The overdiagnosis of depression in non-depressed patients in primary care

Enric Aragone`s a, Josep Lluís Piñol b and Antonio Labad c,d


Background. The underdiagnosis of depression is an important research topic. Nevertheless, overdiagnosis has not been given the importance it deserves by research into the ability of family physicians to diagnose depression correctly.

Objectives. To identify the factors that determine the overdiagnosis of depression by family physicians and to evaluate the clinical significance of this error.

Design. Two-phase cross-sectional study.

Setting. Primary care centres in Tarragona (Spain).

Methods. In the first phase, we screened 906 consecutive patients using Zung’s self-rating depression scale (SDS). In the second phase, all the 209 patients with a positive screening and 97 patients with a negative screening (1 out of 7 randomly) were given the Structured Clinical Interview for DSM-IV Disorders, a series of questionnaires, and the family physician judged whether depression was present. In the 186 patients for whom there were no criteria of major depression or dysthymia, the association of various variables with the physicians’ overdiagnosis of depression was analysed.

Results. The rate of diagnosis of depression in non-depressed patients was 26.5% (95% CI: 19.0–33.9). The factors associated independently with overdiagnosis were the SDS score (OR: 1.05; 95% CI: 1.01–1.10), the Global Assessment of Functioning score (OR: 0.95; 95% CI: 0.90–0.99), previous history of depression (OR: 2.66; 95% CI: 1.12–6.30) and presence of generalized anxiety (OR: 0.42; 95% CI: 0.18–0.97).

Conclusion. Family physicians classify as depressed those patients who do not have the formal signs of depression but who do have antecedents of this disorder or a psychological distress that may be prodromal of future depressive episodes.

Keywords. Depressive disorder, dysthymic disorder, overdiagnosis, primary health care.

Introduction

Depression is the most common and most important mental health problem in primary care.1 However, depression in a large number of people patients attending primary care is not detected by their family physician. The underdiagnosis of depression and its determinants has become a well established topic in research into mental health and primary care. In a previous article,3 we reported that 28% of depressed patients were not detected by their physicians and found that detection was associated with the severity of the depression and the impairment it caused, and the complaint of explicit psychological symptoms. Undetected depressed patients, on the other hand, presented the least serious forms of depression, which had the least impairment.

Underdiagnosis and overdiagnosis are the two sides of a scale: one of the effects of taking measures to increase the sensitivity of family physicians and decrease the number of undetected depressed patients will be to increase the rate of overdiagnosis. This aspect is often not given the value it deserves when analysing the ability of family physicians to detect or diagnose...
depressive disorders in their patients. Nevertheless, it can have important clinical repercussions if an erroneous diagnosis leads to inappropriate therapeutic interventions.

The aim of the present paper is to identify the factors that determine the overdiagnosis of depression and evaluate the clinical significance of this error.

Methods

The study was carried out at 10 Primary Care Centres in the Province of Tarragona, Spain. The centres were chosen to form a heterogeneous sample that is representative of our geographical area.

Twenty-three family physicians took part. Of these, 12 were women and 11 men, and their mean age was 37.4 (between 27 and 50 years old). The mean length of time since they finished their degree was 13 years (between 3 and 26 years) and the mean length of time they had occupied their present post was 3.7 years (between 6 months and 9 years). Seventeen were specialists in family medicine and one was a specialist in internal medicine. Fuller details of the study rationale and design are available elsewhere.4

Study design

We carried out a two-phase cross-sectional study. The first phase consisted of a screening test, and the second phase examined a subsample that was selected in accordance with the screening results to establish the final diagnoses: all the patients with a positive screening, plus 1 out of 7 of the screening negative results chosen at random. The purpose of this over-representation of ‘probable cases’ in the sample for the second phase was to detect the highest number of subjects with depressive disorders. So, from the percentage of ‘probable non-cases’ examined in the second phase, we could correct the estimated populational parameters for the screening.5

Sampling

At each centre and at each consecutive visit, during the study period, the participating family physicians determined the patient’s eligibility to take part in the study by applying the criteria for inclusion (age between 18 and 70 years) and exclusion (language limitation or concurrent illness making it impossible to apply the tests in the study, or psychotic disorder), and asked every eligible patient to participate. If the patient agreed to take part, a research interviewer began the screening and applied the selection rule for the second phase. In the selected patients, the interview in the second phase was conducted immediately afterwards by a second research interviewer.4

The research interviewers were located at the primary care centres during the study period. For the screening phase, they were physicians or nurses skilled in using the recruitment criteria and the screening test. Two interviewers, who were experienced GPs who had received specific training in the structured interview and the scales used, made the second phase assessments.

Measurements

I. First phase (screening). We used Zung’s self-rating depression scale6 (SDS) with a cut-off score of ≥55% that had been validated for our environment.7 The following socio-demographic data were collected: sex, age, marital status (single, married/coupled, divorced/ separated or widowed), education (no studies, primary, lower secondary, upper secondary and university) and social class (I, II, IIIa, IIIb, IV and V of the British Registrar General's Scale).

II. Second phase. We used the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)8 to establish the reference diagnoses for depressive disorders: major depressive episode and dysthymia. We also evaluated the presence of generalized anxiety and panic disorder. The SCID includes the Global Assessment of Functioning (GAF) scale, which measures the effect of mental illness on the general level of psychological, social and labour activity of the patient (Axis V of the DSM-IV). The severity of the somatic comorbidity was measured with the Duke Severity of Illness Checklist (DUSOI)9 and the quality of life related to health was measured with the SF-12 Health Questionnaire.10 This instrument gives two scores as follows: a component of physical health and a component of mental health and it includes an item where the patients assess their own health.

The patients were questioned about the reasons for their consultation and their symptoms, which were classified as somatic or psychological/social (chapters P and Z of the International Classification of Primary Care11). They were asked about the number of primary care visits and the number of days on which they had to restrict their activity because of health problems in the previous 3 months.

The patients’ doctors filled in a questionnaire in which they were asked literally to judge whether ‘the patient was suffering currently from a clinically significant depressive status’. The doctors were unaware of the result of the screening and the psychiatric interview and they had to base their judgement on the content of the current consultation, on the patient’s clinical history and on previous knowledge they had of the patient. The response options were as follows: ‘no’, ‘possibly no’, ‘possibly yes’, ‘yes’ and ‘don’t know’; in this article, we group the responses ‘yes’ and ‘possibly yes’ as ‘identified as depressed’ and the other options as ‘not identified as depressed’.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Confidence</th>
<th>Reference</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>The study was carried out at 10 Primary Care Centres in the Province of Tarragona, Spain.</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twenty-three family physicians took part.</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The first phase consisted of a screening test, and the second phase examined a subsample that was selected in accordance with the screening results to establish the final diagnoses: all the patients with a positive screening, plus 1 out of 7 of the screening negative results chosen at random.</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We used Zung’s self-rating depression scale6 (SDS) with a cut-off score of ≥55% that had been validated for our environment.7</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socio-demographic data were collected: sex, age, marital status (single, married/coupled, divorced/ separated or widowed), education (no studies, primary, lower secondary, upper secondary and university) and social class (I, II, IIIa, IIIb, IV and V of the British Registrar General's Scale).</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We used Zung’s self-rating depression scale6 (SDS) with a cut-off score of ≥55% that had been validated for our environment.7</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The following socio-demographic data were collected: sex, age, marital status (single, married/coupled, divorced/ separated or widowed), education (no studies, primary, lower secondary, upper secondary and university) and social class (I, II, IIIa, IIIb, IV and V of the British Registrar General's Scale).</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Statistical analysis

To estimate the rate of diagnosis of depression in non-depressed patients that could be extrapolated to the total population of patients attended in the participating primary health care centres, we applied expansion weights to the subjects in the second phase sample. In this way, we reproduced the composition of the sample in the first phase. For each patient this weight was the inverse of the probability of passing to the second phase and was obtained by dividing the number of patients in the first phase in each of the strata defined by the screening (positive or negative) by the number of patients in the same strata evaluated in the second phase.7 We analysed the weighted scores using the EPI-Info 6.04 CSAMPLE programme (Center for Disease Control and Prevention, Atlanta).

We selected the patients who had not been diagnosed with current major depression or dysthymia in the psychiatric interview and we analysed the association of various variables with the identification of depression in this group of patients. In the bivariate analysis, we used the chi-square test for the categorical variables and the Mann–Whitney U-test for the continuous variables. The level of statistical significance required was \( P < 0.05 \). To determine which factors were independently linked to the overestimation of depression we carried out multivariate analyses using non-conditional logistic regression with SPSS for Windows (SPSS Inc., Chicago). Population parameters could not be derived from these analyses, so unweighted data were used.

Results

Of the 1050 patients who complied with the prior conditions of eligibility, 23 were excluded because of concurrent illnesses, 6 because of language limitations and 6 because of psychotic disorders. Of the remaining 1015 patients, 906 (89.3%) agreed to take part in the first phase. Of these, 322 (224 positive and 98 negative in the screening) were selected for the second phase and 306 (95%) actually took part. The demographic characteristics of those screened and those interviewed in the screening) were selected for the second phase.5 We analysed the weighted scores using the EPI-Info 6.04 CSAMPLE programme (Center for Disease Control and Prevention, Atlanta).

We selected the patients who had not been diagnosed with current major depression or dysthymia in the psychiatric interview and we analysed the association of various variables with the identification of depression in this group of patients. In the bivariate analysis, we used the chi-square test for the categorical variables and the Mann–Whitney U-test for the continuous variables. The level of statistical significance required was \( P < 0.05 \). To determine which factors were independently linked to the overestimation of depression we carried out multivariate analyses using non-conditional logistic regression with SPSS for Windows (SPSS Inc., Chicago). Population parameters could not be derived from these analyses, so unweighted data were used.

We compared the characteristics of the non-depressed patients identified as depressed by the family physician (false positives: FP) with the patients who had been identified as not depressed by both the family physician and the SCID (true negatives: TN). Among the FPs there was a higher proportion of women, people separated from their partners, widows and widowers, and less educated patients than among the TNs. Social class did not discriminate between FPs and TNs. The mean age of the FPs was only slightly higher but it was statistically significant (51.5 years old compared to 47.0 years old, \( P = 0.047 \)) (Table 2).

Table 3 shows the results for the clinical variables. A third of the FPs had a history of previous depressive episodes, a percentage that is significantly higher than that of the TN group in which one patient of every 10 had a previous history of depression (\( P < 0.001 \)).

A greater proportion of FPs than TNs (21.3 versus 9.9; \( P = 0.030 \)) complied with the criteria for establishing diagnoses of other subthreshold affective disorders: minor depression or mixed anxiety-depression disorder.12 The manifestation of explicit psychological symptoms and the patients’ worse perception of their own health were significantly associated with the identification of depression in a non-depressed patient.

In non-depressed patients, the presence of a panic disorder or a generalized anxiety disorder was not associated with the family physicians’ diagnosis of depression. The FPs had more depressive symptoms (measured with the SDS), a greater functional repercussion due to mental problems (measured by GAF) and a worse score on the SF-12. DUSOI also noted that the organic comorbidity was more severe in this group.

In the multivariate analysis, the history of previous depression (odds ratio: 2.66; 95% CI: 1.12–6.30) and the depressive symptomatology reflected by a higher SDS score (odds ratio: 1.05; 95% CI: 1.01–1.10) were identified as independent factors that were associated with

---

**Table 1** The 2 × 2 contingency table with the SCID diagnoses (current major depression and/or dysthymia, DSM-IV) and the family physicians’ diagnoses in the second phase subsample of patients

<table>
<thead>
<tr>
<th>SCID</th>
<th>Depressed or not</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major depression and/or dysthymia</td>
<td>86 (104)</td>
</tr>
<tr>
<td>Neither major depression nor dysthymia</td>
<td>34 (48)</td>
</tr>
<tr>
<td>Total</td>
<td>120 (152)</td>
</tr>
</tbody>
</table>

Weighted data are shown between brackets.
a greater probability of diagnosing depression in non-depressed patients. The lower functional impact of mental health problems indicated by an increase in the GAF score (odds ratio: 0.95; 95% CI: 0.90–0.99) and the presence of a generalized anxiety disorder prevented the erroneous diagnosis of depression (odds ratio: 0.42; 95% CI: 0.18–0.97).

Discussion

This study is one of only a few that have investigated the overdiagnosis of depression in primary care. The analysis of the primary care physician’s ability to diagnose depression has often been limited to the analysis of underdiagnosis and ignored the other direction in which errors can be made. The sample of patients that we analysed was taken from all the patients who had all sorts of health problems, and this enables us to obtain information about the characteristics of overdiagnosis in conditions that reflect the daily work of the general doctor. We must be careful, however, when generalizing the results. The primary care centres selected cannot strictly be considered representative of those that were not studied or those from other geographical areas. Neither can we guarantee that the physicians who took part in the study are representative of their colleagues who did not take part and we can by no means discount that there may have been a certain bias in their recruitment towards those who have a greater interest in mental health.
The study may have generated a bias towards overdiagnosis. The family physicians were not aware of the results of the screening nor the psychiatric evaluation that the researchers had made of their patients, but they knew that the patients about whom they had been asked for their judgement had qualified for the second phase of the study in which the proportion of depressed patients was higher than in the initial sample of consecutive visits. What is more, the direct question that they were asked about the possible presence of depression in each patient and the fact that the family physicians were aware that their patients would be submitted to a psychiatric evaluation may have artificially increased sensitivity to this diagnosis and correspondingly decreased the specificity.

A certain amount of overdiagnosis is only to be expected in primary care because the constructs we use for depression come from the psychiatric study of patients, where the prevalence is high and the formulation of a diagnosis has a high positive predictive value (PPV). In primary care, a lower prevalence will cause a lower PPV and more FP. What is more, whereas in psychiatry, depressions are well defined and of a certain degree of severity, in primary care there are numerous less severe forms of depression around or under the diagnostic threshold. This means that diagnostic inaccuracy is likely to be more common at this level of care.

**Main findings**
The agreement between the family physician’s diagnosis of depression and the diagnosis provided by the researcher with the SCID is weak. The kappa index is only 0.299, although this is similar to the index mentioned by Tiemens et al., which was 0.29.

Forty per cent of the patients in the second phase of the study who were not suffering from a depressive disorder were positively identified as depressed by their family doctor, and using weighted data to estimate the overdiagnosis in the population of consecutive patients we obtained a rate of 26.5%. Nevertheless, it would be excessively simplistic to attribute this diagnostic error exclusively to the lack of skills or knowledge of the family doctors. They discriminate between two different groups of non-depressed patients. The patients in the FP group had higher levels of psychological distress of the depressive type, and more functional repercussions, and a greater proportion of them stated that psychological symptoms were the reason for their consultation than the patients in the TN group. Their physical health and perception of their own health were also worse. A third of the FPs had a previous history of depression and a further 20% complied with the criteria for a diagnosis of minor depressive disorders.

These results are consistent with those of Klinkman et al. who found that a previous history of depression, a greater presence of depressive/psychological distress symptoms (measured on the CES-D scale) and a worse perception of one’s own health distinguish between FPs and TNs. Tiemens et al. used a more complex plan of analysis and discriminated between various subgroups of FPs. In general, however, their results agree with ours: they found that a greater proportion of FPs had a psychological reason for their consultation, they had a higher level of psychological distress and associated handicap, and a worse opinion of their own health.

**Clinical implications**
In primary care it is common for patients to present with health problems that are not very well defined and, in fact, a precise diagnosis is perhaps less important than deciding on an appropriate course of action. In this study, the physicians take into account clinical aspects to diagnose depression in a particular group of patients, who were not actually depressed. If patients have a previous history of depression but do not comply with the diagnostic criteria required at the time of the study because of the treatment they have been undergoing or because of the evolution of the illness, they can be diagnosed as depressed. A third of the patients in the FP group have these characteristics. It may be appropriate to maintain the diagnosis of depression for these partially or totally recovered patients: patients in remission from a past depressive episode must undergo a clinical follow-up to monitor how they respond and whether they follow the treatment, and to detect the occasional relapse early.

Other patients in the FP group suffer from moderately intense psychological distress. Some of them can be diagnosed as having minor depressive disorders, which may be depressive prodromal states requiring an individualized therapeutic approach or a control over time to detect any eventual worsening in their condition.

This study evaluates the diagnoses that physicians make of depression and we do not have any data on the therapeutic decisions derived from them. Recently, Kendrick et al. made an observational study and mention a lower rate of overdiagnosis (12%). They found that the probability of prescribing antidepressives was associated to the severity of the depression, although almost half of the patients who were prescribed antidepressives were not depressed.

Although we may be able to find plausible justifications for family physicians’ overdiagnosing depression, the other side of the coin is that the erroneous diagnosis of depression in patients with a slight psychological malaise and little functional repercussion leads to the risk of unnecessary and potentially dangerous medicalization.
Acknowledgements

This study has been possible thanks to the generous collaboration of doctors at the Primary Care Centres in Reus: Sant Pere I and II and Riera Miró; the Primary Care Centres in Tarragona: Sant Pere i Sant Pau, Torreforta and Bonavista, and the Primary Care Centres of Salou, El Morell, Alcover and Constantí, who allowed us to access their consultations and to whom we owe the high level of participation of their patients.

Declaration

Funding: the study was supported by grants of the Catalan Society for Family and Community Medicine, the Jordi Gol i Gurina Foundation, the Spanish Primary Care Network and the Carlos III Health Institute of the Spanish Ministry for Health and Consumption (FIS Exp. PI202162). Enric Aragonès received a pre-doctoral grant (Jordi Gol i Gurina Foundation).

Ethical approval: the study protocol was reviewed and approved by the Clinical Research Ethics Committee of the Jordi Gol i Gurina Foundation (11 February 2002).

Conflict of interest: none.

References

2 Davidson JRT, Meltzer-Brody SE. The underrecognition and undertreatment of depression: what is the breadth and depth of the problem? J Clin Psychiatry 1999; 60 (Suppl. 7); 4-9.
6 Zung WWK. A self rating scale for depression. Arch Gen Psychiatry 1965; 12; 63–70.
10 Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34; 220–233.