Screening for depression in high-risk groups: prospective cohort study in general practice

Kim D. Baas, Karin A. Wittkampf, Henk C. van Weert, Peter Lucassen, Jochanan Huyser, Henk van den Hoogen, Eloy van de Lisdonk, Patrick E. Bindels, Claudi L. Bockting, Henricus G. Ruhé and Aart H. Schene

Background
Currently only about half of the people who have major depressive disorder are detected during regular health care. Screening in high-risk groups might be a possible solution.

Aims
To evaluate the effectiveness of selective screening for major depressive disorder in three high-risk groups in primary care: people with mental health problems, people with unexplained somatic complaints and people who frequently attend their general practitioner.

Method
Prospective cohort study among 2005 people in high-risk groups in three health centres in The Netherlands.

Results
Of the 2005 people identified, 1687 were invited for screening and of these 780 participated. Screening disclosed 71 people with major depressive disorder: 36 (50.7%) already received treatment, 14 (19.7%) refused treatment and 4 individuals did not show up for an appointment. As a final result of the screening, 17 individuals (1% of 1687) started treatment for major depressive disorder.

Conclusions
Screening for depression in high-risk populations does not seem to be effective, mainly because of the low rates of treatment initiation, even if treatment is freely and easily accessible.

Declaration of interest
None.

Population-based surveys have revealed that currently only about half of those people who have major depressive disorder are detected by regular healthcare.[1] This seems unacceptable because depression is associated with poor health status and poor quality of life.[2] Screening in general practice populations has been advocated as a solution to this 'hidden morbidity'.[3] However, the positive predictive value of available screening instruments is too low in unselected groups of individuals.[4] Screening in high-risk groups might be a possible alternative in reducing the likelihood of undisclosed depression.[5]

Furthermore, screening without feedback to general practitioners (GPs) has so far been of little value.[6,7] General practitioners have indeed diagnosed more people with major depressive disorder, but this has not translated into more interventions or better outcomes. Therefore, screening should be implemented as part of a disease-management programme with identification of individuals, structured assessment of diagnosis, feedback of the diagnosis to the physician and availability of evidence-based treatments.[7,8]

We defined the following high-risk groups in GP practices: individuals with mental health problems, individuals with unexplained somatic complaints and individuals who frequently attend their GP. Compared with the 10% prevalence of major depressive disorder for those presenting in a primary care setting,[9] the prevalence in these high-risk populations rises to about 20–30%.[10–12] So far, no studies have evaluated the usefulness and effectiveness of this type of selective screening.[1] We looked at how many people with undetected (or undiagnosed) major depressive disorder could be diagnosed and then be treated, when screening in predefined groups at high risk is applied in general practice. In this article the results of the initial screening, as part of a disease-management programme, in three high-risk groups are presented.

Method

Setting

The study was conducted as a prospective cohort study from January 2005 to March 2006 in primary care patients between 18 and 70 years of age, in two regions connected to academic settings in The Netherlands: the Academic Medical Centre in Amsterdam and the University Medical Centre St Radboud in Nijmegen. The study protocol was approved by the institutional ethics review committee of both centres.

Participants

Three high-risk groups were defined and selected from general practice populations.

People with mental health problems

All individuals who presented to their GP with a new mental health problem up to 3 months prior to the selection date were eligible. We selected individuals from the electronic patient database of the participating general practices. A time frame of 3 months was chosen because of the transitory nature of mental health problems. General practitioners in our sample coded all diagnoses or complaints with the International Classification of Primary Care (ICPC)[13,14] classification. With the use of chapters P and Z of the ICPC, individuals were identified with a psychological or social reason for seeing their GP or a mental health diagnosis. To identify all people with possible mental health problems, the electronic patient database was also searched with the predefined free text-words: anxiety, worrying, sadness, stress, feeling down, and insomnia.

People with unexplained somatic complaints

These were identified as follows: people with somatic complaints that could not be explained in biological terms (determined by the GP) and in whom the duration of these unexplained somatic complaints was at least 90 days according to the GP. General practitioners checked their appointment lists during the 4 weeks preceding study allocation and selected individuals fulfilling both criteria.

People frequently attending their general practitioner

The method as proposed by Howe[15] was used to identify frequently attending individuals: the 10% of most frequently consulting women and the 10% of most frequently consulting men in two age groups (18–44 and 45–70 years) in the year preceding study allocation. With this method, differences in gender and age among people who attended frequently were accounted for. We used computerised attendance data from all consultations, home visits and telephone consultations with doctors, nurses and other team members. The highest 10% was determined separately for each GP because of differences in practice styles.

Follow-up procedure

General practitioners received a list of individuals selected by the research team. They excluded those people who were already recognised by the GP as having major depressive disorder, schizophrenia, psychosis or bipolar disorder, or were considered by the GP as not qualified to enter the study because of specific somatic problems (e.g. blindness, terminal illness) or intellectual disability. A last exclusion criterion was having trouble with the Dutch or English language.

Subsequently, an invitation letter signed by their own GP was sent to the people on the final list. The letter describing the study also contained an informed consent form and the screening instrument, the Patient Health Questionnaire (PHQ).[16] If people did not respond within 2 weeks, a reminder was sent. A consecutive diagnostic interview for major depressive disorder (Structured Clinical Interview for DSM–IV Axis I disorders, SCID–I)[17] was performed with participants who had a PHQ–9 score $\geqslant 10$.

Participants diagnosed in our study with any psychiatric disorder by SCID–I, were advised to consult their GP and to discuss the interview results and possible treatment options. General practitioners were also informed about the test results. If individuals did not contact their GP, they were telephoned 2 weeks later, asked why they had not done so and advised to try again. In The Netherlands, treatment for psychiatric disorders is covered by health insurance and is easily accessible.

Measures

The Patient Health Questionnaire (PHQ) is a short self-report version of the Primary Care Evaluation of Mental Disorders (PRIME–MD).[16] The full PHQ screens for the five most common mental disorders using DSM–IV[18] criteria: depressive disorders (major depressive disorder and dysthymia); anxiety disorders (panic disorder and generalised anxiety disorder); alcohol abuse; somatoform disorder; and eating disorders (bulimia nervosa and binge eating disorder). The PHQ–9 is the 9-item sub-scale for major depressive disorder. For identifying major depressive disorder we did not use the diagnostic algorithm, but a score $\geqslant 10$ as this cut-off point proved to be more sensitive.[19,20] Participants with a score $\geqslant 10$ were assessed by telephone clinical

interview (SCID–I) to make a definite diagnosis of major depressive disorder.

The SCID–I is a semi-structured interview designed for diagnosing mental disorders according to DSM–IV criteria.[17] Agreement between diagnosis gained from telephone and live administration of the SCID–I has been found to be excellent (Kappa=0.73 with 90% agreement).[21] The SCID–I was administered by researchers that received SCID–I training from a trained professional. Throughout the study all interviewers had ongoing training sessions supervised by an expert psychiatrist (J.H.). They also had monthly consensus meetings to maximise accuracy and consistency in the administration of the SCID–I. If the results of the mood disorder section of the SCID–I indicated a major depressive disorder, subsequent sections were checked in order to identify people with schizophrenia, bipolar disorder and psychosis. Individuals diagnosed with major depressive disorder were asked if they already received treatment, defined as pharmacotherapy with antidepressants and/or psychotherapy and/or structured and supportive care from the GP or from mental healthcare professionals.

Results

Identification of three high-risk groups

Three health centres with 12 GPs and 15 996 people on their lists participated. In total, 2005 people (12.5%) fulfilled the criteria for: frequently attending their GP ($n = 1258$); mental health problems ($n = 970$) and/or unexplained somatic complaints ($n = 107$) (Fig. 1). A total of 311 people belonged to two or three of the high-risk groups (Fig. 2).

General practitioners excluded 318 (15.9%) individuals from screening: 146 (7.3%) were already known to have major depressive disorder by their GP and 172 (8.6%) were excluded for other reasons: diagnosis of schizophrenia or bipolar disorder ($n = 18$, 0.9%), terminal illness or intellectual disability ($n = 44$, 2.2%), and inability to understand Dutch or English ($n = 110$, 5.5%).

Effectiveness of screening

Of the 1687 people eligible for screening with the PHQ, 826 (49.0%) returned the questionnaire and of these 780 (94.4%) gave informed consent (Table 1). Individuals who gave consent were older, more often female and individuals from the mental health problems group were over-represented. Of the 780 people who gave consent, 226 (29%) had a PHQ–9 score $\geqslant 10$ (Fig. 1); of these, 173 (76.5%) participated in the SCID–I interview. Reasons for non-participation (n=53) despite consent are described in Fig. 1.

Of the 173 SCID–I interviewed participants, 94 (54%) did not meet the criterion for major depressive disorder and 71 (41%) did. Another eight people (5%) met major depressive disorder criteria but also had psychosis and/or bipolar disorder according to DSM–IV. Thus, screening disclosed 79 additional individuals with major depressive disorder (9.6% of 826 who returned the PHQ and 4.7% of 1687 who were invited to participate). Neither age, gender, ethnicity nor high-risk group influenced the chance of diagnosing major depressive disorder, given a positive PHQ screening (Table 2).

Of the 71 individuals identified with major depressive disorder, 36 (51%) already received treatment. Of the 35 (49%) new cases eligible for treatment, 14 (40%) refused treatment after they received information about the diagnosis and treatment from their GP. Despite repeatedly made appointments, another four individuals (11%) did not show up for an appointment with their GP.
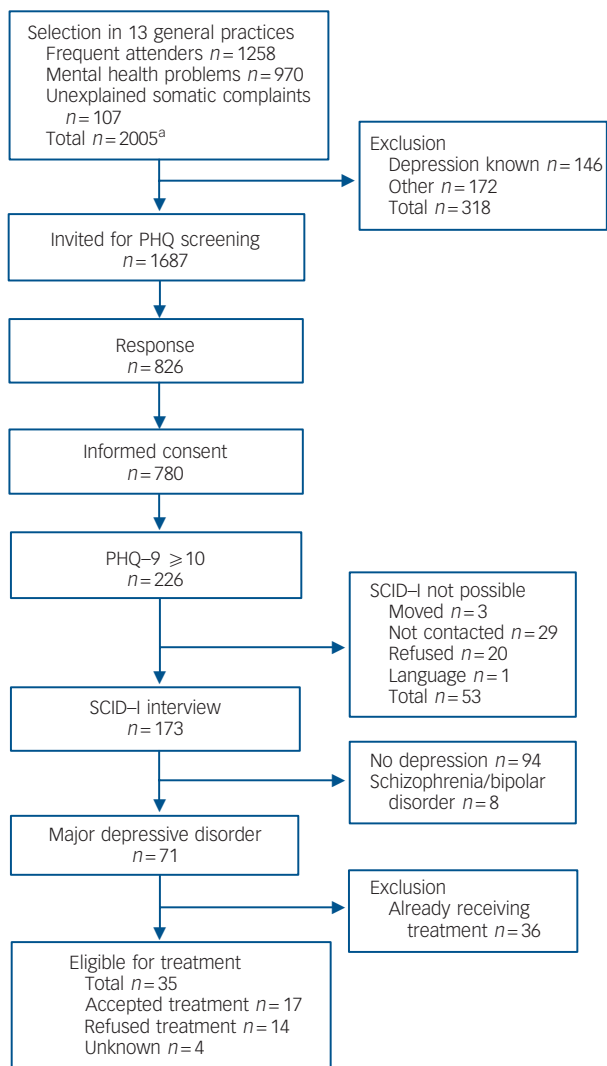
Selection in 13 general practices
Frequent attenders $n = 1258$
Mental health problems $n = 970$
Unexplained somatic complaints
$n = 107$
Total $n = 2005^a$

Exclusion
Depression known $n = 146$
Other $n = 172$
Total $n = 318$

Invited for PHQ screening
$n = 1687$

Response
$n = 826$

Informed consent
$n = 780$

PHQ–9 $\geqslant 10$
$n = 226$

SCID–I not possible
Moved $n = 3$
Not contacted $n = 29$
Refused $n = 20$
Language $n = 1$
Total $n = 53$

SCID–I interview
$n = 173$

No depression $n = 94$
Schizophrenia/bipolar
disorder $n = 8$

Major depressive disorder
$n = 71$

Exclusion
Already receiving
treatment $n = 36$

Eligible for treatment
Total $n = 35$
Accepted treatment $n = 17$
Refused treatment $n = 14$
Unknown $n = 4$

Fig. 1 Flow chart for the screening programme as the first step in a disease-management programme.

PHQ, Patient Health Questionnaire; PHQ–9, Patient Health Questionnaire – nine-item subscale; SCID–I, Structured Clinical Interview for DSM–IV Axis I disorders.
a. Smaller number because of overlap between groups, see Fig. 2.

Fig. 2 Venn diagram showing absolute and relative (% of all) numbers of three high-risk groups (total $n = 2005$).

major depressive disorder is 99 (17/1687), the number needed to screen is 46 (17/780) and the number needed to screen to treat 1 not yet diagnosed major depressive disorder was 118 (17/2005).

Discussion

We have reported on the yield of the initial steps in a disease-management programme for major depressive disorder, consisting of selective screening of patients belonging to three high-risk groups, followed by structured diagnosis and feedback, followed by offering treatment. Among 2005 identified people at high risk, the selective screening diagnosed 71 new individuals with major depressive disorder. Of these, 36 already received some kind of treatment, 14 refused treatment and 17 accepted treatment. We concluded that with the need to screen 118 people to treat 1 additional person with depression, this type of screening has no or only very limited effect.

The effect could have been greater if the number of detected individuals refusing treatment ($n=14$) could have been reduced. General practitioners gave individuals psychoeducation for depression and all of them received a brochure (provided by the Dutch College of General Practice) about depression. We think that this amount of information about depression is adequate

At the final result of the screening and diagnostic phase of our programme, 17 individuals were eligible to receive treatment for major depressive disorder. To treat these 17 people, we had to invite 1687 people for screening, 780 of whom participated in screening. The number needed to invite to treat 1 additional
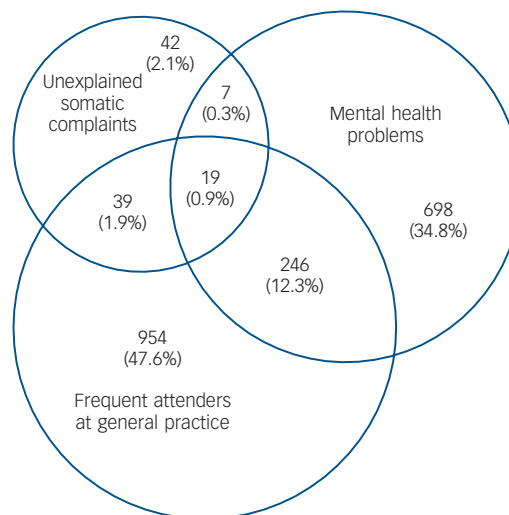
Table 1 Characteristics of those invited to participate who gave informed consent for screening and those that refused to participate

Participant characteristics	Informed consent ($n = 780$)	Non-responders or no informed consent ($n = 907$)	Total ($n = 1687$)	P
Age, years: mean (s.d.)	44.77 (12.72)	41.52 (12.79)	43.02 (12.85)	<0.0001[a]
Female, % (n)	63.9 (499)	53.5 (485)	58.33 (984)	<0.0001[b]
High-risk group, % (n)				
Unexplained somatic complaints	6.2 (48)	5.4 (49)	5.8 (97)	0.50[c]
Mental health problems	49.6 (387)	44.3 (402)	46.8 (789)	0.03[c]
Frequently attend their GP	60.6 (472)	64.9 (589)	62.9 (1061)	0.06[c]

GP, general practice.
a. P represents significance of independent samples t-test.
b. P represents significance of χ^2-test.
c. P represents significance of χ^2-tests of the difference in individuals who gave informed consent v. those that refused participation in the specified high-risk group v. the other two groups (unexplained somatic complaints v. frequently attend their GP and/or mental health problems; frequently attend their GP v. unexplained somatic complaints and/or mental health problems; mental health problems v. frequently attend their GP and/or unexplained somatic complaints).

Table 2 Participants who scored 10 or higher on the PHQ–9: comparison of those with major depressive disorder and those without major depressive disorder according to SCID–I

Participant characteristics	Major depressive disorder (n=79)	No major depressive disorder (n=94)	Total	P
Age, years: % (n)				0.36[a]
18–24	5.4 (4)	10.8 (10)	8.4 (14)	
25–44	48.7 (36)	40.8 (38)	44.3 (74)	
45–64	45.9 (34)	48.4 (45)	47.3 (79)	
Gender, % (n)				0.23[a]
Female	77.2 (61)	69.2 (65)	72.8 (126)	
Male	22.8 (18)	30.8 (29)	27.2 (47)	
Ethnicity, % (n)				0.97[a]
White	78.4 (58)	78.6 (70)	78.5 (128)	
Other	21.6 (16)	21.4 (19)	21.5 (35)	
High-risk group, % (n)				
Unexplained somatic complaints	10.1 (8)	10.6 (10)	10.4 (18)	0.91[b]
Mental health problems	56.9 (45)	58.5 (55)	57.8 (100)	0.83[b]
Frequently attend their GP	55.1 (43)	57.5 (54)	56.4 (97)	0.76[b]

GP, general practitioner; SCID–I, Structured Clinical Interview for DSM–IV Axis I disorders.
a. P represents significance of χ^2-test.
b. P represents significance of χ^2-tests of the difference in SCID–I positive v. SCID–I negative individuals in the specified high-risk group v. the other two groups (unexplained somatic complaints v. frequently attend their GP and/or mental health problems; frequently attend their GP v. unexplained somatic complaints and/or mental health problems; mental health problems v. frequently attend their GP and/or unexplained somatic complaints).

for making a decision about treatment. The people who refused treatment were well informed.

Another issue is the possible health gain that could have been achieved by the treatment of the 17 participants who accepted treatment. Considering that the natural course of depression in general practice is about 3 months,[22] that the severity of depression of these 17 was moderate (mean Hamilton Rating Scale for Depression score 19), that the number needed to treat for anti-depressants is between 6 and 10, although recent papers show that there seems little evidence to support the prescription of anti-depressant medication to individuals other than the most severely depressed,[23] we conclude that treatment of these 17 participants would have resulted in small health gains in comparison with care as usual. Furthermore, it has been shown[24,25] that a number of missed cases are diagnosed at later visits, although others have suggested that some of the low severity cases found by screening are false positives because all diagnostic measures have rating errors.[26]

One can speculate whether our predefined high-risk groups were adequately chosen. We think they were, although we could not replicate the 20–30% major depressive disorder prevalence mentioned above. We found a total prevalence of 18%: 7% (n=146) that were already known by their GP as having major depressive disorder and another 11% with not yet diagnosed or undetected major depressive disorder (extrapolated from the cases found by screening the total high-risk group), figures in accordance with some earlier research.[7,27,28]

A main criticism of screening in primary care is the cross-sectional nature of screening. In a longitudinal study, as discussed above, most of the individuals with undisclosed depression might be identified at a later date. A number of our predefined high-risk individuals were already visiting their GP regularly. Despite this fact, major depressive disorder was not yet diagnosed which at least partly parries this criticism.

Half of the patients did not respond to the letter inviting them to participate despite them being invited by their own GP to enhance participation. In comparison with other screening programmes (e.g. cervical cancer (about 80% response), breast cancer (90–95% response)) this response rate is quite low. This might be caused by the nature of the targeted disorder or the nature of the screening instrument. However, participation of about 50% is not unusual in psychiatric programmes.[29]

Furthermore, a recent survey showed that patient inclusion in research when the GP was the patient's informant about the study was less successful than when the patient received a mailed letter about the study signed by their GP.[30]

Finally, even with the methods we used, participating GPs felt the screening procedure placed an additional high burden on their regular practice and they experienced disturbance of their daily routine.

Of the people who responded, 94.4% gave informed consent. As could be expected, response rates were higher in the group of individuals with mental health problems than in the other groups. Individuals with mental health problems addressed their GP because of the need for treatment.[31] At this time we do not have a valid explanation for the fact that people were willing to go through the complete procedure of screening and diagnosis, but refused treatment. One explanation is that they did not fully understand the screening procedure and/or its possible consequences. Another explanation might be that people who screened positive, but did not agree to treatment, did not accept a psychiatric diagnosis or label or were not in a hurry to seek help.[31] This phenomenon is new and needs to be further explored.

Limitations

A limitation might be that we assumed a 100% sensitivity of the PHQ–9. We did not attempt to detect depression among those people with a score <10. Therefore, we have probably missed some individuals with major depressive disorder. With a 0.88 sensitivity of the PHQ–9 for major depressive disorder, we might have missed 18 individuals with depression, resulting in 4 ($16/71 \times 18$) additional people accepting treatment. This would not have changed our results significantly because the PHQ also measures severity of depression, and with a score smaller than 10 these individuals are likely to have less severe depression. Moreover, higher sensitivity would result in lower specificity, thereby dramatically increasing the number of false positives.[32]

Implications for practice

Screening of high-risk groups for depression in general practice is feasible when the number of newly detected people is the outcome. It is not feasible when the number of individuals that

actually will receive treatment (and might be helped) is the primary outcome. Therefore, in contrast with prior suggestions, the problem is not so much a lack of resources or training and education of GPs, but the unwillingness of individuals to accept treatment for a condition they did not present with to their GP. The key to improving care lies, therefore, not in recognising undisclosed depression only, but in convincing individuals to initiate and thereafter continue with evidence-based treatment for major depressive disorder. Recently, more thorough organisational strategies such as case management have been recommended by the National Institute for Health and Clinical Excellence.[5] Recent reviews and randomised controlled trials have established the effectiveness of these organisational strategies.[33–35] Future studies should focus on the effectiveness of case management in these specific high-risk groups. A non-medical approach that at first abstains from the label of depression has to be considered here.

Screening for depression in high-risk populations in general practice does not appear to be effective even if treatment is freely and easily accessible for two reasons: the total amount of time to execute the screening procedure (including quite a lot of GP time) and the low rates of treatment initiation in combination with the limited health gain per treated case.

Kim D. Baas, MSc, Department of Psychiatry, Academic Medical Center, University of Amsterdam; **Karin A. Wittkampf**, MD, Department of Psychiatry and Department of General Practice, Academic Medical Center, University of Amsterdam; **Henk C. van Weert**, MD, PhD, Department of General Practice, Academic Medical Center, University of Amsterdam; **Peter Lucassen**, MD, PhD, Department of General Practice, Radboud University Nijmegen, Nijmegen; **Jochanan Huyser**, MD, PhD, Department of Psychiatry, Academic Medical Center, University of Amsterdam; **Henk van den Hoogen, Eloy van de Lisdonk**, MD, PhD, Department of General Practice, Radboud University Nijmegen; **Patrick E. Bindels**, MD, PhD, Department of General Practice, Academic Medical Center, University of Amsterdam; **Claudi L. Bockting**, PhD, Department of General Practice, Academic Medical Center, University Amsterdam; **Henricus G. Ruhé**, MD, PhD, **Aart H. Schene**, MD, PhD, Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands.

Correspondence: Aart H. Schene, Department of Psychiatry, Academic Medical Center, (Room PA1-132), Meibergdreef 5 1105 AZ Amsterdam, The Netherlands. Email: a.h.schene@amc.uva.nl

First received 10 Oct 2007, final revision 17 Apr 2008, accepted 17 Jun 2008

References

1 Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. *Cochrane Database Syst Rev* 2005; **2**: CD002792.

2 Rost K, Zhang M, Fortney J, Smith J, Coyne J, Smith GR, Jr. Persistently poor outcomes of undetected major depression in primary care. *Gen Hosp Psychiatry* 1998; **20**: 12–20.

3 Wright AF. Should general practitioners be testing for depression? *Br J Gen Pract* 1994; **44**: 132–5.

4 Williams JW, Jr, Pignone M, Ramirez G, Perez SC. Identifying depression in primary care: a literature synthesis of case-finding instruments. *Gen Hosp Psychiatry* 2002; **24**: 225–37.

5 National Institute for Health and Clinical Excellence. Clinical Guideline 23. Depression: management of depression in primary and secondary care. NICE, 2004 (http://guidance.nice.org.uk/CG23/guidance/pdf/English).

6 Gilbody SM, House AO, Sheldon TA. Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ* 2001; **322**: 406–9.

7 Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, et al. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002; **136**: 765–76.

8 Gilbody S, Sheldon T, Wessely S. Should we screen for depression? *BMJ* 2006; **332**: 1027–30.

9 Bauer M, Bschor T, Pfennig A, Whybrow PC, Angst J, Versiani M, et al. World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for biological treatment of unipolar depressive disorders in primary care. *World J Biol Psychiatry* 2007; **8**: 67–104.

10 Dowrick CF, Bellon JA, Gomez MJ. GP frequent attendance in Liverpool and Granada: the impact of depressive symptoms. *Br J Gen Pract* 2000; **50**: 361–5.

11 Katon W, Von Korff M, Lin E, Lipscomb P, Russo J, Wagner E, et al. Distressed high utilizers of medical care. DSM–III–R diagnoses and treatment needs. *Gen Hosp Psychiatry* 1990; **12**: 355–62.

12 Gill D, Sharpe M. Frequent consulters in general practice: a systematic review of studies of prevalence, associations and outcome. *J Psychosom Res* 1999; **47**: 115–30.

13 Gebel RS, Lamberts H (eds). *ICPC–1 met Nederlandse subtitels*. Nederlands Huisartsen Genootschap, 2000.

14 Lamberts H, Wood M (eds). *ICPC International Classification of Primary Care*. Oxford University Press, 1987.

15 Howe A, Parry G, Pickvance D, Hockley B. Defining frequent attendance: evidence for routine age and sex correction in studies from primary care settings (brief report). *Br J Gen Pract* 2002; **52**: 561–2.

16 Spitzer RL, Kroenke K, Williams JB; and the Patient Health Questionnaire Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999; **282**: 1737–44.

17 First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM–IV Axis I Disorders Patient Edition (SCID–I/P version 2.0). (Translated into Dutch by eds MAC Groenestijn, GW Akkerhuis, RW Kupka, N Schneider, WA Nolen). Swets & Zeitlinger B.V,1999.

18 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorder (4th edn) (DSM–IV)*. APA, 1994.

19 Kroenke K, Spitzer RL, Williams JB. The PHQ–9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; **16**: 606–13.

20 Wittkampf KA, Naeije L, Schene AH, Huyser J, van Weert HC. Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *Gen Hosp Psychiatry* 2007; **29**: 388–95.

21 Simon GE, Revicki D, VonKorff M. Telephone assessment of depression severity. *J Psychiatr Res* 1993; **27**: 247–52.

22 van Weel-Baumgarten E, van den BW, van den HH, Zitman FG. Ten year follow-up of depression after diagnosis in general practice. *Br J Gen Pract* 1998; **48**: 1643–6.

23 Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008; **5**: e45.

24 Kendrick T. Depression in adults: GPs are not so bad at diagnosis (letter). *BMJ* 2008; **336**: 522.

25 Kessler D, Bennewith O, Lewis G, Sharp D. Detection of depression and anxiety in primary care: follow up study. *BMJ* 2002; **325**: 1016–7.

26 Thompson C, Ostler K, Peveler RC, Baker N, Kinmonth AL. Dimensional perspective on the recognition of depressive symptoms in primary care. The Hampshire Depression Project 3. *Br J Psychiatry* 2001; **179**: 317–23.

27 Ormel J, Koeter MW, van den BW, van de WG. Recognition, management, and course of anxiety and depression in general practice. *Arch Gen Psychiatry* 1991; **48**: 700–6.

28 Simon GE, Goldberg D, Tiemens BG, Ustun TB. Outcomes of recognized and unrecognized depression in an international primary care study. *Gen Hosp Psychiatry* 1999; **21**: 97–105.

29 Greil W, Ludwig-Mayerhofer W, Steller B, Czernik A, Giedke H, Muller-Oerlinghausen B, et al. The recruitment process for a multicenter study on the long-term prophylactic treatment of affective disorders. *J Affect Disord* 1993; **28**: 257–65.

30 van der Wouden JC, Blankenstein AH, Huibers MJ, van der Windt DA, Stalman WA, Verhagen AP. Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research. *J Clin Epidemiol* 2007; **60**: 819–24.

31 Arroll B, Goodyear-Smith F, Kerse N, Fishman T, Gunn J. Effect of the addition of a 'help' question to two screening questions on specificity for diagnosis of depression in general practice: diagnostic validity study. *BMJ* 2005; **331**: 884.

32 Henkel V, Mergl R, Kohnen R, Maier W, Moller HJ, Hegerl U. Identifying depression in primary care: a comparison of different methods in a prospective cohort study. *BMJ* 2003; **326**: 200–1.

33 Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organizational interventions to improve the management of depression in primary care: a systematic review. *JAMA* 2003; **289**: 3145–51.

34 Von Korff M, Goldberg D. Improving outcomes in depression (editorial). *BMJ* 2001; **323**: 948–9.

35 Gensichen J, Beyer M, Muth C, Gerlach FM, Von Korff M, Ormel J. Case management to improve major depression in primary health care: a systematic review. *Psychol Med* 2006; **36**: 7–14.