A randomised controlled trial of the effectiveness of a program for early detection and treatment of depression in primary care

A randomised controlled trial of the effectiveness of a program for early detection and treatment of depression in primary care

A. Picardi[1], I. Lega[1], L. Tarsitani[2], M. Caredda[2], G. Matteucci[2], M.P. Zerella[2], R. Miglio[3], A. Gigantesco[1], M. Cerbo[4], A. Gaddini[5], F. Spandonaro[6], M. Biondi[2] and the SET-DEP Group

1 Mental Health Unit, Centre of Epidemiology, Surveillance and Health Promotion, Italian National Institute of Health, Rome, Italy

[2] Department of Psychiatric Sciences and Psychological Medicine, 'Sapienza' University of Rome, Rome, Italy

[3] Department of Statistics, University of Bologna, Italy

[4] National Agency for Regional Health Services, Rome, Italy

[5] Agency for Public Health, Lazio Region, Italy

[6] 'Tor Vergata' University of Rome, Rome, Italy

The SET-DEP Group includes: *Principal Investigator*: A. Picardi. *Co-investigators*: M. Biondi, M. Cerbo, A. Gaddini, A. Gigantesco, F. Spandonaro. *Research Associates*: M. Caredda, A. Cavallo, A. Crescenzi, I. Lega, G. Matteucci, B. Polistena, L. Tarsitani, M.P. Zerella. *Research Assistants*: I. Berardelli, I. Cascavilla, D. Del Re, C. Fini, L. Leoncini. *Primary Care Physicians*: S. Baccarini, F. D'Andrea, M. Di Fonso, R. Dolfi, A. Guerani, D. Mandolini, G. Marri, L. Pagano, M. Pagliarini, M.L. Paoletti, A. Pes, S. Raspa, M. Sabatini, U. Salomone, P. Ventura.

Corresponding author: Angelo Picardi, MD

Italian National Institute of Health

Centre of Epidemiology, Surveillance and Health Promotion

Mental Health Unit

Viale Regina Elena, 299 - 00161 Rome, Italy

Phone: +39 06 49904200 Fax: +39 06 49904182

E-mail: angelo.picardi@iss.it

ABSTRACT

Objective. There is considerable uncertainty about whether depression screening programs in primary care may improve outcomes and what specific features of such programs may contribute to success. We tested the effectiveness of a program involving substantial commitment from local mental health services.

Methods. Prospective, randomised, patient- and evaluator-masked, parallel-group, controlled study. Participants were recruited in several urban primary care practices where they completed the PC-SAD screener and WHOQOL-Bref. Those who screened positive and did not report suicidal ideation (N=115) were randomised to an intervention group (communication of the result and offer of psychiatric evaluation and treatment free of charge; N=56) or a control group (no feedback on test result for 3 months; N=59). After 3 months, 100 patients agreed to a follow-up telephone interview including the administration of the PC-SAD5 and WHOQOL-Bref.

Results. Depression severity and quality of life improved significantly in both groups. Intent-to-treat analysis showed no effect of the intervention. As only 37% of patients randomised to the intervention group actually contacted the study outpatient clinic, we performed a per-protocol analysis to determine whether the intervention, if delivered as planned, had been effective. This analysis revealed a significant positive effect of the intervention on severity of depressive symptoms, and on response and remission rate. Complier average causal effect analysis yielded similar results.

Conclusion. Due to the relatively small sample size, our findings should be regarded as preliminary and have limited generalizability. They suggest that there are considerable barriers on the part of many patients to the implementation of depression screening programs in primary care. While such programs can be effective, they should be designed based on the understanding of patients' perspectives.

Key words: Depression; screening; primary care; mental health services.

Introduction

Depression has major public health implications. It is relatively frequent in the general population [1], is among the leading causes of disease burden worldwide [2], imposes a huge economic burden [3], and is associated with worse prognosis in patients with comorbid physical illness [4, 5].

Primary care (PC) visits represent important opportunities to detect and initiate treatment of depression. Typically, depressed persons in the community are more likely to seek care in general medical settings than in specialty mental health (MH) settings. The prevalence of depression is higher in PC than in the general population [6], and PC physicians are often the first, and not rarely the only, health professionals that are contacted by depressed patients [7].

However, the accuracy of depression recognition by PC practitioners and other non-psychiatrist physicians is not satisfactory [8]. Several factors, such as short consultation time, low clinician confidence and interviewing skills, patient's reluctance to disclose or low awareness of emotional symptoms, difficulty to distinguish somatic symptoms of depression from those of various medical illnesses, and competing demands on the PC physician, may account for the under-recognition of depression in PC [9].

Thus, steps are needed to improve recognition and treatment of depression in PC. As effective treatments and valid case-finding instruments are available, depression meets common criteria for screening. However, the evidence base supporting depression screening in PC is thin [10]. Indeed, screening alone does not result in improved care and outcomes. Although routine feedback of the results of screening to clinicians results in a marginal increase in the rate of diagnosis of depression, increased recognition does not translate into significant improvement in patients' outcomes [11].

Research suggests that, for screening to be useful, it needs to be part of a package of enhanced care. Although no study provided a completely rigorous test of the effectiveness of depression screening programs in PC [10], there is evidence supporting the health benefits of programs in which other staff (e.g., nurse specialists, trained therapists) provide part of the depression care, or efforts are made to enrol patients in specialty MH treatment [12]. Three randomized controlled trials on general adult populations reported that screening programs involving staff support in depression care reduce depressive symptoms to a greater extent than usual care. These studies included elements such as providing PC physicians with a detailed depression treatment protocol during the visit that included patient educational material and logistical support from other staff [13], training of PC physicians and nurses in the detection and management of depression [14], PC physicians training and specialty MH provider participation in depression care [15]. Another study on older adults reported that a screening program combined with logistical support for the PC physician and

the involvement of other staff to provide some depression care was effective in improving depression beyond usual care [16]. However, other studies performed on general adult [17] or older adult [18-20] populations did not find a superiority of depression screening programs over usual care, despite the inclusion of elements such as logistic support for the PC physician, interventions to improve quality of provider's care, and inclusion of other staff to provide some depression care. Overall, while it is reassuring that the largest studies [14-16] yielded positive results, the picture emerging from the available studies is not univocal. Also, considerable uncertainty remains regarding what specific features of a program contribute to its success.

This study aimed at testing the feasibility and effectiveness of a program for early detection and treatment of depression in PC called 'SET-DEP' (Screening and Enhanced Treatment for DEpression in Primary care). The program was designed taking into account three main issues: a) the U.S. Preventive Services Task Force recommendation to screen only when staff-assisted depression care supports are in place; b) the need to impose only a very limited burden on PC physicians; c) the effort to arrange a collaboration between PC and MH services as similar as possible to what would occur in a 'real world' Italian setting. The program involved substantial commitment from a local MH service that provided the resources to ensure the execution of screening and the treatment of all patients who screened positive.

As all previous studies on the topic [10], this study does not provide a completely rigorous test of screening. Such a study should be able to capture all the aspects involved in a screening program that provides benefit over usual care: 1) patients must be approached for screening among all the other duties that a busy clinic must attend to; 2) the patients must agree to be screened; 3) the screening procedure must be able to detect patients with undetected and untreated depression that would not be identified under usual care; 4) effective treatment must be available and accessible, and patients must engage in treatment; 5) treatment must lead to a large enough improvement to justify the costs and possible harms of screening. The present study does not capture all these aspects, because only patients screening positive rather than all screened patients were included, and randomization occurred after screening took place. However, it did address some of these components, particularly the fourth and the fifth, and it can help determine if such a program is beneficial for people screening positive.

Methods

Study setting and procedures

This was a prospective, randomised, patient- and evaluator-masked, parallel-group, controlled study. The participants were recruited at 13 urban general internal medicine PC practices, located in

central Rome, Italy, from January 2009 to June 2010. The study inclusion criteria were age 18-65, and absence of psychosis or severe cognitive impairment as clinically determined. The eligible patients were informed about the study by PC physicians and their secretaries; fliers describing the study were also posted in the waiting room.

Patients willing to participate provided written informed consent to take part in the study according to procedures approved by the Italian National Institute of Health ethical committee. They were given an envelope containing a sociodemographic form and the Primary Care Screener for Affective Disorders (PC-SAD) and WHOQOL-Bref questionnaires to complete. Participants placed the completed questionnaires back in the envelope, and they put it in a transparent drop box located in the waiting room.

The PC-SAD was scored through an automated system by a researcher who was not involved in subsequent assessments. The participants who screened positive and had a PC-SAD result suggestive of suicidal ideation were excluded from the study, were immediately informed of their positive screening result, and were offered psychiatric evaluation and treatment free of charge. By using a computer-generated simple randomisation list, all the other participants with a positive screen were randomly assigned in a 1:1 ratio to either an intervention group (communication of the screening result and offer of psychiatric evaluation and treatment free of charge) or a control group that was temporarily not given feedback on screening. A researcher who was not involved in assessment or treatment prepared and maintained the concealed randomisation list.

Participants assigned to the intervention group were immediately sent a letter informing them that they screened positive, and that they could contact an outpatient clinic that had been specifically set up for the study in order to provide psychiatric evaluation and treatment to study participants, free of charge. This clinic was established at the psychiatric department of the ‘Policlinico Umberto I’ university hospital, which is located in the same health district as the PC practices and thus was easily accessible to participants. After two weeks, the participants who had still not contacted the clinic were sent another similar letter as a reminder.

The participants who contacted the outpatient clinic were offered an appointment to be seen within one week. They underwent a thorough psychiatric examination and were diagnosed according to DSM-IV criteria. All patients who received an Axis I diagnosis, irrespective of the specific diagnosis, were offered free psychiatric treatment as appropriate. Treatment included pharmacotherapy and clinical management (reviewing the patient's clinical status; providing the patient with support and advice if necessary; monitoring medication titration and tapering at the start and at the end of treatment, respectively), and it was maintained for as long as clinically required.

After three months from PC-SAD completion, all randomized patients were contacted and asked to participate in a telephone interview. A psychiatrist who was masked to group allocation performed the interview, which included the administration of the PC-SAD5 and WHOQOL-Bref. The patients with follow-up results suggestive of a mental disorder were advised to contact the outpatient clinic. Thus, all patients randomized to the control group with psychiatric needs were provided an opportunity to be evaluated and treated at no charge.

Instruments

The PC-SAD [21] is a short, self-administered questionnaire, which consists of a 3-item pre-screener including one Dysthymic Disorder (Dys) question and two MDD questions (which reduce respondent burden by terminating the questionnaire if all are negative), a 26-item MDD section, and an 8-item Dys section. The PC-SAD breaks down each DSM-IV symptom of MDD into several simple items, and it is scored using an automated system that integrates the answers mathematically. The scoring algorithm is built in a way that the presence of each symptom can be determined independently from the presence of missing answers to one or more relevant items, provided that at least one item related to the symptom has been answered. The Dys section is scored in a similar way. We used the validated Italian version of the PC-SAD [22].

The 5-item version of the PC-SAD (PC-SAD5) is a short depression severity measure derived from the PC-SAD. In the scoring procedure, each item is given equal weight, and the total score is converted to a 0-100 linear scale, with higher scores indicating greater depression severity. The PC-SAD5 showed satisfactory internal consistency, and displayed significant correlations with the Montgomery Åsberg Depression Rating Scale and the 9-item depression module of the Patient Health Questionnaire. Also, it discriminated between patients with MDD, patients with milder forms of depression, and patients with no mental disorders [23].

The WHOQOL-Bref is a patient-centred instrument developed by the World Health Organization, which explores various dimensions of quality of life (QOL) considered as the most important across cultures and different disease conditions. It includes 26 items tapping 4 main domains (physical domain, psychological domain, social relationships, and environment). Domain scores range from 0 to 100, with higher scores indicating better QOL. While it was designed as a self-report instrument, both assisted and interviewer-based administrations are possible. The WHOQOL project has shown evidence of good psychometric properties for this instrument [24]. We used the validated Italian version [25].

Sample size calculation

The power calculation was based on detecting an effect size of medium magnitude (d=0.55) on the PC-SAD5 [26]. We calculated that a sample size of 104 would result in a power of 0.80 with alpha set at 0.05. To protect against an anticipated loss to follow-up of about 10%, we planned to enrol 114 patients. Additional calculations were based on rates of remission and clinical response, defined as a PC-SAD5 score of less than 15 at follow-up and a decrease in PC-SAD5 score of more than 25%, respectively. The threshold for remission was chosen based on the scores displayed by participants free from mental disorders in the PC-SAD5 validation study [23]. Hypothesizing a remission rate of 30% in the control group, the study (with N=114) would have a power of 0.80 to detect a difference between groups of 27%. Also, hypothesizing a response rate of 50% in the control group, the study would have a power of 0.80 to detect a difference between groups of 26%.

Statistical analysis

All analyses were performed with STATA software, version 12 (Stata Corporation, College Station, Texas, USA). All tests were two-tailed, with alpha set at .05. First, descriptive analyses and simple analyses on completers were performed. Student's t-test and Fisher's exact test were used to test for differences between groups in continuous and discrete variables, respectively; the paired t-test was used to test for within-group differences in changes in continuous variables.

Subsequently, we performed the main analyses of the longitudinal data. In all these analyses, the last observation carried forward method was used to include the patients who did not provide follow-up data.

First, the effect of the intervention was estimated with linear regression models adjusted for baseline PC-SAD5 score and based on the traditional Intent-To-Treat (ITT) and Per-Protocol (PP) analyses.

However, if many patients did not comply with the intervention, these analyses are prone to bias. More advanced methods, such as principal stratification [27], have been proposed to properly define causal effects in the presence of post-treatment variables, such as non-compliance. Individuals are stratified according to potential values of the post-treatment variable. In the case of non-compliance, four strata can be defined: (i) the 'compliers' who would comply to the assigned intervention; (ii) the 'never-takers' who would follow the control intervention whatever their assigned arm; (iii) the 'always-takers' who would follow the experimental intervention whatever their assigned arm; (iv) the 'defiers' who would follow the intervention opposite to their assignment [28]. As these strata – called principal strata – are unaffected by the intervention assignment, they can be viewed as pre-treatment variables, which does not question causal inference. In this study, the patients assigned to

the control condition were prohibited from attending intervention sessions. Therefore, only two principal strata are possible, i.e., the 'compliers' and the 'never-takers'.

The Complier Average Causal Effect (CACE) is the causal difference between the mean potential outcomes of patients receiving the experimental and control intervention for patients who would have complied with the experimental intervention if they had been assigned to it. We estimated the CACE following the method proposed by Jo and Stuart [29]. In parallel with propensity score methods to estimate the treatment effect in observational studies [30, 31], they first estimate a model for the probability of belonging to a principal stratum, termed 'principal score', and use principal score weighting and matching to estimate the CACE. This is an alternative to the instrumental variable approach used by Angrist et al. [28].

In our study, while stratum membership is known for the patients assigned to the intervention group, it is unknown for those assigned to the control condition. Propensity score methods were used to identify individuals in the control group who were likely to be compliers. Then, we could compare outcomes under the treatment and control conditions within the stratum of compliers.

To estimate principal scores (i.e., the probability of being a complier), we used a multiple logistic regression model based on the intervention group members only, predicting compliance given a set of socio-demographic and clinical covariates (sex, age, marital status, work conditions, educational level, and PC-SAD5 score at baseline). Then, we matched the intervention group compliers to control group members with similar principal scores. We considered two controls for each treated member. We used matching within a caliper of 0.2 standard deviation of the logit of the principal score, which is frequently used and recommended for propensity score matching [32]. Then, the CACE was estimated on these selected patients with a linear regression model adjusted for baseline PC-SAD5 score.

A second linear regression model for estimating the CACE was constructed using a technique known as weighting by the odds. The intervention group compliers receive a weight of 1. Control group members receive a weight equal to their principal score converted to the odds scale. The intervention group never-takers, who are not used when estimating the CACE, receive a weight of 0. This weighting serves to make the control group look like the set of the intervention group compliers [33].

Results

A total of 416 patients were included in the study and completed the PC-SAD in the waiting room. Only 7 PC-SAD questionnaires were unusable because of missing answers. Of the 409 participants with a valid PC-SAD, 121 were positive for MDD, either with (N=61) or without (N=60) positivity

also for Dys, and 30 for Dys alone. The remaining 258 participants were negative for both MDD and Dys.

Thirty-six patients with a positive PC-SAD result suggestive of suicidal ideation were excluded from the study and were immediately informed of the results and advised to contact the outpatient clinic. The remaining 115 patients were randomised to either the intervention group (N=56) or the control group (N=59). After three months, 100 patients (87% of randomized patients; intervention group, N=46; control group, N=54) completed the 3-month assessment, while six patients could not be traced and nine patients refused to be interviewed.

No significant differences in baseline PC-SAD5 score and demographic variables were found between patients who completed the follow-up interview and those who did not, except for a higher percentage of women in the former group. The clinical and sociodemographic characteristics of the patients who completed the follow-up assessment are summarised in Table I.

As illustrated in Table II, severity of depression (PC-SAD5 score) and quality of life (scores on all WHOQOL domains) improved significantly in both groups. In the intervention group, 17 patients (37%) contacted the outpatient clinic, whereas 29 did not. The former had significantly better outcome than the latter (e.g., mean decrease in PC-SAD5 score 21.3 vs. 8.9, p=0.046; response rate 82% vs. 48%, p=0.030; mean increase in WHOQOL Psychological domain 2.7 vs. 0.7, p=0.002).

Table III reports the results of ITT, PP, and CACE analyses examining the differences between the intervention group and the control group in the change in PC-SAD5 score. While ITT analysis showed no effect of the intervention, PP analysis revealed a significant effect (p=0.02), with a mean reduction in PCSAD5 score of -10.97 in the intervention group. The CACE analysis yielded similar findings, showing a significant beneficial effect of the intervention on the change in PC-SAD5 score with principal score based on either matching (p=0.01; mean reduction in PCSAD5 score -12.68) or weighting (p=0.003; mean reduction in PCSAD5 score -13.38).

The analyses examining the differences between the intervention group and the control group in the PC-SAD5 response and remission rate yielded analogous results. Whereas ITT analysis did not show a significant effect of the intervention either on response or remission, PP analysis indicated a significant effect both on response (p=0.01) and on remission (p=0.02), with 424% and 368% increase in the probability of response and remission as measured by the odds ratio, respectively. The CACE analysis revealed similar findings, showing a significant beneficial effect of the intervention on response rate with principal score based either on matching (p=0.01) or weighting (p=0.002), as well as on remission rate, with principal score based either on matching (p=0.04) or weighting (p=0.02).

Discussion

This study tested the effectiveness of a program designed to impose a small burden on PC physicians and to guarantee referral to and completion of effective specialized care for all patients with a positive depression screen. The fact that randomisation occurred after the completion of the screener does not reduce the ability of the study to separate the effects of the program from those of usual care in patients screening positive, as the patients allocated to the control group were not informed of the screening results until the follow-up assessment. Even if participating in the study made these patients more alert to the issue of depression and led them to take steps to ascertain the presence of depression on their own, this would have had a conservative effect on the results by improving the outcomes in the control group.

Some issues make it difficult to provide evidence of a beneficial effect of a program for identification and treatment of depression in PC [34, 35]. At the usual levels of prevalence of MDD in PC settings, even a sensitive and specific screener would yield several false positives that cannot, by definition, benefit from screening. Also, not all true positives benefit from screening to the same degree, because patients with more severe forms of depression are usually detected independent of screening.

However, these issues are unlikely to be the main reason for the lack of significant findings in our ITT analysis. In fact, we observed that only a portion of participants allocated to the intervention group actually contacted the psychiatric outpatient clinic. Of importance, these participants showed a significantly better outcome than those who did not contact the clinic. Both the PP and CACE analyses revealed that the program significantly reduced depression severity and led to significantly higher response and remission rates. Therefore, the findings suggest that the program is effective, if delivered as planned. However, many patients allocated to the intervention group did not contact the outpatient clinic, and this ‘diluted’ the effect of the intervention to such a degree that it did not show effectiveness in ITT analysis.

In this study, specialist depression care supports were set in place to ensure accurate diagnosis and effective treatment as needed, and every effort was spent to promote the use of these resources by the participants allocated to the intervention group. They were informed of the screening results and were invited to contact an outpatient clinic that would have provided psychiatric evaluation and treatment as needed, free of charge. Calling the clinic was easy, as there was a dedicated phone line. The clinic was easily accessible, as it was located in the same health district as the PC practices and the appointments were offered within one week. Also, the participants who had still not contacted the clinic after two weeks were sent a reminder letter.

However, in spite of all these efforts, 63% of the participants randomized to the intervention group did not contact the clinic. This finding requires some critical reflections. Previous studies [36, 37] already pointed to the presence of barriers on the part of both physicians and patients to the implementation of depression screening and treatment in routine clinical practice. Our study suggests that there are indeed considerable barriers on the part of many patients to the implementation of such programs, even in the context of a clinical trial with abundant administrative and logistic supports. A mismatch may have occurred between the procedures to improve identification and treatment of depression that we implemented and those desired by some patients.

First, some patients might have preferred to be contacted by the outpatient clinic, rather than being provided contact information. Proactive calls are indeed emerging as an important ingredient to motivate patients to seek further assessment of positive depression screens and promote engagement [38]. Also, the place where the psychiatric assessment would be performed might be another important issue. We decided to arrange psychiatric assessment and treatment in a place separated from the PC practices in order to most closely resemble the actual organization of health care services in Italy. All Italian citizens are registered as patients with a PC physician, who is funded by the Local Health Unit with a fixed allowance per patient. MH services, despite being highly accessible and free of charge, are located separately from PC practices, so most people are not familiar with them. At least a portion of PC patients may therefore view such services with some uneasiness, and mistakenly conceive of these services as healthcare facilities for only severely mentally ill people. Given that the outpatient clinic was located in the psychiatric department of a university general hospital, such feelings towards and beliefs about MH services may have reduced the willingness of participants to be evaluated and treated as needed by the clinic, even though they would not have had to bear any cost. Possibly, more participants would have been willing to be evaluated and treated if we had arranged to provide the MH assessment in the PC practice setting. In this way, they might have become personally acquainted with the psychiatrist and might have been more likely to accept to be treated, if needed, at the outpatient clinic. Another option would be to enhance the role of the PC physician in the management of depression by means of specific training and supervision, while providing the possibility of referral to MH specialists. In the Italian context, such a model of collaborative care intervention was recently found to be more effective than usual care in patients with above-threshold symptoms of depression [39]. Still another alternative would be to provide not only psychiatric assessment, but also psychiatric treatment in the PC setting; however, this would be much more expensive and would not resemble the actual

organization of healthcare services in Italy, where the importance of integrating PC and MH has only recently been grasped [40].

There is currently considerable uncertainty about whether depression screening may improve MH outcomes in PC [41]. While our findings cannot resolve this issue, they may nevertheless be useful to inform future research and secondary prevention efforts. It should be acknowledged that, due to the small sample size, our results have limited generalizability and should be regarded as preliminary and in need of replication with a larger sample. With these limitations in mind, our findings support the view that programs for identification and treatment of depression in PC can be effective, if they include support from MH professionals. However, their effectiveness is not easy to achieve in practice, and is conditional on patient adherence. Therefore, these programs must be designed based on the understanding of the perspectives of the patients, who are key stakeholders in the care they receive. In each specific context, a careful examination of patients' preferences and attitudes is needed to put in place procedures that enhance patients' acceptance and adherence, thus increasing the likelihood of ultimately improving patient outcomes.

Acknowledgements

This study was financially supported by the Italian Ministry of Health in the framework of the 'Programma Ricerca Finalizzata 2006'. We are grateful to Sara Portone, M.Sc., for her valuable technical assistance.

References

1. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, de Girolamo G, Graaf R, Demyttenaere K, Gasquet I, Haro JM, Katz SJ, Kessler RC, Kovess V, Lépine JP, Ormel J, Polidori G, Russo LJ, Vilagut G, Almansa J, Arbabzadeh-Bouchez S, Autonell J, Bernal M, Buist-Bouwman MA, Codony M, Domingo-Salvany A, Ferrer M, Joo SS, Martínez-Alonso M, Matschinger H, Mazzi F, Morgan Z, Morosini P, Palacín C, Romera B, Taub N, Vollebergh WA; ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004;420:21-27.

2. GBD 2013 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition. Lancet 2015;386:2145-2191.

3. Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J Clin Psychiatry 2015;76:155-62.

4. Katon WJ, Rutter C, Simon G, Lin EH, Ludman E, Ciechanowski P, Kinder L, Young B, Von Korff M. The association of comorbid depression with mortality in patients with type 2 diabetes. Diabetes Care 2005;28:2668-72.

5. Frasure-Smith N, Lespérance F. Depression and anxiety as predictors of 2-year cardiac events in patients with stable coronary artery disease. Arch Gen Psychiatry. 2008 Jan;65(1):62-71.

6. Craven MA, Bland R. Depression in primary care: current and future challenges. Can J Psychiatry 2013;58:442-8.

7. Boerma WG, Verhaak PF. The general practitioner as the first contacted health professional by patients with psychosocial problems: a European study. Psychol Med 1999;29:689-696.

8. Cepoiu M, McCusker J, Cole MG, et al. Recognition of depression by non-psychiatric physicians — a systematic literature review and meta-analysis. J Gen Intern Med 2008;23:25–36.

9. Mitchell AJ. Why Do Clinicians Have Difficulty Detecting Depression? In Mitchell AJ, Coyne JC (eds.): Screening for depression in clinical practice. Oxford University Press, New York, 2010 (pp. 57-82).

10. Thombs BD, Ziegelstein RC, Roseman M, Kloda LA, Ioannidis JP. There are no randomized controlled trials that support the United States Preventive Services Task Force Guideline on screening for depression in primary care: a systematic review. BMC Med 2014;12:13.

11. Gilbody SD, Sheldon TD, House AD. Screening and case-finding instruments for depression: a meta-analysis. CMAJ 2008;178:997-1003.

12. 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;151:793-803.

13. Jarjoura D, Polen A, Baum E, Kropp D, Hetrick S, Rutecki G. Effectiveness of screening and treatment for depression in ambulatory indigent patients. J Gen Intern Med 2004;19:78-84.

14. Rost K, Nutting P, Smith J, Werner J, Duan N. Improving depression outcomes in community primary care practice: a randomized trial of the quEST intervention. Quality Enhancement by Strategic Teaming. J Gen Intern Med 2001;16:143-149.

15. Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unu¨tzer J, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. JAMA. 2000;283:212-220.

16. Rubenstein LZ, Alessi CA, Josephson KR, Trinidad Hoyl M, Harker JO, Pietruszka FM. A randomized trial of a screening, case finding, and referral system for older veterans in primary care. J Am Geriatr Soc. 2007;55:166-174.

17. Bergus GR, Hartz AJ, Noyes R Jr, Ward MM, James PA, Vaughn T, et al. The limited effect of screening for depressive symptoms with the PHQ-9 in rural family practices. J Rural Health. 2005;21:303-309.

18. Whooley MA, Stone B, Soghikian K. Randomized trial of case-finding for depression in elderly primary care patients. J Gen Intern Med. 2000;15:293-300.

19. Bosmans J, de Bruijne M, van Hout H, van Marwijk H, Beekman A, Bouter L, et al. Cost-effectiveness of a disease management program for major depression in elderly primary care patients. J Gen Intern Med. 2006;21:1020-1026.

20. Callahan CM, Hendrie HC, Dittus RS, Brater DC, Hui SL, Tierney WM. Improving treatment of late life depression in primary care: a randomized clinical trial. J Am Geriatr Soc. 1994;42:839-846.

21. Rogers WH, Wilson IB, Bungay KM, Cynn DJ, Adler DA. Assessing the performance of a new depression screener for primary care (PC-SAD). J Clin Epidemiol 2002;55:164-175.

22. Picardi A, Adler DA, Abeni D, Chang H, Pasquini P, Rogers WH, Bungay KM. Screening for depressive disorders in patients with skin diseases: a comparison of three screeners. Acta Derm Venereol 2005;85:414-419.

23. Picardi A, Adler DA, Chang H, Lega I, Gigantesco A, Pasquini P, Matteucci G, Zerella MP, Caredda M, Tarsitani L, Biondi M, Rogers WH. Development and preliminary validation of the PC-SAD5, a screener-derived short depression severity measure. J Eur Acad Dermatol Venereol 2012;26:165-171.
4240.

24. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-Bref quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004;13:299–310.

25. de Girolamo G, Rucci P, Scocco P, Becchi A, Coppa F, D'Addario A, Daru E, De Leo D, Galassi L, Mangelli L, Marson C, Neri G, Soldani L. Quality of life assessment: validation of the Italian version of the WHOQOL-Bref. Epidemiol Psichiatr Soc 2000;9:45–55.

26. Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

27. Frangakis CE, Rubin DB. Principal stratification in causal inference. Biometrics 2002; 58(1):21–29.

28. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. Journal of the American Statistical Association 1996; 91:444–455.

29. Jo B, Stuart EA. On the use of propensity scores in principal causal effect estimation. Statistics in Medicine 2009; 28(23):2857–2875.

30. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983; 70(1):41–55.

31. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Statistics in Medicine 2004; 23(19):2937–2960.

32. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharmaceutical Statistics 2011; 10(2):150–161.

33. McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. Psychological Methods 2004; 9:403–425.

34. Coyne JC, Thompson R, Palmer SC, Kagee A, Maunsell E. Should we screen for depression? Caveats and potential pitfalls. Appl Prev Psychol 2000;9:101–121.

35. Palmer SC, Coyne JC. Screening for depression in medical care: pitfalls, alternatives, and revised priorities. J Psychosom Res 2003;54:279-287.

36. Valenstein M, Dalack G, Blow F, Figueroa S, Standiford C, Douglass A. Screening for psychiatric illness with a combined screening and diagnostic instrument. J Gen Intern Med 1997;12:679-685.

37. Valenstein M, Kales H, Mellow A, Dalack G, Figueroa S, Barry KL, Blow FC. Psychiatric diagnosis and intervention in older and younger patients in a primary care clinic: effect of a screening and diagnostic instrument. J Am Geriatr Soc 1998;46:1499-1505.

38. Fernandez y Garcia E, Joseph J, Wilson MD, Hinton L, Simon G, Ludman E, Scott F, Kravitz RL. Pediatric-based intervention to motivate mothers to seek follow-up for depression screens: The Motivating Our Mothers (MOM) trial. Acad Pediatr 2015;15:311-8.

39. Menchetti M, Sighinolfi C, Di Michele V, Peloso P, Nespeca C, Bandieri PV, Bologna M, Fioritti A, Fravega R, Ghio L, Gotelli S, Levantesi P, Ortega MA, Savorani M, Simoni L, Tarricone I, Morini M, Gask L, Berardi D. Effectiveness of collaborative care for depression in Italy. A randomized controlled trial. Gen Hosp Psychiatry 2013;35:579-586.

40. Berardi D, Ferrannini L, Menchetti M, Vaggi M. Primary care psychiatry in Italy. J Nerv Ment Dis 2014;202:460-463.

41. Thombs BD, Ziegelstein RC. Does depression screening improve depression outcomes in primary care? BMJ 2014;348:g1253.

Table I. Sociodemographic and clinical characteristics of participants

	Control Group (N=54)	**Intervention Group (N=46)**
Sex (N, *%*)		
Male	11 *(20)*	5 *(11)*
Female	43 *(80)*	41 *(89)*
Age (mean±SD)	45.9±11.0	48.3±11.0
Marital status (N, *%*)		
Never married	12 *(22)*	18 *(39)*
Married or living with a partner	34 *(63)*	20 *(44)*
Separated, divorced, or widowed	8 *(15)*	8 *(17)*
Education (N, *%*)		
Primary school	1 *(2)*	1 *(2)*
Junior high school	12 *(22)*	6 *(13)*
Senior high school	26 *(48)*	25 *(54)*
University degree	15 *(28)*	14 *(30)*
PC-SAD5 score (mean±SD)	48.7±17.3	48.0±15.6
WHOQOL Psychological domain score (mean±SD)	10.3±2.1	10.2±2.0
WHOQOL Physical domain score (mean±SD)	11.7±2.2	11.6±2.2
WHOQOL Social domain score (mean±SD)	11.1±3.2	10.9±3.0
WHOQOL Environmental domain score (mean±SD)	11.3±2.1	11.1±2.0

Table II. Health outcomes by group

	Control Group (N=54)	**Intervention Group (N=46)**
PC-SAD5 score at follow-up (mean±SD)	37.7±21.4*	34.5±21.8*
PC-SAD5 full remission (N, *%*)	10 *(18)*	12 *(26)*
PC-SAD5 response (N, *%*)	25 *(46)*	28 *(61)*
WHOQOL Psychological domain score at follow-up (mean±SD)	11.9±2.8*	11.8±2.5*
WHOQOL Physical domain score at follow-up (mean±SD)	13.3±2.8*	13.7±2.2*
WHOQOL Social domain score at follow-up (mean±SD)	13.0±3.4*	12.7±3.5*
WHOQOL Environmental domain score at follow-up (mean±SD)	12.6±2.3*	12.5±1.7*

* = significantly improved from baseline, $p \leq 0.001$

Table III. Linear regression adjusted for PC-SAD5 at baseline on PC-SAD5 score at Follow-up: intervention vs control group. ITT, PP, and principal score (CACE estimates) analysis

		Coefficient	**SE**	**p-value**	**95% CI**
ITT	Group	-1.83	3.40	0.59	-8.56 4.90
	PC-SAD5 score (baseline)	0.68	0.10	< 0.001	0.48 0.80
	Constant	7.36	7.39	0.32	-7.29 22.00
PP	Group	-10.97	4.55	0.02	-19.99 -1.94
	PC-SAD5 score (baseline)	0.69	0.10	< 0.001	0.50 0.89
	Constant	5.78	4.99	0.25	-4.10 15.66
CACE (Matching 1:2)	Group	-12.68	4.78	0.01	-22.30 -3.07
	PC-SAD5 score (baseline)	0.74	0.13	< 0.001	0.48 0.99
	Constant	14.28	9.32	1.53	-4.48 33.03
CACE (Weighting)	Group	-13.31	4.28	0.003	-21.84 -4.79
	PC-SAD5 score (baseline)	0.95	0.07	< 0.001	0.81 1.08
	Constant	7.98	5.52	0.153	-3.03 19.00

SE = Standard Error

CI = Confidence Interval