

# Computerised patient-specific guidelines for management of common mental disorders in primary care: a randomised controlled trial

Hollie V Thomas, Glyn Lewis, Margaret Watson, Truda Bell, Ita Lyons, Keith Lloyd, Scott Weich and Deborah Sharp

## SUMMARY

**Background:** A large proportion of people with depression and anxiety go unrecognised by their general practitioner (GP). Case-finding does not appear to be effective on its own.

**Aim:** To compare the effectiveness of case-finding followed by computer-generated patient-specific guidelines with usual care for the management of common mental disorders in primary care.

**Design of study:** Individual patient randomised controlled trial.

**Setting:** Five general practices in Bristol and Cardiff.

**Method:** 762 individuals aged  $\geq 16$  years scoring  $\geq 12$  on the Clinical Interview Schedule Revised were randomised. The experimental intervention required participants to complete a computerised psychosocial assessment that generated a report for the GP including patient-specific treatment recommendations. The control patients were treated as usual with access to locally agreed guidelines.

**Results:** Participants' 12-item General Health Questionnaire (GHQ) score dropped irrespective of treatment allocation. The experimental group had a significantly lower GHQ score at 6 weeks, but not at 6 months. Recovery at 6 months was 3% greater among those receiving the experimental intervention (95% confidence interval [CI] = -4 to 10). Treatment was not significantly associated with quality of life or patient satisfaction.

**Conclusion:** Only small benefits are likely from using case-finding followed by patient-specific guidelines to improve clinical management of common mental disorders in primary care. However, depression and anxiety are important public health problems so the utility of such systems should be further investigated.

**Keywords:** computer-assisted decision making; randomised controlled trial; primary health care; mental disorders.

H V Thomas, BSc, DPhil, lecturer in epidemiology; I Lyons, MB, BS, MRCPsych, former research associate, Department of Psychological Medicine; T Bell, BA, MSc, research associate, Department of Epidemiology, Statistics and Public Health, University of Wales College of Medicine, Cardiff. G Lewis, PhD, FRCPsych, professor of psychiatric epidemiology, Division of Psychiatry; D Sharp, PhD, FRCGP, professor of primary health care, Division of Primary Care, University of Bristol, Bristol. M Watson, BSc(Hons), MSc, PhD, MRPharmS, MRC Fellow, Department of General Practice and Primary Care, University of Aberdeen, Aberdeen. K Lloyd, MD, MSc, MRCPsych, professor of psychiatry, Centre for Health Improvement, Research & Evaluation, University of Wales, Swansea. S Weich, MSc, MD, MRCPsych, professor of psychiatry, Division of Health in the Community, Warwick Medical School, University of Warwick, Warwick.

## Address for correspondence

Glyn Lewis, Professor of Psychiatric Epidemiology, University of Bristol, Cotham House, Cotham Hill, Bristol BS6 6JL.  
E-mail: Glyn.Lewis@bristol.ac.uk

Submitted: 30 January 2004; Editor's response: 23 April 2004; final acceptance: 3 June 2004.

©British Journal of General Practice, 2004, 54, 832-837.

## Introduction

THE common mental disorders of anxiety and depression affect up to a third of all general practice attenders.<sup>1,2</sup> Primary care physicians may only recognise up to 50% of those suffering from these conditions at initial presentation, but a further 25% are likely to be diagnosed during the following 3 years.<sup>3</sup> However, 10-35% of patients may still be unwell and unrecognised at 3 years, leading to prolonged mental ill-health.<sup>3</sup> Improvements in the detection of psychiatric disorders might be achieved using self-administered case-finding questionnaires such as the 12-item, self-administered General Health Questionnaire (GHQ-12), which can be completed and scored easily with a sensitivity of up to 90%.<sup>4</sup> Such questionnaires are not used routinely in clinical practice largely because trials have previously indicated that they are of little benefit on their own.<sup>5</sup> In their systematic review, Pignone *et al* suggest that screening in combination with improved patient management could lead to improvement in clinical outcomes.<sup>6</sup> Based on this review, the US Preventive Services Task Force recommends screening adults for depression in clinical practices that have systems in place to assure accurate diagnosis, effective treatment, and follow-up.<sup>7</sup>

Guidelines for the management of anxiety and depression in primary care are widely available.<sup>8-11</sup> They are most likely to be acted upon if their users participate in their development<sup>12</sup> and if patient-specific reminders regarding treatment are provided at the time of consultation.<sup>13,14</sup> Computer-based clinical decision support systems are capable of combining patient information with treatment guidelines to produce patient-specific prompts.<sup>15</sup> Self-administered computerised assessments of common mental disorders are reliable, unbiased and valid, and can be completed easily by patients within primary care.<sup>16</sup> It is therefore possible to base computerised clinical decision support systems on a standardised assessment of psychiatric disorders, itself administered by computer.

We conducted a randomised controlled trial to evaluate the clinical effectiveness of case-finding followed by feedback of computer-generated patient-specific clinical guidelines to the general practitioner (GP) compared with case-finding and usual care supplemented only by locally-agreed guidelines for the management of symptoms of common mental disorders in primary care.

## Method

### Setting and participants

Five general practices in Bristol and Cardiff participated in the study. Ethical approval was given by the Bro Taf Health Authority and United Bristol Healthcare Trust Local Research

**HOW THIS FITS IN***What do we know?*

Symptoms of depression and anxiety are very common among primary care attenders, but can go unrecognised by general practitioners. This may lead to prolonged ill-health and inappropriate use of resources.

*What does this paper add?*

Case-finding followed by the use of computer-generated patient-specific guidelines was associated with a significantly lower mean General Health Questionnaire (GHQ) score when compared with usual care 6 weeks after randomisation. Given the frequency and public health importance of depression and anxiety, even a small difference in recovery could provide valuable assistance in their management.



Ethics Committees. Participants were consecutive attenders of the selected surgeries, who were aged 16 years and over, had been invited to complete the GHQ-12<sup>4</sup> and scored three or more. The GHQ-12 was completed while waiting to see the GP and scored by the research assistant. The GP then applied the following exclusion criteria: previous diagnosis of psychotic illness, mental handicap or cognitive impairment, language or literacy difficulties, severe or terminal physical illness. Those who remained eligible and gave informed consent were invited by the research assistant to make an appointment to complete the initial assessment using the computer-administered version of the Clinical Interview Schedule Revised (CIS-R).<sup>16,17</sup>

The CIS-R took 10–45 minutes to complete in a consultation room. Sociodemographic and clinical characteristics and quality of life were also measured at this computerised assessment. The CIS-R has been demonstrated to be a reliable and valid measure of psychiatric morbidity in both computer- and interviewer-administered versions.<sup>16,17</sup>

**Assignment and masking**

Individuals scoring  $\geq 12$  on the CIS-R were randomised using a code generated within the computerised assessment that did not allow for any external interference. Allocation was therefore concealed from the researcher and the trial participant. Randomisation was stratified by practice, severity of illness (CIS-R score  $\geq 18$ ), and GP. Due to the nature of the interventions, blinding of GPs was not maintained after the initial concealed random allocation of treatment.

**Protocol**

All participating GPs were provided with clinical practice guidelines based on the ICD10-PhC<sup>10</sup> and invited to contribute to the drafting of these guidelines, which included a list of local voluntary sector and self-help groups.

In addition to the local guidelines, treatment of patients in the intervention group involved feedback to the GP of a computer-generated report providing details of psychiatric symptoms, probable psychiatric diagnosis, social impairment, major life events, likely suicide risk, and patient-specific treatment recommendations. For example, if a patient scored  $>20$  on the CIS-R, antidepressant medication was recommended of a

The results should be used in conjunction with a clinical assessment

Date: XX/XX/XXXX

Subject: X

Somatic symptoms 3\*  
Worry over physical health 1  
Irritability 0  
Poor concentration 2\*  
Fatigue 2\*  
Sleep problems 0  
Depression 3\*  
Depressive ideas 3\*  
Phobias 2\*  
Worry 2\*  
Anxiety 4\*  
Panic 0  
Compulsions 0  
Obsessions 0

Social impairment: things are more difficult but everything gets done

\*Symptoms with scores of 2 or more are probably significant

Total score 22<sup>§</sup>

<sup>§</sup>A total score of 12 or more indicates a clinically significant level of distress

**Probable primary diagnosis is: agoraphobia**

Diagnosis confirmed by: an unreasonably strong fear or avoidance of leaving home, crowded shops, and public transport — usually worse when unaccompanied. The probable secondary diagnosis is: mixed anxiety and depressive disorder. Alcohol misuse score 4: alcohol use unlikely to cause harm. Suicide risk: subject feels hopeless but no suicidal thoughts. Subject has a past psychiatric history.

**Medication**

At this severity antidepressants are likely to be effective: choose a less sedating antidepressant, for example, imipramine, or a selective serotonin reuptake inhibitor (SSRI).

**Psychological management**

Agoraphobia: avoidance of feared situations allows fear to grow stronger. Draw up hierarchy of feared situations. Encourage exposure starting with easiest first. Self-help leaflet.

**Social management**

Mainly worried about family (including spouse or partner). Becomes depressed about own mental health. Problems with grown-up children. Satisfied with current marital relationship. Upset about serious illness or injury that occurred more than 6 months ago.

Figure 1. An example of a computer-generated report provided to the GP as part of the intervention.

type and dose based upon the usual practice of that surgery. If a patient complained of poor sleep, a sedative antidepressant was recommended, or if the patient indicated a desire for support following a recent life event, the phone number of the local self-help group was included on the report given to the GP. An example of a report generated by the computerised version of the CIS-R is presented in Figure 1. The GP received this report on the day of the patient's computer-administered

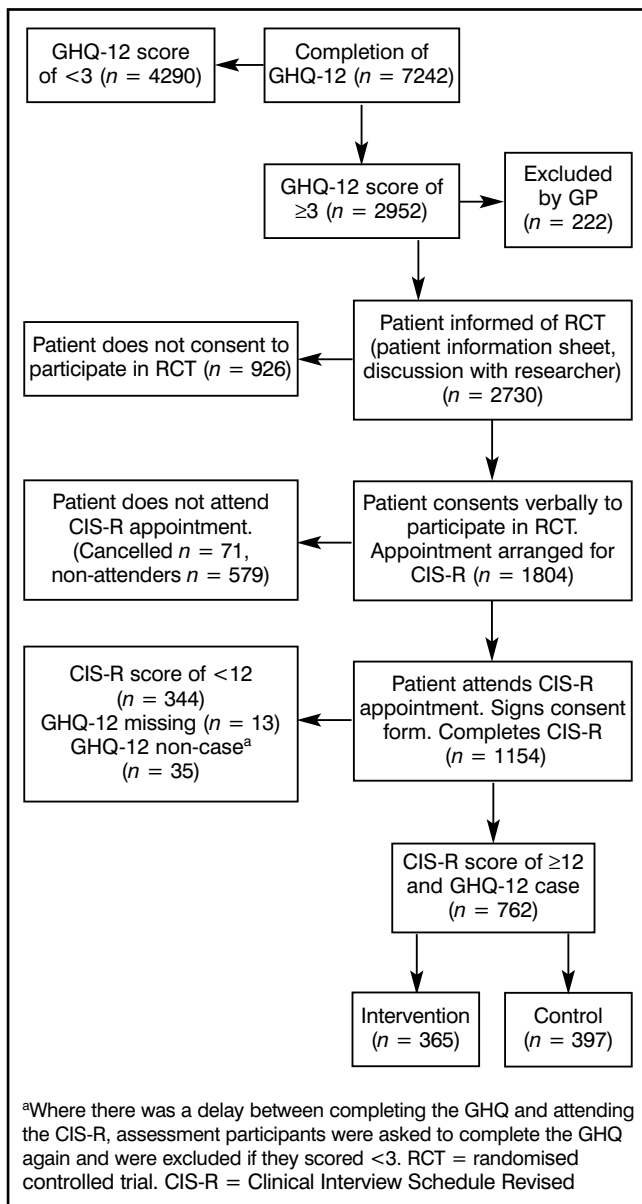


Figure 2. Recruitment and randomisation.

assessment either as a paper copy or directly scanned into the practice's computerised patient notes. The patient was invited to make an appointment with his/her GP within 7 days of the assessment to discuss the results.

### Outcome measures

Outcome was assessed at 6 weeks and 6 months after randomisation via postal questionnaires. Non-responders were reminded by post and telephone. The primary clinical outcome was GHQ-12 score analysed as a continuous variable (Likert method scoring for each of 12 items [0,1,2,3] to provide a total score of 0–36<sup>4</sup>) at 6 weeks after randomisation. A high score represents poor mental health. Recovery from common mental disorder was defined as a score of ≤2 on the GHQ-12 (scored 0,0,1,1) at follow-up,<sup>4</sup> assuming a poor outcome for those with missing data.

Quality of life was measured on a 6-item scale (3 from the Sickness Impact Profile<sup>18</sup> and 3 from the SF-36<sup>19</sup>) resulting in a score of 0–12 with a high score indicating better quality of life. Finally, a single question assessed participants' satisfaction with treatment: 'If a friend were in need of similar help from a GP would you recommend your GP to him/her?'.

### Statistical analysis and sample size

Analysis of variance and covariance was used to study the associations between treatment and GHQ score and quality of life.<sup>20</sup> The associations between treatment and both recovery and patient satisfaction were examined using the  $\chi^2$ -test. A repeated measures analysis of variance investigated the interaction between time and treatment allocation. The main comparative analyses were undertaken only for those subjects with follow-up data. The association between treatment and GHQ score at 6-month follow-up was compared using two methods of managing missing data; either imputations using a regression model based on previous GHQ score, or carrying forward the last known GHQ score. Missing data were imputed using Stata Version 6; all other analyses were performed using SPSS Version 10.

With 750 participants in total (375 in each group) a 10% difference in recovery at 6-week follow-up would be detected between the intervention group and the control group at 80% power and 5% significance (expected prevalence of common mental disorder 30% and 40% respectively). We aimed to recruit 870 individuals to allow for loss at follow-up.

## Results

### Participant flow and follow-up

Full details of the recruitment and randomisation procedure are outlined in Figure 2. Of the 762 randomised participants, 622 (82%) returned questionnaires at 6 weeks and 567 (74%) at 6 months; 527 individuals (69%) returned questionnaires at both follow-up points. The response rate was greater in the control group than in the intervention group at each follow-up point, more notably at the 6-month follow-up (79% versus 70%,  $P = 0.006$ ).

### Characteristics of participants

The characteristics of the experimental intervention and control groups were similar at randomisation (Table 1). Of the 762 randomised participants, 97 (13%) met criteria for an ICD-10 diagnosis of mild to moderate depression only, 74 (10%) met criteria for an ICD-10 anxiety disorder only (generalised anxiety disorder, phobia, or panic disorder), while a further 21 (3%) participants met criteria for both. Overall, 164 participants (22%) reported only depressive symptomatology, 65 (9%) reported only symptoms of anxiety; the majority ( $n = 498$ , 65%) reported both types of symptom.

The average GHQ score at randomisation was similar in the 235 individuals who did not return questionnaires at both follow-up points and the 527 completers (22.1 versus 21.7 respectively,  $P = 0.46$ ). However, non-responders were significantly younger (37.8 years versus 45.2,  $P < 0.001$ ), less likely to be married (50% versus 63%,  $P = 0.003$ ), and less likely to be financially comfortable (9% versus 18%,  $P = 0.006$ ). Non-responders were also less likely at baseline to recommend their GP to a friend (70% versus 79%,  $P = 0.005$ ).

Table 1. Sociodemographic and clinical characteristics at randomisation.

Characteristics	Control n = 397 (SE)	Intervention n = 365 (SE)
Mean age in years	42.4 (0.7)	43.5 (0.8)
Mean CIS-R score	22.9 (0.4)	24.0 (0.5)
Mean GHQ score (Likert scale)	21.6 (0.3)	22.1 (0.3)
Mean quality-of-life score	4.7 (0.1)	4.8 (0.2)
Percentage male	34 (2)	28 (2)
Percentage married/cohabiting	60 (2)	58 (3)
Percentage not in paid employment	45 (2)	44 (3)
Percentage home owners/occupiers	63 (2)	61 (3)
Percentage living comfortably	15 (2)	16 (2)
Percentage car owners	84 (2)	79 (2)
Percentage with long-standing disability/infirmity	66 (2)	61 (3)
Percentage ever prescribed antidepressants	52 (3)	49 (3)

SE = standard error.

### Effect of treatment on GHQ score

On average, all participants experienced a reduction in GHQ score, most notably in the first 6 weeks of follow-up (Table 2). At 6 weeks after randomisation, participants allocated to the computer-generated patient-specific guidelines group had a significantly lower average GHQ score than those receiving usual care (14.8 versus 16.0,  $P = 0.04$ ). This significant effect of treatment on GHQ score was not maintained at 6 months follow-up. Although there was a significant within-subject effect of time ( $F = 239.8$ , degrees of freedom [df] = 2,  $P < 0.001$ ), there was no strong evidence to support an interaction between time and treatment allocation ( $F = 2.57$ , df = 2,  $P = 0.08$ ).

These analyses included participants who had data available at each point of follow-up. Missing GHQ data at 6-month follow-up were estimated using the two methods described in the Methods section; both analyses resulted in almost identical estimates of mean GHQ scores as those presented in Table 2.

### Predictors of outcome and effect modifiers

Financial status (comfort of living, house ownership, car ownership) and previous use of antidepressants were significantly associated with GHQ score at 6 weeks, adjusting for GHQ at baseline. Patient satisfaction at baseline did not significantly affect GHQ score at 6 weeks. However, there was a significant cross-sectional association between

satisfaction at 6 weeks and GHQ score at 6 weeks, adjusting for GHQ at baseline, with satisfied patients having a lower GHQ score (15.1 versus 16.5,  $P = 0.03$ ). Adjusting for any of these potential confounders did little to alter the results presented in Table 2.

Not surprisingly, the strongest predictor of GHQ score at follow-up was the baseline GHQ score, with a significantly greater reduction in GHQ score at 6 weeks seen among those individuals whose baseline score was in the top quarter of the distribution. Furthermore, the baseline GHQ score modified the treatment effect such that the intervention effect appeared stronger in individuals with a low GHQ score (GHQ score of  $< 17$ ; mean GHQ score at 6 weeks was 11.5 in the intervention group versus 13.8 in the control group,  $P = 0.01$ ). Neither diagnosis nor the practice attended by the participant influenced the association between treatment allocation and GHQ score at 6 weeks after randomisation.

### Effect of treatment on recovery

There was not a statistically significant treatment effect on recovery at either follow-up point (Table 3). The 95% confidence interval (CI) around the observed 3% greater recovery rate associated with intervention treatment at 6 weeks ranged from 4% less recovery to 10% greater recovery. Using data for completers only did not alter these conclusions.

### Effect of treatment on quality of life

Baseline quality-of-life scores were available for 745 individuals. Quality-of-life score at follow-up was not significantly associated with treatment allocation (Table 4). On average, all individuals experienced an increase in score over time, indicating an improvement in quality of life. Baseline GHQ score and presence of a long-standing illness, disability, or infirmity were most strongly associated with quality of life at 6 weeks, adjusting for quality-of-life score at baseline. However, adjustments for these variables did not alter the results presented in Table 4.

### Effect of treatment on patient satisfaction

Participants were asked if they would recommend their GP to a friend in need of similar help. The participants had a high degree of satisfaction, with 76% at baseline saying that they would recommend their GP. This level of satisfaction was consistent over the follow-up period (73% and 76% at 6 weeks and 6 months respectively). Treatment group had little effect on satisfaction; for example, at 6 weeks follow-up 72% of participants receiving control treatment and 75% of participants allocated to the experimental intervention were satisfied with their GP ( $P = 0.56$ ).

Table 2. Mean GHQ scores at baseline and at follow-up adjusted for baseline GHQ scores with analysis of covariance.

Treatment	Baseline GHQ score			6-week GHQ score			6-month GHQ score		
	n	Mean	(95% CI)	n	Mean	(95% CI)	n	Mean	(95% CI)
Control	397	21.6	(21.0 to 22.1)	323	16.0	(15.2 to 16.8)	301	14.5	(13.6 to 15.4)
Intervention	365	22.1	(21.5 to 22.8)	287	14.8	(14.0 to 15.6)	244	14.2	(13.2 to 15.2)
					$P = 0.04$			$P = 0.61$	

Analyses based on actual follow-up data. GHQ = general health questionnaire.

Table 3. Effect of treatment on recovery rate (GHQ score of <3) at follow-up.<sup>a</sup>

Treatment	6 weeks Percentage recovered (95% CI)	6 months Percentage recovered (95% CI)
Control	35 (30 to 40)	39 (34 to 44)
Intervention	38 (33 to 43)	35 (30 to 40)
	<i>P</i> = 0.38	<i>P</i> = 0.20

<sup>a</sup>Recovery indicated by GHQ score of <3 at follow-up. Analyses assume poor outcome for patients with missing data. GHQ = general health questionnaire.

## Discussion

### Summary of main findings

Case-finding followed by feedback to GPs of psychiatric assessment and computer-generated patient-specific guidelines was associated with a significantly lower mean GHQ score 6 weeks after randomisation when compared with case-finding followed by locally agreed guidelines under service conditions in primary care. The difference of 1.2 points between groups on the GHQ does not reflect a clinically relevant effect. Although we did not demonstrate a significant treatment effect on recovery from episodes of common mental disorders, we cannot exclude the possibility that up to 10% more (or 4% fewer) patients might recover when patient-specific guidelines are used. Again, although non-significant in this study, there is a suggestion that use of such guidelines may be associated with a faster treatment effect. Given the prevalence and public health importance of depression and anxiety, any potential improvement in speed of treatment response should be investigated further, ideally with the benefit of an economic analysis, before the use of such guidelines is dismissed as a useful adjunct to the treatment of common mental disorders.

### Strengths and limitations

We were aware of the need to screen a very large number of general practice attenders in order to successfully identify sufficient eligible participants. Our greatest problem in recruitment arose because less than half of all eligible attenders agreed to participate in the trial. Although disappointing, this level of participation observed over sustained periods of extensive recruitment is probably a realistic reflection of the limitations inherent to this topic of research and is not unique to this study. This relatively high refusal rate might have been due to reluctance to participate among those with mild symptoms of anxiety and/or depression, who might have attended the practice for somatic complaints.

We did successfully recruit sufficient participants estimated by our *a priori* power calculation but this was based on a 10% difference in prevalence of cases of psychiatric disorders at 6-week follow-up. In fact, we observed a difference of only 3% (95% CI = -4 to 10) and a study powered for this would have needed to be much larger. The possibility of a type 2 error remains.

Among the randomised participants, attrition rates at 6-week and 6-month follow-up were favourable (19% and 26% respectively) in comparison with many randomised controlled trials of antidepressant treatment.<sup>26</sup> Although differential drop-out existed between the randomised groups, it is important to note that completers and non-responders were similar in their baseline GHQ score. Furthermore, we repeated the analyses using two methods of managing missing data, neither of which led to a change in the interpretation of the results.

The design of the trial led to the inclusion of attenders in primary care who might already have been identified by the GP and treated for a psychiatric disorders. This would have tended to reduce any likely therapeutic benefit from the intervention. In addition, we included participants who would not have been recruited into most antidepressant trials, as they had less severe illnesses. For example, only a quarter of our sample met criteria for an ICD-10 diagnosis of mild to moderate depression or anxiety, while 64% scored  $\leq 25$  on the CIS-R at baseline, which is approximately equivalent to a Hamilton Rating Scale for Depression score of <18.<sup>27</sup> Both these factors might have reduced the likely treatment effect, although our subgroup analysis suggested a larger treatment effect in the less severe group.

It should be noted that all trial participants completed the case-finding procedures, and their inclusion in the study would have alerted GPs to the presence of clinically relevant psychiatric morbidity. A greater treatment effect might have been observed in a cluster randomised trial, in which contamination of the treatment process is minimised. However, such a design would have required a much larger study sample.

Finally, as we were limited by the lack of any assessment of whether the intervention actually affected the process of treatment offered by the GP (for example, prescription of antidepressants), it is unclear whether we should even expect any treatment effect to be observed.

### Comparison with existing literature

Previous studies specifically investigating the effect of case-finding followed by feedback and treatment advice on clinical outcome among primary care patients with psychiatric

Table 4. Mean quality-of-life scores at baseline and at follow-up adjusted for baseline scores with analysis of covariance.

Treatment	Baseline QoL score			6-week QoL score			6-month QoL score		
	<i>n</i>	Mean	(95% CI)	<i>n</i>	Mean	(95% CI)	<i>n</i>	Mean	(95% CI)
Control	387	4.7	(4.4 to 4.9)	319	5.8	(5.4 to 6.1)	299	6.2	(5.8 to 6.6)
Intervention	358	4.8	(4.5 to 5.1)	283	5.9	(5.5 to 6.2)	243	6.4	(6.0 to 6.9)
					<i>P</i> = 0.73			<i>P</i> = 0.52	

Analyses based on actual follow-up data. Quality of life based on difficulties arising in social, recreational, and domestic circumstances. A high score indicates better quality of life. QoL = quality of life.

disorders are limited, and have tended to include older patients or patients with more severe symptoms of depression at baseline.<sup>6</sup> The exact interventions, outcome measures, and length of follow-up vary between studies, making interpretation difficult. Thus the effect sizes summarised in the systematic review of Pignone *et al*<sup>6</sup> range from a significant 32% greater recovery at 1-month follow-up among patients receiving intervention to a non-significant 1% greater recovery at 6-month follow-up among the intervention group.<sup>21,22</sup>

A more recent systematic review of management of depression in primary care reports similar difficulties in summarising heterogeneous studies.<sup>23</sup> The authors concluded that those strategies effective in improving patient outcome generally involved more complex interventions that incorporated education, nurse care management, and improved liaison between primary and secondary care. The two trials listed as investigating the effect of computer-based clinical decision support systems for management of more severe depression showed no significant impact on clinical outcome.<sup>24,25</sup> Although patient-specific, the feedback received by GPs in these studies tended to act as a simple reminder of antidepressant dosage and prescription, and number of follow-up visits made.

### Implications for future research

The results from this trial suggest that there may be only small benefits in combining case-finding for psychiatric disorders with individually tailored clinical guidelines generated as a result of a computer-administered assessment. As we still understand relatively little about the most effective treatments for depression and anxiety in primary care settings, we may first need more observational and exploratory studies before devising such systems to assist in the management of clinical dilemmas facing GPs. Future research that aims to investigate the effectiveness of computerised assessment with guidelines would benefit from inclusion of qualitative process measures to help explore GPs' perception of the utility and ease of implementation of such systems, together with perception of barriers to behaviour change. These qualitative measures would provide an understanding of why the intervention was (or was not) associated with any clinical benefit. More relevant patient-specific computerised clinical guidelines might then be devised which, when used within multidisciplinary clinical teams that provide integrated and continuous care, would help improve clinical outcome for a substantial group of patients in primary care.

### References

1. Goldberg D, Huxley P. *Common mental disorders*. London: Routledge, 1992.
2. Kessler D, Lloyd K, Lewis G, Pereira Gray D. Cross sectional study of symptom attribution and recognition of depression and anxiety in primary care. *BMJ* 1999; **318**: 436-439.
3. Kessler D, Bennewith O, Lewis G, Sharp D. Detection of depression and anxiety in primary care: follow up study. *BMJ* 2002; **325**: 1016-1017.
4. Goldberg D, Williams P. *A user's guide to the General Health Questionnaire*. Windsor: NFER-Nelson, 1988.
5. Gilbody S, House AO, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ* 2001; **322**: 406-409.
6. Pignone MP, Gaynes BN, Rushton JL, *et al*. Screening for

depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med* 2002; **136**: 765-776.

7. US Preventive Services Task Force. *Screening for depression: recommendations and rationale*. May 2002. Rockville: Agency for Healthcare Research and Quality. <http://www.ahrq.gov/clinic/3rduspstf/depression/depressrr.htm> (accessed 7 Oct 2004).
8. Paykel ES, Priest RG. Recognition and management of depression in general practice: consensus statement. *BMJ* 1992; **305**: 1198-1202.
9. Effective Health Care. *The treatment of depression in primary care* (Bulletin no 5). Leeds: University of Leeds, 1993.
10. World Health Organisation. Diagnostic and management guidelines for mental disorders in primary care. In: *ICD-10 chapter V primary care version*. Gottingen: Hogrefe and Huber, 1996.
11. Anderson IM, Nutt DJ, Deakin JF. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 1993 British Association for Psychopharmacology guidelines. *J Psychopharmacol* 2000; **14**(1): 3-20.
12. North of England Study of Standards and Performance in General Practice. Medical audit in general practice. I: Effects on doctors' clinical behaviour for common childhood conditions. *BMJ* 1992; **304**: 1480-1484.
13. Haines A, Feder G. Guidance on guidelines. *BMJ* 1992; **305**: 785-786.
14. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993; **342**: 1317-1322.
15. Johnston ME, Langton KB, Haynes RB, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome. A critical appraisal of research. *Ann Intern Med* 1994; **120**: 135-142.
16. Lewis G. Assessing psychiatric disorder with a human interviewer or a computer. *J Epidemiol Community Health* 1994; **48**: 207-210.
17. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardised assessment for use by lay interviewers. *Psychol Med* 1992; **22**: 465-486.
18. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care* 1981; **19**: 787-805.
19. Ware J E J, Sherbourne CD. The MOS 36-Item short-form health survey (SF-36). *Med Care* 1992; **30**: 473-483.
20. Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. *BMC Med Res Methodol* 2001; **1**: 6.
21. Zung WW, King RE. Identification and treatment of masked depression in a general medical practice. *J Clin Psychiatry* 1983; **44**: 365-368.
22. Callahan CM, Hendrie HC, Dittus RS, *et al*. Improving treatment of late life depression in primary care: a randomised clinical trial. *J Am Geriatr Soc* 1994; **42**: 839-846.
23. Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organisational interventions to improve the management of depression in primary care: a systematic review. *JAMA* 2003; **289**: 3145-3151.
24. Simon GE, VonKorff M, Rutter C, Wagner E. Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *BMJ* 2000; **320**: 550-554.
25. Rollman BL, Hanusa BH, Lowe HJ, *et al*. A randomised trial using computerised decision support to improve treatment of major depression in primary care. *J Gen Intern Med* 2002; **17**: 493-503.
26. Hotopf M, Lewis G, Normand C. Putting trials on trial — the costs and consequences of small trials in depression: a systematic review of methodology. *J Epidemiol Community Health* 1997; **51**: 354-358.
27. Hotopf M, Sharp D, Lewis G. What's in a name? A comparison of four psychiatric assessments. *Soc Psychiatry Psychiatr Epidemiol* 1998; **33**: 27-31.

### Acknowledgements

The funding for this project was provided by the Primary Secondary Care Interface initiative of the United Kingdom National Health Service Research and Development Programme (PSI 2-58). We would like to thank the GPs and administrative staff in the practices who participated in this study and the patients who gave their time to the project. The following practices took part: Horfield Health Centre, Bristol; Portishead Health Centre, Portishead; Llanedeyrn Health Centre, Cardiff; North Cardiff Medical Centre, Cardiff; Ashgrove Surgery, Pontypridd. We would like to thank Alison Shaw, Lisa Toy, Vicki Webb, and Liz Hopkin for assistance with data collection, and Nicola Wiles for assistance with data management.