Rethinking recommendations for screening for depression in primary care

Brett D. Thombs PhD, James C. Coyne PhD, Pim Cuijpers PhD, Peter de Jonge PhD, Simon Gilbody DPhil, John P. A. Ioannidis MD DSc, Blair T. Johnson PhD, Scott B. Patten MD PhD, Erick H. Turner MD, Roy C. Ziegelstein MD

Screening for depression in primary care is an issue that is highly contentious and hotly debated, and recommendations have evolved over time. For example, early policy statements from the 1990s in Canada and the United States recommended against screening for depression in primary care or did not find enough evidence to recommend either for or against it. Later, in 2002 in the US and in 2005 in Canada, recommendations were made to screen adults for depression in primary care settings when integrated, staff-assisted systems for assessing and managing depression were available. These “collaborative care” programs typically involve multifaceted systems with central roles for nonmedical specialists, such as case managers, who work with primary care physicians, mental health specialists and others to provide management and follow-up.

In 2009, an updated statement from the United States Preventive Services Task Force reiterated the recommendation that primary care physicians screen patients for depression in the context of integrated systems for managing the condition, but not where such resources are unavailable. In contrast to this position, a 2010 guideline for depression management from the United Kingdom’s National Institute for Health and Clinical Excellence noted a lack of evidence that depression screening would benefit patients and did not recommend routine screening in primary care settings. Critical reflection on key differences in these two recent recommendations is important, both for physicians who must decide whether to screen their patients and for developers of guidelines, such as the Canadian Task Force on Preventive Health Care, which is currently reviewing evidence for an updated recommendation statement. Thus, we critiqued recent recommendations on screening for depression in primary care in the context of current evidence.

What is screening and when should it be recommended?

In the context of depression, the word “screening” has sometimes been used inaccurately to characterize a range of activities, including the use of self-report questionnaires to track the severity of symptoms and the effects of treatment, to detect relapse among patients in which depression has already been detected and treated or to inform clinical consultations provided to all patients regardless of their scores on such questionnaires. None of these activities, however, constitutes screening. Rather, consistent with standard definitions, screening involves the use of questionnaires concerning the symptoms of depression or small sets of questions about depression to identify patients who may have depression but who have not sought treatment and whose depression has not already been recognized by health care providers. Patients identified as possibly having depression need further assessment and, if appropriate, should be offered treatment.

Screening for depression is potentially useful only to the extent that it improves patient outcomes beyond those of standard care. Thus, to be successful, a screening program must identify a substantial number of patients in whom depression...

Key points

- Screening for depression in primary care is recommended in the United States and Canada under certain conditions, but not in the United Kingdom.
- No trials have found that patients who undergo screening have better outcomes than patients who do not when the same treatments are available to both groups.
- Existing rates of treatment, high rates of false-positive results, small treatment effects and the poor quality of routine care may explain the lack of effect seen with screening.
- Developers of future guidelines should require evidence of benefit from randomized controlled trials of screening, in excess of harms and costs, before recommending screening.
sion has not already been diagnosed, engage those patients in treatment and obtain sufficiently positive results to justify the costs and potential harms associated with the program. Potential harms from routine screening for depression are rarely made explicit, but they include the treatment of depression in patients who are incorrectly identified as having the disorder, the treatment of mild symptoms that would often resolve without intervention and, perhaps most importantly, the diversion of scarce resources from other endeavours, such as ensuring better care for patients already identified as having depression. Indeed, the poor quality of routine care for patients with depression in primary care is well-documented. Many patients who are given prescriptions for antidepressant medications stop taking the drugs shortly after treatment begins, and only 20%–30% of people whose depression is treated exclusively in a primary care setting receive adequate care and follow-up.13,14

A number of well-established criteria should be met before screening is recommended.13,14 Generally, it is reasonable to consider screening when the condition in question is important and prevalent, can be effectively treated and cannot be readily detected without screening. Furthermore, screening methods should be accurate and carry a tolerably small risk of false-positive results. Such results could lead to unnecessary diagnostic testing, adverse effects and costs of inappropriate treatment, as well as the sequelae of being incorrectly labelled (e.g., stigma). False reassurance for false-negative results may also require consideration in some circumstances. Ideally, benefits in excess of potential harms should be consistently shown in well-conducted, randomized, controlled trials with sufficiently long follow-up to see important patient-oriented outcomes.

Guidelines from the National Institute for Health and Clinical Excellence

The 2010 guideline statement of the National Institute for Health and Clinical Excellence6 focused on a number of serious concerns about routine screening for depression.15,16 These concerns included the high rate of false-positive results (often more than 50%) returned by screening tools,17 the lack of empirical evidence of benefit to patients, the likelihood that most people identified only by screening would have relatively mild symptoms of depression and often recover without formal intervention, the high cost and large number of resources involved in identifying people who might gain little in terms of improvements to their mental health18 and the diversion of resources away from people with more serious cases of depression whose care is already often inadequate.

Instead of screening, the guidelines recommended that physicians be alert to possible depression, particularly when there is a previous history of the condition or when patients have a chronic physical health problem with functional impairment, and that they ask about symptoms of depression when there is a specific concern.

Recommendation of the US Preventive Services Task Force

The 2009 statement from the US Preventive Services Task Force recommended against depression screening outside of integrated programs of staff-assisted assessment and treatment. Evidence from a recent meta-analysis19 of 11 trials in primary care settings supports this recommendation. Although several of the trials found that screening increased identification or treatment of depression, none found that screening reduced the number of patients with depression or improved depressive symptoms, and the overall effect estimate was virtually zero (standardized mean difference [SMD] –0.02, 95% confidence interval [CI] 0.25–0.20).

In contrast to the guidelines issued by the National Institute for Health and Clinical Excellence,7 the US Preventive Services Task Force called for depression screening in primary care where supports are in place to ensure appropriate diagnosis, treatment and follow-up.4 This recommendation was based on results from four trials, all of which found that patients with depression who were provided staff-assisted programs for the management of their condition had statistically significantly better outcomes than patients who did not receive such an intervention. An earlier meta-analysis of more than 30 trials estimated that interventions that involve collaborative care significantly reduce the symptoms of depression, even if the effect is not large (SMD 0.25, 95% CI 0.18–0.32).20

Evidence evaluating whether screening for depression is effective should be distinguished from evidence evaluating whether interventions that involve collaborative care to treat depression improve outcomes over routine care. Indeed, of the four trials cited by the US Preventive Services Task Force as evidence supporting screening, none actually evaluated screening. In each of the four studies, patients were required to have depressive symptoms or a diagnosis of depression to be eligible for the trial. Patients with depression in
the intervention groups received collaborative care for depression, whereas those in the control groups received only standard care. Thus, whereas the results of the trials suggest that providing collaborative care to patients with depression is better than not providing such care, they do not address whether screening would benefit patients with previously unrecognized depression. Underlining this point, among the three largest studies cited by the task force (studies involving > 100 patients), 44% of patients in one trial were given treatment for depression before their enrolment, and 44% of patients in another of the studies were receiving appropriate care for depression, defined as specialized counselling or treatment using antidepressant medication, before their enrolment; data on pretrial rates of treatment were not provided in the third of these studies, but patients who were already receiving treatment were not excluded from the trial.

There appears to be only one documented attempt to screen and provide collaborative care for depression in a clinical setting. In this study, 1687 patients deemed to be at high risk of depression were sent a screening questionnaire. The patients were invited to participate in the study (completing and returning the questionnaire) via a letter signed by their general practitioner, and a reminder was sent to nonresponders two weeks later. Of the 1687 patients identified, 780 returned the screening questionnaire, 226 of whom had a positive result (28.9%). A total of 173 patients with positive screening results were assessed via a diagnostic interview, and 71 patients with major depression were detected. Of these patients, 36 were already having their depression treated and 18 declined treatment or did not attend their appointment. Thus, only 1.0% (17 of 1687) of patients eligible for screening started treatment for their depression. Put another way, about 100 people had to be invited to be screened for 1 person to receive treatment for depression. Based on a published estimate that about one in seven patients receiving treatment in primary care will do better than those not receiving treatment, about 700 people would need to be approached for screening for one patient to improve clinically. These results highlight the potential resource implications of screening for depression.

**Can we assume that screening will benefit patients?**

We know of no clinical trial in which patients screened for depression had better outcomes than patients who were not screened when the same resources for treatment were available to both groups, as would be the case in primary care settings. Should we nonetheless assume that routine screening would benefit patients? Several factors argue against that assumption.

Screening can generally only succeed to the degree that it identifies patients whose condition is otherwise unrecognized and untreated. Yet, prescription rates for antidepressant medications are already high and are trending upward. Among adults 35 years of age and older in the US, annual rates of use for antidepressant medications increased from 8.3% to 14.1% from 1996 to 2005, with one-third to one-half of prescriptions specifically for psychiatric problems. In a 2005 study from Canada, 7% of a sample of the general population reported current use of antidepressant medications, a higher figure than the estimated prevalence of major depression (4%). Rates of prescriptions for antidepressant medications may be even higher among patients with long-term diseases. In Ontario, the rate of prescriptions for antidepressant medications among patients aged 65 years and older within six months of having an acute myocardial infarction doubled from 8% in 1993 to 16% in 2002.

Beyond treatment with antidepressant medications, province-wide data from Quebec show that 23% of women and 19% of men who consulted a physician in 2003 received a diagnostic billing code covering a depression or anxiety disorder at some point during the year. Indeed, as the frequency of recognition and treatment increases, the yield of screening interventions will decrease.

A consequence of the increasing rate of treatment for depression is that existing studies appear to exaggerate the accuracy of screening questionnaires and the number of otherwise unrecognized patients who would be identified through screening. A recent overview evaluated original studies on the accuracy of depression screening tools that were included in 17 systematic reviews or meta-analyses published from 2005 to 2009. The authors found that 189 of 197 original studies in those reviews (95.9%) included patients with depression that had already been diagnosed or was being treated. Screening, however, is intended to identify new cases. Including patients whose depression has already been diagnosed or treated in studies of screening tools exaggerates the ability of the tools to identify new cases and the number of people who would start treatment as a result of the screening process. Indeed, the authors of the overview estimated that the number of patients with untreated depression who would actually be detected by screening may be less than half the number predicted by existing studies.
In addition, the authors concluded that more than 80% of positive screening results are likely false-positives in primary care settings. This estimate was based on published estimates of sensitivity (85%) and specificity (74%) in primary care and a 10% prevalence of major depression, conservatively assuming that 50% of cases in actual practice are recognized without screening and that published sensitivity figures are overestimated by 10% owing to the inclusion of patients whose condition has already been diagnosed or treated. Thus, when used to identify patients whose condition is undetected and untreated, screening tools may be much less accurate than is usually assumed. Compounding this problem, no studies have estimated the cumulative false-positive rate from periodic screening for depression in primary care, which would certainly be very high.

Recent evidence raises questions about the degree to which standard treatments for depression may benefit patients who are identified through screening, but whose condition is not obvious enough for them to seek help or be otherwise identified. Recent meta-analyses have reported that the effects of antidepressant medications are smaller than previously believed when all studies, rather than just published studies, are included (74 studies, SMD 0.31, 95% CI 0.27–0.35). In addition, when compared with patients who receive a placebo, patients with minor depression or mild symptoms of major depression receive much less benefit, if any, from treatment with antidepressant medications than patients with more severe symptoms.

There is less evidence specific to primary care settings; however, results generally appear to be similar, despite the potentially increased heterogeneity and lower severity of depression in these settings. Consequently, the UK guidelines recommend against using antidepressant medications as a first-line therapy for mild depression in primary care, citing a poor risk–benefit ratio. Psychotherapy is used much less frequently than drug therapy in primary care settings, but estimated effectiveness was generally similar to drug therapies when considering only studies that met minimum quality criteria in eight areas (11 trials, SMD 0.22, 95% CI 0.13–0.31) or accounting for likely publication bias (117 trials, SMD 0.42, 95% CI 0.33–0.51). Authors of a recent meta-analysis found that, in primary care, psychological treatments were effective when patients were referred by their primary care physician (SMD = 0.43, 95% CI 0.28 to 0.58), but not when patients were recruited through screening (SMD = 0.13, 95% CI -0.08 to 0.34). As with antidepressant medications, there is evidence that psychological treatments may be less effective for patients with only mild symptoms of depression.

### What harm does screening pose?

In the absence of any benefit, potential harms of screening for depression should be carefully considered, but the relevant information has rarely been reported in trials. An important consideration is the overall efficiency of screening and the cost of adopting screening programs. At a minimum, the costs of introducing screening into a health care system that is already struggling financially and is unable to effectively address cases of depression that have already been identified would be substantial. Transferring resources to screening activities could lead to a deterioration of the quality of care received by patients with more severe depression who are more clearly in need. In addition, patients who are prescribed antidepressant medications after a positive result from screening, many of whom will not benefit from the treatment, would be unnecessarily exposed to common side effects of the drugs.

In addition, we know little about the potential “nocebo effect” of telling patients who are otherwise not specifically concerned about their mental health that they have depression. A nocebo effect is the opposite of a placebo effect, in that verbal suggestions of a negative outcome can lead to the development or worsening of symptoms. The two primary mechanisms underlying nocebo are conditioning and expectations, and it is possible that both could play a role in linking a message from a physician about depression to subsequent thoughts and behaviour.

### Conclusion

The prevalence of depression and the availability of easy-to-use screening instruments make it tempting to endorse widespread screening for the disease. However, screening in primary care is a resource-intensive endeavour, does not yet show evidence of benefit and would have unintended negative effects for some patients. Evidence from one simulation study using Canadian national data found that the overall burden of depression could be reduced by providing more consistent treatment to reduce symptoms and prevent relapse among people with recurrent disorders, but not by increasing treatment through screening.

We hope that a rigorous review of current evidence will encourage the developers of future guidelines, including members of the Canadian Task Force on Preventive Health Care, to carefully consider their stance on screening for depression. We also hope that, consistent with the 2010 guide-
lines of the National Institute for Health and Clinical Excellence, such developers will conclude that evidence from well-conducted, randomized, controlled trials of the benefit of screening, in excess of its likely harms and costs, is needed before it can be recommended in primary care settings. Specifically, the benefits and harms of screening should be tested in a trial in which all patients identified as having depression should have access to the same integrated care for their condition, regardless of whether they are identified through screening in the intervention group or via physician recognition and referral in a control group. It is possible that such a trial would find that screening benefits patients to a degree that would justify the cost and the harms associated with the process. Until then, however, given the lack of evidence of benefit from screening and the concerns that we have described, it is not reasonable to simply assume that depression screening is good policy.

References

The postfracture

Analysis


Affiliations: From the Departments of Psychiatry; Medicine; Epidemiology, Biostatistics, and Occupational Health; and the School of Nursing (Thombs), McGill University and Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Que.; Behavioral Oncology Program (Coyne), Abramson Cancer Center and Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pa.; Health Psychology Section, Department of Health Sciences (Coyne), University Medical Center Groningen, University of Groningen, the Netherlands; the Department of Clinical Psychology and EMGO Institute (Cuypers), Vrije Universiteit, Amsterdam, the Netherlands; Interdisciplinary Center for Psychiatric Epidemiology (de Jonge), University Medical Center, University of Groningen, the Netherlands; Psychological Medicine and Health Services Research (Gilbody), Hull York Medical School and Department of Health Sciences, University of York, York, UK; Clinical and Molecular Epidemiology Unit (Ioannidis), Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, and Biomedical Research Institute, Foundation for Research and Technology—Hellas, Ioannina, Greece; Institute for Clinical Research and Health Policy Studies (Ioannidis), Tufts Medical Center, and Department of Medicine, Tufts University School of Medicine, Boston, Mass.; Stanford Prevention Research Center (Ioannidis), Stanford University School of Medicine, Stanford, Calif.; the Department of Psychology and Center for Health, Intervention and Prevention (Johnson), University of Connecticut, Storrs, Conn.; the Departments of Psychiatry and Community Health Sciences (Patten), University of Calgary, Calgary, Alta.; the Department of Psychiatry (Turner), Oregon Health and Science University and Portland Veterans Affairs Medical Center, Portland, Ore.; the Department of Medicine (Ziegelstein), Johns Hopkins University School of Medicine, Baltimore, Md.

Contributors: Dr. Thombs drafted the article. All of the authors participated in preliminary discussions to develop the concepts described in the article, agreed upon an outline for the article, revised the article critically for intellectual content and approved the final version of the manuscript submitted for publication.

Funding: Funds from a Meetings, Planning and Dissemination grant from the Canadian Institutes of Health Research (CIHR; grant no. MIN 92725) were used for a meeting of the author group, which led to the preparation of this article. Dr. Thombs is supported by a New Investigator Award from the CIHR and an Establishment de Jeunes Chercheurs award from the Fonds de la Recherche en Santé Québec. Dr. Patten is a Senior Health Scholar with Alberta Innovates, Health Solutions. Dr. Ziegelstein is supported by the National Center for Complementary and Alternative Medicine (grant no. R24AT004641) and the Miller Family Scholar Program of the Johns Hopkins Center for Innovative Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center For Complementary and Alternative Medicine or the National Institutes of Health.