Nicholas Meader, Alex J Mitchell, Carolyn Chew-Graham, David Goldberg, Maria Rizzo, Victoria Bird, David Kessler, Jon Packham, Mark Haddad and Stephen Pilling

Case identification of depression in patients with chronic physical health problems:

a diagnostic accuracy meta-analysis of 113 studies

Abstract

Background
Depression is more likely in patients with chronic physical illness, and is associated with increased rates of disability and mortality. Effective treatment of depression may reduce morbidity and mortality. The use of two stem questions for case finding in diabetes and coronary heart disease is advocated in the Quality and Outcomes Framework, and has become normalised into primary care.

Aim
To define the most effective tool for use in consultations to detect depression in people with chronic physical illness.

Design
Meta-analysis.

Method
The following data sources were searched: CENTRAL, CINAHL, Embase, HMIC, MEDLINE, PsycINFO, Web of Knowledge, from inception to July 2009. Three authors selected studies that examined identification tools and used an interview-based ICD (International Classification of Diseases) or DSM (Diagnostic and statistical Manual of Mental Disorders) diagnosis of depression as reference standard. At least two authors independently extracted study characteristics and outcome data and assessed methodological quality.

Results
A total of 113 studies met the eligibility criteria, providing data on 20 826 participants. It was found that two stem questions, PHQ-9 (Patient Health Questionnaire), the Zung, and GHQ-28 (General Health Questionnaire) were the optimal measures for case identification, but no method was sufficiently accurate to recommend as a definitive case-finding tool. Limitations were the moderate-to-high heterogeneity for most scales and the facts that few studies used ICD diagnoses as the reference standard, and that a variety of methods were used to determine DSM diagnoses.

Conclusion
Assessing both validity and ease of use, the two stem questions are the preferred method. However, clinicians should not rely on the two-questions approach alone, but should be confident to engage in a more detailed clinical assessment of patients who score positively.

Keywords
depression; diagnosis; meta-analysis; primary care

INTRODUCTION

Depression is one of the leading causes of disability and disease burden.[1] It is associated with the most years lost to disability of all diseases worldwide. Identifying depression in patients with chronic physical health problems is important for several reasons. First, a number of studies suggest depression is approximately two to three times as prevalent in such populations, including patients with cancer,[2] chronic heart disease,[3,4] and chronic obstructive pulmonary disease (COPD).[5] Secondly, there appears to be greater disease burden, in terms of healthcare use and functional disability, in people with comorbid depression compared with those with physical health problems alone.[6,7] Thirdly, mortality is greater in several medical conditions when depression is present — heart disease,[8] COPD,[9] stroke,[10] cancer[11] — and in medically ill older adults.[12] Furthermore, morbidity and mortality may diminish with effective treatment of depression.[13,14]

There is convincing evidence that many cases of depression go unrecognised in the general population and in primary care.[15–17] Reasons for under-recognition include a low rate of mood problems as the presenting complaint, infrequent specific enquiry from clinicians, and uncertainty about diagnostic criteria.[18,19] Identifying depression in people with chronic physical health problems may be even more complex, and primary care physicians may be less likely to diagnose depression in this population.[20,21] Reasons for difficulties in raising the issue of depression in consultations are complex.[22] In addition, depressed individuals presenting with somatic complaints are less likely to be detected.[23–26]

Improving case identification for depression has received much attention. For example, the US Preventive Services Task Force recommended screening for depression for all people in primary care (whether they had a physical illness or not), along with the necessary treatment resources for those subsequently identified.[27] In the UK, through the Quality and Outcomes Framework (QOF), GPs are incentivised to ask the case-identification questions of people with diabetes and coronary heart disease.[28] This approach is also advocated in the National Institute for Health and Clinical Excellence (NICE)

N Meader, PhD, systematic reviewer; **M Rizzo**, MSc, research assistant; **V Bird**, BSc, research assistant, National Collaborating Centre for Mental Health, London. **AJ Mitchell**, MD, consultant in liaison psychiatry, Leicestershire Partnership Trust, and honorary senior lecturer in Psycho-oncology, Department of Cancer Studies and Molecular Medicine, University of Leicester, Leicester. **C Chew-Graham**, MD, professor of primary care, University of Manchester, Manchester. **D Goldberg**, DM, FRCP, professor emeritus; **M Haddad**, PhD, RGN, RMN, clinical research fellow, Health Service and Population Research Department, Institute of Psychiatry, King's College, London. **D Kessler**, MD, Walport Clinical Lecturer — primary care, University of Bristol, Bristol. **J Packham**, DM, consultant rheumatologist, Haywood Hospital and Senior Lecturer, Arthritis Research UK Primary Care Centre, Keele University, Keele. **S Pilling**, PhD, joint director, National Collaborating Centre for Mental Health, director of Centre for Outcomes Research and Effectiveness, University College London, and professor of clinical psychology and clinical effectiveness, UCL Research Department of Clinical, Educational and Health Psychology, London.

Address for correspondence
Nicholas Meader, Centre for Reviews and Dissemination, University of York, York, YO10 5DD.
E-mail: nmeader@hotmail.com

Submitted: 19 November 2010; **Editor's response:** 13 January 2011; **final acceptance:** 21 March 2011.

How this fits in

There is strong evidence that the prevalence of depression is raised among patients with long-term conditions and that this comorbidity is associated with adverse outcomes. Inadequate and inaccurate identification of depression has been documented in both primary care and general medical settings. This meta-analysis provides evidence that several brief and feasible depression case-finding approaches can be used as a first assessment for patients with chronic physical health problems, and that two stem questions referring to core depression features appear the most efficient initial approach.

guidelines.[29] However, there is much debate in the literature concerning the effectiveness of screening and case identification.[30] Gilbody and colleagues have shown untargeted screening was not effective in improving the recognition of depression in primary care and general hospital settings.[30] There is also much debate concerning the terminology used in the field. The present study proposes to separate overall accuracy (case identification) into more clinically understandable rule-in and rule-out performance. Rule-in accuracy (positive predictive value) is the ability to correctly identify those with the disorder with minimal false positives, whereas rule-out accuracy (negative predictive value) is the ability to correctly identify those without the disorder with minimal false negatives (missed cases). In order to differentiate from untargeted screening approaches, which appear to be ineffective, this data synthesis will focus on case identification in a population at higher risk of depression (that is, people with chronic physical health problems). This is vital before further case finding is advocated by the QOF for patients with other physical problems.

There are a large number of scales used both in clinical practice and in research studies, few of which have been originally developed for the physically ill. In addition, there are no existing definitive meta-analyses across a comprehensive range of measures. Therefore, a diagnostic accuracy meta-analysis was conducted to assess the sensitivity and specificity of the most widely used case-identification instruments in people who are physically ill.

METHOD

Data sources and searches

The full review protocol can be found in the guideline on depression in people with chronic physical health problems, which was commissioned by NICE.[31] Briefly, a search for studies assessing the validity of case-identification instruments was made using seven electronic bibliographic databases (CENTRAL, CINAHL, Embase, HMIC, MEDLINE, PsycINFO, Web of Knowledge). Each database was searched from inception to October 2009. Additional papers were found by searching the references of retrieved articles, tables of contents of relevant journals, previous systematic reviews and meta-analyses of case identification for depression, written requests to experts, and suggestions made by the members of the Guideline Development Group (comprising clinicians, academics, and service users with expertise in depression and chronic physical health problems).

Study selection

The study included validation studies of mood questionnaires agreed by the authors (see Appendix 1 for further details). The reference standard was diagnoses according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) of the American Psychiatric Association (for example DSM-IV)[32] or International Classification of Diseases (ICD) (for example ICD-10)[33] of the World Health Organization criteria. Studies that did not clearly state the comparator to be DSM or ICD diagnosis of depression, or that did not provide sufficient data to be extracted in the meta-analysis were excluded.

Data extraction and quality assessment

All published studies that met the eligibility criteria were assessed for methodological quality using the Scottish Intercollegiate Guidelines Network (SIGN) checklist for diagnostic studies.[29] Data were extracted independently by at least two authors, and 2×2 tables were constructed, from which the primary outcomes were calculated: that is sensitivity, specificity, and likelihood ratios.

To maximise the available data, the most consistently reported and recommended cut-off points were extracted for each of the scales. There are limitations to this approach, as noted by Furukawa and colleagues,[34] who found that the optimal cut-offs for the General Health Questionnaire (GHQ)-12 and GHQ-28 differed according to the prevalence of depression, and it is likely there are similar problems for most other scales. However, a Bayesian approach makes allowance for variations according to prevalence (see

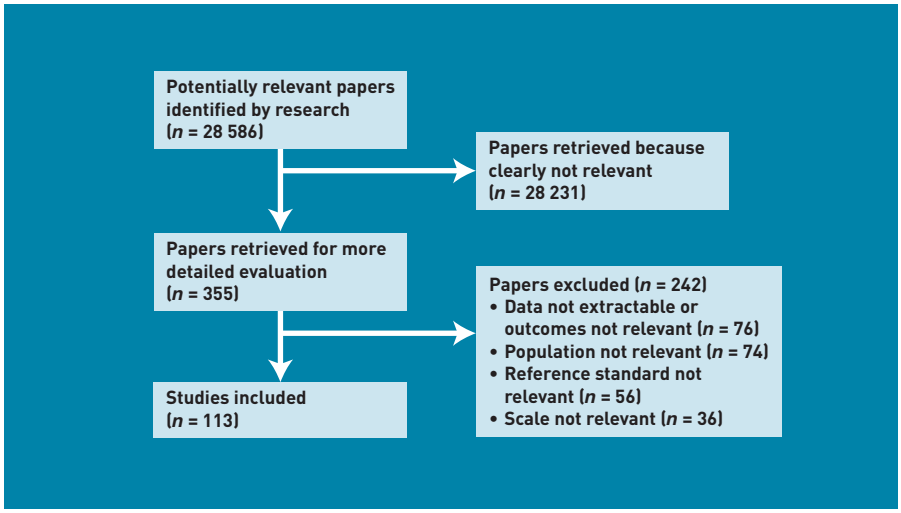

Figure 1. Study flow diagram.

below), therefore seeking to take into account this potential limitation.

Data synthesis and analysis

A bivariate diagnostic accuracy meta-analysis was conducted using Stata (version 10) with the metandi[35] commands, to obtain pooled estimates of sensitivity, specificity, and likelihood ratios. This method was originally developed as a mixed effects regression model for meta-analysis of trials, and modified more recently for studies of diagnostic accuracy.[36,37] Between-study heterogeneity was assessed using the I^2 statistic.[38] In addition, publication bias was assessed by visual inspection of funnel plots, and formal use of Egger's test.[39]

A Bayesian curve analysis was also undertaken; this plots post-test conditional probabilities from all possible pre-test probabilities (prevalence). The area under the Bayesian curve (AUC) for positive results can be used as a statistical comparison of rule-in success and 1 — AUC for negatives results used as an indicator of rule-out success. An area of more than 0.75 can be interpreted as 'satisfactory' and more than 0.80 interpreted as 'good'. If a test achieved more than 0.90 in a rule-in capacity, this was considered sufficient for a recommendation that this tool could be used on its own for case finding.

Additional meta-regression analyses were planned to assess differences in diagnostic accuracy for disease groups. Such analyses were conducted on a scale when there were a minimum of four studies for at least two disease groups.

RESULTS

A total of 113 studies on 20 826 participants met the eligibility criteria of the review (see Figure 1 for full details on study flow information). These studies were both on populations specifically targeted for a chronic physical health problem (such as cancer, heart disease, and stroke), and in general medical settings where all were physically ill and a substantial proportion had a chronic physical health problem. In total, 83 studies specifically targeted people with chronic physical health problems in any setting (Appendix 2). The mean prevalence of depression was 0.25 (95% confidence interval [CI] = 0.05 to 0.61). A further 30 studies were on people in general medical settings, with a mean prevalence of depression of 0.24 (95% CI = 0.04 to 0.52).

Studies recruiting for chronic physical health problem

Sensitivity and specificity. Table 1 provides an evidence summary for the various scales on people recruited for specific chronic physical health problems. There was moderate to high sensitivity for most scales. The tools with the highest sensitivity were the two stem questions (0.98; 95% CI = 0.85 to 0.99), followed by the GHQ-28, Patient Health Questionnaire (PHQ)-9, Beck Depression Inventory (BDI), and BDI non-somatic (Table 1). Sensitivity was lowest for the one-item measure.

The Zung Self Rating Depression Scale had the highest specificity 0.92 (95% CI = 0.68 to 0.98). This was followed by the two stem questions, the Hamilton Depression Rating Scale (HDRS), PHQ-9 and the Centre for Epidemiologic Studies Depression Scale (CES-D); all had high specificity. The lowest specificity was found for the one-item measure and the GHQ-12.

Rule-in (positive predictive value) and rule-out accuracy (negative predictive value). Using Bayesian plots of conditional probabilities to examine rule-in and rule-out performance, only three tools had less than satisfactory rule-in performance, namely the single question: the Geriatric Depression Scale (GDS-30) and GHQ-12. The optimal single tool was the Zung, although it did not reach the *a priori* standard for recommendation when applied alone. For rule-out performance, four methods were not satisfactory. These were the single queston, the Hospital Anxiety and Depression Scale (HADS), GDS-30, and GHQ-12. The optimal tools were the two stem questions and GHQ-28. Overall accuracy was best for the two stem questions, Zung, PHQ-9, and GHQ-28. However, it should be noted that data for the Zung scale were based on just four studies and a relatively small total sample size (n = 190).

Table 1. Evidence summary of scales in studies recruiting for chronic physical illness

Instrument	Total sample size (studies)	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	I^2, %	Prevalence
Patient Health Questionnaire-9	1617 (6)	0.84 (0.69 to 0.92)	0.88 (0.83 to 0.91)	6.77 (4.96 to 9.24)	0.19 (0.10 to 0.37)	93	0.26
Two stem questions (low mood and loss of interest)	1860 (6)	0.98 (0.85 to 0.99)	0.86 (0.70 to 0.94)	6.81 (2.98 to 15.57)	0.02 (0.002 to 0.21)	0	0.20
Beck Depression Inventory (BDI)	3486 (22)	0.83 (0.79 to 0.87)	0.79 (0.74 to 0.84)	3.96 (3.12 to 5.03)	0.21 (0.16 to 0.27)	90	0.20
BDI-non somatic items	946 (8)	0.83 (0.68 to 0.92)	0.79 (0.70 to 0.85)	3.89 (2.66 to 5.69)	0.22 (0.11 to 0.44)	87	0.22
Center of Epidemiological Studies — Depression	1812 (11)	0.77 (0.71 to 0.85)	0.85 (0.80 to 0.90)	5.25 (3.60 to 7.65)	0.27 (0.21 to 0.36)	0	0.23
Geriatric Depression Scale-30	687 (13)	0.79 (0.73 to 0.84)	0.73 (0.67 to 0.79)	2.95 (2.37 to 3.68)	0.29 (0.22 to 0.38)	0	0.28
Geriatric Depression Scale-15	823 (8)	0.84 (0.78 to 0.88)	0.77 (0.73 to 0.81)	3.70 (3.03 to 4.52)	0.21 (0.16 to 0.28)	0	0.29
One item	1940 (11)	0.73 (0.60 to 0.83)	0.77 (0.62 to 0.88)	3.21 (1.98 to 5.21)	0.35 (0.24 to 0.49)	98	0.26
General Health Questionnaire-12	517 (5)	0.81 (0.70 to 0.89)	0.64 (0.52 to 0.75)	2.27 (1.68 to 3.09)	0.29 (0.19 to 0.47)	53	0.16
General Health Questionnaire-28	465 (5)	0.90 (0.79 to 0.96)	0.80 (0.62,0.90)	4.39 (2.31 to 8.35)	0.13 (0.06 to 0.25)	94	0.33
Hospital Anxiety and Depression Scale — Depression	5087 (29)	0.75 (0.67 to 0.81)	0.81 (0.74 to 0.86)	3.90 (2.94 to 5.17)	0.31 (0.24 to 0.41)	99	0.26
Hamilton Depression Rating Scale	985 (11)	0.81 (0.75 to 0.86)	0.85 (0.76 to 0.91)	5.32 (3.24 to 8.71)	0.22 (0.16 to 0.31)	52	0.26
Zung Self Rating Depression Scale	190 (4)	0.78 (0.56 to 0.91)	0.92 (0.68 to 0.98)	9.82 (2.31 to 41.63)	0.24 (0.11 to 0.50)	83	0.30

LR = likelihood ratio.

Meta-regression comparing the diagnostic accuracy for different disease groups was only possible for the BDI and HADS-D. There was no evidence of difference in sensitivity (beta = 0.93, $P = 0.34$) and specificity (beta = 1.56, $P = 0.35$) of the HADS between stroke and cancer patients. There was no evidence of difference in sensitivity (beta = 1.49, $P = 0.60$), but some evidence for differences in specificity (beta = 1.20, $P = 0.02$) of the BDI between heart disease and cancer patients.

Studies in general medical settings

Table 2 summarises the results for general medical settings. There were only three scales that provided sufficient data for analyses. All these scales performed equally well in this setting as compared to populations specifically targeted for chronic physical health problems with a large overlap in confidence intervals.

Sensitivity and specificity. Sensitivity was relatively high in all measures but particularly high in the GDS-15 (0.89; 95% CI = 0.84 to 0.92). Specificity was very similar for the GDS-30, GDS-15, and HADS when used in general medical settings (Table 2).

Rule-in and rule-out accuracy. Using the same methodology for each measure in general medical settings and correcting for prevalence using a Bayesian analysis, the GDS-15 was most successful and the HADS least successful. No method came close to the *a priori* standard for rule-in performance when applied alone. For rule-out accuracy, the HADS was significantly less accurate than the GDS-15 (Area HADS = 0.71, 95% CI = 0.68 to 0.74 versus Area GDS-15 = 0.78, 95% CI = 0.75 to 0.82).

DISCUSSION

Most of the scales performed adequately as case-identification measures for depression, with modest differences in validity coefficients. Most studies targeted chronic physically ill populations rather than general medical settings such as primary care. In order to detect depression in those with chronic physically ill health, the most sensitive instruments appear to be two stem questions, PHQ-9, and GHQ-28. The most specific measure was the Zung. Overall, optimal accuracy was achieved by the two stem questions, Zung, PHQ-9, and GHQ-28. However, it should be noted that estimates on the Zung and GHQ-28 analysis were based on a relatively small sample size; therefore, it is possible that conclusions regarding these scales may change with further data. No method came close to the *a priori* standard for case-finding recommendation when applied alone.

Another important factor to consider when comparing the different measures is the ease of implementation. The Zung is a 20-item scale and therefore is more resource intensive and less likely to be implemented in primary care compared to shorter measures. Taking into account both the psychometric properties and ease of implementation, it would appear the two stem questions may be the preferred measure for case identification in patients with chronic physical health problems. From these data, the authors do not recommend relying upon a single question

Table 2. Evidence summary of scales in general medical settings

Instrument	Total sample size (studies)	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	I^2, %	Prevalence
Geriatric Depression Scale-30	1255 (9)	0.85 (0.78 to 0.90)	0.76 (0.68 to 0.83)	3.55 (2.73 to 4.60)	0.20 (0.15 to 0.27)	98	0.36
Geriatric Depression Scale-15	1108 (8)	0.89 (0.84 to 0.92)	0.74 (0.67 to 0.80)	3.38 (2.65 to 4.31)	0.15 (0.11 to 0.22)	63	0.24
Hospital Anxiety and Depression Scale — Depression	2506 (6)	0.81 (0.73 to 0.87)	0.75 (0.70 to 0.79)	3.22 (2.79 to 3.72)	0.25 (0.18 to 0.36)	89	0.15

LR = likelihood ratio.

alone, and recommend two questions as a minimum initial enquiry. This is consistent with previous pooled data in primary care[40] and cancer settings.[41]

In general medical settings, there were fewer studies, and analysable data were only available for the GDS, GDS-15, and HADS-D. Specificity was similar for all three scales but sensitivity was highest in the GDS-15. Further research is needed to confirm whether the optimal tools in the chronically ill (two stem questions, PHQ-9, and the Zung) perform equally well in general medical samples.

There are several limitations to the results of this systematic review. First, there was moderate to high heterogeneity for most measures. Secondly, there is a paucity of validity studies using the ICD-10 as the criterion standard compared with the DSM-IV, which may favour tools using DSM items, and therefore the authors recommend future examination using this outcome. Thirdly, there were widely used or potentially useful scales that had few or no studies in the physically ill; these include the Montgomery–Asberg Depression Scale (MADRS),[42] and the Clinically Useful Outcome Depression Scale (CUDOS).[43] Further research is needed on these scales for people with chronic physical health problems. Fourthly, there were a number of different semi-structured methods used to determine the interview-based diagnosis, including the Schedules for Clinical Assessment in Neuropsychiatry (SCAN),[44] the Composite International Diagnostic Interview (CIDI),[45] the Structured and Clinical Interview for DSM-III-R (SCID),[46] and the Diagnostic Interview schedule (DIS),[47] all of which may vary in diagnostic accuracy. A further limitation is the lack of cost-effectiveness analyses assessing the cost impact of false positives associated with the use of case-identification measures. However, it should be noted that the cost-effectiveness of case identification is very complex to model and requires a number of assumptions concerning probabilities assigned to events in the depression treatment care pathway, and explicit values of treatment outcomes.[48] Therefore, such issues were considered beyond the scope of this paper.

It should also be acknowledged that the use of case-identification tools may not be translated into real benefit in clinical practice. Case identification may bring limited benefit if there are no effective assessment and treatment services in place, as professionals may be reluctant to make a diagnosis of depression if they have limited resources on which to call.[49] The aim of the NICE guideline for which this review was conducted,[31] is to promote the commissioning of such services. The impact of case finding on the individual consultation may be important, since the use of the PHQ-9 severity questionnaire can cause a tension within the consultation, with GPs struggling to manage formal assessment versus personal enquiry.[50]

From this data synthesis, it appears that there are a number of instruments for the case identification of depression in the medically ill that have similar accuracy. A consideration of both accuracy and acceptability suggests that the two stem questions may be the most efficient initial method, although further validation is needed. We do not recommend the use of a single question used alone. GPs and practice nurses should not rely on the case-finding questions alone; they should be confident to complete an assessment of the patient's mental state and risk, and a pathway within the practice should be in place (particularly when it is the practice nurse who has done the case finding). Resources within the practice should be available to support patients who have depression and a chronic physical health problem, and primary care practitioners should have well-defined links with local primary care mental health services, which should offer appropriate interventions for such patients, including a collaborative care approach as recommended by NICE.[31]

Funding

Stephen Pilling received financial support from the National Institute for Health and Clinical Excellence.

Additional information

The Bayesian plots of conditional probabilities of scale are available on request from the authors.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

David Goldberg developed the General Health Questionnaire. The other authors have declared no competing interests.

Acknowledgements

Thank you to the NICE guideline development groups on Depression and Chronic Physical Health Problems and Depression in Adults for their input during the development of this systematic review.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: **http://www.rcgp.org.uk/bjgp-discuss**

REFERENCES

1. World Health Organization. *Global burden of disease*. Geneva: World Health Organization, 2004.
2. Mitchell AJ, Chan M, Bhatti H, *et al*. Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94 interview-based studies. *Lancet Oncol* 2011; **12(2):** 160–174.
3. Herbst S, Pietrzak RH, Wagner J, *et al*. Lifetime major depression is associated with coronary heart disease in older adults: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychosom Med* 2007; **69(8):** 729–734.
4. Wilhelm K, Mitchell P, Slade T, *et al*. Prevalence and correlates of DSM-IV major depression in an Australian national survey. *J Affect Disord* 2003; **75(2):** 155–162.
5. Wagena EJ, Arrindell WA, Wouters EFM, *et al*. Are patients with COPD psychologically distressed? *Eur Respir J* 2005; **26(2):** 242–248.
6. Stein MB, Cox BJ, Afifi TO, *et al*. Does comorbid depressive illness magnify the impact of chronic physical illness? A population-based perspective. *Psychol Med* 2006; **36(5):** 587–596.
7. Unutzer J, Schoenbaum M, Katon WJ, *et al*. Healthcare costs associated with depression in medically ill fee-for-service medicare participants. *J Am Geriatr Soc* 2009; **57(3):** 506–510.
8. Penninx BW, Beekman AT, Honiq A, *et al*. Depression and cardiac mortality: results from a community-based longitudinal study. *Arch Gen Psychiatry* 2001; **58(3):** 221–227.
9. De Voogd JN, Wempe JB, Koeter GH, *et al*. Depressive symptoms as predictors of mortality in patients with COPD. *Chest* 2009; **135(3):** 619–625.
10. House A, Knapp P, Bamford J, *et al*. Mortality at 12 and 24 months after stroke may be associated with depressive symptoms at 1 month. *Stroke* 2001; **32(3):** 696–701.
11. Satin JR, Wolfgang Linden W, Phillips MJ. Depression as a predictor of disease progression and mortality in cancer patients. A meta-analysis. *Cancer* 2009; **115(22):** 5349–5361.
12. Covinsky KE, Kahana E, Chin MH, *et al*. Depressive symptoms and 3-year mortality in older hospitalized medical patients. *Ann Intern Med* 1999; **130(7):** 563–569.
13. Gallo JJ, Bogner HR, Morales KH, *et al*. The effect of a primary care practice-based depression intervention on mortality in older adults. *Ann Intern Med* 2007; **146(10):** 689–698.
14. Glassman AH, O'Connor CM, Califf RM, *et al*. Sertraline treatment of major depression in patients with MI or unstable angina. *JAMA* 2002; **288(6):** 701–709.
15. Del Piccolo L, Saltini A, Zimmermann C. Which patients talk about stressful life events and social problems to the general practitioner? *Psychol Med* 1998; **28(6):** 1289–1299.
16. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009; **374(9690):** 609–611.
17. Williams JW Jr, Kerber CA, Mulrow CD, *et al*. Depressive disorders in primary care: Prevalence, functional disability, and identification. *J Gen Intern Med* 1995; **10(1):** 7–12.
18. Hickie IB, Davenport TA, Scott EM, *et al*. Unmet need for recognition of common mental disorders in Australian general practice. *Med J Aust* 2001; **175 Suppl:** S18–S24.
19. Verhaak PFM, Schellevis FG, Nuijen J, *et al*. Patients with a psychiatric disorder in general practice: determinants of general practitioners' psychological diagnosis. *Gen Hosp Psychiatry* 2006; **28(2):** 125– 132.
20. Bridges KW, Goldberg DP. Somatic presentation of DSM III psychiatric disorders in primary care. *J Psychosom Res* 1985; **29(6):** 563–569.
21. Rost K, Nutting P, Smith J, *et al*. The role of competing demands in the treatment provided primary care patients with major depression. *Arch Fam Med* 2000; **9(2):** 150–154.
22. Coventry P, Hays R, Dickens C, *et al*. Talking about depression: a qualitative study of barriers to managing depression in people with long term conditions in primary care. *BMC Fam Pract* 2011; **12:** 10.
23. Aragones E, Pinol JL, Labad A, *et al*. Detection and management of depressive disorders in primary care in Spain. *Int J Psychiatry Med* 2001; **34(4):** 331–343.
24. Nuyen J, Volkers AC, Verhaak PFM, *et al*. Accuracy of diagnosing depression in primary care: the impact of chronic somatic and psychiatric co-morbidity. *Psychol Med* 2005; **35(8):** 1185–1195.
25. Pfaff JJ, Almeida OP. A cross-sectional analysis of factors that influence the detection of depression in older primary care patients. *Aust N Z J Psychiatry* 2005; **39(4):** 262–265.
26. Menchetti M, Belvederi M, Murri B, *et al*. Recognition and treatment of depression in primary care. Effect of patients' presentation and frequency of consultation. *J Psychosom Res* 2009; **66(4):** 335–341.
27. O'Connor EA, Whitlock EP, Beil TL, Gaynes BN. Screening for depression in adult patients in primary care settings: a systematic evidence review. *Ann Intern Med* 2009; **151(11):** 793–803.
28. Department of Health. *Quality and Outcomes Framework*. London: Department of Health, 2004.
29. National Institute for Health and Clinical Excellence. *The guidelines manual*. London: National Institute for Health and Clinical Excellence, 2008.
30. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ* 2008; **178(8):** 997–1003.
31. National Institute for Health and Clinical Excellence. *Depression in people with a chronic physical health problem*. London: National Institute for Health and Clinical Excellence, 2009
32. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, 4th Edition (DSM-IV)*. Washington DC: American Psychiatric Association, 1994.
33. World Health Organization. *The ICD-10 classification of mental and behavioural disorders*. Geneva: World Health Organization, 1993.
34. Furukawa TA, Goldberg DP, Rabe-Hesketh S, *et al*. Stratum-specific likelihood ratios of two versions of the General Health Questionnaire. *Psychol Med* 2001; **31(3):** 519–529.
35. Harbord R, Deeks JJ, Egger M, *et al*. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007; **8(2):** 239–251.
36. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002; **21(4):** 589–624.
37. Reitsma JB, Glas AS, Rutjes AWS, *et al*. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; **58(10):** 982–990.
38. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327(7414):** 557–560.
39. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315(7109):** 629–634.
40. Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract* 2007; **57(535):** 144–151.
41. Mitchell A. Are one or two simple questions sufficient to detect depression in cancer and palliative care? A Bayesian meta-analysis. *Br J Cancer* 2008; **98(12):** 1934–1943.
42. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134:** 382–389.
43. Zimmerman M, Chelminski I, McGlinchey JB, *et al*. A clinically useful depression outcome scale. *Compr Psychiatry* 2008; **49(2):** 131–140.
44. World Health Organization. *Schedules for clinical assessment in neuropsychiatry*. Geneva: World Health Organization, 1994.
45. World Health Organization. *Composite international diagnostic interview*. Geneva: World Health Organization, 1990.
46. Spitzer R L, Williams JB, Gibbon M, First MB. The structured clinical interview for DSM-III-R (SCID). I: History, rationale, and description. *Arch Gen Psychiatry* 1992; **49(8):** 624–629.
47. Robins LN, Helzer JE, Croughan J. National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics and validity. *Arch Gen Psychiatry* 1981; **38(4):** 381–389.
48. Valenstein M, Vijan S, Zeber JE, *et al*. The cost utility of screening for depression in primary care. *Ann Intern Med* 2001; **134(5):** 345–360.
49. Burroughs K, Lovell K, Morley M, *et al*. Justifiable depression: how primary care professional and patients view late-life depression? A qualitative study. *Fam Pract* 2006; **23(3):** 369–377.
50. Leydon G, Dowrick C, McBride A, *et al*. Questionnaire severity measures for depression: a threat to the doctor-patient relationship? *Br J Gen Pract* 2011;

61(583): 101–107.

51. Beck AT, Steer RA, Brown GK. *Beck Depression Inventory. 2nd edn: manual.* San Antonio: The Psychological Corporation, 1996.
52. Beck AT, Guth D, Steer RA, *et al.* Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. *Behav Res Ther* 1997; **35(8):** 785–791.
53. Spitzer RL, Kroenke K, Williams JB, *et al.* Validation and utility of a self-report version of the PRIME-MD: the PHQ primary care study. *JAMA* 1999; **282(18):** 1737–1744.
54. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item screener. *Med Care* 2003; **41(11):** 1284–1292.
55. Whooley MA, Avins AL, Miranda J, *et al.* Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med* 1997; **12(7):** 439–445.
56. Goldberg DP, Williams P. *A user's guide to the General Health Questionnaire.* Windsor: NFER-Nelson, 1988.
57. Radloff LS. The CESD Scale: a self report scale for research in the general population. *Appl Psychol Meas* 1977; **1:** 385–401.
58. Yesavage JA, Brink TL, Rose TL, *et al.* Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982; 17(1): 37–49.
59. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. In: Brink TL (ed). *Clinical gerontology: a guide to assessment and intervention.* New York: The Haworth Press, 1986.
60. Zung WWK. A self-rating depression scale. *Arch Gen Psyciatry* 1965; **12:** 63–70.
61. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; **67(6):** 361–370.
62. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23:** 56–62.
63. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134:** 382–389.
64. Zimmerman M, Chelminski I, McGlinchey JB, *et al.* A clinically useful depression outcome scale. *Compr Psychiatry* 2008; **49(2):** 131–140.
65. Aben I, Verhey F, Lousberg R, *et al.* Validity of the Beck Depression Inventory, Hospital Anxiety and Depression Scale, SCL-90 and Hamilton Depression Rating Scale as screening instruments for depression in stroke patients. *Psychosomatics* 2002; **43(5):** 386–393.
66. Agrell B, Dehlin O. Comparison of six depression rating scales in geriatric stroke patients. *Stroke* 1989; **20(9):** 1190–1194.
67. Akechi, TM, Okuyama T, Sugawara Y, *et al.* Screening for depression in terminally ill cancer patients in Japan. *J Pain Symptom Manage* 2006; **31(1):** 5–12.
68. Akizuki N, Akechi T, Nakanishi T, *et al.* Development of a brief screening interview for adjustment disorders and major depression in patients with cancer. *Cancer* 2003; **97(10):** 2605–2613.
69. Aydin IO and Ulusahin A. Depression, anxiety comorbidity, and disability in tuberculosis and chronic obstructive pulmonary disease patients: applicability of GHQ-12. *Gen Hosp Psychiatry* 2001; **23(2):** 77–83.
70. Berard RMF, Boermeester F, Viljoen G. Depressive disorders in an out-patient oncology setting: prevalence, assessment and management. *Psychooncology* 1998; **7(2):** 112–120
71. Berg A, Lonnqvist J, Palomaki H, Kaste M. Assessment of depression after stroke: a comparison of different screening instruments. *Stroke* 2009; **40(2):** 523–529.
72. Blank K, Gruman C, Robison JT. Case-finding for depression in elderly people: balancing ease of administration with validity in varied treatment settings. *J Gerontol* 2004; **59(4):** 378–384.
73. Burke WJ, Nitcher RL, Roccaforte WH, *et al.* A prospective evaluation of the Geriatric Depression Scale in an outpatient geriatric assessment center. *J Am Geriatr Soc* 1992; **40(12):** 1227–1230.
74. Chilcot J, Wellsted D, Farrington K. Screening for depression while patients dialyse: an evaluation. *Nephrol Dial Transplant* 2008; **23(8):** 2653–2659.
75. Chochinov HM, Wilson KG, Enns M, *et al.* 'Are you depressed?' Screening for depression in the terminally ill. *Am J Psychiatry* 1997; **154(5):** 674–676.
76. Costantini M, Musso M, Viterbori P, *et al.* Detecting psychological distress in cancer patients: validity of the Italian version of the Hospital Anxiety and Depression Scale. *Support Care Cancer* 1999; **7(3):** 121–127.
77. Craven J, Rodin G M, Littlefield C. The Beck Depression Inventory as a screening device for major depression in renal dialysis patients. *Int J Psychiatry Med* 1988; **18(4):** 365–374.
78. Cullum S, Tucker S, Todd C, *et al.* Screening for depression in older medical inpatients. *Int J Geriatr Psychiatry* 2006; **21(5):** 476.
79. Diez-Quevedo C, Rangil T, Sanchez-Planell L, *et al.* Validation and utility of the Patient Health Questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. *Psychosom Med* 2001; **63(4):** 679–686.
80. Ertan FS, Ertan T, Kiziltan G, *et al.* Reliability and validity of the Geriatric Depression Scale in depression in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2005; **76(10):** 1447.
81. Forkman T, Vehren T, Boeker M, *et al.* Sensitivity and specificity of the Beck Depression Inventory: how useful is the conventional cut off? *J Psychosom Res* 2009; **67(4):** 347–352.
82. Freedland KE, Rich MW, Skala JA, *et al.* Prevalence of depression in hospitalized patients with congestive heart failure. *Psychosom Med* 2003; **65(1):** 119–128.
83. Furlanetto LM, Mendlowicz MV, Bueno J. The validity of the Beck Depression Inventory-Short Form as a screening and diagnostic instrument for moderate and severe depression in medical inpatients. *J Affect Disord* 2005; **86(1):** 87–91.
84. Galaria I, Casten R, Rovner B. Development of a shorter version of the Geriatric Depression Scale for visually impaired older patients. *Int Psychogeriatr* 2000; **12(4):** 435–443.
85. Gilley DW, Wilson RS. Criterion-related validity of the Geriatric Depression Scale in Alzheimer's disease. *J Clin Exp Neuropsychol* 1997; **19(4):** 489–499.
86. Golden J, Conroy R, O'Dwyer A. Reliability and validity of the Hospital Anxiety and Depression Scale and the Beck Depression Inventory (Full and FastScreen scales) in detecting depression in persons with hepatitis C. *J Affect Disord* 2007; **100(1–3):** 269.
87. Grassi L, Sabato S, Rossi E, *et al.* Affective syndromes and their screening in cancer patients with early and stable disease. *J Affect Disord* 2009; **114(1–3):** 193–199.
88. Hahn D, Reuter K, Harter M. Screening for affective and anxiety disorders in medical patients: comparison of HADs, GHQ-12 and brief-PHQ. GMS Psychosoc Med 2006; **3:** Doc09.
89. Hall A, Hern RA, Fallowfield L. Are we using appropriate self-report questionnaires for detecting anxiety and depression in women with early breast cancer? *Eur J Cancer* 1999; **35(1):** 79–85.
90. Hammer EM, Häcker S, Hautzinger M, *et al.* Validity of the ALS-Depression-Inventory (ADI-12) — a new screening instrument for depressive disorders in patients with amyotrophic lateral sclerosis. *J Affect Disord* 2008; **109(1–2):** 213–219.
91. Harter M, Woll S, Wunsch A, *et al.* Screening for mental disorders in cancer, cardiovascular and musculoskeletal diseases. Comparison of HADS and GHQ-12. *Soc Psychiatry Psychiatr Epidemiol* 2006; **41(1):** 56–62.
92. Harter M, Reuter K, Gross-Hardt K, *et al.* Screening for anxiety, depressive and somatoform disorders in rehabilitation: Validity of HADS and GQH-12 in patients with musculoskeletal disease. *Disabil Rehabil* 2001; **23(16):** 744.
93. Haughey M T, Calderon Y, Torres S, *et al.* Identification of depression in an inner-city population using a simple screen. *Acad Emerg Med* 2005; **12(12):** 1221–1226.
94. Haworth J E, Moniz-Cook E, Clark AL, *et al.* An evaluation of two self-report screening measures for mood in an out-patient chronic heart failure population. *Int J Geriatr Psychiatry* 2007; **22(11):** 1147–1153.
95. Healey AK, Kneebone II, Carroll M, *et al.* A preliminary investigation of the reliability and validity of the Brief Assessment Schedule Depression Cards and the Beck Depression Inventory-Fast Screen to screen for depression in older stroke survivors. *Int J Geriatr Psychiatry* 2008; **23(5):** 531–536.
96. Hedayati SS, Bosworth HB, Kuchibhatla M, *et al.* The predictive value of self-report scales compared with physician diagnosis of depression in hemodialysis patients. *Kidney Int* 2006; **69(9):** 1662–1668.
97. Hermanns N, Kulzer B, Krichbaum M, *et al.* How to screen for depression and emotional problems in patients with diabetes: comparison of screening characteristics of depression questionnaires, measurement of diabetes-specific emotional problems and standard clinical assessment. *Diabetologia* 2006; **49(3):** 469–477.
98. Herrero MJ, Blanch J, Peri JM, *et al.* A validation study of the hospital anxiety

and depression scale (HADS) in a Spanish population. *Gen Hosp Psychiatry* 2003; **25(4):** 277–283.

99. Hickie C, Snowdon J. Depression scales for the elderly. *Clin Gerontol* 1987; **6:** 51–53.
100. Hopko D, Bell J, Armento M. Phenomenology and screening of clinical depression in cancer patients. *J Psychosoc Oncol* 2007; **26(1):** 31–51.
101. Hoyl MT, Alessi CA, Harker JO, *et al.* Development and testing of a five-item version of the Geriatric Depression Scale. *J Am Geriatr Soc* 1999; **47(7):** 873–878.
102. Hughson AVM, Cooper AF, McArdle MC, Smith DC. Validity of the General Health Questionnaire and its subscales in patients receiving chemotherapy for early breast cancer. *J Psychosom Res* 1988; **32(4–5):** 393–402.
103. Ibbotson T, Maguire P, Selby P, *et al.* Screening for anxiety and depression in cancer patients: the effects of disease and treatment. *Eur J Cancer* 1994; **30A(1):** 37–40.
104. Jackson R, Baldwin B. Detecting depression in elderly medically ill patients: the use of the Geriatric Depression Scale compared with medical and nursing observations. *Age Aging* 1993; **22(5):** 349–353.
105. Jefford M, Mileshkin L, Richards K, *et al.* Rapid screening for depression — validation of the Brief Case-Find for Depression (BCD) in medical oncology and palliative care patients. *Br J Cancer* 2004; **91(5):** 900–906.
106. Johnson G, Burvill PW, Anderson CS, *et al.* Screening instruments for depression and anxiety following stroke. *Acta Psychiatr Scand* 1995; **91(4):** 252–257.
107. Katz MR, Kopek N, Waldron J, *et al.* Screening for depression in head and neck cancer. *Psychooncology* 2004; **13(4):** 269–280.
108. Kawase E, Karasawa K, Shimotsu S, *et al.* Evaluation of a one-question interview for depression in radiation oncology department in Japan. *Gen Hosp Psychiatry* 2006; **28(4):** 321–322.
109. Koenig HG, Meador KG, Cohen HJ, *et al.* Screening for depression in hospitalized elderly medical patients: taking a closer look. *J Am Geriatr Soc* 1992; **40(10):** 1013–1017.
110. Kugaya A, Akechi T, Okuyama T, *et al.* Screening for psychological distress in Japanese cancer patients. *Jpn J Clin Oncol* 1998, **28(5):** 333–338.
111. Lam CK, Lim PP, Low BL, *et al.* Depression in dementia: a comparative and validation study of four brief scales in the elderly Chinese. *Int J Geriatr Psychiatry* 2004; **19(5):** 422–428.
112. Lamers F, Jonkers CC, Bosma H, *et al.* Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. *J Clin Epidemiol* 2008; **61(7):** 679–687.
113. Laska AC, Martensson B, Kahan T. Recognition of depression in aphasic stroke patients. *Cerebrovasc Dis* 2007; **24(1):** 74–79.
114. Le Fevre P, Devereux J, Smith S, *et al.* Screening for psychiatric illness in the palliative care inpatient setting: a comparison between the Hospital Anxiety and Depression Scale and the General Health Questionnaire-12. *Palliat Med* 1999; **13(5):** 399–407.
115. Lee AC, Tang SW, Yu GK, Cheung RT. The smiley as a simple screening tool for depression after stroke: a preliminary study. *Int J Nurs Stud* 2008; **45(7):** 1081–1089.
116. Leentjens AFG, Verhey FRJ, Luijckx GJ, *et al.* The validity of the Beck Depression Inventory as a screening and diagnostic instrument for depression in patients with Parkinson's disease. *Mov Disord* 2000; **15(6):** 1221–1224.
117. Leung KK, Lue BH, Lee MB, *et al.* Screening of depression in patients with chronic medical diseases in a primary care setting. *Fam Pract* 1998; **15(1):** 67–75.
118. Lightbody CE, Baldwin R, Connolly M, *et al.* Can nurses help identify patients with depression following stroke? A pilot study using two methods of detection. *J Adv Nurs* 2007; **57(5):** 505–512.
119. Lincoln NB, Nicholl CR, Flannaghan T, *et al.* The validity of questionnaire measures for assessing depression after stroke. *Clin Rehabil* 2003; **17(8):** 840–846.
120. Lloyd-Williams M, Friedman T, Rudd N. Criterion validation of the Edinburgh Postnatal Depression Scale as a screening tool for depression in patients with advanced metastatic cancer. *J Pain Symptom Manage* 2000; **20(4):** 259–265.
121. Lloyd-Williams M, Friedman T, Rudd N. An analysis of the validity of the Hospital Anxiety and Depression Scale as a screening tool in patients with advanced metastatic cancer. *J Pain Symptom Manage* 2001; **22(6):** 990–996.
122. Lloyd-Williams M, Dennis M, Taylor F. A prospective study to compare three depression screening tools in patients who are terminally ill. *Gen Hosp Psychiatry* 2004; **26(5):** 384–389.
123. Love AW, Kisssne DW, Bloch S, Clarke D. Diagnostic efficiency of the Hospital Anxiety and Depression Scale in women with early stage breast cancer. *Aust N Z J Psychiatry* 2002; **36(2):** 246–250.
124. Love A, Grabsch B, Clarke D, *et al.* Screening for depression in women with metastatic breast cancer: a comparison of the Beck Depression Inventory Short Form and the Hospital Anxiety and Depression Scale. *Aust N Z J Psychiatry* 2004; **38(7):** 526–531.
125. Low GD, Hubley AM. Screening for depression after cardiac events using the Beck Depression Inventory-II and the Geriatric Depression Scale. *Soc Indic Res* 2007; **82:** 527–543.
126. Lowe B, Spitzer R, Grafe K, *et al.* Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord* 2004; **78(2):** 140.
127. Lustman PJ, Clouse RE, Griffiths LS, *et al.* Screening for depression in diabetes using the Beck Depression Inventory. *Psychosom Med* 1997; **59(1):** 24–31.
128. Lykouras L, Adrachta D, Kalfakis N, *et al.* GHQ-28 as an aid to detect mental disorders in neurological inpatients. *Acta Psychiatr Scand* 1996; **93(3):** 212–216.
129. Magni G, Schifano F. Assessment of depression in an elderly medical population. *J Affect Disord* 1986; **11(2):** 121–124.
130. McManus D, Pipkin SS, Whooley MA. Screening for depression in patients with coronary heart disease (data from the Heart and Soul Study). *Am J Cardiol* 2005; **96(8):** 1076–1081.
131. McQuillan J, Fifield J, Sheehan P, *et al.* A comparison of self-reports of distress and affective disorder diagnoses in rheumatoid arthritis: a receive operator characteristic analysis. *Arthritis Rheum* 2003; **49(3):** 368–376.
132. Meyer HA, Sinnot C, Seed PT. Depressive symptoms in advanced cancer. Part 1. Assessing depression: The Mood Evaluation Questionnaire. *Palliat Med* 2003; **17(7):** 596.
133. Mitchell AJ, Baker-Glenn E, Thiagarajan S, *et al.* Diagnostic accuracy and utility of the Patient Health Questionnaire (PHQ2 v PHQ9) for major depression in early cancer. *Psychooncology* 2008; **17:** S202–S203
134. Mohr DC, Hart SL, Julian L, *et al.* Screening for depression among patients with multiple sclerosis: two questions may be enough. *Mult Scler* 2007; **13(2):** 215–219.
135. Narding P, Leentjens AFG, van Kooten F, Verhey FRJ. Disease-specific properties of the Hamilton Rating Scale for Depression in patients with stroke, Alzheimer's dementia and Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 2002; **14(3):** 329–334.
136. Neal RM, Baldwin RC. Screening for anxiety and depression in elderly medical outpatients. *Age Ageing* 1994; **23(6):** 461–464.
137. O'Rourke S, MacHale S, Signorini D, Dennis M. Detecting psychiatric morbidity after Stroke: Comparison of the GHQ and the HAD scale. *Stroke* 1998; **29(5):** 980–985.
138. Okimoto JT, Barnes RF, Veith RC, *et al.* Screening for depression in geriatric medical patients. *Am J Psychiatry* 1982; **139(6):** 799–802.
139. Olden M, Rosenfeld B, Pessin H, Breitbart W. Measuring depression at the end of life: is the Hamilton Depression Rating Scale a valid instrument? *Assessment* 2009; **16(1):** 43–54.
140. Ozalp E, Soygür H, Cankurtaran E, *et al.* Psychiatric morbidity and its screening in Turkish women with breast cancer: a comparison between the HADS and SCID tests. *Psychooncology* 2008; **17(7):** 668–675.
141. Parikh RM, Eden DT, Price TR, *et al.* The sensitivity and specificity of the Center for Epidemiologic Studies Depression Scale in screening for post-stroke depression. *Int J Psychiatry Med* 1988; **18(2):** 169–181.
142. Parker G, Hilton T, Bains J, *et al.* Cognitive-based measures screening for depression in the medically ill: The DMI-10 and the DMI-18. *Acta Psychiatr Scand* 2002; **105(6):** 419–426.
143. Passik SD, Kirsch KL, Donaghy KB, *et al.* An attempt to employ the Zung Self Rating Depression Scale as a 'lab test' to trigger follow up in ambulatory oncology clinics. *J Pain Symptom Manage* 2001; **21(4):** 273–281.
144. Patterson K, Young C, Woods S, *et al.* Screening for major depression in persons with HIV infection: the concurrent predictive validity of the Profile of Mood States Depression-Dejection Scale. *Int J Methods Psychiatr Res* 2006; **15(2):** 75–82.

145. Payne A, Barry S, Creedon B, *et al.* Sensitivity and specificity of a two question screening tool for depression in a specialist palliative care unit. Palliat Med 2007; **21(3):** 193–198.
146. Persoons P, Luyckx K, Desloovere C, *et al.* Anxiety and mood disorders in otorhinolaryngology outpatients presenting with dizziness: validation of the self-administered PRIME-MD Patient Health Questionnaire and epidemiology. *Gen Hosp Psychiatry* 2003; **25(5):** 316–323.
147. Picardi A, Adler DA, Abeni D, *et al.* Screening for depressive disorders in patients with skin diseases: a comparison of three screeners. *Acta Derm Venereol* 2005; 85(5): 414–419.
148. Pomeroy I, Clark C, Philp I. The effectiveness of very short scales for depression screening in elderly medical patients. *Int J Geriatr Psychiatry* 2001; **16(3):** 321–326.
149. Poole N, Morgan J. Validity and reliability of the Hospital Anxiety and Depression Scale in a hypertrophic cardiomyopathy clinic: the HADS in a cardiomyopathy population. *Gen Hosp Psychiatry* 2006; **28(1):** 55–58.
150. Rapp SR, Parisi SA, Walsh DA, *et al.* Detecting depression in elderly medical inpatients. *J Consult Clin Psychol* 1988; **56(4):** 509–513.
151. Razavi D, Delvaux N, Farvacques C, Robaye E. Screening for adjustment disorders and major depressive disorders in cancer in-patients. *Br J Psychiatry* 1990; **156:** 79–83.
152. Reuter K, Harter M. Screening for mental disorders in cancer patients — discriminant validity of HADS and GHQ-12 assessed by standardized clinical interview. *Int J Methods Psychiatr Res* 2000; **10:** 86–96.
153. Rinaldi P, Mecocci P, Benedetti C, *et al.* Validation of the five-item geriatric depression scale in elderly subjects in three different settings. *J Am Geriatr Soc* 2003; **51(5):** 694–698.
154. Roger PR, Johnson-Greene D. Comparison of assessment measures for post-stroke depression. *Clin Neuropsychol* 2009; **23(5):** 780–793.
155. Rovner BW, Shmuely-Dulitzi Y. Screening for depression in low-vision elderly. *Int J Geriatr Psychiatry* 1997; **12(9):** 955–959.
156. Sagen U, Vik TG, Moum T, *et al.* Screening for depression and anxiety after stroke: comparison of the Hospital Anxiety and Depression Scale and the Montogomery Asberg Depression Rating Scale. *J Psychosom Res* 2009; **67(4):** 325–332.
157. Scheinthal SM, Steer R, Giffin L, *et al.* Evaluating geriatric medical outpatients with the Beck Depression Inventory-FastScreen for medical patients. *Aging Ment Health* 2001; **5(2):** 143–148.
158. Schein RL, Koenig HG. The center for epidemiological studies-depression (CES-D) scale: assessment of depression in the medically ill elderly. *Int J Geriatr Psychiatry* 1997; **12(4):** 436–446.
159. Serrano-Duenas M, Solledad SM. Concurrent validation of the 21-item and 6-item Hamilton Depression Rating Scale vs the DSM-IV diagnostic criteria to assess depression in patients with Parkinson's Disease: an exploratory analysis. *Parkinsonism Relat Disord* 2008; **14(3):** 233–238.
160. Shinar D, Gross CR, Price TR, *et al.* Screening for depression in stroke patients: the reliability and validity of the Center for Epidemiologic Studies Depression scale. *Stroke* 1986; **17(2):** 241–245.
161. Silberman CD, Laks J, Capitao C, *et al.* Recognizing depression in patients with Parkinson's Disease. *Arq Neuropsiquiatr* 2006; **64(2B):** 407–411.
162. Singer S, Danker H, Dietz A, *et al.* Screening for mental disorders in laryngeal cancer patients: a comparison of 6 methods. *Psychooncology* 2008; **17(3):** 280–286.
163. Sivrioglu EY, Sivrioglu K, Ertan T. Reliability and validity of the Geriatric Depression Scale in detection of post-stroke minor depression. *J Clin Exp Neuropsychol* 2009; **31(8):** 999–1006.
164. Stafford L, Berk M, Jackson H. Validity of the Hospital Anxiety and Depression Scale and Patient Health Questionnaire-9 to screen for depression in patients with coronary artery disease. *Gen Hosp Psychiatry* 2007; **29(5):** 417–424.
165. Strik J, Honig A, Lousberg R, *et al.* Sensitivity and specificity of observer and self-report questionnaires in major and minor depression following myocardial infarction. *Psychosomatics* 2001; **42(5):** 423–428.
166. Tang WK, Ungvari GS, Chiu HFK, *et al.* Screening post-stroke depression in Chinese older adults using the Hospital Anxiety and Depression Scale. *Aging Ment Health* 2004; **8(5):** 397–399.
167. Tang W, Chan S, Chiu H, *et al.* Can the Geriatric Depression Scale detect poststroke depression in Chinese elderly? *J Affect Disord* 2004; **81(2):** 153–156.
168. Thekkumpurath P, Venkateswaran C, Kumar M, *et al.* Screening for psychological distress in palliative care: performance of touch screen questionnaires compared with semistructured psychiatric interview. *J Pain Symptom Manage* 2009; **38(4):** 597–605.
169. Turner JA, Romano JM. Self-report screening measures for depression in chronic pain patients. *J Clin Psychol* 1984; **40(4):** 909–913.
170. Upadhyaya A, Stanley I. Detection of depression in primary care: comparison of two self-administered scales. *Int J Geriatr Psychiatry* 1997; **12(1):** 35–37.
171. Vahter L, Kreegipuu T, Gross-Paju K. One question as a screening instrument for depression in people with multiple sclerosis. *Clin Rehabil* 2007; **21(5):** 460–464.
172. Vargas H, Matsuo T, Blay S. Validity of the Geriatric Depression Scale for patients seen at general outpatient clinics. *Clin Gerontol* 2007; **30:** 65–78.
173. Walker J, Postma K, McHugh GS, *et al.* Performance of the Hospital Anxiety and Depression Scale as a screening tool for major depressive disorder in cancer patients. *J Psychosom Res* 2007; **63(1):** 83–91.
174. Watnick S, Wang PL, Demadura T, *et al.* Validation of 2 depression screening tools in dialysis patients. *Am J Kidney Dis* 2005; **46(5):** 919–924.
175. Weintraub D, Oehlberg K, Katz I, *et al.* Test characteristics of the 15-item Geriatric Depression Scale and Hamilton Depression Rating Scale in Parkinson disease. *Am J Geriatr Psychiatry* 2006, **14(2):** 169–175.
176. Wilhelm K, Kotze B, Waterhouse M, *et al.* Screening for depression in the medically ill: a comparison of self-report measures, clinician judgment, and DSM-IV diagnoses. *Psychosomatics* 2004; **45(6):** 469.
177. Williams LS, Brizendin EJ, Plue L, *et al.* Performance of the PHQ-9 as a screening tool for depression after stroke. *Stroke* 2005; **36(3):** 635–638.
178. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychotherapy* 1987; **150:** 782–786.

Appendix 1. Full list of instruments considered

1. Beck Depression Inventory (BDI):
   - BDI-II[51]
   - BDI Cognitive-Affective scale[51]
   - BDI Fast Screen[52]
2. Patient Health Questionnaire (PHQ):
   - PHQ-9[53]
   - PHQ-2[54]
3. Two stem questions:[55]
   - These are similar to the PHQ-2 except that the scoring system is dichotomous ('yes' or 'no') rather than the Likert scale used for both PHQ-9 and PHQ-2, and the period of reported low mood or loss of interest is 1 month rather than 2 weeks
4. General Health Questionnaire:[56]
   - GHQ-12
   - GHQ-28
5. Centre of Epidemiological Studies – Depression (CES-D)[57]
6. Geriatric Depression Scale (GDS)
   - GDS-30[58]
   - GDS-15[59]
7. Zung Self Rating Depression Scale[60]
8. Hospital Anxiety and Depression Scale — Depression[61]
9. Hamilton Depression Rating Scale HDRS[62]
   - Both 17– and 21–item versions
10. Montgomery–Asberg Depression Rating Scale (MADRS)[63]
11. Clinically Useful Depression Outcome Scale[64]
12. One-item measures of depression
13. Edinburgh Postnatal Depression Scale[178]

Appendix 2. Summary characteristics of included studies

Study	Country	Patients	Female	Mean age, years	Physical health condition	Instrument
Aben *et al*, 2002[65]	Netherlands	171	91	68	Stroke	BDI HADS-D HDRS
Agrell and Dehlin, 1989[66]	Sweden	40	22	80	Stroke	CES-D GDS HDRS Zung
Akechi *et al*, 2006[67]	Japan	205	68	61	Cancer	1-item Two stem questions
Akizuki *et al*, 2003[68]	Japan	275	164	52	Cancer	HADS-D 1-item
Aydin and Ulusahin, 2001[69]	Turkey	100	Not reported	Not reported	Tuberculosis and COPD	GHQ-12
Berard *et al*, 1998[70]	South Africa	100	87	50	Cancer	HADS-D BDI
Berg *et al*, 2009[71]	Finland	100	32	55	Stroke	BDI, HDRS
Blank *et al*, 2004[72]	US	125	Not reported	77	Various (general medical)	GDS GDS-15 CES-D Two stem questions
Burke *et al*, 1992[73]	US	67	44	77	Various (general medical)	GDS
Chilcot *et al*, 2008[74]	UK	41	16	53	Dialysis	BDI
Chochinov *et al*, 1997[75]	US	197	103	Not reported	Cancer	Two stem questions
Costantini *et al*, 1999[76]	Italy	132	132	53	Breast cancer	HADS-D
Craven *et al*, 1988[77]	Canada	99	36	51	Renal dialysis	BDI
Cullum *et al*, 2006[78]	UK	618	371	80	Medically ill (general medical)	GDS-15
Diez-Quevedo *et al*, 2001[79]	Spain	1003	451	43	Medically ill (general medical)	PHQ-9
Ertan *et al*, 2005[80]	Turkey	109	36	67	Parkinson's Disease	GDS
Forkman *et al*, 2009[81]	Germany	126	37	51	Heart disease	BDI BDI-non somatic
Freedland *et al*, 2003[82]	US	682	327	66	Congestive heart failure	BDI
Furlanetto *et al*, 2005[83]	US	155	98	50	Medically ill (general medical)	BDI non-somatic items
Galaria *et al*, 2000[84]	US	70	41	77	Visual impairment	GDS GDS-15
Gilley and Wilson, 1997[85]	US	93	42	70	Medically ill (general medical)	GDS
Golden *et al*, 2007[86]	US	88	23	Not reported	Hepatitis C	BDI BDI non somatic items HADS-D
Grassi *et al*, 2009[87]	Italy	109	83	55	Cancer	HADS-D
Hahn *et al*, 2006[88]	Germany	204	98	50	Medically ill (general medical)	GHQ-12 HADS-D
Hall *et al*, 1999[89]	UK	266	266	Not reported	Breast cancer	HADS-D
Hammer *et al*, 2008[90]	US	39	19	58	Amyotrophic lateral sclerosis	BDI
Harter *et al*, 2006[91]	Germany	206	103	48	Musculoskeletal diseases	GHQ-12 HADS-D
Harter *et al*, 2001[92]	Germany	569	285	54	Medically ill (general medical)	GHQ-12 HADS-D
Haughey *et al*, 2005[93]	US	226	226	40	Medically ill (general medical)	Two stem questions
Haworth *et al*, 2007[94]	US	88	15	70	Heart failure	GDS-15 HADS-D
Healey *et al*, 2008[95]	UK	49	28	79	Stroke	BDI non-somatic
Hedayati *et al*, 2006[96]	US	98	44	57	Haemodialysis	CES-D BDI
Hermanns *et al*, 2006[97]	Germany	376	148	52	Diabetes	BDI CES-D
Herrero *et al*, 2003[98]	Spain	385	181	38	Medically ill (general medical)	HADS-D
Hickie *et al*, 1987[99]	US	39	Not reported	Not reported	Medically ill (general medical)	GDS
Hopko *et al*, 2007[100]	US	33	25	54	Cancer	HAM-D BDI CES-D
Hoyl *et al*, 1999[101]	US	74	2	74	Medically ill (general medical)	GDS-15
Hughson *et al*, 1988[102]	UK	75	75	51	Cancer	GHQ-28
Ibbotson *et al*, 1994[103]	UK	513	282	Not reported	Cancer	HADS-D GHQ-28

Appendix 2 continued. Summary characteristics of included studies

Study	Country	Patients	Female	Mean age, years	Physical health condition	Instrument
Jackson and Baldwin, 1993[104]	US	59	Not reported	77	Medically ill (general medical)	GDS GDS-15
Jefford *et al*, 2004[105]	US	100	Not reported	Not reported	Cancer	1-item
Johnson *et al*, 1995[106]	Australia	204	Not reported	71	Stroke	GDS HADS-D GHQ-28
Katz *et al*, 2004[107]	Canada	60	13	61	Cancer	BDI HADS-D CES-D
Kawase *et al*, 2006[108]	Japan	305	Not reported	62	Cancer	1-item
Koenig *et al*, 1992[109]	US	109	0	74	Medically ill (general medical)	GDS GDS-15
Kugaya *et al*, 1998[110]	Japan	128	48	61	Cancer	HADS-D
Lam *et al*, 2004[111]	Hong Kong	100	56	69	Medically ill (general medical)	HADS
Lamers *et al*, 2008[112]	Netherlands	713	350	71	Medically ill (general medical)	PHQ-9
Laska *et al*, 2007[113]	Sweden	89	40	74	Stroke	MADRS
LeFevre *et al*, 1999[114]	UK	79	35	70	Cancer	HADS-D
Lee *et al*, 2008[115]	China	253	94	Not reported	Stroke	GDS-15
Leentjens *et al*, 2000[116]	Netherlands	53	Not reported	67	Parkinson's disease	BDI HDRS MADRS
Leung *et al*, 1998[117]	Taiwan	50	Not reported	54	Chronic medical disorders	Zung
Lightbody *et al*, 2007[118]	UK	28	Not reported	72	Stroke	MADRS
Lincoln *et al*, 2003[119]	UK	143	70	66	Stroke	BDI GHQ-28
Lloyd-Williams *et al*, 2000[120]	UK	100	56	57	Cancer	EPDS 1-item
Lloyd-Williams *et al*, 2001[121]	UK	100	56	57	Cancer	HADS-D
Lloyd-Williams *et al*, 2004[122]	UK	74	37	68	Cancer	1-item EPDS
Love *et al*, 2002[123]	Australia	303	303	Not reported	Cancer	HADS-D
Love *et al*, 2004[124]	Australia	227	227	52	Breast cancer	HADS BDI non-somatic
Low *et al*, 2007[125]	Canada	119	30	63	Myocardial infarction or angina	BDI GDS
Lowe *et al*, 2004[126]	Germany	501	167	42	Medically ill (general medical)	HADS-D PHQ-9 PHQ-2
Lustman *et al*, 1997[127]	US	172	83	48	Diabetes	BDI
Lykouras *et al*, 1996[128]	Greece	107	57	43	Neurological disorder	GHQ-28
Magni *et al*,1986[129]	Italy	220	109	76	Medically ill (general medical)	GDS
McManus *et al*, 2005[130]	US	1024	184	67	Chronic heart disease	PHQ-9 PHQ-2 Two stem questions CES-D
McQuillan *et al*,2003[131]	US	415	344	58	Rheumatoid Arthritis	CES-D
Meyer *et al*, 2003[132]	US	45	Not reported	Not reported	Cancer	1-item
Mitchell *et al*, 2008[133]	UK	129	Not reported	58	Cancer	PHQ-2 1-item
Mohr *et al*, 2007[134]	US	260	190	51	Multiple sclerosis	Two stem questions 1-item
Narding *et al*, 2002[135]	Netherlands	44	16	70	Stroke	HDRS
Neal *et al*, 1994[136]	UK	45	28	77	Medically ill (general medical)	GDS GDS-15
O'Rourke *et al*, 1998[137]	UK	105	Not reported	68	Stroke	HADS-D
Okimoto *et al*, 1982[138]	Japan	55	Not reported	Not reported	Medically ill (general medical)	Zung
Olden *et al*, 2009[139]	US	439	239	66	Cancer	HDRS
Ozalp *et al*, 2008[140]	Turkey	208	208	51	Cancer	HADS-D
Parikh *et al*, 1988[141]	US	80	40	58	Stroke	CES-D
Parker *et al*, 2002[142]	Australia	302	175	47	Medically ill (general medical)	HADS-D BDI non-somatic
Passik *et al*, 2001[143]	US	60	58	31	Cancer	Zung
Patterson *et al*, 2006[144]	US	310	37	40	HIV	BDI non-somatic

Appendix 2 continued. Summary characteristics of included studies

Study	Country	Patients	Female	Mean age, years	Physical health condition	Instrument
Payne *et al*, 2007[145]	US	167	Not reported	Not reported	Cancer	Two stem questions 1-item
Persoons *et al*, 2003[146]	Netherlands	97	64	48	Otolaryngology	PHQ-9
Picardi *et al*, 2005[147]	Italy	141	79	38	Dermatology	GHQ-12 PHQ-9
Pomeroy *et al*, 2001[148]	US	87	52	78	Medically ill (general medical)	GDS GDS-15 1-item
Poole *et al*, 2006[149]	US	115	47	43	Heart disease	HADS-D
Rapp *et al*, 1988[150]	US	150	48	69	Medically ill (general medical)	GDS BDI
Razavi *et al*, 1990[151]	Belgium	210	140	55	Cancer	HADS-D
Reuter and Harter, 2000[152]	Germany	188	51	54	Cancer	HADS-D GHQ-12
Rinaldi *et al*, 2003[153]	US	181	Not reported	79	Medically ill (general medical)	GDS-15
Roger *et al*, 2009[154]	US	67	35	71	Stroke	CES-D HDRS GDS-15
Rovner *et al*, 1997[155]	US	70	41	77	Visual impairment	GDS
Sagen *et al*, 2009[156]	Norway	104	43	65	Stroke	HADS-D MADRS
Scheinthal *et al*, 2001[157]	US	75	42	74	Medically ill (general medical)	GDS-15 BDI non-somatic
Schein *et al*, 1997[158]	US	76	35	70	Medically ill (general medical)	CES-D
Serrano-Duenas *et al*, 2008[159]	Ecuador	115	33	70	Parkinson's disease	HDRS
Shinar *et al*, 1986[160]	Israel	27	16	56	Stroke	CES-D
Silberman *et al*, 2006[161]	Brazil	46	19	68	Parkinson's disease	MADRS BDI
Singer *et al*, 2008[162]	Germany	250	23	Not reported	Cancer	HADS-D
Sivrioglu *et al*, 2009[163]	Turkey	85	53	59	Stroke	GDS
Stafford *et al*, 2007[164]	Australia	193	39	64	Heart disease	HADS-D PHQ-9
Strik *et al*, 2001[165]	US	206	49	59	Myocardial infarction	HADS-D HDRS BDI
Tang *et al*, 2004[166]	China	100	45	74	Stroke	GDS HADS-D GDS-15
Tang *et al*, 2004[167]	China	60	Not reported	Not reported	Stroke	GDS HADS-D
Thekkumpurath *et al*, 2009[168]	UK	150	86	70	Cancer	GHQ-12
Turner *et al*, 1984[169]	US	40	20	47	Chronic pain	Zung BDI BDI-non somatic
Upadhyaya *et al*, 1997[170]	UK	72	35	71	Medically ill (general medical)	HADS-D
Vahter *et al*, 2007[171]	Estonia	134	Not reported	44	Multiple sclerosis	1-item
Vargas *et al*, 2007[172]	Portugal	484	276	70	Medically ill (general medical)	GDS
Walker *et al*, 2007[173]	UK	361	238	Not reported	Cancer	HADS-D
Watnick *et al*, 2005[174]	US	62	20	63	Dialysis	PHQ-9 BDI
Weintraub *et al*, 2006[175]	US	148	Not reported	71	Parkinson's disease	GDS-15 HDRS
Wilhelm *et al*, 2004[176]	Australia	212	117	Not reported	Medically ill (general medical)	BDI non-somatic
Williams *et al*, 2005[177]	US	296	219	59	Stroke	PHQ-2 PHQ-9