treatment guidelines for mild depressive episodes. The only guidelines that make specific recommendations for mild depressive episodes are the National Institute for Health and Clinical Excellence (NICE) Guidelines (National Collaborating Centre for Mental Health, 2007) for treatment of depression in primary and secondary care settings, which are an example of a union of categorical and dimensional approaches. They use a stepped care approach that recognizes the different treatment requirements for different levels of severity of depression as defined in the ICD-10.

Our findings show that in terms of QoL, the ICD-10 classification of depressive episodes into three levels (mild, moderate and severe) is only validated for the two extremes. We found significant differences between controls and mild depression, scant differences between mild and severe, and almost no difference between mild and moderate, and moderate and severe episodes of depression.

The classification system used to diagnose depressive disorders is very important to determinate the epidemiology of the disorders and planning of treatment management strategies in health resourcing—and if the classification does not have a proper validation, it is doomed to failure. Current ICD-10 classification of depression severity has only been validated in clinical samples, including inpatients (Kessing, 2004a).

Limitations
Depression and quality of life

Our sample comes from a multicentre study across five countries, so some results could have been explained by country differences rather than by severity or type of depression; this is why we controlled by country on all the analyses, together with potential effects on demographic characteristics across samples.

The number of subjects with “severe MD with psychotic features” (n=6) was so small that we were unable to analyse them as a different group and test the hypothesis that the presence of psychotic symptoms could add an aggravating factor to the depressive episode. This is the only weighted symptom in the current ICD classification, when present automatically the level of depression is severe.

Conclusions

In conclusion, the results of the present work support the relationship between level of depression and worse QoL, but provide indirect evidence against the three thresholds suggested by the ICD-10 for differentiating among different levels of depression: only the threshold between “normal” and “depression” is clearly established, and among levels of depression, only the extreme poles appeared clearly distinguished in terms of QoL.

We feel that in the re-emergence of this nosology, an expert consensus is an important element for establishing diagnosis guidelines; however, these need to be empirically validated before introducing major changes, which are possibly needed, into the current classification system.
Acknowledgments

This study was partially supported by the Instituto de Salud Carlos III, Centro de Investigación en Red de Salud Mental, CIBER-SAM Spain.
Depression and quality of life

Table 1. Comparison of SF-36 mean scores according to the severity of depressive symptoms

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>F*</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning</td>
<td>92.0 (15.9)</td>
<td>73.6 (25.2)</td>
<td>71.0 (27.7)</td>
<td>67.4 (34.0)</td>
<td>12.3</td>
<td>.066</td>
</tr>
<tr>
<td>Role limitations, physical</td>
<td>89.4 (26.5)</td>
<td>54.4 (40.7)</td>
<td>55.3 (43.8)</td>
<td>51.4 (42.7)</td>
<td>23.5</td>
<td>.119</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>83.2 (21.6)</td>
<td>59.0 (27.9)</td>
<td>54.5 (31.3)</td>
<td>52.0 (31.1)</td>
<td>27.8</td>
<td>.141</td>
</tr>
<tr>
<td>General health perception</td>
<td>82.4 (18.2)</td>
<td>51.8 (23.6)</td>
<td>49.7 (24.1)</td>
<td>47.6 (26.8)</td>
<td>60.2</td>
<td>.257</td>
</tr>
<tr>
<td>Vitality/energy</td>
<td>72.2 (17.0)</td>
<td>32.8 (16.9)</td>
<td>27.2 (18.5)</td>
<td>19.4 (16.9)</td>
<td>151.8</td>
<td>.465</td>
</tr>
<tr>
<td>Social functioning</td>
<td>94.9 (15.6)</td>
<td>56.4 (27.0)</td>
<td>47.8 (29.3)</td>
<td>38.9 (27.4)</td>
<td>104.2</td>
<td>.376</td>
</tr>
<tr>
<td>Role limitations, emotional</td>
<td>92.0 (23.5)</td>
<td>44.0 (39.2)</td>
<td>31.1 (36.4)</td>
<td>15.4 (28.0)</td>
<td>97.6</td>
<td>.358</td>
</tr>
<tr>
<td>Mental health</td>
<td>82.8 (12.8)</td>
<td>46.7 (15.0)</td>
<td>37.9 (17.8)</td>
<td>31.8 (16.9)</td>
<td>235.9</td>
<td>.576</td>
</tr>
</tbody>
</table>

*All Fs had p < .001
Depression and quality of life

Table 2.
Comparisons of SF-36 mean scores between depressed individuals with single or recurrent depression

<table>
<thead>
<tr>
<th>SF-36</th>
<th>Single (N = 227)</th>
<th>Recurrent (N = 76)</th>
<th>p</th>
<th>Hedges’ g (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>70.3 (28.8)</td>
<td>75.3 (24.3)</td>
<td>0.367</td>
<td>0.18 (.44/ -.08)</td>
</tr>
<tr>
<td>Role limitations, physical</td>
<td>53.0 (43.0)</td>
<td>57.3 (40.4)</td>
<td>0.172</td>
<td>0.10 (.36/ -.16)</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>53.9 (29.8)</td>
<td>61.1 (29.3)</td>
<td>0.068</td>
<td>0.24 (.50/ -.02)</td>
</tr>
<tr>
<td>General health perception</td>
<td>48.9 (24.5)</td>
<td>54.6 (24.3)</td>
<td>0.079</td>
<td>0.23 (.49/ -.03)</td>
</tr>
<tr>
<td>Vitality/energy</td>
<td>27.3 (17.7)</td>
<td>30.1 (19.4)</td>
<td>0.244</td>
<td>0.15 (.41/ -.11)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>49.1 (28.2)</td>
<td>50.6 (30.1)</td>
<td>0.702</td>
<td>0.05 (.31/ -.21)</td>
</tr>
<tr>
<td>Role limitations, emotional</td>
<td>33.6 (38.7)</td>
<td>32.0 (34.2)</td>
<td>0.747</td>
<td>-0.04 (-.22/ .30)</td>
</tr>
<tr>
<td>Mental health</td>
<td>40.1 (17.7)</td>
<td>41.2 (18.2)</td>
<td>0.623</td>
<td>0.06 (.32/ -.20)</td>
</tr>
</tbody>
</table>

Numbers between brackets are standard deviations.
Depression and quality of life

Reference List


Depression and quality of life


23. ----- (2003b) Levels of disability in major depression: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). J.Affect.Disord., 77, 53-64.

24. ----- (2003a) Levels of disability in major depression: findings from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). J.Affect.Disord., 77, 53-64.


Depression and quality of life


Depression and quality of life

