CMAJ OPEN

Research

Screening for depression: a systematic review and meta-analysis

Homa Keshavarz PhD, Donna Fitzpatrick-Lewis MSW, David L. Streiner PhD, Maureen Rice MLIS, Usman Ali MD MS, Harry S. Shannon PhD, Parminder Raina PhD

Abstract

Background: The Canadian Task Force on Preventive Health Care has a guideline on screening for depression among adults 18 years of age or older at average or high risk for depression. To provide evidence for an update of this guideline, we evaluated the literature on the effectiveness of screening for depression in adults.

Methods: For the period 1994 to May 23, 2012, we searched the following electronic databases: MEDLINE, Embase, PsycINFO, the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews. Randomized controlled trials, observational studies and systematic reviews with evidence for the benefits or harms of screening for depression were eligible for inclusion. We performed screening for relevance, extraction of data, analysis of risk of bias and quality assessments in duplicate. We used the generic inverse variance method to conduct a meta-analysis. To determine confidence in the effect, we analyzed the results according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system.

Results: Five quasi-experimental studies (before–after design with a nonrandomized control group) met the inclusion criteria for this review. These studies reported on the effect of community-based screening for depression, with follow-up on the risk of suicide completion, for older residents in regions of rural Japan with high suicide rates. Meta-analysis showed that the screening program had a protective effect on the overall incidence of suicide completion (ratio of rate ratios [RRR] 0.50, 95% confidence interval [CI], 0.32–0.78). When sex was considered, the RRR indicated a significantly lower rate of suicide among women (RRR 0.37, 95% CI 0.21–0.66) but not among men (RRR 0.67, 95% CI 0.35–1.27). The overall GRADE rating applied to this evidence indicated very low quality. No studies addressing the harms of screening for depression met the inclusion criteria for the review.

Interpretation: There is very limited research evidence allowing conclusions about the effectiveness of screening for depression in either average-risk or high-risk populations.

Depression is a complex mental illness that is associated with disability and reduced quality of life for the person with the disorder; it also poses a substantial societal burden. The prevalence of past-year episodes of depression in the Canadian population has been estimated to vary from 5% to 8.2% annually.[1,2] In 2005, the Canadian Task Force on Preventive Health Care published a guideline on screening for depression among adults (18 years or older) at average or high risk for depression.[3] In 2013, the task force released an updated guideline.[4] The systematic review on which this paper is based provided evidence for that 2013 update.

The World Health Organization (WHO) Psychological Problems in General Health Care study,[5] released in 1996, reported that primary care physicians diagnosed major depression in only 42% of adult patients who had the condition. The potential benefits of screening for depression in adults include improved detection of major depressive disorder, dysthymia and subsyndromal depression. Improvements in detection can lead to earlier treatment, and treatment of major depressive disorder in adults is thought to result in improved outcomes, such as better quality of life, better work life and minimization of the risk of suicide.[6] One argument against screening is that screening instruments have low positive predictive value, meaning that many people with a positive screening result do not have depression.[7,8] Although a previous review found no literature specifically evaluating the harms associated with screening for depression and related disorders,[9] those with positive screening results who do not have the disorder may be exposed to stigmatization and further psychologic testing, as well as unnecessary psychologic and pharmacologic treatment regimes.

In preparing to update the guideline, the Canadian Task Force on Preventive Health Care undertook a de novo review, given the guideline's focus on a comparison between screening for depression in people with no apparent symptoms versus no

Competing interests: None declared.

This article has been peer reviewed.

Correspondence to: Parminder Raina, praina@mcmaster.ca, or Donna Fitzpatrick-Lewis, fitzd@mcmaster.ca

***CMAJ Open* 2013. DOI:10.9778/cmajo.20130030**

screening. Our review thus differed from the reviews by Pignone and colleagues[10] and O'Connor and colleagues,[9] which served as the evidentiary base for the 2009 US Preventive Services Task Force screening recommendations for adults. Those reviews included studies in which all members of the population were screened, the comparisons being treatment versus no treatment or feedback (providing depression score to the patient or physician) versus no feedback. The review by Gilbody and colleagues[11] was also outside the scope of our review, because it was a review of depression screening tools.

In the current systematic review, we explored the benefits and harms of screening for depression in asymptomatic adults 18 years of age or older from the general population (at average risk for depression) and in adults at high risk for depression in both primary care and other outpatient settings.

Methods

The search strategy was developed by a librarian (M.R.) experienced in searches for systematic reviews. We searched several electronic databases, specifically MEDLINE, Embase, PsycINFO, the Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews, for the period 1994 to May 23, 2012. The MEDLINE search was updated in April 2013, 6 weeks before publication of the guideline, to identify any recent, potentially relevant randomized controlled trials. The search was broad, the only limitations being date of publication, research subjects (limited to humans) and language (English or French) (see Appendix 1 at www.cmajopen.ca/content/1/4/E159/suppl/DC1). In addition, we searched the grey literature (primarily Canadian sources) up to May 2012, using a number of keyword terms for depression and screening.

Eligible studies were those involving adults at least 18 years of age from unselected populations or high-risk groups (see Appendix 2 at www.cmajopen.ca/content/1/4/E159/suppl/DC1). The intervention of interest was routine screening, and we considered studies of any design that compared screening with no screening. The study settings were primary care or, for high-risk groups, specialty clinics. The outcomes of interest were quality of life, rates of suicidality (attempts or ideation), all-cause mortality, depression-related mortality, rates of hospital admission and changes in symptoms of depression (treatment response or remission). The harms of interest were psychological stress (labelling, anxiety or stigma), false positive results, false negative results, decreased day-to-day functioning and increase in symptoms (see Appendix 3 at www.cmajopen.ca/content/1/4/E159/suppl/DC1).

To determine factors indicating possible higher risk of depression, we searched the websites of reputable sources of information about depression, including the Mayo Clinic (Rochester, Minn.), the Canadian Mental Health Association, the Centre for Addiction and Mental Health (Toronto, Ont.), the Mood Disorders Society of Canada and one published source.[12] From these sources, we collated a list of risk factors (see Appendix 2), which we used to identify studies involving high-risk populations.

Study selection and data extraction

Pairs of reviewers experienced in systematic review methodology and statistics (PhD researchers [including H.K.], a PhD student and a research assistant; pairings varied) independently screened all identified citations for relevance, inclusion criteria and study quality and performed the data extraction. Potentially relevant citations went through 2 levels of title and abstract screening. The first screening was broad, to eliminate citations that were obviously not on topic; the second screening had more specific exclusion criteria, such as age less than 18 years and study not representing primary research on screening. Any citation deemed potentially relevant was retrieved for full review. Pairs of reviewers (as described above) independently reviewed the potentially relevant studies, and for any studies excluded at this stage, agreement about the reason for exclusion was required. All disagreements were resolved through discussions.[13] Reference lists of on-topic systematic reviews retained for analysis were searched to ensure that we considered all primary studies meeting our inclusion criteria.

Quality assessment

We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system to determine the quality of the evidence. This widely used system has been endorsed by over 40 major organizations, including the WHO, the US Centers for Disease Control and Prevention, and the US Agency for Healthcare Research and Quality.[14] GRADE considers 5 criteria (design, consistency, directness, precision and reporting bias) to rate the quality of evidence as high, moderate, low or very low; these ratings represent an assessment of the likelihood that further research will lead to changes in the estimate of effect.[15] We assessed the risk of bias with the Newcastle–Ottawa tool.[16] Two of the authors (H.K. and D.F.L.) independently assessed the evidence according to these criteria and reached agreement on the ratings and the overall quality of the summary statistics.

Statistical analysis

In the included studies, data were obtained before (baseline) and after implementation of the intervention or control measure. Two of the 5 identified papers[17,18] included 2 control groups; the remaining 3 studies each had 1 control group.[19-21] Four out of the 5 papers presented data using adjusted incidence rate ratios (IRR) and one reported adjusted odds ratios. This variation in data presentation necessitated calculation of the ratio of rate ratios (RRR) for each group (i.e., ratio of postimplementation rate to preimplementation rate in the geographic area where the intervention was applied divided by the corresponding ratio of postimplementation rate to preimplementation rate in the control area). We used Cochrane Review Manager software (RevMan, version 5.1, Nordic Cochrane Centre of the Cochrane Collaboration, Copenhagen, Denmark) to conduct the meta-analysis.

We calculated a weighted intervention effect across studies using data for overall population and data stratified by age and sex. An RRR of less than 1.0 indicates a reduction in the outcome IRR. We calculated standard errors for logarithms of rate ratios and 95% confidence intervals (CIs) for the RRR values, assuming that the number of events in each study area in each

period followed a Poisson distribution. We used the generic inverse variance method with a random-effects meta-analysis model, because all of the included studies had been done by the same team of authors working with the same research design. We used the Cochrane Q ($\alpha = 0.10$) and the I^2 statistic to quantify statistical heterogeneity among studies, where $p < 0.10$ indicated a high level of statistical heterogenity.[13]

Results

Study selection and characteristics

Our search identified 14 226 potentially relevant citations (Figure 1). Of these, 12 694 were excluded after screening of titles and abstracts. We retrieved a total of 1532 papers and assessed them against the inclusion criteria; 1527 of these papers were excluded. Five quasi-experimental studies (before–after design with a nonrandomized control group), all with the same first author, met the inclusion criteria and provided the evidence for the review questions (Table 1).[17–21]

Average-risk populations

The first question of interest for this review was, "What is the evidence for the benefit (i.e., improvement in clinical outcomes) of screening for depression in asymptomatic adults (18 years of age or older) from the general population, in either primary care or other outpatient settings?" No studies of

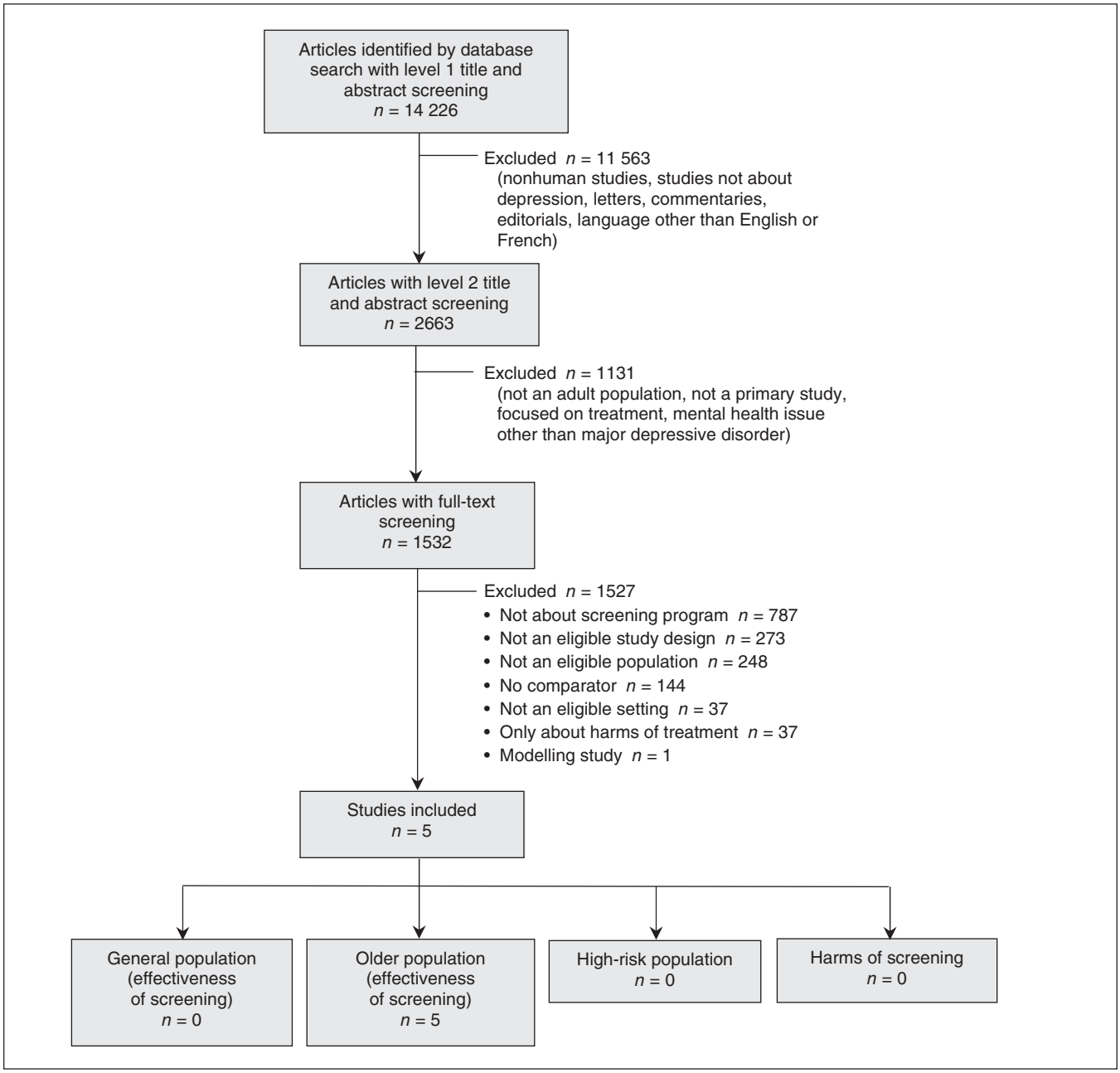

Figure 1: Identification and evaluation of studies for a systematic review of screening for depression.

screening for depression in the average-risk population as a whole met the inclusion criteria of this review. The 5 included studies focused on community-based screening for depression among older people (i.e., age ≥ 60 or age ≥ 65, depending on the study).[17–21] These studies were conducted in rural regions of Japan, where suicide rates among older people ranged from 50 to 418 per 100 000 among women and from 113 to 326 per 100 000 among men.[17–21] Oyama and associates[17–21] developed a universal suicide prevention program, which included a screening component adapted from the WHO World Mental Health Survey.[24] The program involved screening for depression, follow-up with mental health care or psychiatric treatment, and psychoeducation in the community setting. The control communities had similar demographic characteristics and were in the same geographic region as the intervention communities, but they received no components of the pro-

Table 1 (part 1 of 2): Characteristics of studies included in a meta-analysis of the benefit of screening for depression

	Study population			Outcomes	
Study	Description	Definition	Evaluation	Definition	Results
Oyama et al.[19] (5-yr quasi-experimental study in Matsudai, Japan [rural])	Total person-years: 11 567 for intervention, 15 055 for control Age, mean: NR Age, range: ≥ 65 yr Age, median: NR Sex, female: 57.6% Ethnicity: Japanese Education: NR Dx: major and minor depression	Older (≥ 65 yr) residents in 6 rural municipalities of southwest and central Japan Intervention: mental health workshop, referral to general practitioner or follow-up interview with PHN Exclusions: severely disabled or hospitalized cases were excluded from the study	Screening instrument: SDS Other rating: RDC Confirmatory exam: ICD-9 No. of follow-ups: 10 No. of stages: two 10-yr	Main outcome: Change in risk of completed suicide Age-adjusted IRRs of completed suicide before and after intervention or control	Main outcome: Risk of completed suicide in intervention area reduced by 70% among women, no significant change among men Intervention: IRR 1.02 (95% CI 0.49–2.13) for men and 0.30 (95% CI 0.14–0.67) for women Control: No significant change
Oyama et al.[20] (10-yr quasi-experimental study in Yasuzuka, Japan [rural])	Total person- years: 9791 for intervention, 16 032 for control Age, mean: NR Age, range: ≥ 65 yr Age, median: NR Sex, female: NR Ethnicity: Japanese Education: NR Dx: major and minor depression	Older (≥ 65 yr) residents of agricultural rural area in Japan with high suicide rate Intervention: (a) public health education from 1991 to 2000 and (b) screening for depression with follow-up from 1991 to 1997 Exclusions: NR	Screening instrument: SDS Other rating: RDC Confirmatory exam: ICD-9 No. of follow-ups: 7 No. of stages: two 10-yr	Main outcome: Change in risk of completed suicide Age-adjusted IRRs of completed suicide before and after intervention or control	Main outcome: Risk of completed suicide in intervention area reduced by 64% among women, no significant change among men Intervention: IRR 0.51 (95% CI 0.22–1.19) for men and 0.36 (95% CI 0.14–0.93) for women Control: No significant change
Oyama et al.[17] (10-yr quasi-experimental study in Joboji town, Japan [rural])	Total person- years: 9721 for intervention, 17 166 for control Age, mean: NR Age, range: ≥ 65 yr Age, median: NR Sex, female: 50.8% Ethnicity: Japanese Education: NR Dx: depression (unspecified)	Older (≥ 65 yr) residents of agricultural rural area in Japan with high suicide rate Intervention: 2-step depression screening performed by PHN and psychiatrist with follow-up by psychiatrist every 3 yr in targeted district of intervention municipality, health education and emphasis on suicide taboo every year in 10-yr period from 1990 Exclusions: Older people receiving social welfare	Screening instrument: SDS Other rating: SADD Confirmatory exam: ICD-9 No of follow-ups: 10 No. of stages: three 5-yr	Main outcome: Change in suicide rates Age-adjusted IRRs of completed suicide before and after intervention or control	Main outcome: Risk of suicide in intervention area reduced by 73% among men and by 76% among women during implementation decade (relative to pre-implementation decade) Intervention: IRR 0.27 (95% CI 0.08–0.88) for men and 0.24 (95% CI 0.11–0.52) for women Control: No significant change

Continued

Table 1 (part 2 of 2): Characteristics of studies included in a meta-analysis of the benefit of screening for depression

	Study population			Outcomes	
Study	Description	Definition	Evaluation	Definition	Results
Oyama et al.[18] (5-yr quasi-experimental study in Nagawa town, Japan [rural])	Total person- years: 1982 for intervention, 16 754 for control Age, mean: NR Age, range: ≥ 65 yr Age, median: NR Sex, female: 59%–60.8% Ethnicity: Japanese Education: NR Dx: depression (unspecified)	Older (≥ 65 yr) residents of agricultural rural area in Japan with high suicide rate Intervention: 2-step screening for depression and follow-up by PHN, mental health workshop 3 or 4 times a year, group activity program once a month Exclusions: NR	Screening instrument: SDS Other rating: RDC Confirmatory exam: ICD-9 No. of follow-ups: 6 No. of stages: two 6-yr	Main outcome: Change in risk of completed suicide Age-adjusted IRRs of completed suicide before and after intervention or control	Main outcome: Risk of suicide in intervention area reduced by 74% among women, no significant change among men Intervention: IRR 0.48 (90% CI 0.10–2.31) for men and 0.26 (90% CI 0.07–0.98) for women Control: No significant change
Oyama et al.[21] (5-yr quasi-experimental study in 6 rural municipalities of the Sanpachi Second Medical Zone, Japan [rural])	Total person-years: 28 838 for intervention, 27 633 for control Age, mean: NR Age, range: ≥ 60 yr Age, median: NR Sex, female: 57.5% Ethnicity: Japanese Education: NR Dx: depression (unspecified)	Older (≥ 60 yr) residents living in 6 rural municipalities of Sanpachi Second Medical Zone of Japan (mostly agricultural region with high suicide rate) Intervention: (a) health education and (b) screening for depression with follow-up, using community resources of primary care and public health nursing Exclusions: NR	Screening instrument: CES-D, DSS, SDS, GDS-5 Other rating: CIDI Confirmatory exam: ICD-10 No. of follow-ups: 2 No. of stages: two 2-yr	Main outcome: Change in risk of completed suicide Age-adjusted IRRs of completed suicide before and after intervention or control	Main outcome: Risk of suicide in intervention region reduced by 61% among men; no significant change among women Intervention: IRR 0.39 (90% CI 0.18–0.87) for men and 0.49 (90% CI 0.19–1.22) for women Control: No significant change

Note: CES-D = Center for Epidemiologic Studies Depression Scale, CI = confidence interval, CIDI = Composite International Diagnostic Interview, DSS = Depression and Suicide Screen, Dx = diagnosis, GDS-5 = 5-item Geriatric Depression Scale,[22] ICD = International Statistical Classification of Diseases, IRR = incidence rate ratio, NR = not reported, PHN = public health nurse, RDC = Research Diagnostic Criteria, SADD = schedules of Standardized Assessment of Patient with Depressive Disorders, SDS = Self-rating Depression Scale.[23]

gram. The duration of the included studies ranged from 4 to 20 years (over the period 1978 to 2006). The overall aim of the studies was to evaluate the effectiveness of the community-based depression screening program over both the short term and the long term. The outcome of interest was completed suicides, determined from registrations of suicides at local public health centres. The diagnoses in the registry were based on the International Classification of Diseases, Ninth Revision, in which confirmed and probable suicides were grouped together.

All 5 studies used a pre- and post-implementation design. In all studies, more than 60% of men and more than 80% of women within the targeted groups of residents participated in the program during the implementation period.

All 5 studies involved implementation of the suicide prevention program, which had a 2-step process for screening and follow-up for depression. In the first step, older residents within the selected communities were called with an invitation to participate in an educational health workshop on the signs of and possible treatments for depression and suicide risk and also on how to use mental health services. Following the workshop, those who agreed to participate in the program completed the Japanese version of the Zung Self-rating Depression Scale,[23] a 20-item scale that measures affective, psychologic and somatic symptoms associated with depression (used in all 5 of the included studies[17–21]), or the 5-item Geriatric Depression Scale[22] (used in 1 of the included studies[21]). Those who did not attend the workshop were contacted the following day and asked to participate in the program. Examiners then visited all those who agreed to participate and conducted the suicide prevention program according to the same procedures. There were several examiners, including psychiatrists and public health nurses.

In the second step, public health nurses conducted a mental health assessment for each enrolled participant who had a positive screening result on the Self-rating Depression Scale. The nurses used Japanese translations of a standardized assessment for patients with depressive disorders[25] and made a clinical decision about whether a medical examination by a psychiatrist was necessary.

The meta-analysis of the target population involved 70 053 person-years and 65 completed suicides in the intervention groups and 113 324 person-years and 145 completed suicides in the control groups during the respective implementation periods. On the basis of the information provided in the included studies (specifically, average population sizes over 5 years and average percentage of people over the age of 65), we estimated that the overall sample sizes were 18 311 for the intervention groups and 19 736 for the control groups. The studies reported 6 sex- and age-specific target population groups (men and women aged 65–74, 75–84 and ≥ 85), with the exception of one study,[21] which used age groups 60–69, 70–79 and ≥ 80. All 5 studies presented data stratified by age, sex and time periods for baseline and program implementation.

Study	Log RRR	SE	Weight (%)	RRR (95% CI)
Oyama et al.[17]	–1.23	0.43	21.5	0.29 (0.13–0.68)
Oyama et al.[19]	–0.31	0.41	23.2	0.73 (0.33–1.64)
Oyama et al.[20]	–0.19	0.38	26.0	0.83 (0.39–1.74)
Oyama et al.[18]	–0.99	0.68	9.9	0.37 (0.10–1.41)
Oyama et al.[21]	–1.07	0.46	19.4	0.34 (0.14–0.84)
Overall (I^2 = 21%)			100.0	0.50 (0.32–0.78)

Test for overall effect: $Z = 3.06$ ($p = 0.002$)

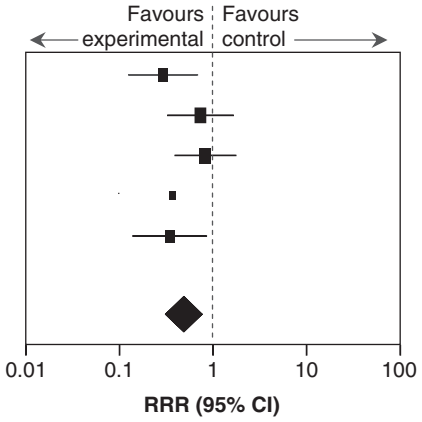

Figure 2: Meta-analysis of the effect of community-based suicide prevention programs, including screening for depression, on suicide rates reported in cohort studies. A rate ratio (RR) less than 1.0 indicates a benefit of suicide prevention programs. CI = confidence interval; RRR = ratio of rate ratios (rate ratio for intervention divided by rate ratio for control), where RRR less than 1.0 indicates a benefit of suicide prevention programs; SE = standard error.

Study	Log RRR	SE	Weight (%)	RRR (95% CI)
Men (non-adjusted)				
Oyama et al.[17]	–1.36	0.69	8.6	0.26 (0.07–0.99)
Oyama et al.[20]	0.42	0.54	14.1	1.52 (0.53–4.39)
Oyama et al.[19]	–0.18	0.57	12.7	0.84 (0.27–2.55)
Oyama et al.[18]	–0.30	1.01	4.0	0.74 (0.10–5.36)
Oyama et al.[21]	–0.91	0.62	10.7	0.40 (0.12–1.36)
Overall (men) (I^2 = 21%)			50.1	0.67 (0.35–1.27)
Women (non-adjusted)				
Oyama et al.[17]	–1.15	0.56	13.1	0.32 (0.11–0.95)
Oyama et al.[19]	–0.46	0.61	11.1	0.63 (0.19–2.09)
Oyama et al.[18]	–1.58	0.94	4.7	0.21 (0.03–1.30)
Oyama et al.[20]	–0.87	0.57	12.7	0.42 (0.14–1.28)
Oyama et al.[21]	–1.25	0.70	8.4	0.29 (0.07–1.13)
Overall (women) (I^2 = 0%)			49.9	0.37 (0.21–0.66)
Overall (I^2 = 0%)			100.0	0.51 (0.34–0.75)

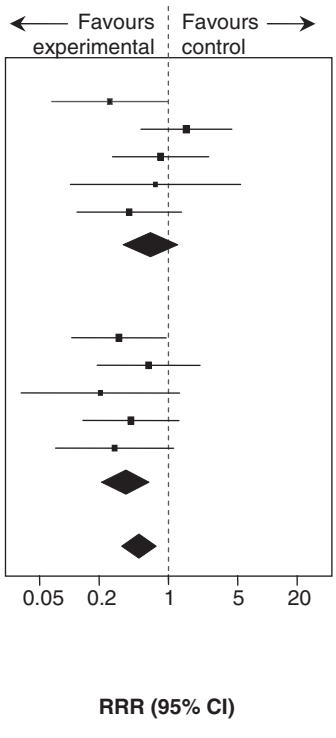

Figure 3: Meta-analysis of the effect of community-based suicide prevention programs, including screening for depression, on suicide rates by sex, as reported in cohort studies. A rate ratio (RR) less than 1.0 indicates a benefit of suicide prevention programs. CI = confidence interval; RRR = ratio of rate ratios (rate ratio for intervention divided by rate ratio for control), where RRR less than 1.0 indicates a benefit of suicide prevention programs; SE = standard error.

The outcome measure in each study was an IRR based on binary data (i.e., suicide/no suicide, calculated for both implementation and control groups before and after the intervention). There was no significant heterogeneity among the studies (I^2 = 21%, χ^2 = 5.04, p = 0.28). When the data for men and women were analyzed separately, there was no significant heterogeneity among the studies (for men, I^2 = 21%, χ^2 = 5.07, p = 0.28; for women, I^2 = 0%, χ^2 = 1.41, p = 0.84). Publication bias could not be assessed because the number of included studies was small.

The difference between pooled IRRs and corresponding 95% CIs for completed suicide was calculated using the generic inverse variance weighting method for the overall study population and for women and men separately. The pooled data from the 5 studies[17–21] showed a statistically significant reduction in the number of completed suicides after implementation of the community-based depression screening program (RRR 0.50, 95% CI 0.3–0.78) (Figure 2). RRRs also indicated a significant reduction in the suicide rate among women (RRR 0.37, 95% CI 0.21–0.66) but no significant effect among men (RRR 0.67, 95% CI 0.35–1.27) (Figure 3).

High-risk populations

We found no studies that examined the benefits of screening high-risk populations (defined using the factors in Appendix 2) versus not screening.

Harms of screening

The second question of interest for this review was "What is the evidence for harm (i.e., decline in clinical outcomes) of screening for depression in asymptomatic adults from the general population, in either primary care or other outpatient settings?" We found no studies meeting our inclusion criteria that could help to answer this question.

GRADE rating

According to the GRADE system for assessing quality, observational evidence (including evidence from studies with a cohort design) begins with a "low" rating. We downgraded the rating because of indirectness, given that the included studies all involved older populations in a rural Japanese setting, who are unlikely to be representative of Canadians. We also downgraded the evidence because the studies used community-based depression screening programs that incorporated education and treatment; as such, their results cannot be attributed solely to the screening component of the programs. Thus, the overall GRADE rating applied to this evidence was very low quality (see Table 2).

Interpretation

We found no direct evidence for benefit of screening in the average-risk population; rather, we identified 5 studies of older populations conducted by the same primary researcher in rural Japan. Although these 5 studies met the inclusion criteria for our review, their results provide limited evidence on the effectiveness of screening for depression in the average-risk population or high-risk groups. The potential generaliz-

Table 2: GRADE evidence profile for the effect of community-based suicide prevention programs, including screening for depression, on the incidence of suicide

Quality assessment							No. of patients/person-years		Effect			
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Screening (older persons)	Control	Relative (95% CI)	Absolute	Quality	Importance
Overall (follow-up 4–20 yr; assessed with community-based depression screening)												
5[17–21]	Observational	No serious risk of bias*	No serious inconsistency†	Very serious‡	No serious imprecision§	None¶	65/70 053 (0.09%)	145/113324 (0.13%)	RRR 0.50 (0.32–0.78)	1 fewer per 1000 (range 0 fewer to 1 fewer)	Very low	Critical

Note: CI = confidence interval, RRR = ratio of rate ratios.
*The quality assessment tools identified a few concerns (e.g., selection of non-exposed cohort, blinding and reporting of withdrawals and drop-outs); however, the evidence was not downgraded for these reasons.
†Heterogeneity statistics were not significant: τ^2 = 0.05, χ^2 = 5.04, df = 4 (p = 0.28); I^2 = 2%.
‡Directness was downgraded because of concerns about population characteristics. The included papers all involved older populations in rural areas of Japan, which are unlikely to be representative of Canadians at average or high risk for depression. Directness was downgraded further because of concerns regarding community-based screening for depression: The studies included in the analysis evaluated the effectiveness of community-based programs to screen for depression, which incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in the community setting. As such, any observed reduction in suicide rates could not be attributed solely to the screening component of these programs.
§The number of events was small (< 300, which is the threshold rule-of-thumb value for dichotomous outcomes); however, with regard to the specific outcome, the evidence was not downgraded.
¶Funnel plot of the comparison indicated potential asymmetry and thus potential publication bias. However, the number of papers (n = 5) was too small to assess publication bias with confidence (≥ 10 papers being the threshold rule-of-thumb value).

ability of the findings of these studies should be considered with caution, as Japan has a national suicide rate much higher than that in Canada or the United States. Among Japanese women 75–84 years of age, for whom benefit of screening was observed in the included studies, the suicide rate is more than 7 times higher than among Canadian women of the same age group (23.4 v. 3.3 per 100 000, respectively).[26] In addition, the geographic regions included in the study had average rates of suicide much higher than even the Japanese average.[17–21] We can draw no conclusions about the potential harms of screening for depression, as we found no studies of such harms that met our inclusion criteria.

These results are consistent with previous guidelines and evidence reviews. The 2009 systematic evidence review of the US Preventive Services Task Force[27] found no evidence of any benefit of screening for depression in the absence of treatment programs. The lack of direct evidence to support general screening programs has also been recognized by the National Institute for Health and Care Excellence[28] and the Scottish Intercollegiate Guidelines Network;[29] neither of these organizations recommend screening of asymptomatic people in the general population. The National Institute for Health and Care Excellence guideline for people with chronic illness recommended that physicians remain alert to the possibility of depression,[30] and another guideline for perinatal women[31] recommended screening women postpartum, yet those recommendations were based on indirect evidence of a benefit of treatment, rather than direct evidence of effectiveness of screening or case-finding for depression.

Limitations

The findings of this review are affected by the limitations of the literature search and of the studies that were included. Because of resource limitations, we limited our search to papers written in English or French, as those could be assessed by the team. It is possible that we missed papers written in other languages.[32] We chose 1994 as the start date for the search, as that was the publication date for the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders*, which changed the definition of major depression.[33] The studies that were reviewed here evaluated the effectiveness of community-based depression screening programs that incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in community settings in rural Japan that had higher-than-average suicide rates. As such, the observed reduction in suicide rates or recovery from depression cannot be attributed solely to the screening component of these programs. As well, given that the program involved community psychoeducation, it is likely that people in the area were more aware of depression and suicide, which may have altered the reporting of deaths as suicide.

Conclusion

The ultimate goal of screening for depression is to reduce associated morbidity and mortality. This review found limited evidence to estimate the effectiveness of screening for depression in primary care among individuals at average risk for depression, no evidence for screening in high-risk populations and no evidence of the harms of screening. Randomized controlled trials comparing screening and no screening should help to clarify these issues. Future research must have a broader demographic, geographic and cultural scope. Trials on the effectiveness of screening among people who are at increased risk of major depressive disorder are also needed to help in the early diagnosis and treatment of those most likely to be affected by depression. More evidence is needed on the harms of screening for depression (e.g., false positive rates) and the related potential for unnecessary, and possibly harmful, diagnostic and therapeutic procedures. Finally, more research on the most effective method of screening for depression in relation to clinically important outcomes is needed in populations with increased risk of depression.

References

1. Patten SB, Wang JL, Williams JV, et al. Descriptive epidemiology of major depression in Canada. *Can J Psychiatry* 2006;51:84-90.
2. Vasiliadis HM, Lesage A, Adair C, et al. Do Canada and the United States differ in prevalence of depression and utilization of services? *Psychiatr Serv* 2007; 58:63-71.
3. MacMillan HL, Patterson CJ, Wathen CN, et al. Screening for depression in primary care: recommendation statement from the Canadian Task Force on Preventive Health Care. *CMAJ* 2005;172:33-5.
4. Canadian Task Force on Preventive Health Care, Joffres M, Jaramillo A, Dickinson J, et al. Recommendations on screening for depression in adults. *CMAJ* 2013;185:775-82.
5. Sartorius N, Ustun TB, Lecrubier Y, et al. Depression comorbid with anxiety: results from the WHO study on psychological disorders in primary health care. *Br J Psychiatry Suppl* 1996;Jun:38-43.
6. National Collaborating Centre for Mental Health. *Depression: management of depression in primary and secondary care* [NICE Clinical Guidelines, no. 23]. London (UK): National Institute for Health and Clinical Excellence; 2004. Rep. no.: CG90.
7. National Collaborating Centre for Mental Health. *Depression. The treatment and management of depression in adults (updated edition)* [Clinical Guideline 90]. London (UK): National Institute for Health and Clinical Excellence; 2009. Available: http://publications.nice.org.uk/depression-in-adults-cg90 (accessed 2013 Dec. 16).
8. Palmer SC, Coyne JC. Screening for depression in medical care: pitfalls, alternatives, and revised priorities. *J Psychosom Res* 2003;54:279-87.
9. O'Connor EA, Whitlock EP, Gaynes BN, et al. *Screening for depression in adults and older adults in primary care: an updated systematic review* [Evidence synthesis no. 75. AHRQ report no. 10-05143-EF-1]. Rockville (MD): Agency for Healthcare Research and Quality; 2009.
10. Pignone MP, Gaynes BN, Rushton JL, 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.
11. Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. *Cochrane Database Syst Rev* 2005;(4):CD002792.
12. Patten SB, Juby H. *A profile of clinical depression in Canada*. Calgary (AB): Department of Community Health Sciences and Psychiatry, University of Calgary; 2008. Available: http://hdl.handle.net/1880/46327 (accessed 2013 Apr. 18).
13. Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 ed. New York (NY): Cochrane Collaboration and Wiley-Blackwell; 2008 [updated 2011].
14. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; (336):924-6.
15. Shaffer HJ, Vander Bilt J, Hall MN. Gambling, drinking, smoking and other health risk activities among casino employees. *Am J Ind Med* 1999;36:365-78.
16. Wells G, Shea B, O'Connell D, et al. *The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Ottawa (ON): Ottawa Hospital Research Institute; 2000.
17. Oyama H, Koida J, Sakashita T, et al. Community-based prevention for suicide in elderly by depression screening and followup. *Community Ment Health J* 2004;40:249-63.
18. Oyama H, Ono Y, Watanabe N, et al. Local community intervention through depression screening and group activity for elderly suicide prevention. *Psychiatry Clin Neurosci* 2006;60:110-4.
19. Oyama H, Fujita M, Goto M, et al. Outcomes of community-based screening for depression and suicide prevention among Japanese elders. *Gerontologist* 2006; 46:821-6.

20. Oyama H, Goto M, Fujita M, et al. Preventing elderly suicide through primary care by community-based screening for depression in rural Japan. *Crisis* 2006; 27:58-65.
21. Oyama H, Sakashita T, Hojo K, et al. A community-based survey and screening for depression in the elderly: The short-term effect on suicide risk in Japan. *Crisis* 2010;31:100-8.
22. Hoyl MT, Alessi CA, Harker JO, et al. Development and testing of a five-item version of the geriatric depression scale. *J Am Geriatr Soc* 1999;47:873-8.
23. Zung WW. A self-rating depression scale. *Arch Gen Psychiatry* 1965;12:63-70.
24. Kessler RC, Ustun TB. The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res* 2004;2:93-121.
25. Gastpar M. The ICD-9 and the SADD-criteria for depression. *Acta Psychiatrica Scand Suppl* 1983;310:31-41.
26. Shah A. A replication of the relationship between elderly suicides rates and elderly dependency ratios: a cross-national study. *J Inj Violence Res* 2010;2:19-24.
27. O'Connor EA, Whitlock EP, Beil TL, et al. Screening for depression in adult patients in primary care settings: a systematic evidence review. *Ann Intern Med* 2009;151:793-803.
28. National Collaborating Center for Mental Health. *The NICE guideline on the management and treatment of depression in adults (updated edition)* [National Clinical Practice Guideline 90]. Leicester (UK): British Psychological Society; 2010. Available: www.nice.org.uk/nicemedia/live/12329/45896/45896.pdf (accessed 2013 Dec. 16).
29. Scottish Intercollegiate Guidelines Network (SIGN). *Management of perinatal mood disorders* [Report no. 127]. Edinburgh (Scotland): SIGN; 2012.
30. National Collaborating Centre for Mental Health. *Depression in adults with a chronic physical health problem: treatment and management* [NICE Clinical Guidelines, no. 91]. London (UK): National Institute for Health and Care Excellence; 2009.
31. National Collaborating Centre for Mental Health. *Antenatal and postnatal mental health: clinical management and service guidance* [NICE Clinical Guidelines, no. 45]. London (UK): National Institute for Health and Clinical Excellence; 2007.
32. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of metaanalyses? *J Clin Epidemiol* 2000;53:964-72.
33. Wakefield JC, Schmitz MF, Baer JC. Did narrowing the major depression bereavement exclusion from DSM-III-R to DSM-IV increase validity? Evidence from the National Comorbidity Survey. *J Nerv Ment Dis* 2011;199:66-73.

Affiliations: McMaster Evidence Review and Synthesis Centre and Department of Clinical Epidemiology and Biostatistics (Keshavarz, Rice, Ali, Shannon, Raina); McMaster Evidence Review and Synthesis Centre and School of Nursing (Fitzpatrick-Lewis); Department of Psychiatry and Behavioural Neurosciences (Streiner), McMaster University, Hamilton, Ont.; Department of Psychiatry (Streiner), University of Toronto, Toronto, Ont.

Contributors: Homa Keshavarz, Donna Fitzpatrick-Lewis, David Streiner, Maureen Rice, Usman Ali, and Parminder Raina were involved in conceptualizing the study, analyzing the data and writing the manuscript. Harry Shannon contributed to analyzing the data and provided manuscript revisions. Homa Keshavarz and Donna Fitzpatrick-Lewis were also involved in designing the methods, collecting the data and coordinating the project. Maureen Rice performed the literature search. All of the authors approved the final manuscript.

Funding: This review was funded by the Canadian Institutes of Health Research. The views expressed herein are the opinions of the authors and do not necessarily represent the views of the Canadian Institutes of Health Research.

Acknowledgements: The authors thank Catherine Salmon, Sara Kaffashian and Robert Stevens for reviewing the titles and abstracts. Leah MacDonald formatted the large evidence review, as well as the manuscript submitted for publication. Members of the Canadian Task Force on Preventive Health Care, including Michel Joffres, chairperson of the Depression Screening Guideline Working Group, contributed to protocol development and/or discussion of findings. Lesley Dunfield and Alejandra Jaramillo, Public Health Agency of Canada officers, were involved with performing the review and creating the guideline. Parminder Raina holds a Tier 1 Canada Research Chair in Geroscience and the Raymond and Margaret Labarge Chair in Research and Knowledge Application for Optimal Aging.

Supplemental information: For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/1/4/E159/suppl/DC1

Editor's note: This paper is based on a full systematic review entitled "Screening for depression," which can be found at: http://canadiantaskforce.ca/guidelines/depression/