

## Management of depression in UK general practice in relation to scores on depression severity questionnaires: analysis of medical record data

Tony Kendrick, professor of primary medical care,<sup>1</sup> Christopher Dowrick, professor of primary medical care,<sup>2</sup> Anita McBride, research fellow,<sup>1</sup> Amanda Howe, professor of primary care,<sup>3</sup> Pamela Clarke, research assistant,<sup>2</sup> Sue Maisey, research associate,<sup>3</sup> Michael Moore, senior lecturer,<sup>1</sup> Peter W Smith, professor of social statistics<sup>4</sup>

<sup>1</sup>University of Southampton Primary Medical Care Group, Aldermoor Health Centre, Southampton SO16 5ST

<sup>2</sup>University of Liverpool School of Population, Community and Behavioural Sciences, University of Liverpool, Liverpool L69 3GB

<sup>3</sup>University of East Anglia School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ

<sup>4</sup>Southampton Statistical Sciences Research Institute, University of Southampton, Southampton SO17 1BJ

Correspondence to: T Kendrick  
A.R.Kendrick@Southampton.ac.uk

Cite this as: *BMJ* 2009;338:b750  
doi:10.1136/bmj.b750

### ABSTRACT

**Objective** To determine if general practitioner rates of antidepressant drug prescribing and referrals to specialist services for depression vary in line with patients' scores on depression severity questionnaires.

**Design** Analysis of anonymised medical record data.

**Setting** 38 general practices in three sites—Southampton, Liverpool, and Norfolk.

**Data reviewed** Records for 2294 patients assessed with severity questionnaires for depression between April 2006 and March 2007 inclusive.

**Main outcome measures** Rates of prescribing of antidepressants and referrals to specialist mental health or social services.

**Results** 1658 patients were assessed with the 9 item patient health questionnaire (PHQ-9), 584 with the depression subscale of the hospital anxiety and depression scale (HADS), and 52 with the Beck depression inventory, 2nd edition (BDI-II). Overall, 79.1% of patients assessed with either PHQ-9 or HADS received a prescription for an antidepressant, and 22.8% were referred to specialist services. Prescriptions and referrals were significantly associated with higher severity scores. However, overall rates of treatment and referral were similar for patients assessed with either measure despite the fact that, with PHQ-9, 83.5% of patients were classified as moderately to severely depressed and in need of treatment, whereas only 55.6% of patients were so classified with HADS. Rates of treatment were lower for older patients and for patients with comorbid physical illness (including coronary heart disease and diabetes) despite the fact that screening for depression among such patients is encouraged in the quality and outcomes framework.

**Conclusions** General practitioners do not decide on drug treatment or referral for depression on the basis of questionnaire scores alone, but also take account of other factors such as age and physical illness. The two most widely used severity questionnaires perform inconsistently in practice, suggesting that changing the recommended threshold scores for intervention might

make the measures more valid, more consistent with practitioners' clinical judgment, and more acceptable to practitioners as a way of classifying patients.

### INTRODUCTION

Since April 2006 the United Kingdom general practice contract quality and outcomes framework (QOF) has provided incentives to general practitioners to measure the severity of depression with a validated questionnaire at the start of treatment in all diagnosed cases.<sup>1</sup> The aim is to improve the targeting of treatment, particularly antidepressants, to patients with moderate to severe depression, in line with guidelines.<sup>2,3</sup> The rationale is that doctors' global assessments of depression severity do not agree well with valid and reliable self reported measures of severity in terms of cut-off levels for case identification,<sup>4-7</sup> resulting in overtreatment of mild cases and undertreatment of moderate to severe cases.<sup>7,8</sup>

The three recommended measures of depression severity are the 9 item patient health questionnaire (PHQ-9),<sup>9</sup> the depression subscale of the hospital anxiety and depression scale (HADS),<sup>10,11</sup> and the Beck depression inventory, 2nd edition (BDI-II).<sup>1,12,13</sup> In principle, a higher score on these measures indicates greater severity requiring greater intervention. However, the QOF guidance also recommends that clinicians consider the degree of associated disability, history of depression, and patient preference when assessing the need for treatment rather than relying completely on the questionnaire score.<sup>1</sup>

These measures were designed for slightly varying purposes, and none of them is a "gold standard" measure of depression severity. HADS is a short screening instrument designed to identify patients with a greater probability of depression, who should then be further assessed with a more extended measure or a clinical interview, rather than a measure of severity in itself.<sup>10</sup> BDI-II, on the other hand, was designed as a longer measure of severity,<sup>12</sup> and PHQ-9 was developed as both a screening instrument and a severity

measure.<sup>14</sup> Data on the completion of these measures from the National Health Service Information Centre showed that they were used in a mean of 91% of diagnosed cases across all UK practices in 2007-8, up from 81% in 2006-7.<sup>15</sup> The accuracy and utility of the measures has been questioned, however, suggesting that, even if they use the questionnaires, practitioners may ignore the scores when deciding about treatment or referral.<sup>16</sup>

The aim of this study was to examine the general practice management of patients with depression who completed severity questionnaires, to determine whether the use of the measures was consistent with the rationale for their introduction—specifically whether rates of treatment with antidepressant drugs and referrals for psychological or psychiatric treatment differed in line with patients’ scores on the measures. Other potentially important predictors of intervention—including demographic factors, history of depression, and concurrent physical illness—were also examined to explore whether these factors seemed to influence rates of treatment or referral as well.

The study was conducted in three sites that, between them, served inner city, suburban, and rural areas with varying levels of deprivation, presence of ethnic minorities, and availability of primary care mental health workers and psychological therapies. This allowed us to explore the effect of site on rates of treatment and referral.

Qualitative data on the use of the measures were also collected through interviews with general practitioners and patients and are reported separately.<sup>17</sup>

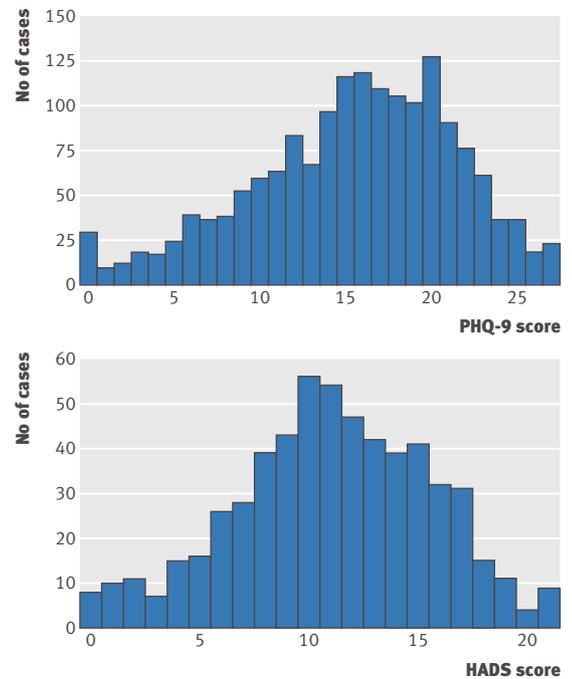
**METHODS**

**Setting**

The study was conducted in three primary care trusts—Southampton City, Liverpool, and Norfolk. All general practices within Southampton City Primary Care Trust were approached. Practices approached in Liverpool were either members of the Mersey Primary Care R&D Consortium (and therefore had an interest in research) or were members of the Matchworks Consortium, an informal association of practices with an interest in mental health. Practices approached in

**Table 1** | Characteristics of general practice patients with depression that was assessed with the patient health questionnaire, 9 item version (PHQ-9), or with the depression subscale of the hospital anxiety and depression scale (HADS). Values are numbers (percentages)

Characteristic	Assessed with PHQ-9 (n=1658)	Assessed with HADS (n=584)
Male	593 (35.8)	208 (35.6)
Age ≥65 years	264 (15.9)	60 (10.3)
History of depression	624 (37.6)	188 (32.2)
Diabetes	147 (8.9)	32 (5.5)
Coronary heart disease	192 (11.6)	30 (5.1)
Other chronic physical illness	749 (45.2)	214 (36.6)
Recruited in Liverpool	739 (44.6)	253 (43.3)
Recruited in Southampton	656 (39.6)	294 (50.3)
Recruited in Norfolk	263 (15.9)	37 (6.3)



**Fig 1** | Distributions of depression severity scores among general practice patients with depression that was assessed with the patient health questionnaire, 9 item version (PHQ-9) (n=1658) or with the depression subscale of the hospital anxiety and depression scale (HADS) (n=584)

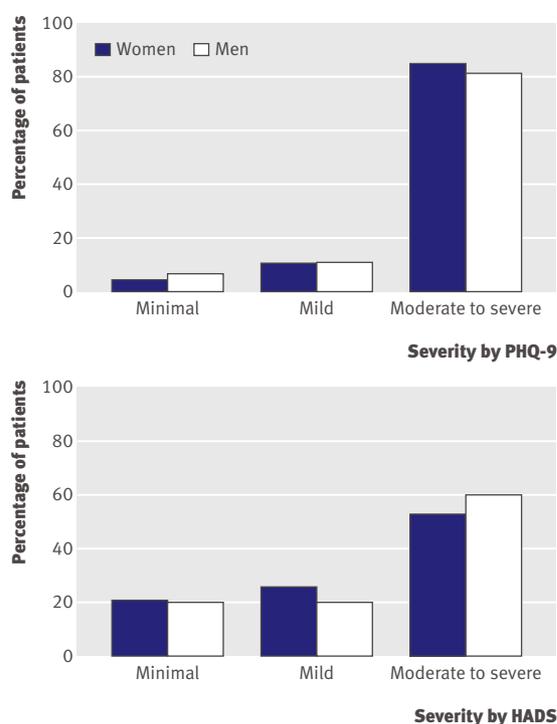
Norfolk were those which had previously indicated an interest in participating in research. Interested practices were asked if they would provide anonymised data on all patients they had assessed with the depression severity measures between April 2006 and March 2007 inclusive.

**Data collection**

With the help of practices’ clerical staff and local primary care trusts’ pharmaceutical advisers, anonymised data were extracted from the computerised medical records of patients for whom a depression severity questionnaire score had been recorded between April 2006 and March 2007. Data extracted included patients’ scores recorded on the questionnaires; age and sex; concurrent physical illness; history of depression; and subsequent management within three months of completion of the questionnaire scores (including follow-up appointments with general practice staff, antidepressant drug treatment, and referrals for psychological, psychiatric, or social services). Data were entered into a Microsoft Excel database and transferred into the statistical programs SPSS, version 15, and Stata, version 10, for analysis.

**Sample size**

We estimated that the incidence of new diagnoses of depression in the year would be around 1%,<sup>2</sup> a mean of around 60 per practice, but assumed conservatively that only about half of the affected patients would be assessed with a severity measure. We aimed to detect a



**Fig 2** | Distribution of depression severity among general practice patients with depression that was assessed with the patient health questionnaire, 9 item version (PHQ-9) (n=1658) or with the depression subscale of the hospital anxiety and depression scale (HADS) (n=584)

15% difference in the proportion of patients treated with antidepressants between those with mild depression and those with moderate to severe depression. On the basis of these assumptions, we aimed to gather data on about 560 patients with mild depression and 280 with moderate to severe depression, from a minimum of 28 group practices providing a mean of 30 patients each: this would provide 90% power at the 5% level of significance to detect the 15% difference, allowing for an intracluster correlation coefficient of 0.05 based on the levels typically found in primary care studies.<sup>18</sup>

## Analysis

Rates of follow-up, treatment, and referrals were analysed to determine associations with the severity scores recorded. We compared patients in three categories—minimal depression, mild depression, and moderate to severe depression—using a  $\chi^2$  test with an adjustment to allow for clustering within practices. The threshold scores for mild depression and for moderate to severe depression were 5 and 10 respectively for the patient health questionnaire (PHQ-9),<sup>9,14</sup> 8 and 11 respectively for the depression subscale of the hospital anxiety and depression scale (HADS),<sup>10</sup> and 14 and 20 respectively for the Beck depression inventory (BDI-II).<sup>13</sup> The effects of the other factors were investigated using logistic regression models which included, when significant, a random effect to allow for clustering within practices. Originally, our plan was to analyse data on all patients together, regardless of which instrument was used to categorise the severity of their depression.

## RESULTS

The numbers of practices agreeing to take part were 15 (54%) of the 28 approached in Liverpool, 13 (34%) of 38 approached in Southampton, and 10 (38%) of 26 approached in Norfolk. Overall, 2294 patients from the 38 practices had a depression severity score recorded in their records, including 1658 for PHQ-9, 584 for HADS, and 52 for BDI-II. The mean number of patients assessed per practice was 60.4, twice as many as we had anticipated. PHQ-9 was used in 32 practices (13 in Liverpool, 10 in Southampton, and 9 in Norfolk) and HADS in 21 (8 in Liverpool, 9 in Southampton, and 4 in Norfolk). Both measures were used in 15 practices. The number assessed with the BDI-II was considered too small for any meaningful analysis. Table 1 shows the characteristics of patients assessed by the other two measures. The sample assessed with HADS included fewer older patients, fewer patients with recurrent depression, and fewer with concurrent physical illness.

**Table 2** | Management of general practice patients with depression by severity category according to the patient health questionnaire, 9 item version, (PHQ-9) or with the depression subscale of the hospital anxiety and depression scale (HADS). Values are numbers (percentages)

Item of management received by patient	Depression severity by PHQ-9				Depression severity by HADS			
	Minimal (n=85)	Mild (n=189)	Moderate to severe (n=1384)	Total (n=1658)	Minimal (n=121)	Mild (n=138)	Moderate to severe (n=325)	Total (n=584)
Follow-up appointment in $\leq 4$ weeks	41 (48.2)	107 (56.6)	917 (65.8)	1058 (63.8)**	77 (63.6)	98 (71.0)	223 (68.6)	398 (68.2)
Prescription for an antidepressant	23 (27.1)	92 (48.7)	1195 (86.7)	1310 (79.0)***	64 (52.9)	108 (78.3)	292 (89.9)	464 (79.5)***
Referral to mental health or social services:								
Counsellor	6 (7.1)	18 (9.5)	197 (14.2)	221 (13.3)*	9 (7.4)	10 (7.2)	48 (14.8)	67 (11.5)
Primary care mental health worker	3 (3.5)	5 (2.6)	84 (6.1)	92 (5.5)	4 (3.3)	4 (2.9)	9 (2.8)	17 (2.9)
Psychology	1 (1.2)	2 (1.1)	22 (1.6)	25 (1.5)	1 (0.8)	6 (4.3)	14 (4.3)	21 (3.6)
Social services	0	1 (0.5)	8 (0.6)	9 (0.5)	0	1 (0.7)	1 (0.3)	2 (0.3)
Psychiatry	2 (2.4)	7 (3.7)	78 (5.6)	87 (5.2)	2 (1.6)	4 (2.9)	14 (4.3)	20 (3.4)

\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  for significance of differences between severity categories ( $\chi^2$  test, with adjustment for clustering by practice).

**Table 3** | Bivariate associations for general practice patients with severity of depression measured with the patient health questionnaire, 9 item version, (PHQ-9)

Item of management received by patient	Odds ratios (95% CI) for bivariate association									
	Severity†		Sex‡	Age§	History of depression¶	Diabetes¶	Coronary heart disease¶	Other chronic physical illness¶	Recruitment centre††	
	Mild	Moderate to severe							Southampton	Norfolk
Follow-up appointment in ≤4 weeks	1.40 (0.84 to 2.35)	2.06 (1.33 to 3.19)**	0.87 (0.70 to 1.08)	0.77 (0.57 to 1.05)	1.01 (0.73 to 1.39)	0.86 (0.57 to 1.29)	0.60 (0.38 to 0.96)*	0.77 (0.70 to 1.09)	2.07 (1.53 to 2.80)***	1.04 (0.69 to 1.56)
Prescription for an antidepressant	2.56 (1.23 to 5.30)*	17.04 (7.97 to 36.43)***	0.73 (0.53 to 1.00)	0.27 (0.20 to 0.38)***	1.74 (1.22 to 2.49)**	0.26 (0.19 to 0.37)***	0.21 (0.14 to 0.33)***	0.47 (0.34 to 0.65)***	5.84 (2.93 to 11.65)***	1.08 (0.64 to 1.84)
Referral to mental health or social services:										
Counsellor	1.32 (0.65 to 2.67)	2.29 (1.23 to 4.25)*	1.23 (1.01 to 1.50)*	0.32 (0.22 to 0.46)***	1.21 (0.84 to 1.78)	0.39 (0.25 to 0.66)**	0.28 (0.15 to 0.52)***	0.70 (0.56 to 0.87)**	0.72 (0.45 to 1.15)	0.64 (0.36 to 1.13)
Primary care mental health worker	1.39 (0.44 to 4.33)	2.19 (0.76 to 6.26)	0.95 (0.70 to 1.30)	0.27 (0.17 to 0.44)***	1.06 (0.64 to 1.77)	0.55 (0.29 to 1.05)	0.09 (0.03 to 0.31)***	0.71 (0.56 to 0.90)**	0.77 (0.45 to 1.31)	0.22 (0.03 to 1.44)
Psychology	0.74 (0.38 to 1.46)	1.77 (0.43 to 7.27)	1.22 (0.88 to 1.70)	0.35 (0.19 to 0.67)**	1.56 (0.93 to 2.60)	0.45 (0.17 to 1.18)	0.25 (0.08 to 0.74)*	0.73 (0.45 to 1.20)	0.36 (0.10 to 1.34)	2.63 (0.73 to 9.43)
Social services	0.90 (0.27 to 2.96)	1.36 (0.22 to 8.34)	1.20 (0.44 to 3.25)	0.46 (0.12 to 1.70)	3.00 (0.82 to 10.92)	0.42 (0.06 to 3.22)	0.31 (0.45 to 2.19)	0.47 (0.23 to 0.93)*	0.22 (0.78 to 0.63)**	0.14 (0.02 to 1.09)
Psychiatry			1.44 (0.47 to 4.43)	0.66 (0.06 to 7.69)	0.83 (0.25 to 2.78)			0.15 (0.01 to 1.81)	7.96 (1.03 to 61.0)*	2.82 (0.23 to 33.92)
	1.60 (0.42 to 6.12)	2.48 (0.64 to 9.67)	2.32 (1.67 to 3.22)***	0.75 (0.36 to 1.60)	1.12 (0.64 to 1.95)	0.11 (0.02 to 0.73)*	1.11 (0.63 to 1.95)	0.85 (0.54 to 1.33)	1.04 (0.62 to 1.76)	0.53 (0.24 to 1.20)

For missing entries the odds ratio was either zero or infinity as a result of a zero cell count.  
 \*P<0.05; \*\*P<0.01; \*\*\*P<0.001 for significance of differences between categories. Clustering by practice taken into account for calculating 95% confidence intervals.  
 Comparisons: †Mild or moderate to severe v minimal. ‡Men v women. §≥65 years v <65 years. ¶Yes v no. ††Southampton or Norfolk v Liverpool.

Figure 1 shows the distribution of scores for the two measures. The distribution was roughly normal for both measures, although the PHQ-9 scores were slightly more positively skewed. The mean (standard deviation) PHQ-9 score was 15.5 (6.0) and the mean HADS score 11.1 (4.6).

Figure 2 shows the proportions of patients in three categories of depression severity as defined by their scores—minimal depression (scores 0-4 with PHQ-9, 0-7 with HADS); mild (scores 5-9 with PHQ-9, 8-10 with HADS); and moderate to severe (≥10 with PHQ-9, ≥11 with HADS). With PHQ-9, 1384 (83.5%) of the 1658 patients assessed were categorised as having moderate to severe depression (for which active intervention is generally recommended), compared with only 325 (55.6%) of the 584 patients assessed with HADS. On finding these differences in the proportions of patients classified as having moderate to severe depression, we decided to change our plan of analysis. Instead of analysing the data altogether regardless of the instrument used, we carried out separate analyses for the two main instruments.

Table 2 shows the proportions of patients, within each category of depression severity according to PHQ-9 and HADS, in receipt of a follow-up appointment within four weeks; a prescription for an antidepressant; and a referral to a counsellor, primary care mental health worker, psychology, social services, or psychiatry. Overall 1774 (79.1%) of 2242 patients assessed with either measure received a prescription for an antidepressant, and 512 (22.8%) were referred to specialist services. Prescriptions for antidepressants were significantly associated with greater severity of

depression for both measures. Follow-up appointments, referrals for counselling, and any referral to specialist services were also significantly associated with greater severity for patients assessed with PHQ-9, but not for those assessed with HADS.

Table 3 shows odds ratios for bivariate associations between items of management and patient characteristics for those assessed with PHQ-9. Follow-up within four weeks, a prescription for antidepressants, and any referral to specialist services were significantly more likely with moderate to severe depression compared with minimal depression. Men were more likely to be referred to psychiatry than women, but referrals and antidepressant prescriptions were significantly less likely for older patients or for patients with diabetes, coronary heart disease, or other chronic physical illness.

Table 4 shows odds ratios for bivariate associations between items of management and patient characteristics for those assessed with HADS. Again, a prescription for antidepressants was significantly more likely with increasing severity of depression, and there were trends towards more referrals with increasing severity and towards fewer antidepressant prescriptions for older patients and for those with concurrent physical illnesses.

Tables 3 and 4 also show that in Southampton practices antidepressant prescriptions were significantly more likely whereas referrals to psychology were significantly less likely, compared with the Liverpool and Norfolk practices. Referrals to primary care mental health workers were significantly more

likely in Norfolk practices, at least among those assessed with HADS.

Table 5 shows the results of logistic regression analyses of factors associated with antidepressant prescriptions or any referral to mental health or social services to account for confounding of the various patient and recruitment centre factors. Three factors remained consistently significant across the two groups of patients. A prescription for an antidepressant was significantly more likely for patients with moderate to severe depression on either measure, while patients aged  $\geq 65$  years were less likely to be referred to specialist services, and patients in Southampton practices were more likely to be prescribed antidepressants than patients at the other two sites.

## DISCUSSION

### Main findings

The distribution of the questionnaire scores we found represents the right hand end of the distribution which is found when all general practice patients are screened.<sup>19</sup> General practitioner diagnosis of depression is associated with greater severity,<sup>4,20</sup> and the distribution we found is consistent with the purpose of the quality and outcomes framework (QOF) indicator, which encourages the use of questionnaires to assess the severity of depression that has already been diagnosed and is being considered for treatment rather than to screen patients for undiagnosed depression.

Around 80% of patients assessed with either the patient health questionnaire (PHQ-9) or the hospital anxiety and depression scale (HADS) received

prescriptions for antidepressants, and around 20% were referred to specialist services. In our logistic regression analysis, a greater severity of depression according to either measure was significantly associated with an increased likelihood of being prescribed an antidepressant, and being referred to specialist services was significantly more likely for patients with the highest scores on HADS. These findings are in line with the rationale for the introduction of the measures. They also accord with previous research showing that treatment of depression in general practice is related to greater severity of symptoms.<sup>21</sup> However, we found other factors to be associated with treatment and referral, and rates of treatment and referral were not consistent for categories of severity when we compared the two measures of depression.

### Strengths and weaknesses of the study

We recruited 38 general practices, 10 more than the 28 we had aimed for, and found twice as many patients per practice had been assessed with the severity questionnaires than we had anticipated, which meant we had a much larger sample overall than we expected. This increased the potential power of the sample. However, this was not a random sample of general practices as practices had to be willing to take part, so we probably recruited practices including doctors with a greater interest in the assessment of depression. They may not be representative of all UK general practices, although they were recruited from a range of locations including inner city areas, relatively affluent suburbs, and rural areas across the three recruitment sites,

**Table 4** | Bivariate associations for general practice patients with severity of depression measured with the depression subscale of the hospital anxiety and depression scale (HADS)

Item of management received by patient	Odds ratios (95% CI) for bivariate association									
	Severity†		Sex‡	Age§	History of depression¶	Diabetes¶	Coronary heart disease¶	Other chronic physical illness¶	Recruitment centre††	
	Mild	Moderate to severe							Southampton	Norfolk
Follow-up appointment in $\leq 4$ weeks	1.40 (0.76 to 2.58)	1.25 (0.84 to 1.86)	1.33 (0.89 to 2.01)	1.01 (0.46 to 2.23)	0.55 (0.36 to 0.84)**	1.21 (0.58 to 2.50)	1.30 (0.43 to 0.96)	0.82 (0.51 to 1.33)	2.57 (1.74 to 3.81)***	1.55 (1.00 to 2.42)
Prescription for an antidepressant	3.21 (1.52 to 6.74)**	7.89 (3.51 to 17.68)**	1.08 (0.50 to 2.34)	0.52 (0.23 to 1.06)	1.26 (0.64 to 2.46)	0.47 (0.18 to 1.20)	0.36 (0.16 to 0.81)*	0.77 (0.40 to 1.48)	2.20 (1.57 to 3.06)**	0.51 (0.37 to 0.70)**
Referral to mental health or social services:										
Counsellor	0.97 (0.21 to 4.52)	2.15 (0.74 to 6.29)	0.63 (0.32 to 1.24)	0.12 (0.01 to 1.30)	1.49 (0.74 to 2.99)	0.50 (0.19 to 1.34)	1.20 (0.20 to 7.06)	1.11 (0.52 to 2.34)	0.54 (0.27 to 1.06)	
Primary care mental health worker	0.87 (0.15 to 5.24)	0.83 (0.19 to 3.56)	0.75 (0.29 to 1.95)	1.17 (0.27 to 5.09)	3.12 (0.45 to 21.71)	2.39 (0.51 to 11.18)	1.16 (0.29 to 4.68)	1.22 (0.43 to 3.38)	0.12 (0.04 to 0.41)**	11.30 (3.70 to 34.35)***
Psychology	5.45 (0.61 to 48.76)	5.40 (0.37 to 78.29)	0.90 (0.68 to 1.20)		0.84 (0.48 to 1.44)			0.40 (0.21 to 0.75)**	3.82 (1.10 to 13.27)*	
Social services:										
Psychiatry	1.78 (0.24 to 13.20)	2.68 (0.77 to 9.35)	0.59 (0.35 to 1.02)	0.45 (0.05 to 3.92)	0.90 (0.54 to 1.50)			0.42 (0.09 to 1.92)	2.06 (0.74 to 5.71)	

For missing entries the odds ratio was either zero or infinity as a result of a zero cell count.

\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  for significance of differences between categories. Clustering by practice taken into account for calculating 95% confidence intervals.

Comparisons: †Mild or moderate to severe v minimal. ‡Men v women. § $\geq 65$  years v  $< 65$  years. ¶Yes v no. ††Southampton or Norfolk v Liverpool.

including a range of levels of deprivation and differing proportions of minority ethnic patients. The somewhat higher response rates among practices in Liverpool and Norfolk were probably because the practices approached there were members of groups interested in research or mental health or both.

Our sample included too few patients assessed with the Beck depression inventory for meaningful analysis of that measure, so we were unable to include it. Originally, we had intended to analyse the total sample together, but when we found a marked difference between the samples assessed with PHQ-9 and with HADS in the proportions of patients classified as having moderate to severe depression, and therefore potentially in need of treatment, we decided we should keep the analyses separate for the two instruments. The considerably larger than expected sample size favoured this separate analysis, but the number assessed with HADS was only a third of the number assessed with PHQ-9, which meant that within the HADS sample we had less power to identify significant associations between scores and treatment and referral, increasing the risk of a type II error.

**Findings in relation to other studies**

Our logistic regression analyses showed that factors other than measures of depression severity were independently associated with treatment or referral. Patients aged ≥65 years were less likely to be referred to specialist services, which is in line with previous research suggesting undertreatment of older people with depression, recently highlighted by Age Concern.<sup>22</sup>

The bivariate associations between older age and lower rates of antidepressant treatment among patients assessed with PHQ-9 seemed to be accounted for by the presence of diabetes or coronary heart disease when these were considered together in the logistic regression. It seems likely that more older patients were administered the PHQ-9, not because they were being considered for treatment in ordinary surgeries, but

because they had been screened for depression routinely as part of practice care for diabetes or coronary heart disease, as screening in those conditions is encouraged in the quality and outcomes framework.<sup>1</sup> If so, it is possible that general practitioners tended to discount depression detected on screening when compared with depression presented to them by patients, perhaps because of an unwillingness to medicalise distress and label patients who were not complaining of depression. However, we cannot be sure this was the reason, and ideally we would need to compare rates of treatment and referral between patients with and without diabetes or coronary heart disease in a sample where all patients had been screened for depression to determine whether this explanation is correct.

Whatever the reason, it is perhaps surprising that patients with diabetes, coronary heart disease, and other physical illness tended to be treated and referred less often, despite the fact that they are known to be at higher risk of depression, and that depression is associated with a worse prognosis for the physical illnesses.<sup>23-25</sup> It seems that encouraging general practitioners to screen for depression in these patients does not lead on to them being more likely to receive treatment than patients without diabetes or coronary heart disease. It is possible that the practitioners were concerned about possible drug side effects affecting the comorbid physical problems. However this would not explain the lower rates of referral for psychological treatments. Another possibility is that patients with comorbid physical conditions and multiple medications may be reluctant to accept either treatment or referral and even more so if they have been detected by screening rather than presenting symptoms themselves.

A history of depression was associated with a greater likelihood of drug treatment, at least among patients assessed with PHQ-9, which suggests that patients with a history may be more willing to consider medication in light of previous experience of the illness, or general

**Table 5 | Logistic regression analyses of patient and centre factors predicting drug treatment and referral for general practice patients with severity of depression measured with the patient health questionnaire, 9 item version, (PHQ-9) or the depression subscale of the hospital anxiety and depression scale (HADS)**

Item of management received by patient	Adjusted odds ratios (95% CI) for bivariate association									
	Severity†		Sex‡	Age§	History of depression¶	Diabetes¶	Coronary heart disease¶	Other chronic physical illness¶	Recruitment centre††	
	Mild	Moderate to severe							Southampton	Norfolk
<b>Patients assessed with PHQ-9††</b>										
Prescription for an antidepressant	1.61 (0.86 to 3.00)	8.75 (4.97 to 15.44)***	0.94 (0.70 to 1.27)	0.79 (0.53 to 1.18)	1.51 (1.11 to 2.06)**	0.57 (0.37 to 0.87)*	0.54 (0.35 to 0.83)**	0.82 (0.60 to 1.11)	5.37 (3.12 to 9.24)***	1.37 (0.81 to 2.30)
Referral to mental health or social services	1.17 (0.53 to 2.58)	1.63 (0.81 to 3.30)	1.39 (1.09 to 1.79)**	0.50 (0.31 to 0.80)**	0.99 (0.77 to 1.27)	0.51 (0.29 to 0.89)*	0.40 (0.23 to 0.71)**	0.82 (0.64 to 1.06)	0.62 (0.39 to 0.97)	0.68 (0.41 to 1.14)
<b>Patients assessed with HADS</b>										
Prescription for an antidepressant	2.87 (1.64 to 5.03)***	6.55 (3.84 to 11.18)***	1.09 (0.68 to 1.73)	1.23 (0.55 to 2.74)	1.42 (0.86 to 2.35)	0.52 (0.22 to 1.24)	0.68 (0.26 to 1.78)	0.93 (0.58 to 1.61)	1.69 (1.06 to 2.69)*	0.54 (0.25 to 1.17)
Referral to mental health or social services	1.39 (0.68 to 2.84)	2.35 (1.26 to 4.37)**	0.60 (0.38 to 0.95)*	0.18 (0.05 to 0.60)**	1.35 (0.85 to 2.15)	0.69 (0.22 to 2.20)	1.98 (0.61 to 6.44)	0.76 (0.47 to 1.25)	0.64 (0.39 to 1.07)	1.60 (0.66 to 3.89)

\*P<0.05; \*\*P<0.01; \*\*\*P<0.001 for significance of differences between categories. Comparisons: †Mild or moderate to severe v minimal. ‡Men v women. §≥65 years v <65 years. ¶Yes v no. ††Southampton or Norfolk v Liverpool. ‡‡Significant random effect included to allow for clustering by practice.

practitioners may be more willing to prescribe for them, or both.

There were differences in treatment and referral rates between recruitment sites, which may be explained by local circumstances. Patients in Southampton practices were more likely to be prescribed antidepressants than patients at the other two sites, which is consistent with prescribing analyses and cost data from the Prescription Pricing Authority<sup>26</sup> showing that Southampton practices were relatively high prescribers of antidepressants. The average daily quantity use of selective serotonin reuptake inhibitors was 2150.15 per 1000 “age, sex and temporary resident originated prescribing units” in Southampton City Primary Care Trust for the second quarter of 2006 compared with 1529.69 for England (personal communication Julie Mulvihill, prescribing support technician, Southampton City). At the time covered by our study data, psychological services and counselling had longer waiting lists in Southampton than in Liverpool and Norfolk, which may explain the differences in treatment rates, as general practitioners probably prescribe antidepressants more often if specialist psychological treatment is less readily available.

#### Meaning of the study and implications for practice

The proportion of patients categorised as having moderate to severe depression was 83.5% for those assessed with PHQ-9 compared with only 55.6% for those assessed with HADS. Despite this large difference in classification of patients, the general practitioners prescribed antidepressants to a similar proportion of patients assessed with either measure, which suggests that the score on the measure was not the defining factor in the decision to prescribe. It is likely that the practitioners usually decided on treatment on the basis of their clinical judgment, taking into account other factors besides the symptom count—

such as the degree of associated impairment, a history of depression, and previous treatment with antidepressants. This explanation is supported by the findings of the qualitative interviews of participating general practitioners which suggested that they usually used their clinical judgment to decide on treatment and were only occasionally influenced by the questionnaire scores.<sup>17</sup> Future research could address this issue by asking practitioners to rate their certainty of diagnosis and need for treatment before administering the instruments.

Research suggests that general practitioners’ categorical assessments of patients often differ from those of questionnaire measures.<sup>47</sup> The question then is which is more accurate in terms of predicting the need for treatment, general practitioners’ clinical judgment or questionnaire measures? Previous research into concomitant administration, to a single sample of patients, of the two measures included in this study found that a greater proportion of the sample was classified as having moderate to severe depression according to PHQ-9 than according to HADS.<sup>27</sup>

It should be emphasised that neither of the two measures is an optimum measure of the severity of depression, and scores above the recommended cut-off values give only an indication that a particular patient is likely to have major depressive disorder. Recent validation studies against more extensive diagnostic assessments have suggested that the accuracy of the measures in predicting major depressive disorder could be improved by using a more conservative cut-off score of 12 rather than 10 with PHQ-9<sup>6,28</sup> and a less conservative cut-off score of 9 rather than 11 with HADS.<sup>6</sup> In our sample 76.1% of patients scored  $\geq 12$  on PHQ-9 and 72.6% of patients scored  $\geq 9$  on HADS, so if these two cut-off values had been used a similar proportion of patients would have been classified as having moderate to severe depression by the two instruments. These proportions would also be more in line with the general practitioners’ rates of treatment, and so changing the recommended cut-off scores in these ways might make the scores more valid, more consistent with practitioners’ clinical judgment, and therefore more acceptable to practitioners as a way of classifying patients.

Both PHQ-9 and HADS were used in some of the participating practices, so even within group practices it seems that doctors may differ in terms of their instrument of choice. This is an argument for retaining the option of using a number of severity measures within the quality and outcomes framework, as long as greater consistency in classifying mild and moderate depression can be achieved.

We thank the participating practices and their staff, and the PCT pharmaceutical advisers for their help with data collection. We also thank Simon Gilbody, Geraldine M Leydon, Robert Peveler, Deborah Sharp, and Andre Tylee for advice on the design, conduct, and interpretation of the study.

**Contributors:** TK and CD devised the idea of the study and designed the methods. TK raised the funding. AMcB, PC, and SM were responsible for implementing the study. PWS led the analysis with TK, with contributions from CD, MM, and AH. TK prepared the first draft of the manuscript, and all

#### WHAT IS ALREADY KNOWN ON THIS TOPIC

Under the quality and outcomes framework, UK general practitioners are rewarded for using validated questionnaire measures of depression severity at the outset of treatment in all diagnosed cases

While general practitioners are using the questionnaires in more than 90% of diagnosed cases, qualitative evidence suggests they doubt the validity of the measures and use their clinical judgment to decide about treatment regardless of patients’ questionnaire scores

#### WHAT THIS STUDY ADDS

Prescriptions for antidepressants and referrals to specialist services were significantly associated with higher depression severity scores, but other factors were independently associated with treatment and referral including patient age and concurrent physical illness and geographical area

The patient health questionnaire classified 83.5% of patients as moderately to severely depressed and in need of treatment, compared with only 55.6% of patients assessed with the hospital anxiety and depression scale, but overall rates of treatment and referral were similar for patients assessed with either measure

Changing the recommended severity thresholds for intervention might make the measures more valid, more consistent with practitioners’ clinical judgment, and therefore more acceptable to practitioners as a way of classifying patients

authors contributed to each section of the final draft of the manuscript. TK is guarantor for this study.

**Funding:** This study was funded by an unrestricted educational grant from Lilly, Lundbeck, Servier, and Wyeth pharmaceuticals. It also received funding from Southampton City Primary Care Trust and the Mental Health Research Network, West, North West, and East Anglia hubs. None of the above bodies had any role in study design; the collection, analysis, and interpretation of data; the writing of the paper; or the decision to submit this paper for publication. The study was sponsored by the University of Liverpool.

**Competing interests:** The study was funded by pharmaceutical companies that manufacture antidepressants. TK has received fees for presenting at educational meetings from Lilly, Lundbeck, Wyeth, and Pfizer pharmaceuticals. TK and CD are members of the mental health expert panel of advisers for the UK general practice contract quality and outcomes framework, which recommended the inclusion of the incentives for using depression severity questionnaires in the contract.

**Ethical approval:** Permission for the study was obtained from Liverpool Paediatric Research Ethics Committee, reference 07/Q1502/23, and approval was obtained from local ethics committees and primary care trust research governance offices at all three sites.

- 1 British Medical Association, NHS Employers. *Revisions to the GMS contract 2006/07. Delivering investment in general practice*. London: Department of Health, 2006.
- 2 National Collaborating Centre for Mental Health. *Depression: management of depression in primary and secondary care. Clinical guideline 23*. London: National Institute for Health and Clinical Excellence, 2004.
- 3 Anderson IM, Nutt DJ, Deakin JFW. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 1993 British Association for Psychopharmacology guidelines. *J Psychopharmacol* 2000;14:3-20.
- 4 Dowrick C. Case or continuum? Analysing general practitioners' ability to detect depression. *Primary Care Psychiatry* 1995;1:255-7.
- 5 Williams JW, Noel PH, Cordes JA, Ramirez G, Pignone M. Is this patient clinically depressed? *JAMA* 2002;287:1160-70.
- 6 Lowe B, Spitzer RL, Grafe K, Kroenke K, Quenter A, Zipfel S, et al. Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord* 2004;78:131-40.
- 7 Kendrick T, King F, Albertella L, Smith PWF. GP treatment decisions for patients with depression: an observational study. *Br J Gen Pract* 2005;55:280-6.
- 8 Kendrick T, Stevens L, Bryant A, Goddard J, Stevens A, Raftery J, et al. Hampshire Depression Project: changes in the process of care and cost consequences. *Br J Gen Pract* 2001;51:911-3.
- 9 Kroenke K, Spitzer RL, Williams J. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606-13.
- 10 Zigmond AS, Snaith RP. The hospital anxiety and depression rating scale. *Acta Psychiatr Scand* 1983;67:361-70.
- 11 Wilkinson M, Barczak P. Psychiatric screening in general practice: comparison of the general health questionnaire and the hospital anxiety depression scale. *J R Coll Gen Pract* 1988;38:311-3.
- 12 Beck AT. *Beck depression inventory-II: manual*. 2nd ed. San Diego: Harcourt Brace, 1996.
- 13 Arnau R, Meagher MW, Norris MP, Bramson R. Psychometric evaluation of the Beck depression inventory-II with primary care medical patients. *Health Psychol* 2001;20:112-9.
- 14 MacArthur Initiative on Depression in Primary Care. Use of the PHQ-9 for treatment selection and monitoring (determining a severity score). 2008. [http://depression-primarycare.org/clinicians/toolkits/materials/forms/phq9/severity\\_scoring/](http://depression-primarycare.org/clinicians/toolkits/materials/forms/phq9/severity_scoring/)
- 15 NHS Information Centre. Quality and outcomes framework database. 2008. [www.ic.nhs.uk/qof](http://www.ic.nhs.uk/qof)
- 16 Jeffries D. Ever been HAD? *Br J Gen Pract* 2006;56:392.
- 17 Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, Clarke P, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ* 2009;338:b663.
- 18 Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004;57:785-94.
- 19 Thompson C, Kinmonth A-L, Stevens L, Peveler RC, Stevens A, Ostler KJ, et al. Effects of a clinical practice guideline and practice based education on detection and outcome of depression in primary care: Hampshire depression project randomised controlled trial. *Lancet* 2000;355:185-91.
- 20 Thompson C, Ostler K, Peveler RC, Baker N, Kinmonth A-L. Dimensional perspective on the recognition of depressive symptoms in primary care. *Br J Psychiatry* 2001;179:317-23.
- 21 Dowrick C, Buchan I. Twelve month outcome of depression in general practice: does detection or disclosure make a difference? *BMJ* 1995;311:1274-6.
- 22 Age Concern. *Undiagnosed, untreated, at risk. The experiences of older people with depression*. London: Age Concern, 2008.
- 23 Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. The prevalence of comorbid depression in adults with diabetes. *Diabetes Care* 2001;24:1069-78.
- 24 Williams JW Jr, Katon W, Lin EH, Nöel PH, Worchel J, Cornell J, et al. The effectiveness of depression care management on diabetes-related outcomes in older people. *Ann Intern Med* 2004;140:1015-24.
- 25 Barth J, Schumacher M, Herrmann-Lingen C. Depression as a risk factor for mortality in patients with coronary heart disease: a meta-analysis. *Psychosom Med* 2004;66:802-13.
- 26 Prescription Pricing Authority. Prescribing analysis and cost (PACT) data. 2007. [www.ppa.org.uk](http://www.ppa.org.uk)
- 27 Cameron IM, Crawford JR, Lawton K, Reid I. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *Br J Gen Pract* 2008;58:32-6.
- 28 Gilbody SM, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ-9 and CORE-OM. *Br J Gen Pract* 2007;57:650-2.